

Chapter 12

Facebook and Google as Offices of Censorship



On May 15, 2018, Facebook continued its springtime campaign to restore its reputation in the aftermath of the Cambridge Analytica scandal. A “Transparency” report was published, which included statistics on the extent of content removal, organized by category. Let’s look at for instance the category “Graphic Violence”: “In Q1 2018, we took action on a total of 3.4 million pieces of content, an increase from 1.2 million pieces of content in Q4 2017. This increase is mostly due to improvements in our detection technology, including using photo-matching to cover with warnings photos that matched ones we previously marked as disturbing. These actions were responsible for around 70% of the increase in Q1.”¹ The numbers may seem high, but they only tell half the story. Another graph in the report shows that 71.56% of the 1.2 million users were tracked by Facebook itself, until user complaints started flooding in; in the first quarter of 2018, this figure rose to 85.6%. The fact that the number of removals tripled means that content removed because of user complaints rose from 341,000 to almost 500,000, in absolute figures, despite the decrease in percentage. So, the increase can be attributed not only to better tracking equipment, but also to more complaints favored. These numbers are a testimony to content removal on a disproportionately large scale, also known as censorship.

¹ Facebook “Community Standards Enforcement Preliminary Report.”

In other categories, the numbers are even higher. The category “Pornographic nudity and sexual activity” remains constant over the two quarters: 21 million posts censored in each quarter. “Terrorist propaganda” increased from 1.1 to 1.9 million cases, out of which 99.5% were removed before even appearing, that is, as acts of pre-censorship. “Hate speech” went up from 1.6 to 2.5 million cases over the course of the two quarters. Out of these cases, only 23.6% and 38%, respectively, were found by Facebook itself. The majority of these were identified by flagging users, so the company goes to great lengths to accommodate users’ sense of violation. In those Q4 and Q1, respectively, 727 and 936 million cases of spam were deleted, while 694 and 583 million false accounts were shut down. The total number of posts removed increased over the two quarters, rounding a billion, which amounts to more than 10 million per day—the vast majority of them spam.

Given the speed of the procedure, some questions should be asked: How accurate or indicative can these numbers be? Is there a reason to believe that all the removals can be attributed to someone actually noticing the alleged norm-breaking content? And if not, do some—or maybe even many—removals happen as a result of accusation alone, that is, without going through the actual content? Of the content-based complaint categories, nudity-and-sex is the most frequent one. This could explain why in some cases the category seems most liable to be used politically by users to have their opponents silenced. There seems to be systematic use of the complaints option. If a group of people agree to complain against someone voicing something, it seems fairly easy to have that person thrown off of Facebook. It also seems that the plausibility of the complaint filed is not always given the highest attention.

In 2016, the Council of Ex-Muslims of Britain claimed that 19 different Facebook groups or sites organized by Arabic ex-Muslims or freethinkers had either already been shut down or underwent attacks via organized abuse of the

flagging system.² Thus, it seems that Islamist groups (or even governments in the Middle East?) use the flagging system, in an organized manner, in order to remove democratic Muslim or anti-Islamist sites from Facebook. In the conservative online magazine *American Thinker*, it has been claimed that such shutdowns often happen in the following way: Massive complaints of pornography are filed by many complainants at the same time against a given page, which is then shut down.³ From the look of it, the reason is that sheer number of complaints is taken as an indication of the complaint's justification, and/or that the pressure on the staff is so high that not all cases can be properly handled. Among the deleted accounts in 2013 were "Ban Islam"; "Islam Against Women"; "Islam Free Planet". The interesting thing is that the majority of pages hit in this way do not contain pornography at all, since they are in fact politico-religious pages. Experience seems to suggest that sex complaints are easily accepted, so that large amounts of complaints almost automatically will trigger the blocking of the targeted Facebook page, with no review of whether there is even sex on the page. Such abuse may comprise anything from spontaneous actions to systematic flagging of political opponents, and such cases are obviously invisible to the Facebook stats, where systematic weeding out of democratic voices in the Middle East is then represented in the stats simply as removed pornography.

No one knows the extent of coordinated abuse of the flagging feature. Gillespie mentions cases like "Operation Smackdown", organized by a group of YouTube users to attack pro-Muslim content on the platform by complaining against it for featuring acts of terrorism. The attack was orchestrated with a long list of videos to target, detailed instructions on how to file complaints and a Twitter account

²"Facebook: Stop Censoring Arab Ex-Muslims and Freethinkers NOW" *Council of Ex-Muslims of Britain*. 02-20-16.

³Murphy, P. "Blasphemy Law Comes to Facebook" *American Thinker*. 06-27-13.

featuring the dates on which the videos were to be attacked. This operation was active from 2007 to 2011.⁴ Obviously, surrendering an important part of the removal process to the users' own reporting activity is dangerous, since user groups can abuse the feature to foment their own agendas. We have not been able to find clear estimates of how widespread this low-intensity online culture war is. Notes Tarleton Gillespie: "There is evidence that strategic 'flagging' has occurred and suspicions that it has occurred widely."⁵

Thanks to the flagging system, Facebook's own removal reports may thus hide censorship and let the company off the hook. Evidence suggests that Facebook's current set of rules and statistics does not contain the whole truth. In many cases, the enforcement of the policy is not consistent with equality before the law — sometimes criticism of Islam is removed with greater enthusiasm than, for example, anti-Semitism or criticism of the state of Israel. In 2016, New York-based Jewish website *the algemeiner* quoted Amos Yadlin, former chief of Israel Defense Forces, an Israeli intelligence service, for saying that "The most dangerous nation in the Middle East acting against Israel is the state of Facebook." Yadlin, who now heads the Institute for National Security Studies in Tel Aviv, continued: "It has a lot more power than anybody who's operating an armed force. Unlike before, there's no longer an existential military threat facing Israel. Rather, it's a strategic threat."⁶ Since then, Facebook and Israel seem to have reached an agreement to remove "incitement" from the platform, but the details of the agreement are not known to the public.⁷ However, as mentioned, there is no reason to expect that the many removals taking place will remain consistent,

⁴Gillespie (2018) p. 92.

⁵Ibid.

⁶Sherman, E. "Ex-IDF Intel Chief: 'State of Facebook' Greatest Mideast Threat to Israel" *the algemeiner*. 01-31-16; translated into English from the Hebrew website *nrg*.

⁷Kaye, D. "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression" *UN Human Rights Council*. 04-06-18. Chapter 20.

and certainly not over time, as Facebook may be easily influenced by lobbyists, campaigns and pressure from both Israeli and Arab sides.

Similarly, Catholic associations in the US have complained about their Facebook accounts being shut down. Facebook probably did not take a classic and uncomfortable fact from the history of religion into account: that many of the large religions practice, as a natural and central custom, insults, mockery and ridicule of other religions, or worse: they may have a strong tradition for calls to violence against the followers of other religions or against infidels—sometimes such practices even take place in the sacred texts of certain religions.

