PRIVATE PLATFORMS

Stanford | ETHICS TECHNOLOGY & PUBLIC POLICY

PART ONE

I. Three strikes

The end for Alex Jones, when it came, was surprisingly swift.

On July 24, 2018, YouTube removed four videos produced by the notorious right-wing conspiracy theorist and shared on his channel, Infowars. Within three weeks, Jones had been kicked off almost every major social media and online distribution platform. It was a rapid and dramatic downfall—not least because Jones had been sharing extremist viewpoints and harassing people online for decades.

Like most social media sites, YouTube users can flag content they believe to be breaking the site's community guidelines [See Exhibit 1, YouTube Community Guidelines] and alert the site's moderators. At YouTube, more than 10,000 paid workers spread around the world monitor huge volumes of these user requests every day and decide what to do with them.

Users had begun using this system to flag content produced by Jones, suggesting that he had broken the rules with a string of incendiary videos. Offensive bombast from Infowars was nothing new, given Jones's incessant string of outlandish and often ludicrous claims: these include his bizarre theory that the murder of 20 children and six teachers at Sandy Hook Elementary School in 2012 was a government-sponsored hoax; regular accusations that Barack Obama "smells of sulfur" and is the global head of Al-Qaeda; and that Hillary Clinton is a "demon from hell" who secretly runs a child abuse ring out of a Washington, DC pizza restaurant.

The new controversy, though, centered on hate speech and violence, rather than crackpot conspiracies. Two of the four Infowars videos contained anti-Muslim sentiments, while a third contained verbal attacks on transgender people. A fourth showed a cap-wearing middle aged man (not Jones) confronting an aggressive young boy, eventually grabbing the child by the throat and forcefully throwing him to the ground. Jones had titled the video "How to prevent liberalism."

Receiving the reports, YouTube's moderation team consulted the company's rulebook and executives to decide whether Jones had violated their policies. Eventually the decision came down: the four videos were in contravention of the rules, and they were deleted from the site.

On the surface, it was just another step in YouTube's ongoing wrestling match with Infowars and others who pushed the boundaries when it came to abiding by the community standards of large tech platforms. (As recently as February 2018, YouTube had issued Jones a "strike" warning for harassment and bullying in videos about Sandy Hook.) But this time it tapped into something new.

Three days later, on July 27th, Facebook suspended his profile page. The same day, Spotify removed some episodes of his podcast. They both cited reasoning similar to YouTube's. Two days after that, podcast service Stitcher removed his show from its platform completely, saying he had "harassed or allowed harassment" of private individuals on multiple occasions.

Over the next two weeks, as more complaints piled up, all these sites permanently deleted Jones's content and accounts, along with similar decisions from Apple, Vimeo, Pinterest, and many more.

Even Twitter—which had come in for regular criticism for being reticent to remove offensive content—joined the fray. At first, Twitter CEO Jack Dorsey publicly stated that Jones hadn't broken

the site's rules.[1] However, when evidence was shown to the contrary,[2] the company went ahead and suspended Jones's accounts. Eventually it went even further, and banned him from the platform permanently.

After 20 years of causing outrage with only marginal resistance, Alex Jones had been deplatformed – removed from nearly every major social media and mainstream content network – in the space of a few days.

Such swift action from so many corners inevitably generated its own set of conspiracy theories, with Jones and his followers railing against a perceived plot by the tech industry to silence their truth-telling and infringe on their rights. In fact, it appears that it was a snowball effect: each decision affected the next, and each ban made it easier for other platforms to take stronger action.

"When [Apple] moved, it was, like, OK, we shouldn't just be sitting on this content and these enforcement decisions," Facebook CEO Mark Zuckerberg told *The New Yorker*.[3] "We should move on what we know violates the policy. We need to make a decision now."

==

Do individuals have rights to speech on privately-owned tech platforms? Some people were concerned at what banning Infowars meant. Ben Wizner, director of the ACLU's Speech, Privacy and Technology project warned that hate speech was "extremely subjective" and that any ban was "worrisome." Although private companies are not subject to the First Amendment – and therefore are free to ban users for whatever reason they see fit – such policies were open to abuse, he said.

"Governments at least purport to be acting solely in the public interest," he told Huffington Post.[4] "But platforms are making these decisions based on what's in their financial interest. So their interest might be in avoiding controversy, but do we want the most important speech platforms in the world to avoid controversy?"

The question of financial incentives was a real one. YouTube maintains a close relationship with advertisers, and has a long history of stopping questionable sites from making money from ads – a move YouTubers call "demonetization." The concern over Jones's videos had already led some advertisers to pull their money from YouTube, and a fresh eruption of outrage was potentially leading the platform into a direct and public conflict between profit and free speech.[5]

YouTube's moderation team, though, said they were simply abiding by the rules they applied to every account – and that Jones's repeated transgressions were being held to the same standard of civil discourse that they applied across the site.

"We have longstanding policies against child endangerment and hate speech," the company said in a statement.[6] "We apply our policies consistently according to the content in the videos, regardless of the speaker or the channel. We also have a clear three strikes policy and we terminate channels when they receive three strikes in three months."

---

[1] https://twitter.com/jack/status/1026984242893357056
[2] https://money.cnn.com/2018/08/09/media/twitter-infowars-alex-jones/index.html
[3] https://www.newyorker.com/magazine/2018/09/17/can-mark-zuckerberg-fix-facebook-before-it-breaks-democracy
[4] https://www.huffingtonpost.com/entry/aclu-alex-jones-social-media-ban_us_5b7afce7e4b0a5b1febdc797
[5] https://money.cnn.com/2018/03/03/technology/youtube-ads-infowars-alex-jones/index.html
[6] https://techcrunch.com/2018/07/25/youtube-punishes-alex-jones-channel-for-breaking-policies-against-hate-speech-and-child-endangerment/

But even if the rules were simple, the impact of following them was not. A ban for bad behavior might make sense in isolation, but when your website is one of the largest content networks in the world – and when others take your decisions as a cue to make their own rulings – how do you consider the impact? When a website wields such influence, what is its duty to the political and public sphere? What at first looked like a conspiracy theorist getting his comeuppance had, in fact, exposed some of the critical ethical conflicts of social networking.

II. Everything in moderation

The action that took down Alex Jones might have been rapid and apparently consistent, but the reality is that there is still no agreed code of conduct online. At different points in the history of the web, different sites and services have developed their own rules and guidelines, which they continue to abide by. And while they tend to respond most consistently to national laws – generally around prohibitions on agreed standards for illegal activity like posting images of child abuse – most other guidelines are based on where services themselves choose to draw the line.

In America, where the internet originated under the direction of the Department of Defense, those rules are underscored by the constitutional framework. Chiefly this is driven by the First Amendment, which guarantees freedom of religion, assembly, petition, and speech. But while it enshrines protections for citizens, it prevents only the government from interfering with their activities ("Congress shall make no law…"), not individual service providers. Private companies, operating on private networks, are not – and never have been – subject to the First Amendment.

In the early days of the web, that meant the rules for sites, hosts, and forums were dictated by the instincts of the folks in charge, often the individuals who owned and operated them. Sometimes they were focused on keeping sites clean of pornography, or deleting content about gambling, or blocking copyright violations – or simply preventing speech that they found offensive. But the open and distributed nature of the online world meant that users could and would regularly move between services if they bumped up against rules that they didn't like.

"As web-hosting services like Angelfire and Tripod and community spaces like Geocities emerged, many did have 'community guidelines' or terms of service that prohibited certain kinds of content," writes **Tarleton Gillespie** in his book Custodians of The Internet.[7] "But even as these platforms grew, they were hardly an oligopoly… there was nearly infinite competition. If one host proved too tame, a user could simply find another willing to cater to more illicit materials."

Still, guidelines and moderation were not just about enforcing the tastes of a site's owners: they were a way of creating and reinforcing bonds between members of these small communities. And, crucially, they were enforceable because the scale of the internet was still relatively small. Even the most popular forums, websites, and usenet groups in the early days of the internet were miniscule by today's standards. GeoCities, one of the biggest such services, boasted of breaking the 1 million user barrier in 1997. (In 1998, around 17 percent of people in developed countries had access to the internet; 10 years later it stood at 61 percent. Today it is in excess of 80 percent.[8])

And even if these rules were imposed by individual site owners and service providers for arbitrary, personal reasons, they had legal protection over their decisions. Most jurisdictions recognized early on that it was almost impossible to control the actions of an individual online, and introduced rules and frameworks to protect services from being seen as the originator of illegal material. In the U.S,

[7] Gillespie, *Custodians of the Internet*, p27
[8] http://www.itu.int/ITU-D/ict/statistics/ict/

Section 230 of the Communications Decency Act of 1996 [See Exhibit 2, CDA 230] gave protections for distributors of information online.

"No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider," the legislation stated. The buck, essentially, stopped with the originator of the information, not the owner of the service on which that content is disseminated.

This solved lots of early questions about liability. After all, site operators had worried, if one service blocks its users from posting infringing material, but another did not (or could not) did that mean they could be held liable for the actions of users. Section 230 decided, effectively, that the telephone exchange is not responsible for you making prank calls: you are.

Similarly, the Digital Millennium Copyright Act of 1998, underscored Section 230 provisions, and also offered sites "safe harbor" from digital infringements such as copyright violation. As long as web hosts and Internet Service Providers were not actively encouraging customers to act illegally, they could not be considered the publishers of that material. A neo-Nazi website, even if it broke the law by threatening acts of violence, wasn't the legal responsibility of the site's webhost – any more than it was the responsibility of the power company that supplied electricity to the fascist's home.

These legal frameworks became central to how social networks and private platforms viewed themselves. And they found an ally in another central pillar of internet belief.

==

On February 8, 1996, **John Perry Barlow** decided to rebel. In some ways that wasn't so surprising: Barlow, then 48 years old, was a garrulous raconteur of cyberspace with a colorful history that included being a Wyoming cowpoke, a tripped-out lyricist for the Grateful Dead, and a co-founder of the online civil liberties group the Electronic Frontier Foundation.

With his eclectic taste and Western swagger, he knew how to make an impression. But this moment presented him a particularly unique opportunity to create a stir. Barlow was sitting in the snowy fishbowl of Davos, Switzerland, surrounded by the elite attendees of the World Economic Forum. Perhaps nowhere on the planet was there a more concrete representation of the global order. This was a cauldron of international economic power, a demonstration of the influence of authority. So, Barlow thought… what better place to stick it to The Man?

He grabbed a computer, logged on, and posted a wild, psychedelic manifesto to the web. The "Declaration of the Independence of Cyberspace"[9] echoed his belief that the internet was a startling new realm where the old rules didn't apply. The digital world, he argued, existed outside the remit of national regulators, away from the influence of governments, and beyond the reach of the law. The text was a gigantic fuck you to the existing system.

"Governments derive their just powers from the consent of the governed. You have neither solicited nor received ours," he wrote, bristling with counter-cultural swashbuckle.[10] "Your legal concepts of property, expression, identity, movement and context do not apply to us."

His statement captured the sense in the 1990s that the internet was a kind of Wild West that was ungoverned and, perhaps more excitingly, ungovernable.

==

---

[9] https://www.eff.org/cyberspace-independence
[10] https://www.eff.org/cyberspace-independence

Although Barlow's beliefs pushed at the edges of a different way of viewing the world, they also relied on the very structures whose power he refused to recognize. After all, two of his central obsessions – the internet and freedom of speech – were both very American inventions. The internet was the result of decades of investment by the US Department of Defense, while free speech was a concept only legally enshrined for the first time by the United States Bill of Rights.

The First Amendment to the US Constitution, instituted in 1791, establishes the right of Americans to freedom of religion, freedom of speech, and assembly. Internet users liked to focused on the Amendment's second clause, which prevented Congress from making laws aimed at "abridging the freedom of speech, or of the press."

These restrictions only applied to the actions of the government, and not private companies, but many free speech advocates took Barlow's cry to heart and adopted the First Amendment as a de facto creed: Don't limit what we say. Even if the First Amendment did not legally apply to private companies, the spirit of free speech could.

But as time passed, and the internet became more financially valuable and centralized, the balance shifted away from Barlow's anarchic urgings and towards a more consumer-focused universe. The forces of individualism and chaos started rubbing up against the forces of regulation and order. As sites and communities became larger and more influential, they had to deal with problems that made simple concepts like unfettered free speech and safe harbor messy and complicated.

"Pretty early on, we all agreed that Flickr was a business, it wasn't a country," says **Heather Champ**. She was community manager of the photography network, which pioneered many features that are now fundamental to social media, for five years. "We had lots of arguments with users around the idea of what is censorship."[11]

Jason Shellen, an early product manager at Blogger, remembers being caught in the ethical crossfire. He joined the site after it had been bought by Google in 2003, and quickly found that content moderation questions created a very particular kind of internal conflict.

"I remember fielding phone calls from principals in schools saying, 'Listen, James said something terrible about Billy on this blog, and we need you to take it down,'" he says.[12] "It was like, 'Wait, we need to do what?'"

"We had many, many meetings. There were people who took a more conservative approach, who thought we needed to do all sorts of moderation of what happened on the site. And then there was definitely a more ISP approach, which was 'OK, this is the internet. What will be will be.'"

It was the same all around the web, with sites – and the people who ran them – having to decide where they drew the line. Flickr was careful to develop rules that were aimed at keeping the service civil and friendly. It included extensive moderation and filtering to make sure that users weren't going to come across unexpected pornography, and community management techniques that often led to creepy or unpleasant users being banned from the service.

"People understood that they needed to participate by the rules," says Champ. "I guess I have very strong feelings that it really is the responsibility of the people who are running a community – the business entity – to create a safe space for the people who are participating in the community."

Others took different approaches.

---

[11] Interview with Heather Champ, October 2018
[12] Interview with Jason Shellen, October 2018

MySpace, which was the biggest social network between 2005 and 2009, encouraged good behavior by connecting everyone back to a central authority – every user's first friend on the service was founder Tom Anderson, better known as "MySpace Tom." Some used simple techniques taken from online forums that screened out certain words or phrases in order to try and prevent service violations: YouTube, for example, applied filters to comments posted under videos to remove offensive language.[13]

Many sites, however, took a more laissez-faire approach to community protection, preferring a position that reflected the internet's anarchic beginnings and a dedication to principles of free speech.

For years Reddit, the home of countless forums and large amounts of what it referred to "odious" or "distasteful" material, rarely intervened to police the content it hosted. Co-founder **Alexis Ohanian** once suggested America's founding fathers would like the site because it was "a bastion of free speech on the World Wide Web."14

The notorious message board 4Chan, meanwhile, was even more permissive. It became the playground of trolls who delighted in trying to be as offensive as possible, reveling in the anonymity the site offered them by throwing around language and images that might get them banned elsewhere. (Founder Christopher Poole, who now works for Google in Japan, gave up control of the site in 2015 and has since walked back some of his views on the site's permissive approach to speech.)

Then there was Twitter, which also appeared to place free speech at its core. Although its founders now say its position was never strict, Twitter's policy approach was shaped by early staffers like general counsel **Alex Macgillivray**, who had previously been part of Google's team and defined himself as a dedicated proponent of free speech. Macgillivray's claim that Twitter was the "free speech wing of the free speech party" became a part of company lore.

"One of our core values is 'defend and respect the users voice'," he said in a 2012 interview.[15] "No one wants a pen that has an opinion about what you write. Everyone wants a pen that will write what you tell it to write."

This idea – that any attempt to moderate or enforce speech behaviors on users of a network would inevitably break the tool and infringe on the right to free expression – spread around the company, and many parts of the industry. Even if the First Amendment wasn't legally binding, then adherence to the ideal of free speech could nevertheless be adopted. Moderation came second to expression at almost every turn.

===

Facebook, by comparison, had a deeply tortured relationship with moderation. In its very earliest days, the site took no responsibility whatsoever for anything posted on it – a nod to Barlow's declaration, even if it was unintentional. It soon became apparent, however, that there should probably be limits. Some images invaded people's privacy, for example, and a system that promoted or shared those images with other users could land Facebook in legal trouble. This realization generated a small set of guidelines and a flagging system, which covered elements like nudity and the posting of private information.

---

[13] https://www.youtube.com/watch?v=hO367o3JOhI
[14] [https://www.forbes.com/sites/kashmirhill/2012/02/02/reddit-co-founder-alexis-ohanians-rosy-outlook-on-the-future-of-politics/
[15] https://gigaom.com/2012/09/02/twitter-is-a-speech-loving-tech-company-the-amac-interview/

But the thin rulebook that Facebook started with was not enough. As the site expanded, it faced continual challenges to its internal guidelines and had to keep evolving them accordingly. [See Exhibit 3, Facebook Content Moderation Guidelines] Groups including breastfeeding mothers, journalists, Islamic hardliners, and conservative American pundits all found loopholes or exceptions that the existing rules couldn't cope with. And so Facebook continued to redraw its policies, until it had a byzantine set of regulations that were intended to create consistency but, in fact, caused vast levels of confusion.[16]

Here, as an example, are some of the rules contained in one of Facebook's moderation manuals, a 5,000 word, 22-page document that became public in spring 2018.[17]

- Videos of dying or wounded people are banned if they contain dismemberment, visible internal organs, or burned bodies.
- Videos of abortions are allowed, as long as there is no nudity.
- Images of cannibalism and throat-slitting are acceptable, but Facebook will place a warning screen on them.
- Convicted murderers are able to have accounts and pages, as long as they have not killed more than four people in a single incident.

News that was entirely false, meanwhile, was deemed acceptable – or at least too difficult to make decisions on.

"Reducing the spread [of fake news]... is a responsibility that we take seriously," says the document, but "we also recognize that this is a challenging and sensitive issue."[18]

It also lays out the resulting compromise: producers of fake stories could expect to have their distribution reduced, rather than deleted. The manual was complicated and often contradictory.

"When we were building that framework, we did a lot of tests," **Monica Bickert**, who leads global policy at Facebook, told Radiolab in 2018.[19] "And we saw sometimes that it was just too hard for our reviewers to implement a more detailed policy consistently. They just couldn't do it accurately."

Still, while there was a range of interpretation about the need to moderate content, the general bias across the growing social network industry was largely to favor free speech—even if it made some people feel uncomfortable.

Companies like Flickr viewed their product as a community, but most saw themselves creating the tools for communication, as with Macgillivray's pen. They were building network of friends, a connection between two strangers, a space for other people to publish. Wasn't that powerful enough?

III. Protest and counter-protest

When **Laura Gómez** joined Twitter in September 2009, the site was bouncing with energy off the back of a historic moment.

Earlier in the summer, company engineers had been preparing for scheduled downtime to fix some of the technical issues that had been plaguing it. But, just hours before the shutdown was due to happen, Twitter received a surprising email. It came from **Jared Cohen**, a young official at the U.S. State Department who was particularly focused on the role technology companies had to play in American foreign policy. He asked Twitter to stay online – to delay maintenance – so that it could be used by protesters in Iran.

[16] https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence
[17] https://gizmodo.com/here-s-everything-that-s-banned-on-facebook-all-on-one-1825495383
[18] https://gizmodo.com/here-s-everything-that-s-banned-on-facebook-all-on-one-1825495383
[19] https://www.wnycstudios.org/story/post-no-evil

Just a few days previously, Iran had held a presidential election. The incumbent, conservative Mahmoud Ahmadinejad, had been returned for a second term with 62 percent of the vote. The result surprised many, and claims of voting irregularities caused millions of people to hit the streets in protest. Reports at the time[20] called it "the biggest unrest since the 1979 revolution," and observers wondered if the president might be toppled.

Tensions were increased one day after the result, when Ahmadinejad announced that foreign media outlets like the BBC would be censored, and implemented a filter on Facebook so that it could not be used to spread information. Twitter, however, was not blocked—and so protesters started using it to organize themselves. The State Department worried that downtime would have a damaging effect on the demonstrations.

"This was just a call to say: 'It appears Twitter is playing an important role at a crucial time in Iran. Could you keep it going?" said PJ Crowley, the assistant secretary of state, shortly afterwards.[21] "We are proponents of freedom of expression... Information should be used as a way to promote freedom of expression."

Evan Williams, who was Twitter's CEO at the time, says he relied on his policy advisers for guidance—including Jason Goldman, who later went on to be the first Chief Digital Officer at the White House under President Obama.

"I remember having meetings, like, 'Oh, shit. What do we do about this?'," he says.[22] "What does it mean if we're taking direction from the State Department? Does that undermine our potential to be a global, sort of neutral, player?"

In the end, Twitter complied and pushed its shutdown back into the middle of Iran's night.

Gómez says that when she joined shortly afterwards, the perception inside the company was one of jubilation. This wasn't a triumph of technology – it was, after all, "a bare, basic technology" she says – but a "triumph of humanity."

"People that didn't have voices [could] actually tweet and talk about their experiences. It was very much idealistic," she says. "That called to me... especially knowing that there was a tool for people in oppressive regimes that can have a voice."

Gómez, a Mexican-American immigrant who had studied the often-turbulent world of Latin American politics, took on the task of leading internationalization team (that is, making the site available to users in many other languages). Internationalization was a crucial element of the company's plan to grow the number of users it had, but the experience of the Iranian protests suggested that its influence would be closely tied to news and politics, too.

Things would get there faster than anyone anticipated.

===

In December 2010, Tunisian street seller Mohamed Bouazizi set himself on fire in protest at police corruption and mistreatment in his hometown of Sidi Bouzid. His action sparked outrage throughout the country. Widespread anger at the 23-year rule of President Zine El Abidine Ben Ali suddenly given shape, and violent protests began taking place. In January 2011, after weeks of sustained demonstrations, Ben Ali resigned and fled the country.

---

[20] http://english.aljazeera.net/news/middleeast/2009/06/2009613172130303995.html
[21] https://www.nytimes.com/2009/06/17/world/middleeast/17media.html
[22] Interview with Evan Williams, October 2018

Similar protests sprang up in other nations, in a movement that became known in the West as the Arab Spring. This widespread revolution seemed reliant on social media, as rebels and dissidents from different countries were connected to each other and to their local supporters online. Twitter, Facebook, blogs, and messaging apps gave people the ability to see what kinds of protest were happening elsewhere, and the means to organize themselves. The kind of information that was previously restricted to mainstream media (which were susceptible to government interference) or to private whisper networks (which limited the spread of ideas) was suddenly in the hands of the people, in what often amounted to a public forum.

By 2012, governments and leaders had been overthrown not just in Tunisia, but also Egypt, Libya and Yemen. Civil uprisings had taken place in Bahrain and Syria, while protests of various sizes had affected almost every nation across North Africa and the Arabian peninsula.

Egypt was one of the most turbulent scenes. Using video blogs, Facebook groups, text messaging, and word of mouth, demonstrators gathered in their millions demanding the overthrow of President Hosni Mubarak. The President, who had ruled for 30 years, responded with military suppression, violence, curfews, and a communications shutdown. Eventually, after several weeks, he was forced out of office.

Governments across the region, and beyond, trembled.

The Arab Spring was galvanizing for those in the technology industry who saw their work as a mission to encourage communication. After all, these tools were fundamentally democratic and expressive—two ideals enshrined in the U.S. Constitution. Still, some worried that there would be natural limits or human consequences to social media tools being used this way.