There is an increasing number of cases where Facebook in fact removes seemingly legitimate political views, such as support for Russia or support for Trump’s more unusual bills. In January 2018, Uffe Gardel, a Danish Eastern Europe journalist, reported on a peculiar experience. He describes it: “I participated in a passionate debate on my own Facebook page: the topic was the Russia-backed war in Eastern Ukraine. We participated around five users, all of us Danes: two pro-Russian views and three pro-Ukrainian views. We debated in a lively and matter-of-fact way. Suddenly, not a word came from one of the pro-Russian participants. He did not respond when addressed. Not a word from him, and moreover his previous posts were suddenly gone.”⁸ Gardel was surprised that his debate opponent Jesper Larsen suddenly withdrew from the debate. When he returned, Larsen wrote that Facebook had informed him that his posts had been deleted as spam. A new test post from him was deleted in a matter of seconds only. However, it was not spam, but a short comment featuring a link to Ukrainian television. Could it be thinkable that Facebook had begun removing pro-Russian posts? Maybe after the ongoing Russian bot campaigns interfering in American politics had become known?

⁸Gardel, U. “Når Facebook censurerer” *Journalisten*. 01-11-18.

Gardel contacted the Danish branch of Facebook, whose representative Peter Andreas Münster explained: “The point here is that ‘real’ people can easily risk triggering our anti-spam systems if they post stuff very frequently and very quickly.” No information was provided on whether the removal was influenced by user complaints. What also remains unclear is whether Jesper Larsen had posted hyperactively, and whether Facebook’s explanation is trustworthy, given the fact that Facebook is the only source of this information. As Gardel adds, this is not the only recent case of political content leading to deletion. He quotes Danish writer and debater Suzanne Bjerrehuus, who was sanctioned with a three-day quarantine from Facebook that same winter. She had posted the following comment on a series of gang rapes in the Swedish city of Malmö: “Brutal and abhorrent violence and then they get away with it. The police are powerless. [...] The Swedes ought to break with those politicians who have ruined Sweden.” Facebook’s “hate speech” clause was only made public a couple of months later—but at the time, Bjerrehus received the explanation that her post was in breach of the company’s ban on “posts attacking people based on race, ethnic background, national origin, religious affiliation, sexual orientation, gender or disability.” The many separate problems of this clause aside, it is in fact peculiar that her opinion should fall within the scope of that clause. Gardel rightly states that one needs not agree with Larsen’s or Bjerrehuus’ views in order to find the removal of their statements extraordinarily problematic. He concludes that tech giants such as “[...] Facebook have gained such a strong position that regulation is needed. A still increasing proportion of the Danish debate on public matters is now taking place on Facebook. Facebook pages become actual media, which are then enrolled in the Danish Media Ethical Commission. In some cases, established web media use Facebook’s debate forums to control user comments; currently, this is what’s happening with the newspapers of media outlet Syddanske Medier. These media organizations end up in fact leaving parts of their editing rights in the hands of

Facebook, and this alone should alarm everyone in the publishing industry.” Concludes Gardel admonishingly: “In any case, it must now be clear that we cannot have both a safe Internet and a free network. And it’s an old truth that he who gives up freedom for security is at risk of losing both.” We support this outcry—playing on Benjamin Franklin’s classic words—to the fullest.

There is much to suggest that content with different political motivations is removed. Stories and documentation abound online of strange omissions, excessive removal and inconsistencies in Facebook’s censorship.⁹ However, it should come as no surprise that the removal does not have the same consistency as a court bound by precedence, given that the control is so speedy, comprehensive, reckless and carried out by legally untrained employees. At the congressional hearing in April 2018, Senator Ted Cruz (R) was critical, as he himself had experienced the tendency of Facebook to remove conservative more eagerly than liberal content, more republican than democratic. However, just because content critical of Islam is sometimes removed and content critical of Christianity is not, it does not follow that such a bias is systematic and a sign of double standards. Given the vast amount of content deletions, one does not exclude the other. Only whistleblowing or deep statistical surveys would be able to uncover explicit or implicit double standards. In December 2015, Israeli NGO Shurat HaDin¹⁰ did a little experiment: They created two parallel Facebook pages entitled *Stop Israel!* and *Stop the Palestinians!* with identical setups, designs and rhetoric. Facebook shut down the anti-Palestinian page, but not the anti-Israeli one. The organization has since sued Facebook.

⁹See e.g. Tobin, A. & Varner, M. & Angwin, J. “Facebook’s Uneven Enforcement of Hate Speech Rules Allows Vile Posts to Stay Up”, *ProPublica*. 12-28-17.

¹⁰Melnick, O. “Facebook’s Hate Speech Double Standard” *WND*. 01-11-16.

In 2016, technology site *Gizmodo* featured an article¹¹ based on statements from former Facebook employees who claimed that employees who edited incoming news content for the Facebook “Trending” column routinely removed conservative news, e.g. news on Republicans Ron Paul and Mitt Romney. The column claims that it algorithmically reflects “topics that have recently become popular on Facebook.” But several of Facebook’s former news curators, as they were called in the organization, also told *Gizmodo* that they were instructed to artificially “inject” stories into the trending news feed, even though they were not popular enough to even be there—in some cases the stories had no following whatsoever. These former curators, who all worked on contract, also said that they were told not to include news about Facebook, not even in the trending feature. An anonymous former employee kept a protocol of news stories that were buried in this way. *Gizmodo* therefore concluded that Trending on Facebook works like a plain opinion-based newspaper, except it maintains a surface of neutrality. Top management at Facebook rejected all these allegations as false.

Nevertheless, the allegations about what was happening on Trending Topics seem to have affected Facebook. As early as January 2015, the company had announced a campaign against the volume of “fake news” that abounded in the column. In August 2016, shortly after the revelation in *Gizmodo*, Facebook dismissed the 26 editors who had fed the news column and replaced them with an automatic algorithm. The algorithm would ensure that the news stories featured also reflected their actual popularity on the platform. However, this step did nothing short of opening the floodgates for viral spread of false news. Apparently, the company had overestimated ability or willingness of users to identify and reject false news. Just two days after the new algorithm was put to use, a false story about a Fox News journalist made it high up on the list: “Breaking: Fox News Exposes Traitor Megyn

¹¹Nunez, M. “Former Facebook Workers: We Routinely Suppressed Conservative News” *Gizmodo*. 05-09-16.

Kelly, Kicks Her Out for Backing Hillary.”¹² According to a *Washington Post* survey covering a three-week period in September 2016, five fake and three highly misleading news stories ranked high on the Trending Topics section of four different Facebook accounts (of course, there may have been even more on other accounts because of the personalization of each account). In January 2018, in the aftermath of the chronic problems that had turned Facebook into a main supplier of “fake news” during the presidential campaign, the company attempted to shift the news feed balance from journalistic news to local news from “friends.” In June, a further step was taken when Facebook announced the complete elimination of Trending Topics.¹³

At the same time, the company met new problems due to a policy introduced on May 24, 2018. The policy gives political ads a special label and collects such ads in a separate archive containing information on their ad budget, number of users who have seen them, etc.¹⁴ This goes for ads related to candidates and elections, but also political issues such as “abortion, arms, immigration and foreign policy.” The intention was, of course, to increase transparency around political ads. However, newspapers and media associations, led by *New York Times*, protested fiercely over the fact that their articles *about* politics were given the same categorization and labeling on Facebook, the “political ad” warning. Media representatives argued that since the media pays to have such articles promoted as a way of selling their own product, Facebook must respect the boundary between political ads on the one hand and ads for quality journalism *about* politics on the other, instead of trying to erase it. After this, *New York Times* and other leading media stopped paying to place their content on the platform. At the same time, reports came out

¹²Solon, O. “In firing human editors, Facebook has lost the fight against fake news” *The Guardian*. 08-29-16.