"We always celebrated the Arab Spring, but we never thought about what happens after that," says Gómez. "What happens to the people who decided to stand up?"

====

Johnetta Elzie remembers exactly where she was when she stood up.

It was the morning of August 9, 2014 that she saw a rumor on Twitter that a local black teenager, Michael Brown, had been shot in the street by a white police officer in the St Louis suburb of Ferguson. The story was deeply personal: Elzie had grown up a couple of streets away from the scene, and one of her friends had been killed by the St Louis police earlier that year. By the evening, she went to see and joined a gathering of angry locals outside the police station.

"I wouldn't even call the first few weeks of what happened in St. Louis protests, because the National Guard was present," she says.[23] "In my mind, and a few other people who were there, you would call it a military occupation."

Quickly, she was part of a growing movement that grabbed national and international attention, and one of a handful of faces at the forefront of the protests. Ground organization mainly happened over Whatsapp, while social media became a channel for the outside world to observe and join in what was happening (her social media accounts grew rapidly as she reported what she saw on the streets.) Hashtags like #BlackLivesMatter became rallying calls, spreading the message, and organizing protests across social media.

But Elzie quickly discovered that while social networks could help foment protest and gain visibility for what was happening in Ferguson, they could be used to attack her. She became a magnet for all kinds of abuse.

---

[23] Interview with Johnetta Elzie, November 2018

Angry, racist trolls harassed her on every platform, even before her name and identity became public. Sites including Twitter and Facebook became battlegrounds, almost impossible to use because of the volume of vitriol pointed in her direction. False information or misleading stories began to be published about her and other visible leaders of the movement. Every time she or her fellow activists were mentioned in the media, their inboxes would be flooded with derogatory messages and attacks.

"Just the fact that I was out there was a problem for white racists," she says. "Once my name was found, it was just typical internet harassment, like phone calls at old phone numbers, there was people searching to see what my previous job history was… My mentions, my email. Anything that was public."

Today she tries to put those hurtful attacks in the past.

"A lot of this stuff, honestly I've tried to move on from," she says. "Because if I start thinking about it, I would lose my mind."

It was clear that the intention was to scare her – and an entire movement of protesters – into silence. What wasn't always clear was who was behind it. While #BlackLivesMatter was getting some very public opposition from conservative political figures, and many were clearly angry individual racists, other attacks seemed to come from anonymous accounts, or were obviously part of an army of robots being pressed into service.

What Elzie and others had encountered was the reality of modern political warfare. It's what researchers like **Samuel Woolley and Philip Howard**, of the Oxford Internet Institute, label *computational propaganda*, which they call "one of the most powerful new tools against democracy." "Regimes use political bots, built to look and act like real citizens, in efforts to silence opponents and to push official state messaging," they have written.[24] "Political campaigns, and their supporters, deploy political bots—and computation propaganda more broadly—during elections in attempts to sway the vote or defame critics. Anonymous political actors harness key elements of computation propaganda such as false news reports, coordinated disinformation campaigns, and troll mobs to attack human rights defenders, civil society groups, and journalists."

Activists had used the power of social networks to spur a movement. Their attackers had turned these tools of free expression into weapons for cheap suppression.

===

Was this turn of events predictable? The ability of protesters and counter-protesters to use the exact same systems for radically different ends certainly hit many social network employees as a shock.

Evan Williams, who was ousted as the CEO of Twitter in 2010, says he was certainly not prepared—and that the outcomes were not easy to predict.

"It's actually a good thing that there are services on the internet that maximize openness—and that is what we were trying to do at Twitter," he says. "What we underestimated is abuse and misinformation, and all those things that come as the cost of that."

"We create new things, they have effects we don't fully understand," he adds. "Some of them are good, some of them are bad. And there's really no way you can predict the effects. People act like you can, and I think it's just extremely naïve. You don't know how these things are going to work, and who they're going to affect."

[24] Woolley & Howard, The Computational Propaganda Research Project, p. 6.

Laura Gómez, however, says it was more a case of looking in the wrong direction.

"The political will always circle back into that dark," she says. "We always celebrated all the goodness, but we never really dealt with the consequences."

Others had sounded very clear warnings, contrary to what some industry insiders might say. **Zeynep Tufekci**, a sociologist at University of North Carolina, Chapel Hill, was among those who had long cautioned that the duality of social networks was inherent in the way they operated.

As a Turkish citizen, she'd seen President Recep Tayyip Erdoğan clamp down on dissidents using digital tools, and said that dictators and oppressive governments would become adept at every new platform that came along. And she knew that the same tools praised for helping protesters to organize would also be used to surveil and attack them, exactly what Elzie and the Ferguson protesters had experienced.

"One person telling you that your political views are stupid or even treasonous is almost certainly free speech, at least by U.S. First Amendment standards," she writes in her book *Twitter and Tear Gas*.[25] "However, it is another matter when tens of thousands of people attack your political views, interspersed with a random scattering of more serious threats. Such an avalanche creates fear and paralysis, especially since those targeted on social media are sometimes also targeted by government."

It was a pattern that was repeated again and again over subsequent months and years, often escalating deadly situations and endangering lives.

In Syria, Russian digital disinformation campaigns used the social web to undermine the White Helmets, a humanitarian rescue group disliked by the regime of Bashar al-Assad. Myanmar's military, meanwhile, used Facebook to fuel the persecution of the Rohingya Muslim minority, forcing more than 700,000 people to flee their homes and causing the United Nations to call for an investigation into genocide.

And in the Philippines, social platforms have been deployed by supporters of President Rodrigo Duterte to attack his enemies, including the Catholic Church, judges on the country's Supreme Court, and the press. The Committee to Protect Journalists says 40 cases of murdered Filipino reporters remain unsolved—more than any other country on its watchlist.[26]

"I don't think it's inaccurate to say that we were optimizing for freedom of expression," says Williams, who now runs the publishing network Medium, and remains on the board of Twitter.

"There are certainly things we would have done differently had we known more—that's the case with every single business. But I don't beat myself up about 'Goddamn, we got Donald Trump elected, so we should have done things differently.'"

Yet real-world experiences – of Elzie, of Egyptian protesters, of Tufekci, and of every voter in the 2016 U.S. Presidential Election – show that while these companies were optimizing for free expression, the result was almost the opposite of what they intended. Instead of promoting speech and protecting the public from government interference, social networks were now allowing dangerous governments and their supporters to run organized online campaigns that directly attacked the press and drowned out the political speech of ordinary citizens.

So where did things go wrong? Perhaps it's worth taking a step back.

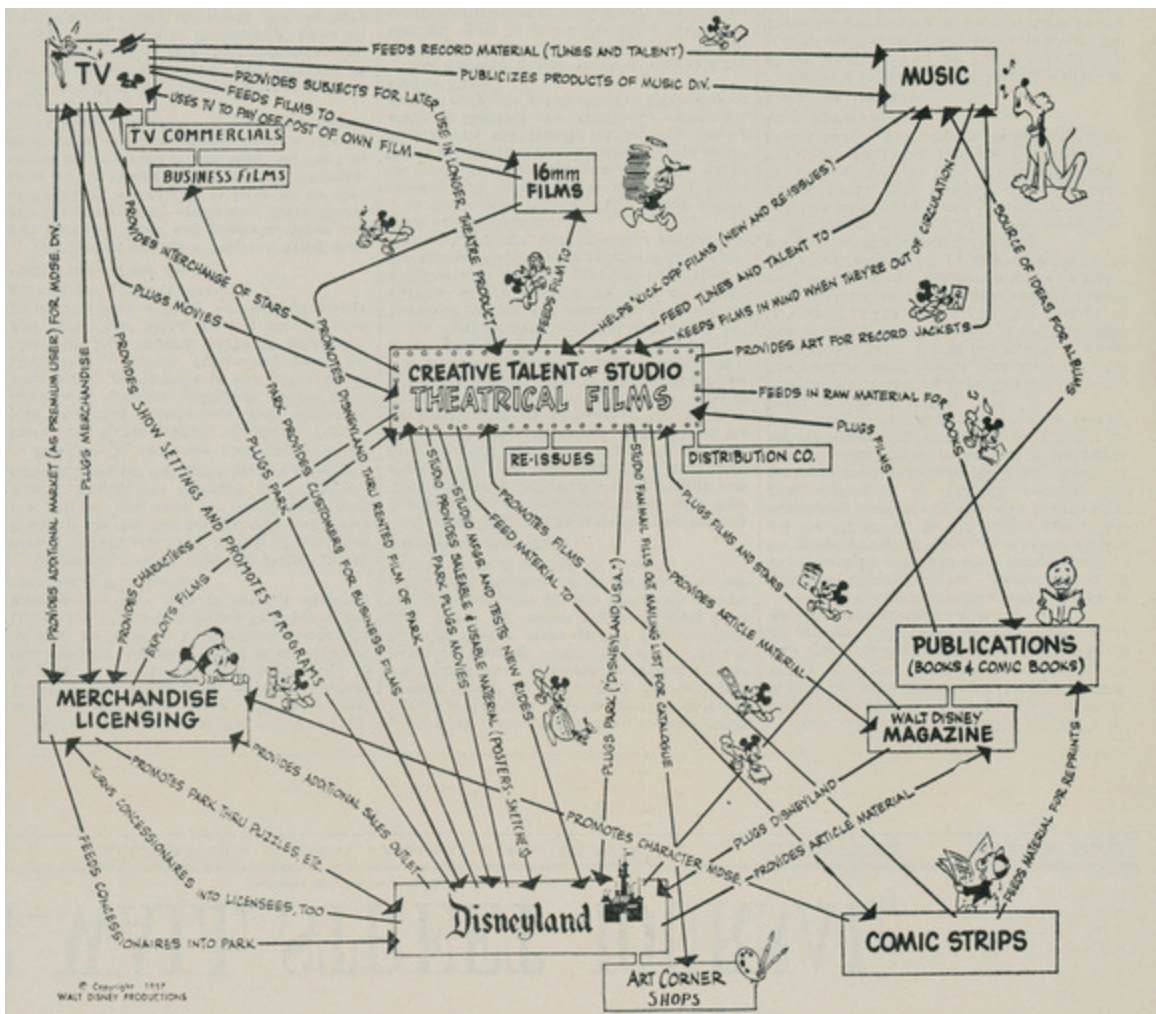25 Tufekci, *Twitter and Tear Gas*, p. 179
26 The Impunity Index, 2018

PART TWO: BUSINESS MODELS

I. Platforms

In early 1957, Walt Disney had a problem. He had a vision for the company he'd built—a strategy that he needed investors to understand—but he didn't quite have the words to explain exactly how large that vision was. By this point, he was already massively successful: The Disney Studios in Burbank, California had first created and then commanded the animated film industry for a generation, with a run of hits from Snow White to Peter Pan. His forays into TV were still just a few years old, but already heralded by viewers (and advertisers) across America. Two years earlier he had opened the Disneyland theme park in nearby Anaheim. While some shareholders had worried about its prospects, the site was an immediate smash hit, drawing in 3.6 million visitors in its first year alone.

But now, 56 years old and less involved with the company's day-to-day business of animation, he was stretching his ambitions in a different direction. He imagined Disney not just as a series of movies, or TV shows, or experiences. He saw it as a network—a platform for everything that Disney could do. And since explaining it was hard, he turned to what he knew best: He picked up his pen and started to draw.

What emerged was a corporate strategy for the whole of Disney, boiled down to a diagram.

It was a complex network of divisions all feeding into each other and creating a virtuous cycle of branding, merchandising, and intellectual property. At the center was Creative Talent, building new projects and earning money both from original work and re-issues. Around the edges were newer media, like TV, which was fast becoming a cornerstone of the business through both original programming and commercials. The Music division had its part to play, too, pulling material from movies and television and pushing it back out into the market to "keep films in mind when they're out of circulation." Books, comic strips, and magazines formed a publishing arm, while Merchandise Licensing created a flourishing secondary industry of toy-builders, game-creators, puzzle-makers, and clothing manufacturers who created profit and advertised Disney's characters and properties at the same time. And every single element fed into Disneyland itself, creating a venue where the experiences forged across all those other media could be realized in physical form.

Designing a large-scale business full of inter-related units—what we now call a conglomerate—was not necessarily new. But it wasn't as common in the media industry, or as clearly crystallized as Walt Disney's 1957 platform laid out. In fact, this model still informs Disney's strategy more than 60 years later, even if it has had to gobble up Marvel, Lucasfilm, and Pixar along the way, as well as build theme parks in Florida, Tokyo, Paris, Hong Kong, and Shanghai. Today, each element of Disney's business feeds into the others, enabling it to grow, to find new opportunities, to reinforce its successes and minimize its failures. What started out as an animation studio has become the world's most successful independent media company—a global entertainment platform that is valued in the hundreds of billions of dollars.

===

Thirty years after Disney drew his interconnected business strategy, another company – and another visionary CEO – laid out his own complex plan for industry dominance.

It was 1984, and a young **Bill Gates** was nearly a decade into his career as the founder and chief executive of Microsoft. Gates had dropped out of Harvard to build his business, which was privately held, growing fast, and had a great reputation. But Microsoft faced potentially deadly competition from rivals like Lotus, Ashton-Tate and most of all, IBM, the maker of the hardware which most companies used. In the early days of computing, the most successful companies created every element of the stack, not just the physical hardware, but the operating systems that powered the machines, and each bit of software that would run on them.

Gates had different plans. Eyeing IBM's decision to open its hardware architecture – allowing anyone to build compatible programs or accessories – he saw an opportunity to position Microsoft not just as the biggest player in software, or a major provider of services for other companies, but as the best enabler of other businesses, full stop.

He outlined his thoughts in a strategy memo that aimed to "take over the world with Windows."[27]

The idea? Create a technical baseline for others to build on, gain traction and build economic value. Unlike Disney's platform, which strictly controlled the manufacturers and franchisees who were given access to Mickey Mouse and friends, Microsoft's platform took a different approach.

Through a set of well-defined standards and parameters, almost anyone could theoretically create Microsoft-compatible programs, distribute them, sell them, and profit from them. Microsoft often had very thin relationships with other software makers—and sometimes no knowledge at all of what they were doing—but that didn't matter, because the platform meant that every single one of them,

[27] https://ubiquity.acm.org/article.cfm? id=1022357

and therefore their customers, too, became invested in the success of Microsoft's ecosystem over its rivals. Each installation of Windows became a platform for others to succeed – the gasoline of computing.

This approach worked stupendously well. By 1995, Microsoft had outmaneuvered all its rivals, gained control of around 90% of the global computer operating system market, and Gate, just 40 years old, had become the richest person in the world. His platform strategy became a model for how to dominate a market, and was picked up by software and technology companies that followed.

Buoyed by this example, technology companies have followed a Microsoft-like playbook again and again and again. Sometimes they do it by creating brand new markets; sometimes by injecting themselves into existing ones and upsetting the status quo.

Amazon started by digitizing one small corner of retail, book sales, before expanding into a gigantic retail platform that puts buyers at the center, whether they are purchasing from Amazon itself or from the countless small retailers selling their own goods in its marketplace. Etsy, eBay, and others followed similar patterns in other areas of shopping.

Apple and later Google, meanwhile, created markets for mobile apps, programs built by independent coders and downloadable by anybody with an Android handset or iPhone.

But not every platform requires the ownership and sale of physical products – the books, clothes, or phones. Just as Microsoft showed, some of the most successful platform businesses are those which do not really own anything at all. Uber took on the taxi industry by finding ways to connect two sides of its market, drivers and riders, without ever requiring it to actually buy the cars that customers would hail. Airbnb, meanwhile, has directly challenged hotel chains like Marriott and Hilton by connecting people with rooms or homes to rent with people who need somewhere to stay, but it does not own rooms itself – it merely connects the two sides of the market together.

This approach makes expansion much faster and cheaper, because it doesn't require the time or heavy capital involved in building hotels or buying fleets of cars.

So where do social networking companies fit into this arrangement? Many drew inspiration from the eBays and the Amazons, mediators of connections in a virtual marketplace. And looking at the standards set by Microsoft and the technology companies that followed, it is no surprise that they naturally saw the platform strategy as key to their own business success.

But there is a significant difference between those services which ultimately provide consumer goods – whether it's Mickey Mouse merchandise, or a room to sleep in, or even a piece of downloadable productivity software – and an array of digitized information that carries civic and political significance. After all, buying a car or a carton of milk has few direct political consequences; creating a spreadsheet rarely upsets the geopolitical order. Transmitting a tweet, on the other hand, can threaten nuclear conflict.

So what was their marketplace? What platforms were they building?

One view was that social media companies provided a market between people who wanted to say things and people who wanted to see those things, an exchange using attention as its main currency. But that wasn't the only way to look at these new platforms. Others saw the potential in a "social operating system" – well-defined parameters that allowed people to build software and services based on people's information and links to each other. It was a familiar concept, given a new lease of life.

II. Networks

On May 24, 2007, **Mark Zuckerberg** took the stage at the San Francisco Design Center to explain what that could look like – his version of taking over the world with Windows. In a presentation to more than 750 developers and journalists, he introduced two new concepts: the "social graph," a term for the complicated mathematical map of relationships between people; and the Facebook Platform, a set of standards that would allow developers to use that social graph and build their own software. Build software on top of Facebook using its APIs, he said, and Facebook will give you access to its user base and all the information you need—a move straight out of the Microsoft playbook.

"Until now, social networks have been closed platforms," he said.[28] "Today, we're going to end that."

He was right, to some degree. Facebook wasn't the first to do so (Flickr, Twitter, and others had all released their own APIs by this point) but it was going against the grain. The biggest network, MySpace, took a dim view of outside developers and tried to build most of its products in-house. That was the angle much of the media coverage focused on when Zuckerberg made his announcement—TechCrunch reported at the time that Facebook was "the anti-MySpace."[29]

The result of Zuckerberg's announcement was a flowering of applications, from photo management to quizzes to games; more than 3,000 in the first six months.[30] On one level, this looked like classic multi-sided platform strategy. Just as Microsoft had laid out its standards and become the operating system of businesses, Facebook would become the operating system for the social web.

Not everyone believed this, however – including the pioneer of software platform strategy himself. Microsoft had invested $240m in Facebook just a few months after the Platform launch (one of the critical events that fueled the company's hyper-growth) but **Chamath Palihapitiya**, who ran Facebook's platform, remembers Bill Gates looking at the launch skeptically.

"That's a crock of shit," Palihapitiya recalls Gates saying at the time.[31] "This isn't a platform. A platform is when the economic value of everybody that uses it exceeds the value of the company that creates it. Then it's a platform."

Gates was correct: In truth, very few of the apps developed for Facebook seemed to generate a lot of value. The main beneficiary was Facebook itself, which found a new way to keep users on its site instead of going elsewhere on the web, as well as collecting the data that these new apps harvested.

And, if Facebook was capturing most of the economic value, there was also the question of whether Facebook was *responsible* for these services, or merely the common carrier on which they operated. Platforms were agnostic, at least in the definition put forward by Section 230. They were not the originators or creators of content, merely the distributors of it. Therefore they had no legal exposure when the worst content crossed acceptable boundaries. But social networks muddied the waters, because they did not merely distribute, their guidelines and systems made decisions about which kinds of posts would be distributed. Some kinds of content are ranked algorithmically higher than others: a hit YouTube video is more likely to be re-surfaced to other users, while an offensive post that garners little engagement is likely to be suppressed. This curated, editorialized approach certainly has common ground with a Microsoft-style platform, where certain partners were given more attention and promotion. But it also has a great deal in common with Disney-style franchises

[28] https://newsroom.fb.com/news/2007/05/facebook-unveils-platform-for-developers-of-social-applications/
[29] https://techcrunch.com/2007/05/24/facebook-launches-facebook-platform-they-are-the-anti-myspace/
[30] https://www.wired.com/2007/09/ff-facebook/
[31] http://haystack.vc/2015/09/17/transcript-chamath-at-strictlyvcs-insider-series/

and media ownership, where choices are made and ownership accepted. Many content owners were, like Bill Gates, skeptical that Facebook's total control over content would generate any meaningful value, especially for the owners themselves. This issue of publisher versus platform was going to come up more often in the future.

So if they weren't building software platforms, what was it these companies were creating?

In fact, they *were* building marketplaces – just not ones for ordinary consumers, or professional users, or even for app developers. Their market was advertisers.

Just as Disney had outlined in 1957, social networks were building complex media distribution systems which they could monetize, and the easiest way to do that was by copying the main model available to media companies: selling ad space.

This wasn't a revelation. Google had already made itself the internet's most successful company by creating a hugely valuable advertising business, and Wired had created the first banner ads a decade before. But most social networks hadn't originally intended to go the same way: instead, they had started with no business model at all, hoping that they would discover a secret new way to generate revenue further down the line.

But as pressure grew to turn their size into cash—mainly from investors who had pumped millions of dollars into the social networking space—they decided to create platforms that were perfect for advertisers. So a cycle began: Creating the best platform for advertisers meant capturing the most amount of user attention; Capturing more attention meant getting more users. It all boiled down to one thing: Growth.

III. Growth is God

Instagram didn't have a lot to back up its valuation when it was purchased by Facebook for $1 billion in 2012. It was just two years old, revenue-free, and had a team of 13 employees. What it did have, though, was what everyone in Silicon Valley was looking for: dramatic, runaway growth. In those two years, the photo sharing app had attracted 30 million users, numbers that outstripped even Facebook's own rapid early expansion.[32]

Instagram CEO **Kevin Systrom** had known how important growth was all along. He'd worked at Google and Odeo (the company that eventually morphed into Twitter) and lead growth and marketing efforts at a small travel startup called NextStop. He knew that Instagram would flourish if he could make it grow faster than its competitors, so much of Instagram's early strategy was focused on tricks to encourage growth.

How would it manage that? At first, the app's photo filters helped people make their pictures look better—this was when phone cameras were still in their infancy, so even professional pictures sometimes looked like amateur snaps. But this didn't make it radically different from rivals like Hipstergram.

To get to the next phase, Systrom decided to piggy-back on the existing social networks by making it the fastest way to upload your pictures to a range of sites. A little smart branding helped encourage new signups and give it a leg-up over its photo app rivals.

And that's when Systrom saw a much bigger opportunity – and an entirely different category of competition. Instagram wouldn't win if it simply rode on the back of those other social services, he

---

[32] https://finance.yahoo.com/news/number-active-users-facebook-over-years-214600186--finance.html

thought: it would just be the most popular tool. In order to really prosper, it had to be a social service of its own.

"I remember hearing from people. They're like, 'Wow, I just got a like,'" he told LinkedIn co-founder Reid Hoffman on Hoffman's Masters of Scale podcast.[33] "And I was like, 'Yeah, because it's a network.' And they're like, 'Oh, I didn't realize that. I thought it was just a camera app.'"

===

Not every network is worth a billion dollars, but the concept behind Systrom's success is relatively straightforward. A network effect is, essentially, the idea that a service or product doesn't simply get bigger as more people use it, it actually improves. This improvement encourages more people to use it, which improves it further, which encourages even more people to use it, and so on and so on.