¹³Kastrenakes, J. “Facebook will remove the Trending topics section next week” *The Verge*. 06-01-18.

¹⁴Constine, J. “Facebook and Instagram launch US political ad labeling and archive” *TechCrunch*. 05-24-18.

showing that people had less confidence in news coming from social media than from all other media, and that news consumption via Facebook was declining (hardly surprising in light of the suppression of “real” news earlier that year).¹⁵ CEO of *New York Times* Mark Thompson called Facebook’s categorization a “threat to democracy.” In an angry debate, he accused Campbell Brown, Head of Global News Partnerships at Facebook, of supporting the enemies of quality journalism.¹⁶ It is rather ironic that *real* news was turned into political ads as the result of an attempt to make political ads explicit and the extent of them public — thus hoping to eradicating “dark ads” in the form of targeted political ads visible only to their receiver. The case also shows Facebook’s ongoing conflict with the media. Despite its many attempts at forming alliances, by categorizing journalism as ads Facebook unilaterally launched a new and secretly developed policy, without having consulted their supposed media allies beforehand. In July 2018, researchers from New York University demonstrated that from May to July, the first two months of the new ads archive, Facebook’s largest political advertising client was ... Donald Trump.¹⁷

Only a few days later, another scandal broke: Zuckerberg “accepted” Holocaust denial and claimed that it “deserved” its place on the platform. This made him the target of a veritable shitstorm in both the offline and online media. However, he had not used these words. He was interviewed for an hour and a half by Kara Swisher of *recode* on the topic of Facebook’s “annus horribilis”;¹⁸ an interview unsurprisingly circling around news, “fake news”, disinformation, etc. The irony is that in the interview, Zuckerberg goes to great lengths to defend free speech on his platform: “There are

¹⁵“Digital News Report 2018.” Last visited 07-30-18: <http://www.digitalnewsreport.org>.

¹⁶Moses, L. “How The New York Times’ Mark Thompson became the latest thorn in Facebook’s side” *DigiDay*. 07-11-18.

¹⁷Frenkel, S. “The Biggest Spender of Political Ads on Facebook? President Trump” *New York Times*. 07-17-18.

¹⁸Swisher, K. “Zuckerberg: The Recode interview” *Recode*. 07-18-18.

really two core principles at play here. There's giving people a voice, so that people can express their opinions. Then, there's keeping the community safe, which I think is really important. We're not gonna let people plan violence or attack each other or do bad things. Within this, those principles have real trade-offs and real tug on each other." If what is meant by "attack" is real, violent attack, there is hardly a single free speech supporter out there who will disagree with this—parallel to the limits to freedom of expression drawn at "incitement to imminent lawless action".¹⁹ In the following sentence, however, Zuckerberg changes course: "In this case, we feel like our responsibility is to prevent hoaxes from going viral and being widely distributed." He then goes on to present the news that verifiably "fake news" must be downgraded in the news feed, but not removed from it. The journalist does not comment on Zuckerberg's mix-up of misinformation and planning violence but instead asks him why "fake news" should be downgraded and not simply eliminated entirely. To this, Zuckerberg again defends freedom of expression: "... [A]s abhorrent as some of this content can be, I do think that it gets down to this principle of giving people a voice." This prompted the journalist to give an example of a post she felt should just plainly be removed: the claim that "Sandy Hook never happened" (a tragic school shooting in 2012, which later became the subject of an *InfoWars* conspiracy theory claiming that the event never took place but was staged by

¹⁹The classic American phrase from 1919 famously states that the limit is "clear and present danger" (from iconic Supreme Court Justice Oliver Wendell Holmes' turn of phrase in *Schenk v the US*). The current phrase goes like this: "[...] the constitutional guarantees of free speech and free press do not permit a State to forbid or proscribe advocacy of the use of force or of law violation except where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action" – a quote from the *Brandenburg v Ohio* case of 1969. In the case *Hess v Indiana*, the Supreme Court made clear that unless statements made "... were intended to produce, and likely to produce, imminent disorder, those words could not be punished by the State on the ground that they had a 'tendency to lead to violence.'" Cf. Freedom Forum Institute: "Incitement to Imminent Lawless Action": 05-12-08.

anti-gun activists). Zuckerberg defended why this conspiracy theory was not removed from Facebook — and then went on to compare it to the Holocaust: “I’m Jewish, and there’s a set of people who deny that the Holocaust happened. I find that deeply offensive. But at the end of the day, I don’t believe that our platform should take that down because I think there are things that different people get wrong. I don’t think that they’re *intentionally* getting it wrong, but I think ..” Then he is interrupted by the journalist. He continues: “It’s hard to impugn intent and to understand the intent.” He is certainly right about that — but he is just as certainly wrong in claiming that Holocaust deniers innocently make a mistake the same way he himself could make a mistake publicly. Deflating Holocaust deniers’ motives in this way caused a scandal, and the day after he had to pull back: “I personally find Holocaust denial deeply offensive, and I absolutely didn’t intend to defend the intent of people who deny that.”²⁰

In the midst of this media shitstorm, many people thought it obvious that conspiracy theories such as the one about Sandy Hook should be deleted as a matter of routine. But these people were never able to come up with a clear principle as to where to draw the line between conspiracy theories, lies, false statements, satire, irony, quotations, and random mistakes, and how such a line should then be monitored. Although Zuckerberg went quite far to defend freedom of expression, he expressed himself in covert and unclear fashion, confusing violent attacks with false statements and presenting a half-baked theory that people’s sincerity—which is not easy to measure—should act as thermometer to assess whether statements should be allowed, downgraded, or altogether removed. It is not comforting that a man in his position is unable to express himself more clearly, and it calls for a reminder of a classic warning: hazy words cover up hazy thoughts.

²⁰Swisher, K. “Mark Zuckerberg clarifies: ‘I personally find Holocaust denial deeply offensive, and I absolutely didn’t intend to defend the intent of people who deny that.’” *recode*. 07-18-18.

The defense of free speech did not last long, however. On August 6, 2018, after weeks of heated public discussion, Facebook blocked four accounts belonging to Alt-right talk show host Alex Jones and his *InfoWars* podcast shows. Apple was the first tech giant to block *InfoWars* and was quickly followed by Facebook, YouTube, Pinterest and even YouPorn, all on the same day in what may resemble a coordinated action. This is probably the biggest act of online censorship to date—Jones had 1.4 million followers on Facebook and 2.5 million on YouTube. In a public statement, Facebook argued that Jones’ pages were removed for “glorifying violence, which violates our graphic violence policy, and using dehumanizing language to describe people who are transgender, Muslims, and immigrants, which violates our hate speech policies.”²¹ The banning process is ugly and devoid of principles: as usual, no one is given information on exactly which statements are deemed unacceptable. Only a few weeks earlier, Zuckerberg had even defended Jones’ presence. Jones was punished for accumulating “too many strikes”; but there is nothing about how many that is, and which strike was the final blow. If Jones really fell under the scope of Facebook’s “hate speech” policy, why had it not happened before? For years he had presented his huge audience with grotesque opinions on the platform.²² The other tech giants made similar references to “hate speech” rules. Rumor spread fast that Apple’s action made the other tech giants follow suit because Apple threatened to throw them out of its App Store, which has strict regulations. Commentator Brian Feldman wrote: “What the *InfoWars* decisions represent is a capitulation—not to censors, not to the public, not to the deep state, but to the only entity left that has any real power over Facebook

²¹Shaban, H., Timberg, C. and Stanley-Becker, I.: “YouTube, Apple, Facebook and Spotify escalate enforcement against Alex Jones” *Washington Post*. 08-06-18.