These effects are visible in many markets, from group purchasing (where bulk goods become cheaper as more people sign up to buy together) to insurance (where premiums drop for everyone as risk is spread across a larger pool). But nowhere are network effects more visible than on the social web.

As more people sign up to a service like Facebook or Instagram, so the ability to see and hear from more people goes up. And because being seen and heard is the fundamental action of the network, almost every user has a vested interested in its growth. Everyone becomes a salesperson for the system, and it develops a kind of built-in virality.

With companies that didn't have a business model – which counted for a large number of the so-called Web 2.0 services – the only real metric they had to measure success was growth. It was something founders and venture capitalists obsessed over.

And the cash kept pouring in. Investors threw more money into services that hoped to get big enough to create their own sense of gravity. The possibilities expanded as acquisitions began happening. Flickr was purchased for around $30m in 2005; MySpace was bought by Rupert Murdoch for $580m in 2006. At that point, Facebook had raised more than $40 million, and was valued at around $500 million. What seemed to be a high watermark was simply the beginning: the most "successful" businesses posted stats that went from millions of users to tens of millions to hundreds of millions in short order.

All of this meant that growth teams became the high priests of scale. They were asked to do anything they could to increase numbers, and when they were successful they were treated like stars (Mark Zuckerberg called Facebook's growth team "the most important product feature that we ended up building.").[34]

And as if that wasn't encouragement enough, this growth-centric orientation got a boost from a familiar ally: freedom of speech.

Growth at all costs seemed to be at odds with a strong-handed moderation policy. Moderation was expensive and put the brakes on user engagement, which therefore reduced growth and therefore hampered the value of the company. Better for everyone—or at least, for the networks themselves—to let it be a free-for-all and keep expanding policies to prioritize growth. The equation was simple: Free speech meant more growth, which meant more money. That was a mission everyone could get behind.

[33] https://mastersofscale.com/kevin-systrom-how-to-keep-it-simple-while-scaling-big/
[34] https://www.facebook.com/YCombinator/videos/how-to-build-the-future-with-mark-zuckerberg/10153993693829107/

That perspective was most clearly expressed by senior Facebook vice president **Andrew Bosworth** in an internal memo from 2016 titled "The Ugly" in which he laid out a credo in which connecting people together—i.e. building Facebook—becomes a greater good.

"Maybe it costs a life by exposing someone to bullies. Maybe someone dies in a terrorist attack coordinated on our tools. And still we connect people," he wrote in June 2016.[35]

The timing was no coincidence: there had been a recent swirl of news about the way Facebook could be used for ill. Controversy had erupted earlier that month after a man in Texas stabbed his girlfriend and posted a photograph of her dead body online. Facebook did not remove the image for 36 hours. And, just days before, the father of an American woman killed in a terror attack in Paris filed a federal lawsuit accusing social media platforms of providing "material support" to ISIS.[36]

It was also the middle of the contentious 2016 Presidential election campaign, with Donald Trump and Hillary Clinton confirmed as the candidates but several months before the final vote. Facebook was undoubtedly going to face some turbulent times and accusations of bias as the date of the election drew near. Bosworth's commentary seemed like an attempt to rally troops faced with these aggressive outside forces.

"The ugly truth," continued Bosworth, "is that we believe in connecting people so deeply that anything that allows us to connect more people more often is *de facto* good."

Bosworth has since said he was merely playing devil's advocate, and that neither he nor senior leadership figures (including Zuckerberg) actually agreed with his point. But the reality is that this view—that networking is speech, and more speech is more growth—was prevalent across the industry. It fostered a convenient, intimate connection between the philosophical and the commercial.

This neat dovetailing of priorities justifies a cycle in which a social network expends fewer resources while generating more attention and engagement, all in service of building an advertising marketplace. It's an internal narrative that allowed companies to justify all manner of action and, perhaps more importantly, inaction.

===

Sandy Parakilas believes this addiction is a core problem with the model of social networks. And he should know: He has worked at Facebook and Uber, two of the most aggressive, growth-focused companies to emerge from Silicon Valley.

He gives these businesses credit for some of their screw ups: they often made errors because they simply couldn't wait to make an impact. (Facebook's early motto, notoriously, was "move fast and break things.")

"One of the key things to understand about super-duper high-growth companies like Facebook or Uber is they often make mistakes by omission," he says.[37] "They move so fast that they don't put what you might consider proper process in place."

He gives one small example: Uber was still using QuickBooks – usually the accountancy tool of choice for small businesses – to do all its number-crunching, even when it was generating millions of dollars in revenue. A robust infrastructure wasn't a priority as Uber scaled up; any effort to create

[35] https://www.buzzfeednews.com/article/ryanmac/growth-at-any-cost-top-facebook-executive-defended-data
[36] https://www.theguardian.com/media/2016/jun/16/family-california-student-killed-paris-attacks-sue-google-facebook-twitter
[37] Interview with Sandy Parakilas, October 2018

a more formal accounting system would have slowed things down in a company that was sprinting desperately towards some invisible finish line.

But there are differences in degree, and then there are differences in kind. Errors of omission are acceptable when they don't affect the fundamentals of a business, he says. But it is something else when those omissions deliberately put a company at odds with its customers. In Silicon Valley's chase to create companies that expand rapidly, even exponentially – typified with the "hockey stick" growth charts that leave venture capitalists besotted – the curve is the end.

"If you were to define the two big problems, one is the business model, and the second is the growth mindset," he says. "We've created a system where we measure companies by their growth in users, and we don't really look too deeply. This was much less of an issue when the big technology companies were Apple and Microsoft, and the companies were selling products directly to users."

If you are either of those companies, says Parakilas, "there is a real incentive to protect those consumers—because they're the ones who pay you. Whereas if you are getting paid by advertisers, your interests are aligned with the advertisers… whose job is, fundamentally, to manipulate users. Which is a kind of strange thing, when you think about it."

Parakilas now works for the Center for Humane Technology, a group of Silicon Valley veterans who are trying to undo some of the damage they were party to in the last 20-30 years. They describe the problems created or exacerbated by technology – our addiction to phones, political fragmentation, mental health issues and elevated stress – and try to find ways to alleviate or eradicate them.

For social networks, the first problem was not thinking about the problems that these systems might cause. The next mistake was not hearing the voices of those who were predicting what might happen. But the unforgivable error, according to Parakilas, was to willingly adopt a position that the results *couldn't* be known precisely because the financial incentives made it easier to look in the other direction.

"When those externalities hit they can do huge, huge damage," he says. "The companies have been effectively told by Wall Street not to care and to look the other way because it's only rewarded in focusing on the top line growth."

This may be the major reason it took years to tackle notorious and dangerous misinformation peddlers like Alex Jones. It's why the systems built for advertisers could also be used to collect data on millions of people and manipulate them for political ends, as with Cambridge Analytica.

And an obsession with growth can also help explain the public sphere has been warped into a spectrum of political aggression with anger, trolling and brigading at one end, and the genocide at the other.

PART THREE: PUBLIC DEBATE

Surprise results

On June 23rd, 2016, the United Kingdom managed to not only surprise the world, but also itself. In a referendum on the UK's membership of the European Union, the majority of politicians and the leaders of all main parties—including Prime Minister David Cameron—had campaigned to remain with the status quo. And yet, when the results came in, they swung narrowly for Brexit: a 52% to 48% decision to leave Europe.

The result created shockwaves immediately, but for many people the most astounding thing was the discovery that the vote had been in the balance. Large numbers of citizens, many who did not vote, were dismayed to see a result they thought was simply impossible. For those who were left reeling, the public debate had never been about whether Brexit was going to *win*—it was how soundly the idea would be defeated.

Some pollsters attributed the surprise partly to the "shy Tory" effect,[38] where people don't tell the truth about their voting intention because they fear public admonition. But others suggested a different reason.

"Remain supporters expressed astonishment (mostly to each other on Facebook) that they never heard from leave campaigners," wrote Emily Bell, director of the Tow Center for Digital Journalism at Columbia University.[39] "Posing again the perennial unanswered question of whether we are ever more living in a 'filter bubble'."

The term was first been coined in 2011 by Eli Pariser, the founder of MoveOn.org and viral news site Upworthy. He used it to describe the way that algorithmically-determined information can end up creating closed systems that warp perception by giving people more of what they want.

Because sites like Google and Facebook chase maximum engagement, they are therefore trained to bring us more of what we want to see, and so we see only sources that confirm our existing biases. While such bubbles exist in other areas of life—Democrats are unlikely to watch Fox News, and die-hard second amendment devotees are unlikely to have friends who detest guns—the digital sphere's bubbles appear at a different layer of the system (everyone uses the same Facebook, but everyone's Facebook is perfectly tailored to their own tastes.)

It seemed a perfect way of explaining the enclosed, curated information environment that gave people more of what they wanted to hear and changed their idea of what the public debate actually was.

"Modern truth is not just outpaced by fiction," said Bell. "It can be bypassed altogether as part of a sound political strategy or as a central requirement of a media business plan."

Filter bubbles are often made worse because many users don't realize they exist—a 2015 study suggested that around 60% of Facebook users don't realize there is any filtering or curation on the site at all.[40] And these problems seem to be further exacerbated by the echo chamber effect—where a closed system allows a limited set of ideas to get repeated and re-emphasized, amplifying their power. (Imagine a sound inside a closed room bouncing backwards and forwards until your ears are ringing.)

---

[38] https://www.telegraph.co.uk/news/2016/06/24/eu-referendum-how-right-or-wrong-were-the-polls/
[39] The truth about Brexit didn't stand a chance in the online bubble, *The Guardian*, 2016
[40] Eslami, et al "I always assumed I wasn't really that close to her" 2015

In addition, while filter bubbles are partly the result of user choice—you pick your friends in a much more precise, and your friends often say only the things you want to hear—they are also the result of design and editorial decisions made by the platforms themselves. Algorithmic ranking, where a system changes the visibility of posts it thinks will increase or reduce user engagement, is common. Some sites, like Reddit, prioritize user ranking (upvotes and downvotes) and make them visible. Others use a complex and deliberately concealed algorithmic ruleset to understand what users are more likely to click on, and then show them more of those things.

Both Google and Facebook are notorious for constant tweaks and changes to their systems (showing you more activity from your closest friends, reducing the visibility of sites that use certain keywords, and so on) which ripple through the information economy with metronomic consistency.

But these filters have operated with very little public accountability or oversight, not least because of the legal protections offered by the state. Since these algorithms are made by humans, though, some have argued that they are essentially editorial determinations—made by social networks to try and increase engagement or spread certain kinds of information over others.

In some countries, the line between publisher and utility is drawn sharply along exactly these kinds of lines—where changing or prioritizing content (or, indeed, not deleting it) is seen as as a publishing activity. But in the United States, the debate over whether technological platforms are responsible for the material published on their sites—whether they are the telephone line or the prank caller—was long settled by CDA 230. The courts ruled that a website operator can make changes to content without assuming the liabilities of a publisher, essentially granting it a kind of First Amendment protection.

However, this construction has started to come under fire. In 2016, a lawsuit against Twitter accused it of breaking opt-in communication laws by delivering unwanted tweets in messages to cellphones which used recycled numbers (the previous owner had signed up to receive tweets; the new owners did not opt-in, but still received messages.) Section 230, which it attempted to use as a defense, was deemed not to apply. In June 2017, Yelp was ordered to remove user-written reviews that were found to be defamatory: Section 230, again, was not a protection.

The Electronic Frontier Foundation, among others, remains concerned that courts have not fully clarified the level of oversight and editing at which a service becomes an "information content provider" rather than a utility. And even if Section 230 *does* apply in the eyes of the courts, there appears to be a difference between the legal and the moral obligations of companies who have outsized influence on our political and public spheres.

The ecosystem of computational propaganda, astroturfing campaigns, foreign interference, fake news, filter bubbles, and echo chambers combine with the power of private platforms to create an atmosphere for "public debate" in which both the ideas of "public" and "debate" are arguable.

This wasn't exactly how people imagined it would be. Nicholas Negroponte, the founder of MIT's famous Media Lab, had made the compelling and inspiring argument in 1995 that personalized news was an inevitable consequence of the digital revolution. This system, he wrote in his book *Being Digital*, was a customized stream of information that you chose to receive.

"Instead of reading what other people think is news and what other people justify as worthy of the space it takes, being digital will change the economic model of news selections, make your interests play a bigger role, and, in fact, use pieces from the cutting-room floor that did not make the cut on popular demand... Call it *The Daily Me*."[41]

---

[41] Negroponte, *Being Digital*, 1995, p153

While Negroponte wasn't an unquestioning cheerleader for the concept, he thought more control for consumers was a net positive—and that we would choose personalization levels based on our mood and context.

"Imagine computer display of news stories with a knob that, like a volume control, allows you to crank personalization up or down," he writes. "You could have many of these controls, including a slider that moves both literally and politically from left to right to modify stories about public affairs. These controls change your window onto the news, both in terms of its size and its editorial tone."[42]

Little more than a decade later, social networks and social media turned *The Daily Me* into a reality. Except we didn't choose to modify our own feeds; we let the companies choose for us. And that meant we opted for filter bubbles, for echo chambers, for more of what we already knew. We opted for the obscure incentives and perverse economic desires of social networks. We willingly handed control of the public conversation to private companies, and gave those private companies immunity.

II. The public sphere

The idea of political debate happening in the public sphere goes back a long way—before the media industry, even before the invention of writing, and certainly a long, long way before social networks. It was a core concept of Greek democracy, and a central pillar of civic society over the last 300 years. And, as we have seen, it was part of the founding myth of the United States.

As such, there have been laws to protect, govern, and regulate speech and assembly in public spaces for almost as long. For most of the modern era, most political discussion has happened in these public spaces—in face to face conversations or debates—or in mediated, regulated private spaces such as newspapers, magazines, or television and radio broadcasts.

But the online world, with its own rules and different basis for responsibilities, has created a different set of actors. Social networks, when they arrived on the scene, were quick to claim they were part of this broader public sphere—the new guardians of a town square where open speech is encouraged in the service of finding political solutions.

"A lot of people come to Twitter and they don't actually see an app or a service—they see what looks like a public square," said Jack Dorsey, looking back on his company's role in a 2018 interview. "And they have the same sort of expectations of a public square. And that is what we have to make sure that we get right, and, also, make sure that everyone feels safe to participate in that public square."[43]

This phrasing is important to note. By seeing themselves this way, social networks were explicitly choosing to side with a media-driven concept of speech, rather than identifying with service-driven technologies like the telephone or the pen. Even if they claimed to be mere tools when the law came knocking, they were taking the role previously occupied by publishers and distributors. After all, it is not a pen's job to make people feel safe to write, or to safeguard democracy.

Still, these technology companies were not the first private organizations to operate in the public sphere. Media outlets of all shapes and sizes—first newsprint, then radio and television—have dominated the conversation for generations. And even if there are rules of engagement today demand journalists strive for "fairness" or "objectivity", the press has not always been disinterested in the outcomes of public debate.

---

[42] Negroponte, *Being Digital*, 1995, p154
[43] Jack Dorsey on Twitter's Role in Free Speech and Filter Bubbles, *Wired*, 2018

Many influential press barons have been ruthless in compromising fairness in order to generate revenue – Joseph Pulitzer's New York World was a pioneer of sensational yellow journalism before he decided to sponsor prizes for journalistic excellence (a moment of legacy-spinning second only to Alfred Nobel's decision to offset his invention of an explosive that could blow people limb from limb by instituting an international peace prize.)

And few of them have shied away from using their power to direct or reframe public debate in their favor. William Randolph Hearst's newspaper and magazine empire often focused on promoting isolationist policies, and Rupert Murdoch built a global media empire of outlets united by a right-wing, pro-business agenda.

Yet the technical limitations of media – including their need to build generalist audiences and engage in competition with rivals – created a kind of para-public space for conversation. In #Republic, Cass Sunstein describes the serendipity of the newsstand.

"While shopping at the local grocery store, you might see the cover of Time or Newsweek, and the story – about a promising politician, or a new risk, a surprising development in Europe –might catch your attention, so you might pick up the issue and learn something even if you had no interest in advance," he writes.[44]

Such intellectual serendipity may sound like a rose-tinted view of the past from where we sit now, but it is also somewhat accurate. Similarly, the evening news bulletin – a standard of media consumption for many years – used its variety as a strength: You would not know the topics that the news would cover (or if you did know, you went to the evening news for a shared experience of them) but the ambition of those producing it was usually to leave you better informed. But it was the result of an almost accidental monopoly on the public sphere that was owned by media organizations.

"I think there's a great nostalgia for the way things are done amongst journalists which is entirely rational," says Alan Rusbridger, the former editor-in-chief of The Guardian.[45] "So, if you were a journalist, it was really lovely to be in control of the message, and the system of distribution, and the supply of information, and the supply of money. There was nothing not to like."

It is no surprise that the most trusted gentlemen broadcasters of this age—Edward Murrow, Walter Cronkite, and others—echo the paternalistic world in which they existed. The height of their influence happened around the same time as the corporate economy reached new heights of power during the 1950s and 1960s (Economist Marina Whitman describes this as the phase in which "America's large corporations became private institutions endowed with a public purpose.")[46]

Just as media organizations created public spaces by fostering loosely-knit communities, so there was a general trend towards corporations assuming a state-like presence, building homes and schools, providing a social safety net and giving moral guidance to their worker-citizens.

That began to shift thanks to societal changes and a revolutionary new technology that allowed people to tune out things they didn't want to hear. It wasn't social networking, though: it was cable news. In a 2017 study on echo chambers and political partisanship, the researchers Erik Peterson, Sharad Goel and Shanto Iyengar point out that the foundation of political division we see today was formed in the 1970s and 1980s. A plurality of outlets was already building a broadcast version of The Daily Me.

"While the period in question encompasses multiple societal changes—greater ethnic and religious diversity, a declining manufacturing sector, and heightened income inequality," they write, "It was

[44] Sunstein, *#republic*, 2017, p42
[45] Interview with Alan Rusbridger, November 2018
[46] Whitman, *New World, New Rules: The Changing Role of the American Corporation*, p34

also a time of seismic changes in the media environment. 24-hour cable news channels emerged as competitors to network news. The availability of cable television in the 1970s provided partisans their first real opportunity to obtain news from like-minded sources."[47]

This was the world into which the social network emerged; one where people were eager to dive into hyper-partisanship—where they were not just stumbling into an echo chamber, they were primed for it.

Today's situation has become even more complex because social networks have emerged at a different layer in the media stack—the distributor—which puts them in competition with the news media itself.

Journalistic outlets, still major contributors to public debate, are now dependent on social networking for distribution. Two thirds of Americans now use social media to get their news, while nearly half of Facebook users rely on that site alone for their news.[48] And yet social networks are also cannibalizing the main revenue source—and doing it much more efficiently. In large part, says Alan Rusbridger, because they are able to provide services that people want.

"I don't think it's possible to go back," he says. "I think there's a certain inevitability that if you give people that power of taking part, they will take it. And there was no stopping that once it was unleashed."

===

This is another form of ethical quandary that technology companies struggle with.

When presented with these deficiencies, social networks will often try to deflect the question or point to initiatives to counter their impact. For example, YouTube's parent company Google can point to the Google News Initiative, which is offering $300 million to the news industry[49] to "help journalism thrive in a digital age."[50]

Or, when faced with the question of whether Twitter chose growth over curbing abuse, Evan Williams quickly pointed to things that Twitter has done that were not in service of growth.

"One quick example was the DM [direct message] rule on Twitter—that you had to be following a person for them to send you a message," he says. "That was contradictory to growth."

Similarly, Mark Zuckerberg's immediate reaction to being presented with the idea that Facebook damaged U.S. politics in the 2016 Presidential Election was to deny it was possible.

"Personally, I think the idea that fake news on Facebook, which is a very small amount of the content, influenced the election in any way — I think is a pretty crazy idea," he said. "All the research we have suggests that this isn't really a problem."[51]

But that is disingenuous. Social networks, particularly Facebook, are incredibly influential, and that influence is massively important to them as systems and as businesses. Without the ability to influence consumers, a social network's marketplace for advertisers is worthless, and the billions that these companies earn for their investors and their employees would disappear in a wisp of smoke. (Zuckerberg has since revised his comments.) It happens in part because a CEO is responsible for more people than just the users of their services.

[47] Peterson, et al, Echo Chambers and Partisan Polarization, 2017
[48] News Use Across Social Media Platforms 2017, Pew Centers, 2017
[49] Google Offers News Industry $300m Olive Branch, *Columbia Journalism Review*, 2018
[50] Introducing the Google News Initiative, 2018
[51] Zuckerberg: The idea that fake news on Facebook influenced the election is 'crazy', *The Verge*, 2016

"When companies go from being privately controlled puppies to sitting on a public stock market, they are beholden to the street," says danah boyd, principal researcher at Microsoft Research and the founder of Data & Society.[52]

"The majority of [Mark Zuckerberg's] employees experience their financial strength through the fact that they are compensated through stocks. So, if he wants to keep his talent and not lose them to competitors in a really hot market, he has to be able to pay them in stock going up. He wants the stock to go up, he's got three options: He could make more money per user, he can extend the number of users, or he can diversify his profit and loss infrastructure. He's not really succeeded at the latter, he's really maxed out on the number of users, and once you start talking about making more money per user, it starts to get really creepy really fast. That's a really ugly bind that doesn't end well."

===

There are other good reasons for this tortured understanding of speech and influence, beyond an addiction to money and a simple wish to avoid blame when things go wrong. Social networks might profess to be creating a public square where speech is sacred, but the reality is that they have not been practicing it very well.

Several times in the last 20 years, sites of all sizes—including Tumblr, Instagram, Pinterest, and Blogger—have chosen to develop rules that ban content promoting self-harm and eating disorders. So-called "pro-ana" sites and accounts (the term is short for "pro-anorexia") have been viewed with disdain and concern, and social networks have consistently chosen to silence such users in order to protect others—even though there is little evidence that such bans actually work.[53]

Inside their own offices too, these companies struggle with the concept of unfettered speech. Facebook's complex moderation rules allow users to say "black children should not be allowed in our town",[54] but when an unnamed employee wrote "All Lives Matter" on an office wall at Facebook—effectively erasing the previous "Black Lives Matter" text—Zuckerberg wrote to the company's entire staff to say that the incident was under investigation.

"We've never had rules around what people can write on our walls… [but] crossing out something means silencing speech, or that one person's speech is more important than another's," he wrote.[55]

There are plenty of other kinds of "censorship" that platforms engage in. Spam blocking, for example, is a problem that platforms spent significant time and energy on (most laws against spam are focused on the invasion of privacy that is involved in receiving unsolicited emails, not on the actual mass dissemination of messages.)

Other kinds of manipulation include giving prominent users extra kinds of support, or additional tools that encourage them to post more often to speed up a site's viral loops and network effects. Laura Gómez, who had the unedifying task of verifying the Twitter account of dictatorial Venezuelan president Hugo Chavez while leading the site's internationalization efforts, says that the process of giving a bigger bullhorn was among its biggest mistakes.