²²Facebook’s statement on the Alex Jones case: “Enforcing our community standards”: 08-06-18.

and YouTube: Apple.”²³ The exception was Twitter. For eight days they hesitated, until finally blocking *InfoWars*. But it was only a week long suspension and it was not for “hate speech”, but more explicitly for encouraging violence, as Jones had urged his followers to have their “battle rifles” ready to fight mainstream media.²⁴

Commentators point out that Jones and his supporters will see the ban as evidence of their claim of a coordinated political attack against them—and that it will only strengthen Jones’ position as a right-wing martyr.²⁵ Jones and his followers will most likely regroup in underground networks, separated from the general public. Predictably, Jones was infuriated by the removal: “We’ve seen a giant yellow journalism campaign with thousands and thousands of articles for weeks, for months misrepresenting what I’ve said and done to set the precedent to de-platform me before Big Tech and the Democratic Party as well as some Republican establishment types move against the First Amendment in this country as we know it.”²⁶ Jones is especially known for spreading provable untruths. Most famous were “Pizzagate”, the claim that Hillary Clinton ran a pedophile ring out of a Washington pizzeria, or Jones’ assertion that the Sandy Hook school shooting in 2012 never took place. Jones went on to harass parents of children killed at the massacre—not to mention spreading fake conspiracy theories about teenagers who survived the Parkland school shooting in Florida of February 2018. Recently, he has tried to convince the public that the Democrats wanted to start a civil war on the 4th of July.

²³Feldman, B. “The Only Pressure Facebook Understands Comes from its Megaplatform Rivals” *New York Magazine*. 08-06-18.

²⁴Kang, C. and Conger, K. “Twitter Suspends Alex Jones and Infowars for Seven Days” *New York Times*. 08-14-18.

²⁵See e.g. Lapowsky, I.: “Why Big Tech’s Fight Against InfoWars is Unwinnable” *Wired*. 08-06-18.

²⁶Shaban, H., Timberg, C. and Stanley-Becker, I.: “YouTube, Apple, Facebook and Spotify escalate enforcement against Alex Jones” *Washington Post*. 08-06-18.

There is no doubt that Jones has repeatedly peddled abominable lies, false accusations and bizarre conspiracy theories. Still, Facebook's argument for banning Jones is not based on his "fake news" but on the murkier concept of "hate speech"—presumably in an attempt to avoid taking the seat of judge between true and false. As stated by several observers, nevertheless, "hate speech" is not only a vague, politicized and subjective category. It is also full of double standards, because it does not equally protect all groups defined by race, gender, religion and so on.²⁷ It is an all-purpose category with no clear limits, so it can be stretched to accuse points of view that are simply not liked. As pointed out by Robby Soave, no one will miss *InfoWars*—the serious issue raised by this event is that completely unclear rules and procedures now govern the removal of content on the giants' platforms.²⁸ The Jones case seems to be triggered by public pressure, and one thing remains particularly unclear: are there also plans to crack down on other right-wing extremists with similar views but fewer supporters operating on the same platforms out of the public eye? There is no shortage of those. Perhaps a less problematic cure against characters like Jones would be to bring the removal criteria closer to existing US legislation. That would make it possible to intervene, assisted by proper authorities, against clearly illegal acts such as slander, libel, threats and harassment. In Jones' rhetoric alone, there is more than enough of these.²⁹

²⁷ Shapiro, B. "What Tech Giants' Alex Jones Ban Got Wrong" *National Review*. 08-07-18.

The fact that Facebook now express a wish to be inspired by the International Covenant on Civil and Political Rights (ICCPR) from the 1960s is puzzling, inasmuch as that law helped many countries introduce hate speech laws (see Chapter 11) and the United States was among the countries that chose *not* to comply with ICCPR.

²⁸ Soave, R. "Banning Alex Jones" *Reason*. 08-07-18.

²⁹ Cf. French, D. "A Better Way to Ban Alex Jones. *New York Times*. 08-07-18.

All in all, the presentation of news through Facebook has been characterized by recurring problems with “fake news”; mixing up news with ads, political bias and content deletion, not to mention improvised reactions to public and political pressure. Various remedy initiatives have not produced any successful cure, neither of a human nor algorithmic form (see Ch. 11 for a discussion of fact checkers).

The secretive, opaque and shifting removal procedures obviously make tech giants subject to political pressure from international top players who wish to influence the removal policy—not to mention journalistic Kremlinology trying to interpret what is really going on behind the scenes, based on small signs, rumors and stand-alone issues. When Zuckerberg met with Angela Merkel in Berlin during the European migrant crisis of 2015, she apparently encouraged him to crack down harder on “hate speech”, to which he is said to have made the following response: “Yeah.” This was interpreted by some media as Facebook committing itself to suppressing critical news coverage of migrants in Europe.³⁰

The spring of 2019 was characterized by an increasing effort to censor different types of Facebook content, particularly “fake news” and “hate speech.” In March, after repeated criticism from journalists and lawmakers, Facebook announced that it was diminishing the reach of anti-vaccine posts.³¹ Later the same month, Facebook announced it would ban white nationalist content.³² These developments clearly indicate the ad hoc character of the company’s removal policy, without clear principles.

A radical change in the overall Facebook vision was announced by Mark Zuckerberg March 6th: “As I think about the future of the internet, I believe a privacy-focused communications platform will become even more important than

³⁰David, J. E. “Angela Merkel caught on hot mic griping to Facebook CEO over anti-immigrant posts” *CNBC*. 09-27-15.

³¹Matsakis, L. “Facebook will crack down on anti-vaccine content” *Wired*. 03-07-19.

³²Stack, L. “Facebook announces new polity to ban white nationalist content” *New York Times*. 03-27-19.

today's open platforms. Privacy gives people the freedom to be themselves and connect more naturally, which is why we build social networks.”³³ Zuckerberg defined privacy in terms of six headlines: privacy of interactions in selected communication types; end-to-end encryption extending from WhatsApp over the whole platform; extended possibility for posting information for shorter periods of time only; increased safety; interoperability in the sense of communication ability across Facebook's different platforms; increased protection of data in countries violating human rights. This general declaration of intent, of course, is an attempt to preempt oncoming government regulation. Simultaneously, it is striking that the core business model seems all but untouched by the new principles—nothing is said about data sharing and ad targeting.³⁴ Late March, however, Zuckerberg gave in to looming regulation in a surprising op-ed in *Washington Post* where he called for some sort of external regulation. In the face of increased political pressure, Zuckerberg now chose the tactics of delimiting regulation to four specific areas: harmful content, election integrity, privacy and data portability. “I've come to believe that we shouldn't make so many important decisions about speech on our own. So we're creating an independent body so people can appeal our decisions,” he said.³⁵ The degree of independence of a voluntary Facebook-invented body will of course be a matter of contention. Zuckerberg, however, also envisions some kind of cross-platform authority to standardize removal practices over the internet at large. Again, Facebook's sudden willingness to abandon responsibility can be seen as a preemptive move as against threatening anti-trust initiatives.