"We never really dealt with the consequences of having those leaders on the platform," she says. "It was the number one thing—if I could go back I would be like, what does the leadership mean and

52   Interview with danah boyd, November 2018
53   Yeshua-Katz and Martins, Communicating stigma: the pro-ana paradox, *Health Communication* 2013
54   By Facebook's Logic, Who Is Protected From Hate Speech? Wired, 2017
55   Inside Facebook: Mark Zuckerberg's 2 Years of Hell, *Wired*, 2018

who gets to decide who is in it? And why don't they get suspended if they do wrong?"

Even a ranked feed, which prioritizes certain kinds of content over others instead of a purely chronological stream of information, could be seen as a kind of censorship. Instagram and Twitter have both caused controversy after implementing such changes. While these shifts are ostensibly done to increase perceived utility, they were largely aimed at increasing the amount of time a user would spend on the service. After all, the logic goes, giving people more of what they want is more useful.

But much like echo chambers and filter bubbles, which involve both user choice and platform design, it is impossible to know exactly what combination of goes into producing a typical activity feed—and therefore what the difference is between censorship and custodianship. This is in large part because the makeup of such systems is closely-guarded. In many cases, in fact, companies see this recipe as the secret to their success. Take Google's search algorithm, for example—perhaps the mother of all algorithmic ranking systems online. It is based on a simple premise of "backlinking," in which each link to a certain page from the rest of the web is counted as a vote on that site's validity. That straightforward concept has become more complex over time through more obvious decisions (some sites' outbound links are ranked more authoritatively than others) and other, more obscure ones (for example, linguistic diversity or content "quality".) The combination of cryptic, constantly shifting rules, has created a huge market in search engine optimization for those wanting to gain more visibility in Google's listings.

It's the same pattern seen whenever there is an attention market online: Content producers try to understand algorithmic ranking because they are chasing traffic or activity; Websites keep changing their algorithms to increase their effectiveness and generate cash from the content producers who want to boost their own position. And there it is again: growth and money—the business model dictating the boundaries that the product chooses to respect.

A network full of spam is annoying for users, but despite this problem billions of people continue to use email. For a social network, however, spam is a real pain—the network begins to lose its value to advertisers because it contains lots of misleading data. A system that truly adopts free speech principles, rather than adopting the fig leaf of commercial sovereignty, would struggle to get advertisers at all.

This is something most insiders are aware of, but are only now starting to reckon with.

Fifteen years ago Jason Shellen would receive phone calls from school principals worried about stories posted on Blogger—and he'd respond by batting their requests away with safe harbor and free speech justifications. Today, he wonders whether those principles even hold.

"Should free speech, the way we define it in the United States, should that extend online in quite the same way? Because there are no self-harm provisions for the First Amendment, right?" he says. "I do sometimes think that if people have enough power or influence, that maybe there should be a tad more accountability."

PART FOUR: WHAT HAPPENS NEXT

Redesign

Tim Wu has some pretty strong opinions. That's no surprise: he's a professor at Columbia Law School, opinion writer for the New York Times, a vocal proponent of Net Neutrality, and once ran for Lieutenant Governor of New York.

But Wu takes the arguments about free speech and turns them around, like a well-practiced kid assembling a Rubik's Cube in a few seconds. Forget the idea that the First Amendment is no longer the best way to frame issues of speech—he says it's possible that it may be entirely the wrong way to look at things.

In a paper with the provocative title "Is The First Amendment Obsolete?"[56] Wu acknowledges that the internet has radically changed things, in unexpected ways.

"A range of speech control techniques has arisen from which the First Amendment, at present, provides little or no protection," he writes.

America's founding fathers were looking at a marketplace of ideas in which the ability to be a speaker was a rare commodity, and constantly under threat. Today, it is not speakers who are a scarce resource in need of protection, argues Wu: it's the ability to be heard.

This is why the most dangerous threats to speech are not other opinions, but the attempts to drown out and silence dissent. Technology and social networks have created a new kind of silencing speech—the brigading experienced by protesters in Ferguson, or the silencing of Egyptian demonstrators, or the surveillance experienced by Turkish dissidents. These chilling effects are aimed at stopping dissent before it evens begins.

How does Wu suggest we change this? He puts forward a set of possible solutions, some of them more achievable than others. He suggests, among other things, "the development of better social norms, adoption of journalistic ethics by private speech platforms, or action by the political branches. For good measure, he posits a fresh assessment of "how the First Amendment handles efforts to promote healthy speech environments and protect listener interests."

Even if these answers seem esoteric, the underlying point is how important it is to recognize how the nature of networks, and the nature of data, has changed the nature of the threat. As Zeynep Tufekci puts it:

"Censorship during the internet era does not operate under the same logic it did under the heyday of print or even broadcast television. To be effective, censorship in the digital era requires a reframing of the goals of censorship not as a total denial of access, which is difficult to achieve, but as a denial of attention, focus, and credibility."[57]

Understanding this shift may be difficult, but Alan Rusbridger says that it is worth remembering that it has, ultimately, happened rapidly.

"I feel sort of oddly sympathetic to Zuckerberg, when I see the problems that journalism hasn't solved after 300 years," he says. "We can't really agree what the public interest looks like. We can't agree on what the business model looks like. An American journalist and a British journalist will not be

---

[56] Wu, Is The First Amendment Obsolete?
[57] Tufekci, *Twitter and Tear Gas*, 2017, p226-228

able to agree whether subjectivity or objectivity is a better way of truth. There are many things that journalism has failed to settle in 300 years of trying, and Zuckerberg and his chums have unleashed something which we're sort of expecting after 15 years to have come up with an answer. It just seems to be completely unreasonable."

So if we accept that the First Amendment's intention was to promote a healthy environment for speech to occur, and to protect the process of democratic deliberation from interference, and that networks need to keep working at improving that, then… what next?

It may first be necessary to understand that the rules of engagement are radically different than many, including legislators and service operators themselves, believe. As **Renee Di Resta** puts it,[58] the establishment of community guidelines, terms of service, and other protections have created a "digital maginot line." That is, a defense that only works if everybody adopts the same approach to warfare and established convention. When the rules are ignored – by a ground army which simply outflanks the line by invading from a different direction, or by information warriors who do not care whether they overstep the moral borders – the protections become immediately invalid.

Some services have found other ways to rebalance the scales, a little at least. In August 2014, Reddit had become the home for a leak of more than 500 nude celebrity photographs. First appearing on 4Chan and Imgur – two sites that Reddit had close relationships with – the images appeared to have been lifted from private storage services. Victims included actor Jennifer Lawrence and musician Ariana Grande. More than 100,000 people rapidly subscribed to the forum in which the images were being shared, and despite protests from many quarters, Reddit took a week to close it down.

Soon after, the site decided to ban "revenge porn"—intimate images posted online without the knowledge of their subjects. The site's approach revolved around taking two ideas that Reddit users held in high regard, and pitting them against each other: the right to privacy and the right to free speech. While many users were vocal about their right to say whatever they liked (even if it was egregious or offensive), even more agreed that a user's right to privacy should be the site's prime directive. The decision was unpopular with a noisy proportion of the site's users, but the logic was ultimately difficult to protest.

Next up was hate speech: the site shut down five subreddits that were considered grossly offensive and venues for harassment. One of them, /fatpeoplehate, had more than 150,000 subscribers. While many users continued to rebel (forcing out CEO **Ellen Pao** as a result) a group of researchers later decided to investigate what these decisions had actually meant for the site. What they found was clear: banning just a tiny portion of Reddit's most offensive forums had significantly increased the quality of debate on the site.

"Many more accounts than expected discontinued their use of the site; and, among those that stayed active, there was a drastic decrease (of at least 80%) in their hate speech use," said the authors. "Though many subreddits saw an influx of r/fatpeoplehate and r/CoonTown "migrants," those subreddits saw no significant changes in hate speech use. In other words, other subreddits did not inherit the problem."[59]

The ban worked. It's not clear whether it simply moved the behavior to other sites, but Reddit itself continued to grow—at various points in 2018, it was ranked as the 6th most popular website in the world.

---

[58] https://www.ribbonfarm.com/2018/11/28/the-digital-maginot-line/
[59] Chandrasekharan et al, You Can't Stay Here: The Efficacy of Reddit's 2015 Ban, 2017

Does this suggest there are ways for private platforms to take action? Doing so may require putting community standards above growth and short-term profit. It is true that many insiders at social networking companies are working hard to deploy policies to achieve similar results, and the leaders of such platforms make regular promises that they are working on it. (Many remain skeptical that the insulated leaders of Silicon Valley will act unless they are very personally affected, however.)

Caterina Fake, one of the co-founders of Flickr—which notably treated its network like a community, rather than a commodity—says the evidence is clear that good policy is worthwhile.

"I've gone through many of the distinctions of what makes an online community and social media, and social media makes it possible to sell people's attention," she says.[60] "That, I think, is the crucial distinction between it and online communities. You cannot be a part of an online community without participating in the community itself."

What good policy is will depend on the site and its users, she says—but above all, the decision-making process should be human and transparent.

"It is an art and not a science," she says. "You cannot algorithmically decide who is a good actor and who is not. Somebody has to make a call, somebody has to make an ethical judgment. And whether or not those people are people with a face... or it's a bunch of anonymous dudes in some backroom writing algorithms that decides who is right is who is wrong, those, I think are very different models of software design."

II. Regulate

Another option available to anyone addressing the overwhelming power and influence of social platforms is to impose some form of government regulation. After all, just because Facebook claims more than 2 billion users – bigger than any single country in the world – that doesn't mean it is actually a sovereign nation. John Perry Barlow's middle finger to authority didn't quite play out as he'd hoped, and Facebook is still subject to the laws of the countries it operates in.

In some cases, countries apply their own restrictions on speech, and companies that have significant business operations in those countries choose to abide by those rules. That's one reason for the complicated set of rules around content moderation—nudity or homosexuality might be socially acceptable in some nations, but illegal in others. In Germany and France, for example, there are specific laws developed after the Holocaust which are intended to prevent the rise of neo-Nazi movements.

In Singapore, regulations are more strict. The government maintains close control over the traditional media, and has traditionally tried to exert similar influence over social networks. But while it has rarely intervened by deleting or preventing speech, it has pursued legal action against many who offend its sensibilities: teenage blogger Amos Yee was imprisoned twice for "hate speech"—which is how it characterized his criticism of religion – and eventually fled the country in 2016 and claimed asylum in the U.S.

Not all of these interventions are culturally acceptable to American-based companies or Western users, and there are rightful concerns that localized rules result in pre-emptive censorship where companies struggling to keep up with billions of users adopt a lowest common denominator approach.

---

[60] Interview with Caterina Fake, November 2018

In 2012, Twitter began adopting a country-by-country system for taking down content—it would comply with a legal request or ruling from foreign governments if it was made in a nation where the site had employees or significant business operations. Two years later, under pressure from the Russian government, Twitter blocked the account of a Ukranian far right wing political party. Yet it had no employees there, and was therefore under no legal obligation to do so.

"We're sad to announce that Twitter has now caved," wrote Eva Galperin of the Electronic Frontier Foundation.[61] "Pravy Sektor's account is plainly political. If Twitter won't stand up for political speech in a country where independent media is increasingly under attack, what *will* it stand for?"

This certainly produces challenges when the state is acting against the best interests of its citizens and against the free flow of information that is the heart of the internet. In the Philippines, for example, several major social networking companies have significant numbers of employees (many working on moderation). President Duterte has demonized and attacked those dealing in free information, such as journalists and news organizations, while admittedly using sympathetic armies of online trolls to hound his rivals.

And in China, the government has shifted its approach away from isolation and outright bans on speech by dissidents, and toward a set of tactics that are themselves computational propaganda. This includes not *removing* dissenting speech but drowning it out or suppressing it through large-scale astroturfing campaigns.