Only a couple of weeks later, Facebook announced a series of new measures to ensure “integrity” in its much-debated

³³Zuckerberg, M. “A Privacy-Focused Vision for Social Networking” *Facebook*. 03-06-19.

³⁴Lapowsky, I & Thompson, N. “Facebook's pivot to privacy is missing something crucial” *Wired*. 03-06-19.

³⁵Zuckerberg, M. “Mark Zuckerberg: The Internet needs new rules. Let's start in these four areas” *Washington Post*. 03-30-19.

news feed—of which the most important device was the so-called Click-Gap.³⁶ It will influence the determination of the ranking of a given post in the feed. The idea is to limit the dissemination of websites which are deemed disproportionately viral on Facebook in comparison with the net as a whole. A given news content of that sort will be limited in reach on Facebook. The controversial and contested news feed is sought domesticated by a conformity measure making it a mirror of the average traffic on the internet. This means that it provides no security against viral matters which are popular also outside of Facebook. During little than one month, Facebook announced a handful of new initiatives, most of all giving evidence of increasing panic in the head office of 1 Hacker Way.

Facebook has always had a very comprehensive removal policy. By contrast, Google has held the free speech banner quite high. For example, in 2010 Google pulled out of China after several years of hacking attempts and pressure to enact censorship. But Google has also been accused of outright censorship. In 2016, Robert Epstein, a professor of psychology and Google critic, came up with an overview of at least nine different blacklists at work in Google's content filtering.³⁷ Especially Epstein's first blacklist is relevant—it

³⁶Dreyfuss, E. & Lapowsky, I. "Facebook is changing news feed (again) to stop fake news" *Wired*. 04-10-19.

³⁷Epstein's nine blacklists read as follows:

1. The autocomplete blacklist, which automatically blocks the guesses that follow when certain keywords are entered.
2. The Google Maps blacklist—maps with disputed geographic areas, which without explanation are not shown—military zones, wealthy people who paid to have their land exempt.
3. The YouTube blacklist—for instance in Pakistan featuring what the government demands to be removed from the platform.
4. The Google Account blacklist—blocking users who have not complied with the "Terms of Service," which can typically be terminated "at any time" and with no real appeal option.
5. The Google News blacklist, which has been accused of leaving out news critical of Islam (see below).
6. The Google AdWords blacklist—certain words cannot appear in ads, or entire industries whose ads Google does want on the platform.

concerns the autocomplete feature, which was introduced in 2008 and which works by completing entered keywords with a variety of suggestions generated by an algorithm. This feature blocks, for example, obscene words. But Epstein also found political effects in the auto-complete feature. For example, when writing the word “Lying” during the American election campaign in 2016, what followed was “Ted” (Trump’s nickname for Ted Cruz, “Lyn’ Ted”), but when writing “Crooked,” what was then suggested was *not* “Hillary” (Trump’s nickname for Hillary Clinton, “Crooked Hillary”) — and thus Google served as a protection of Clinton but not of Cruz. Others, however, have pointed out politically conflicting biases—if someone wrote “Feminism is” or “Abortion is,” then the suggestions that came up were “cancer” and “sin,” respectively.³⁸ After Google was incriminated for caving in to censorship in hardliner Islamic countries, the company has been accused of favoring a rosy description of Islam, also in other countries. It does indeed cause concern that in Denmark, in June 2018, when googling “Islam is,” the first four suggestions are “Islam isimleri,”³⁹ “Islam is Peace,” “Islam is ...” and “Islam is a peaceful religion.” By comparison, the first four auto-suggestions for “democracy is” are “bad,” “dead,” “failing” and “not good,” respectively.

-
7. The Google AdSense blacklist—concerning websites paid by Google for their skill at attracting users to ads—in these cases Google is accused of withdrawing from the agreements right before payments are due.
 8. The search engine blacklist—it sends search results to the bottom ranks, potentially ruining companies affected.
 9. The quarantine list—blocking anything from individual users to entire sections of the Internet, sometimes taking a very long time to be restored.

See Epstein, R. “The New Censorship” *US News*. 06-22-16.

³⁸Solon, O. & Levin, S. “How Google’s search algorithm spreads false information with a rightwing bias” *The Guardian*. 12-16-16.

³⁹“Islam isimleri” is Turkish for “Islamic names.” The auto-completion suggestions are of course personalized and relative to the searcher and time.

A central battlefield within Google is its crucial ranking system. It determines which search results end up at the top of the results list. As described above, it has been personalized since 2009. But there is also a long and growing list of other conditions, pressures and forces that influence rankings. It is believed that more than 200 different principles now govern rankings, including how old a website is, the length of its URL, a special preference for YouTube links, emphasis on local websites in a geographical area, and many more; some innocent, others causing suspicion.⁴⁰ It has become a large independent industry to try to “trick” Google’s ranking criteria, enabling companies and others to pay to rank high on search lists—called Search Engine Optimization, SEO. The method makes use of various tricks, such as creating lots of artificial links between websites that one would like to see promoted, repeating keywords throughout a text, automatically copy-pasting from—and then adding a few changes to—already successful sites, which are then given other titles, and much more (“spamdexing”). In 2016, SEO was already a \$70 billion industry in and of itself. Google is said to be constantly struggling to make its ranking principles more sophisticated and coordinated in efforts to eliminate the possibility of capitalizing on the system in this way—in which is, of course, an infinite arms race with increasingly sophisticated responses from SEO companies. But if companies can game Google’s ranking algorithms and place interested customers on the top of the list, this can be used by political interests in the same way as commercial ones. *The Guardian* has thus mapped out how extreme right-wing sites in particular seem to have figured out how to take advantage of the complicated ranking procedures to come up high on the search rankings.⁴¹

⁴⁰Dean, B. “Google’s 200 Ranking Factors: The Complete List (2018)” *Backlinko*. 05-16-18.

⁴¹Cadwalladr, C. “Google, democracy and the truth about internet search” *The Guardian*. 12-04-16.

But these principles can also be politically influenced by the company itself.⁴² Already in 2002, critics of Scientology were being removed from search results.⁴³ In 2009, searches on then-First Lady Michelle Obama resulted in high ranking of a photo where she had been morphed with a monkey. At first, Google refused to do anything, referring to the company's neutrality policy, but after much criticism in the media, the image was removed and replaced with an explanation as to why.⁴⁴ Political battles in recent years seem to have intensified political censorship. In August 2017, neo-Nazi site *Daily Stormer* was deleted from the web-hosting platform GoDaddy because it had mocked one of the victims of the Charlottesville riots. The site shifted to Google, which blocked it after only three hours. *Daily Stormer* is undoubtedly a detestable site, but once again the removal conflicted with Google's tradition of claiming full neutrality regarding content, including political content.

In November 2017, company CEO Eric Schmidt announced that Google would downgrade Russian propaganda in its ranking system: "We're working on detecting this kind of scenario ... and de-ranking those kinds of sites", Schmidt said, in response to a question at an event in Halifax, Canada. "It's basically RT and Sputnik. We're well aware and we're trying to engineer the systems to prevent it."⁴⁵ The Cato Institute, a libertarian think tank, immediately took note of this initiative and asked whether Google itself was really concerned about

⁴²The Black feminist Safiya Noble highlights cases of algorithmically driven data failures that are specific to people of color and women and argues that marginalized groups are problematically represented in erroneous, stereotypical, or even pornographic ways in search engines. See Noble (2018).

⁴³Hansen, E. "Google pulls anti-Scientology links" *Cnet*. 04-22-02.

⁴⁴Google's explanation was: "Sometimes Google search results from the Internet can include disturbing content, even from innocuous queries. We assure you that the views expressed by such sites are not in any way endorsed by Google"; cf. Ahmed, S. "Google apologizes for results of 'Michelle Obama' image search" *CNN*. 11-25-09.

⁴⁵Hern, A. "Google plans to 'de-rank' Russia Today and Sputnik to combat misinformation" *The Guardian*. 11-21-17.

Russian influence, or if the company was rather acting on overt or maybe hidden political pressure from the US Government.⁴⁶ The Cato Institute referred to a recent congressional hearing where Senator Dianne Feinstein (D), a senior member of the Senate Select Committee on Intelligence, had blamed Google Vice President Kent Walker for not responding to Russian propaganda long ago: “.. I think we’re in a different day now, we’re at the beginning of what could be cyberwar, and you all, as a policy matter, have to really take a look at that and what role you play.” As noted by The Cato Institute, such a political imposition would not only violate the First Amendment guarantee of free speech; it would also violate Google’s judicially protected freedom to be in charge of managing its service, that is, prioritizing search results.⁴⁷

There are two conflicting problems at play here: government intervention in the tech giant’s freedom of expression, but also potentially Google’s own opaque, politicized prioritization of search results—depending on which of the two explanations is the correct one (one of them, of course, needs not exclude the other). In the first case, the perspective is that we need to settle for getting only search results approved by the US government. In the second case, we must settle for search results in alignment with Google’s political stance or the company’s voluntary or involuntary permissiveness in the face of pressure groups. The Cato Institute, focusing on a narrow definition of freedom of expression as closely linked to the actions of governments, seems to have no issue with the latter of the two scenarios just described. However, it seems to us that the prioritization of search results should in all cases adhere to openly available criteria, so that users are informed and aware of any political biases—if not the very searches themselves ought to be governed by principles of fairness and neutrality.

⁴⁶Samples, J. “Censorship Comes to Google” *Cato Liberty*. 11-21-17.

⁴⁷Sterling, G. “Another Court Affirms Google’s First Amendment Control Of Search Results” *Search Engine Land*. 11-17-14.

Curiously enough, Google has been accused of favoring both critique of Islam and defense of Islam in its rankings. This might seem contradictory, but only on the surface. Both charges can actually be correct at different times, because if the platform has been pressured to modify the algorithm to de-rank one of the two, then other will be consequently favored. A similar accusation is based on a survey from *The Guardian*,⁴⁸ which showed that a search for “Jews” would direct the searcher towards radical anti-Semitic websites, the same way as entering “did the Hol” would lead to websites denying the Holocaust. The latter is probably because most serious studies of the Holocaust do not even question the fact that the event took place and so do not contain the letter sequence “did the Hol”, making it a non-factor in the search ranking. However, in suggested videos on Google-owned YouTube, there is a tendency to prefer extreme results based on search words (see below), probably because extreme videos generate more clicks and are therefore better for advertising. If ranking algorithms are indeed set up to prioritize extreme views over moderate ones, it is ironic that they reflect the conscious targeting strategy of the Russian troll factories: not to support particular positions favored in the West, but to spread disagreement, controversy and disintegration in Western societies by supporting extremism across the political spectrum.

In March 2018, Google announced a new policy in the fight against “fake news”: the creation of a “Disinfo Lab” meant to downgrade or remove misinformation among search results and rank serious journalism high. The intention behind this initiative is commendable, but the effects of it are yet to be assessed. We remain skeptical as to whether Google, even with its huge economic muscle, could be able to create a

⁴⁸Cadwalladr, C. “Google is not ‘just’ a platform. It frames, shapes and distorts how we see the world” *The Guardian*. 12-11-16. Google seems to have reacted to this accusation by simply removing the “suggestions” when for instance entering “Jews are”, “Americans are”—but not when entering “Danes are”, which still prompts suggestions such as “reserved”, “cold”, “unfriendly” but also the happiest people in the world.

clearinghouse for truth that would surpass the existing networks of media, courts and universities. It is also hard to imagine such a lab as operating without political bias, if not favoring a particular party, then because its values will be based on a “Californian” outlook and the implicit platitudes of the *Zeitgeist*.

The idea of fact checking as something that can take place quickly and effectively is counteracted by the simple fact that there will always be important cases that remain undecided and moot—and even more so by the fact that some truths we take for granted today will be overthrown by new evidence tomorrow. That is, if indeed this new evidence is given the opportunity to come forward and is not fact-checked away in a flash. Take the process of Danish transitional justice after World War II.⁴⁹ Back then, a number of Nazi authors were punished for expressions made during wartime. This took place by recently adopted, retroactive legislation: they were sentenced for actions that were not criminal at the time of the deed. What is worse, when Harald Tandrup—writer and journalist at Danish Nazi daily newspaper *Fædrelandet* (*The Fatherland*)—was sentenced to three years in prison, a piece of “fake news” acted as crucial evidence. In the beginning of World War II, more than 8,000 Polish officers were rounded up and executed near Katyn, outside the city of Smolensk, Russia. After the massacre was discovered in 1943, Tandrup advanced in the Nazi press the scandalous assertion that the Soviet Army perpetrated the massacre. Everyone knew that Nazis were responsible for the Katyn massacre, so Tandrup’s assertion was deemed Nazi propaganda. In 1952, however, an American commission of inquiry found that the Soviet Union was in fact behind it, and only in 1990 did the country, through Mikhail Gorbachev, admit that Soviet troops had indeed been responsible for the massacre. It turned out Tandrup had been right all along. One is of course free to believe that the freedom of expression and rule of law for Nazi suspects presents no major problem, despite the fact that abominable

⁴⁹ Cf. See Mchangama and Stjernfelt (2016) p. 665ff.

persons and their views are exactly what principles should be tested against. In the context of this book, the example goes to show that claims—even when put forward by a coherent and serious group of people who are entirely sure of its veracity—may later be debunked if new evidence comes to light. But there is only room for such knowledge gains if there is no commission or algorithm performing “fact checks”; removing such evidence from the public sphere long before any thorough investigation can take place.

Compared to Facebook, Google seems more seriously concerned with freedom of expression, for example, in its year-long infight with Chinese censorship, but that has not prevented the company from making deals with a number of countries on local modifications of the algorithms aimed at removal of specific content. Currently, Google is resuming relations with China after the break in 2010, and possible censorship consequences of this new development remain unclear. Increasing rumors hint at the development of a specially censored search engine called Dragonfly which will automatically remove content based on dictates from the Chinese government; these rumors are taken seriously to the point that Senators from both US parties have asked Google’s top leaders to explain.⁵⁰ This has also led to more than a thousand Google employees protesting against the management’s plans.⁵¹ Critics fear that if Google and China agree on such an arrangement, it might form a model for censored Google searches in a number of other countries, such as Pakistan, Iran, Saudi Arabia, etc. Whether Google, like Facebook, has already caved in to Pakistani requirements remains disputed. In the spring of 2018, Swedish newspaper *Expressen* began a campaign against the tech giant, claiming that as a publishing entity, Google was responsible for spreading hatred and that it should be subjected to censorship. This prompted the

⁵⁰ Yuan, L. & Wakabayashi, D. “Google, seeking a return to China, is said to Be Building a Censored Search Engine” *New York Times*. 08-01-18.

⁵¹ Associated Press: “More than 1,000 Google Workers Protest Censored China Search” *Washington Post*. 08-17-18.

Swedish government to call Google up for a meeting. The government was represented by Justice Minister Morgan Johansson (Social Democratic Party) and Minister of Digitization Peter Eriksson (The Green Party). They expressed concerns that on the platform, Google allowed “illegal” and “harmful” content which could affect the Swedish elections. Note that the two legislators did not restrict their concerns to illegal content. Google promised to modify the algorithm and hire more staff to ensure that threats and hate were removed from search results and YouTube videos.⁵² Other Swedish newspapers such as *Göteborgs-Posten* and *Ystads Allehanda* warned against the *Expressen* campaign and the government initiative, stating that “spring-cleaning” Google could be extremely damaging to freedom of expression.⁵³

The bottom line is, however, that we know that Google has knowingly used the ranking algorithm in several cases to prioritize or deprioritize political and other content—but we do not know anything about how often it happens or of the principles behind it. The ranking algorithm and its constant development and sophistication remains part of Google’s innermost secret DNA. But one wonders if a de facto monopoly on a piece of public infrastructure such as Google should be based on principles entirely opaque to the public, or if algorithms should instead be publicly accessible and subject to discussion.

In August 2018, the possible lopsidedness of Google’s searches was questioned again, now as part of the Alex Jones case. President Trump posted one of his infamous tweets, based on an article by Paula Bolyard in *PJ Media*. She had searched for “Trump news” on Google, looked at the first 100 results and claimed that the 96 of them linked to left-wing media — based on a definition that ranked virtually all main-

⁵²“Censorship by Google” *Wikipedia*.

⁵³Boström, H. “Hatet mot Google” *GP*. 19-03-18; “Rensa nätet försiktigt” *Ystads Allehanda*. 03-12-18. Last visited 08-03-18: <http://www.ystadsallehanda.se/ledare/rensa-natet-forsiktigt/>

stream media as “left-wing.”⁵⁴ Three days later, Trump posted a stream of tweets: “Google search results for “Trump News” shows only the viewing/reporting of Fake News Media. In other words, they have it RIGGED, for me & others, so that almost all stories & news is BAD. Fake CNN is prominent. Republican/Conservative & Fair Media is shut out. Illegal? 96% of results on “Trump News” are from National Left-Wing Media, very dangerous. Google & others are suppressing voices of Conservatives and hiding information and news that is good. They are controlling what we can & cannot see. This is a very serious situation-will be addressed!”⁵⁵ The statement was backed by the White House, which announced government checks against Google and the other tech giants. This naturally caused a heated debate, indicating that such control would squarely violate the freedom of expression protection of the First Amendment. Also, a very reasonable objection was made: if one labels all media who have criticized Trump as left-wing, it is no surprise that one reaches conclusions like that of Bolyard. At the same time, an equal number of search results for and against a given subject cannot count as a criterion of fairness: Should a search on “Flat Earth” then return an equal number of websites claiming the Earth is flat versus round?

Google’s own response to Trump’s criticism showed, nevertheless, how difficult it was for the company to come up with a clear defense: “When users type queries into the Google Search bar, our goal is to make sure they receive the most relevant answers in a matter of seconds. Search is not used to set a political agenda and we don’t bias our results toward any political ideology. Every year, we issue hundreds of improvements to our algorithms to ensure they surface high-quality content in response to users’ queries. We con-

⁵⁴Bolyard, P. “96 Percent of Google Search Results for ‘Trump’ News Are from Liberal Media Outlets” *PJ Media*. 08-25-18.

⁵⁵Cit. fra Wemple, E. ”Google gives Trump a look at reality. Trump doesn’t like it” *Washington Post*. 08-28-18. The President is not himself a computer user, so it is believed that his numbers come from Bolyard’s article.

tinually work to improve Google Search and we never rank search results to manipulate political sentiment.”⁵⁶ What is meant by the fluffy words “relevant” and “high quality”, on which this argument relies? Google’s weak response shows that, in a sense, the company had it coming, exposing itself to attacks like Trump’s: when the company operates with an opaque, increasingly complicated algorithm, it is no wonder that it calls out for conspiracy theories, and with Google’s history of actual political manipulation with searches, it is hard to muster much trust in the company’s defense. Trump’s idea of state censorship is terrifying and unconstitutional, but his confused tweet contains the correct observations that Google defines what we see, and its workings are not transparent.

Interestingly, the accusations of liberal bias among the tech giants caused a group of internal critics on Facebook to surface, who pointed to a left-wing trend among the company staff. Brian Amerige, top Facebook engineer, wrote in an internal memo: “We are a political monoculture that’s intolerant of different views. We claim to welcome all perspectives, but are quick to attack—often in mobs—anyone who presents a view that appears to be in opposition to left-leaning ideology.”⁵⁷ There is little doubt that Amerige’s observations apply also to staff at the other tech giants in famously liberal Silicon Valley. Whether this imbalance is reflected in the product is of course another question. But the combination of the staff’s bias and the lack of transparency in the companies’ procedures makes them a natural target for conspiracy theories such as Trump’s, theories which—for that very same reason—the companies have more than a hard time repudiating.

In July 2018, three representatives from Google, Facebook and Twitter were summoned to testify before the House Judiciary Committee about the content moderation proce-

⁵⁶ Cit. fra Ohlheiser, A. og Horton, A. “A short investigation into Trump’s claims of ‘RIGGED’ Google results against him” *Washington Post*. 08-28-18.

⁵⁷ Cit. fra Conger, K. & Fraenkel, S. “Dozens at Facebook Unite to Challenge Its ‘Intolerant’ Liberal Culture” *New York Times*. 08-28-18.

dures of the companies. During the hearing, the tech giants were repeatedly accused of censoring conservative voices. An interesting thing about this hearing was that it became increasingly apparent that several legislators present did not understand how beneficial certain legislations on technology has been to these companies, and whose benefits are only recently being seriously questioned. Tech giants have always been able to enjoy full freedom from responsibility when it comes to the communication of their users. They remain under the political and legal radar because of the Safe Harbor Act—also known as Section 230—of 1996. The law is extremely convenient for tech giants. Firstly, it ensures that platforms that provide access to content are not accountable for the the expressions and actions of users on the platforms. This means that the platform providers do not have to control what their users are doing. Secondly, the second part of the law includes the decisive detail that if the platforms actually do decide to control what their users are expressing or doing, they do not lose their protection under Safe Harbor. This means that if a platform removes or moderates content, it will not suddenly be categorized as a publisher with associated responsibilities.⁵⁸ At the time, the second part of the law was considered an encouragement for tech companies to take on the difficult task of limiting online pornography or other unwanted content without being held responsible if the task seemed impossible to solve. But with Section 230, the principle that control implies liability was dissolved. The law—captured by the phrase “you have the right, but not responsibility”—leaves legislators without political leverage because it immunizes the tech companies, regardless of whether they restrict and censor user communication or not.

The law gives rise to some confusion because its second part is less known. During the hearing, Congressman Matt Gaetz (R) questioned whether tech companies can claim to be exempt from liability under Section 230 while at the same time asserting their freedom of expression with reference to

⁵⁸ Gillespie (2018) p. 30.

the First Amendment, which guarantees publishers the right to freely restrict content on their platforms. Gaetz's reasoning was that calling upon the Section 230 protection necessarily means giving up the right to be a publisher.⁵⁹ But this reasoning is a sign that the law was misunderstood, and Gaetz is not the only one. Due to the aforementioned detail in the second part of the law, it does not prescribe neutrality, which is the underlying premise of Gaetz's criticism.

Supporters of Section 230 have raised serious concerns due to increasing criticism of the law among Members of Congress combined with its widespread misinterpretation. One of the alarmists is Eric Goldman, a leading researcher on Section 230. He points out that the First Amendment prohibits the government from intervening in freedom of expression, and that this protection applies to both private companies and publishers as well as tech companies. Goldman says: "Private entities can engage in censorship. We call that editorial discretion." He then added: "When the government tells publishers what they can and can't publish, that's called censorship."⁶⁰ It is Goldman's point that by threatening to compromise the moderation practices of tech companies, congress is likely guilty of committing the very censorship they accuse the tech companies of doing. This is, however, a drastic warning, as tech companies would very much prefer not to be categorized as publishers. In order to highlight this dilemma, we point to the comprehensive responsibilities that the European Union is beginning to impose on tech companies — measures which do not leave the freedom of expression of users any better off. An example of this was a decision by the European Court of Human Rights in the case *Delfi v Estonia* in 2015. It was concluded that an Estonian website could actually be held responsible for reader comments in a debate posted on the forum without that being a violation of Article 10 of the European Human Rights Declaration on

⁵⁹Lapowsky, I. "Lawmakers Don't Grasp the Sacred Tech Law They Want to Gut" *Wired*. 07-17-18.

⁶⁰Op. cit.

Freedom of Expression — a somewhat excessive publisher responsibility.

This mess suggests that it is difficult to modify and adjust Section 230 without any clear definition of “tech giants”: Are they publishers, distributors, a public sphere, or something entirely different? This is one of the major problems with tech companies—they do not fit into existing categories. Practically all tech giants make their own content policies and police their platforms themselves. With 230 in hand, the giants have the freedom to arbitrarily decide when, to what extent, and why they should take responsibility for their users’ content. As previous chapters of this book have shown, this freedom to restrict often goes far beyond what is legally required. Often, it is just the result of economic strategy. One thing is certain: Section 230 is outdated.

Section 230 was adopted as part of the Communications Decency Act of 1996. Not only does it pre-date Facebook, Twitter and Google, but also platforms such as MySpace, Friendster and Napster. The point is that the law is in no way designed for social media, which did not exist in 1996.⁶¹ It does not take into account Google’s ranking algorithm that prioritizes or downgrades specific content, YouTube’s filtering technology which, despite claims to the contrary, could identify copyrighted material, or Facebook’s personalized news feed algorithm and removal handbook. And, most particularly, it does not take into account the emerging monopoly status of tech giants that has developed in the course of the 2010s. Tech giants are a hybrid of many existing business categories, which makes it extremely difficult to carry out a political and legal review of the nature of platforms’ responsibilities. There is quite simply no clear point from which to consider them in existing legal terms. In April 2018, a study by the Pew Research Center showed that over half of Americans support tech companies that take the initiative to limit false information in the fight against misinformation,

⁶¹ Gillespie (2018) p. 33-34.

“even if it limits public freedom to access and transmit information.”⁶² This is not the right way forward.

In the fall of 2018, a new wave of censorship swept through the main tech giants. In September, Twitter adopted new guidelines under the nauseating motto “Be Sweet When You Tweet”⁶³. It prohibits “[...] content that dehumanizes others based on their membership in an identifiable group, even when the material does not include a direct target” and adds a version of the standard list of selected groups to be granted special protection: “Dehumanization: Language that treats others as less than human. Dehumanization can occur when others are denied of human qualities (animalistic dehumanization) or when others are denied of their human nature (mechanistic dehumanization). Examples can include comparing groups to animals and viruses (animalistic), or reducing groups to a tool for some other purpose (mechanistic). Identifiable group: Any group of people that can be distinguished by their shared characteristics such as their race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, serious disease, occupation, political beliefs, location, or social practices.”⁶⁴ The description of “dehumanization” is extremely vague and wide-ranging and it is obviously it can be used to stifle much standard political debate—such as claims that such and such political group has been instrumentalized by lobbyists.

In October, a long-held internal Google memo with the title “The Good Censor” was leaked.⁶⁵ The memo is a blend

⁶²Mitchell, A., Grieco, E. & Sumida, N. “Americans Favour Protecting Information Freedoms Over Government Steps to Restrict False News Online” *Pew Research Center*. 04-19-18.

⁶³Matsakis, L. “Twitter Releases New Policy on ‘Dehumanizing Speech’” *Wired*. 09-25-18.

⁶⁴Twitter: “Creating New Policies Together”09-25-18. Last visited 12-18-2018: https://blog.twitter.com/official/en_us/topics/company/2018/Creating-new-policies-together.html.

⁶⁵Bokhari, A. (upload) “The Good Censor – GOOGLE LEAK?” Last visited 12-18-2018: https://www.scribd.com/document/390521673/The-Good-Censor-GOOGLE-LEAK#from_embed. More details:

of interviews and contributions from academics, journalists, and cultural critics, arguing for narrowing the scope of Google's traditional free speech stance. Here, the introduction of censorship is portrayed as a balance between free speech and the protection of users from "harmful conduct." The memo discusses whether users can be protected the users against negative phenomena like bots, trolling, and extremism while still being a platform for all voices. The memo does not yet conclude in terms of a new rulebook, but clearly the tendency goes in the direction of less rather than more freedom of expression. The general, hard-to-solve tension between liberty and security is the same conundrum as encountered by all the tech giants in the wake of the Alex Jones case in August 2018.

Moderation, content deletion, censorship by the tech giants—call it what you will—is undoubtedly here to stay. The Internet must be policed for criminal activities such as threats, harassment, extortion, incitement to violence, organization of violence, or forming terrorist cells. However, it does not follow that control should spread from such illegal activities to a wide variety of other types of content. It is also not obvious that the control, as is the case today, should remain hidden. Finally, there is no reason such control and its principles should be the privilege of tech companies themselves. The tech giants, often relying on a simplified and romantic idea of representing a "community" of common values, must realize that their vast populations of users are highly complex and represent strong, often opposing currents and values that also exist and act offline. The companies should instead realize that such contradictions are real, and not only the result of poor communication which will magically disappear through mantras such as "connecting people." Their task is rather to make available the many widely different, incompatible positions and values and provide a forum for serious clashes to take place and develop in a clear and unfeigned manner, one

Statt, N. "Leaked Google Research Shows Company Grappling with Censorship and Free Speech" *The Verge* 10-10-18.

that is free of violence and free of crime. This means forming public spaces rather than “communities”—and bringing the policies closer to ordinary, transparent standards of free expression.

However, strong trends are unfortunately heading in a completely different direction. With the law in hand, this tendency seeks to expand the deletion practice and responsibilities of the tech giants and thus—somewhat unwittingly—hand them even *more* power over the public.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