What other legal remedies are available? Other laws include well-known tools for tackling dangerous levels of influence, such as antitrust. But there is, perhaps unsurprisingly, some resistance to this idea.

Is it even possible for social networks to be monopolistic? After all, Facebook and Twitter and Snap and so many others compete in the same kinds of spaces, and they all continue to exist. Even more fundamentally, aren't tweets, Facebook updates, MySpace posts, Instagram photos and snaps essentially the same thing? They're just strings of text or links or images, and even if they may have constraints like Twitter's character limit, or Instagram's need for a photograph. They are all competing for similar kinds of user attention with similar products.

The counterbalance to this might be to say that while these monopolies might not have fully developed, they are being built. In fact, they are almost inevitable, as the logical consequence of platform strategies and network effects. After all, since the value of a platform goes up as it gets bigger and more people use it, the platform owner gains a distinct advantage. They are then able to use their power to get even more powerful, creating a winner-takes-all dynamic that often raises concerns with antitrust regulators. (Witness Facebook's acquisition of Instagram and WhatsApp, identified as two of its potential future rivals.)

Even if that is accurate, though, some economists argue that this dominance is actually better for society overall. This is because it improves the efficiency of investment—what is known as "increasing returns to scale." Platforms which have control of the market might appear uncompetitive, but their dominance makes it easier for the market itself to grow.

In their book Big Is Beautiful, **Robert Atkinson and Michael Lind** suggest that this state of affairs is not just inevitable—it is actually preferable.

---

[61] Twitter Steps Down From The Free Speech Party, Electronic Frontier Foundation, 2014

"Having ten, or even frankly three, aviation firms each investing $15 billion to develop a 787-like jet would be a waste of societal resources," they argue, "As would having ten firms producing PC operating systems software, since all would have to invest considerable amounts in software programming, but each would have on average one tenth of the sales compared to just one firm."[62]

In fact, this is exactly how things developed during the early days of home computing. Dozens of companies – Acorn, Apple, Atari, Commodore, Hewlett-Packard, Olivetti, Sharp, Sinclair, Tandy, and many more – made competing hardware and operating systems. IBM up-ended this model by using an open architecture that allowed almost anyone to make hardware add-ons and compatible software, and Microsoft capitalized by taking it a step further and becoming the dominant operating system provider. Within a few years of this shift, nearly all other manufacturers were extinct or had switched to producing IBM or Microsoft-compatible products.

The consolidation did not restrict innovation. Conversely, it actually allowed consumers to be more confident that they were investing their money correctly, which boosted sales, which lowered the risk for new software businesses, which created many new jobs, which increased market potential, which encouraged more consumers, and so on. The result was positive overall, even if– in fact, *because* – the diversity of providers was dramatically reduced.

There is economic truth behind this argument for monopolistic behavior, but it also ignores the cultural impact of such dominance. When so much public debate is conducted over these networks, and platforms are able to tweak and shift what people see with a keystroke, the appearance of a natural monopoly presents an existential threat to political life and social cohesion.

The history of anti-trust almost parallels our general societal views on monopolies and their benefits. In the early days of American industry, influential companies such as Standard Oil and JP Morgan were able to grow and control vast swathes of the economy. What was good for them, it seemed, was good for the world. But when it became apparent that they had a potentially dangerous level of control, and were maintaining their dominance through coercion and price discrimination, laws were introduced to break up the worst offenders and restrict their influence and expansion. This approach continued through to the 1970s, and often affected the technology industry—one of the last major anti-trust actions was to break up the Bell System, at that time the single choke point for almost all communications services in America, and the biggest corporation in the nation's history.

By the 1980s and 1990s, however, attitudes towards shifted. As Tim Wu outlines in his book *The Curse of Bigness*, beliefs softened as a consumerist viewpoint became dominant. If low prices were a sign of healthy competition, then any actions that kept prices down—even if it came from a monopolistic company—were de facto good. This led to extreme levels of concentration in, industries like air travel and media, as fragmented markets rolled up into highly consolidated ones. It's no coincidence that the remnants of Standard Oil, like Exxon and Mobil re-merged in 1999. It also became extremely difficult in situations where the marginal cost of products was zero, like software. When Microsoft began bundling its Internet Explorer browser with its operating system and demanding licensing fees from computer manufacturers, it faced regulatory lawsuits in the U.S. and Europe, aimed at its predatory tactics, anti-competitive behavior, and lack of interoperability.

After an initial ruling against Microsoft in 1999, the court considered remedies and ruled in 2000 that the company should be broken up into two components, although a year later the Department of Justice reversed this decision and decided to settle instead. Attitudes did not shift entirely equally on

---

[62] Atkinson and Lind, *Big is Beautiful*, 2018, p210

both sides of the Atlantic, however. Europe had pursued its own case against Microsoft, and levied a range of huge fines—first €497 million, then another €280.5 million, then another €860 million—for the company's behavior and subsequent non-compliance. [See Exhibit 4, Speech by Neelie Kroes, European Commissioner for Competition Policy]

Although Microsoft remained intact, the damage was already done: Internet Explorer, while dominant, began to drift. The company faced new challenges from a resurgent Apple and a new rival in Google, both of which established dominance in areas that Microsoft was weak in, such as mobile. But while they have steered its behavior somewhat, these anti-trust actions have not prevented Microsoft from being successful: in 2018 it was been valued at more than double its 2000-era high. Intervention and punishment has just shifted the company's focus towards other core areas.

Smashing dominant and monopolistic companies into their constituent parts, or fining them in lieu of such action, is not the only possible regulatory enforcement however. Following Europe's example, some see the most viable regulatory idea as a kind of enforced interoperability.

Leaning on the computing concept of "multihoming," where a computer is connected to more than one network to increase reliability, some have proposed the idea that your social network data should be portable. Imagine being able to not just quit Facebook, but to export your list of friends so that you can take it to another service. Just like the early days of the web, where people would roam around to find a host whose approach to censorship was acceptable, users could look around to see which network was the best fit and go there.

It would certainly be less onerous than breaking apart, and let consumers make their own choices. Decentralization is already the idea behind some alternatives, such as the open source system Mastodon – a Twitter-like product in which users can not only create their own accounts, but their own networks with specific rules.

This question of interoperability has come up several times in the past. Around 2007, a number of companies "OpenSocial alliance" to promote a common set of standards for social networks to allow users to leap from one to another. The effort was backed in large part by Google, which had struggled to compete against Facebook with its own social networking products (over the years it tried and failed with products including Orkut, Google Buzz, and Google+.) But without any real gravity backing it, the effort struggled. Facebook, acting on its own and hell bent on growth, was able to outrun their plans and simply continue providing something that people enjoyed using more in the short term. The alliance went precisely nowhere.

Today, even after the long series of abuses of politics manipulation, personal abuse, and data scandals, users show few signs of really adopting a different stance. Mastodon gets a flood of users every time Twitter makes a public mistake (around 80,000 people signed up after Twitter equivocated on Alex Jones, for example.) But the same incentives remain: Facebook's stock has plummeted around 25% since the Cambridge Analytica revelations came to light in 2018… but it is still worth more now than it was in 2017, and many times what it was worth when it first hit the public stock market.

One fundamental difficulty with tackling this monopolistic behavior is that while we are led to believe that social networks are in competition with each other, they aren't really competing over the same set of goods.

That's because, for networked businesses like Facebook and Twitter, the product is not really the post, or the tweet, or the photograph, even if it seems that way. The product is the network—the connections between users, the links between posts, the avalanche of data created with every interaction, every share, every like. And each company has a virtual monopoly over its own data. Exporting your friends list, or your contacts, or your connections, will only give a very basic approximation of your value as a user.

"I think it's possible that something could be built that is decentralized and serves this function for people," says Evan Williams. But he says imagining this data layer as a protocol is at odds with the business interests of every existing social networking company.

"The biggest strategic mistake I made while I was CEO [of Twitter], was not putting the kibosh on the API sooner. People thought it would lead to innovation but actually it didn't lead to innovation, because protocols can't change... Our innovation layer became stilted, we became stilted."

And just as forms of openness are seen as antithetical to progress, so other forms of regulatory oversight are resisted. The technology industry spends vast amounts lobbying against onerous regulation and promising to fix itself. Today technology companies account for seven of the world's 10 most valuable businesses, with the big five from America (Apple, Amazon, Alphabet, Microsoft and Facebook) worth more than the next 13 companies combined.[63]

That is a lot of gravity to overcome.

III. Reimagine

We live in a world where small decisions quickly add up and become big, complex problems – often in ways we struggle to understand. YouTube's decision to delete a video by Alex Jones appears logical and relatively consistent, and the decision by other sites to do the same is equally straightforward. They have tools to confront rule violations—algorithmic deranking and moderation – but Jones could continue to have exact a negative toll on each service, even if he was subject to restriction or shadowbans. In this way, even the eventual, drastic deplatforming of Infowars – a sequence of decisions that almost entirely removed his work from popular sites – is logical.

But what happens when it's not a notorious conspiracy theorist in the firing line? What if it's a private individual, or a protected class, or a religious minority in an oppressive state? The obsession with creating rules and systems struggle whenever they are up against real life examples. There are always loopholes and backdoors and fringe cases to consider.

Lawyers and engineers, two dominant organizational professions of modern life, have created a set of interdependent systems to try and account for all these ethical questions. And yet the systems and behaviors continue to evolve, and the problems continue to grow.

As a result, private platforms seem caught between a set of competing and often irreconcilable concepts. Are they influential brokers of the public, or merely tools for communication? publishers or platforms? Are they censors or free speech advocates?

Instead of trying to come up with ways out of this labyrinth, they tend to adopt whichever guise is most beneficial to the company itself.

When touting their power and influence – including to their customers, advertisers – these platforms position themselves as public squares. When faced with criticism that they allow the public square to

[63] https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-value/

become polluted in ways that would be unacceptable in real life, these platforms claim that they are not really influential at all.

When asked to restrict or moderate the information propagated on their platforms, they claim to cleave to the democratic intent of American free speech principles. And when required to justify why they cannot remain tolerant of all kinds of speech – when the patience runs out for Alex Jones – they remind us that they are private entities and not organs of the state.

danah boyd, the Microsoft researcher, points out that many of the troubles we now see in the digital sphere exist because they were conceived as theories of zeroes and ones, not necessarily grounded in the real world, in the physical and often messy friction between humans.

When John Perry Barlow made his Declaration of Cyberspace Independence, she says, "you get these visions of community that are divorced from the physical realities of experience. You get some really strange dynamics there."

She makes a point of highlighting an old-style industrialist like Henry Ford.

"He was not an altruist – he was bastard – but he understood, without being told, that there was a thing called a second bottom line. He understood that he had to stabilize Detroit and Dearborn in order to have a functioning company. That meant he needed to build schools, and he needed to build roads, and he needed to build ways of getting everybody employed in the community because it wasn't just about keeping individuals employed at his factories, it was making sure that all their neighbors were employed too so they could actually stabilize the entire community so people could be happy. It wasn't a second bottom line. It was a business necessity."

Technology companies, on the other hand, have no such business necessity because their work is not rooted in the physical. "It's no longer about stabilizing community," she says. "Instead it's about a fantasy of community, a fantasy of identity-based communities. They don't know how to design for people for whom the locality mattered, for people who religion matters, for people who might feel disenfranchised and decide to go after and use these tools for this empowerment."

Caterina Fake, the Flickr co-founder, is now an investor herself, and has seen her own fair share of fast-growing companies. After moving on from Flickr, she spent eight years on the board of Etsy—itself a popular retail platform—and experienced rapid expansion as a co-founder Hunch, a networked Q&A service which was bought by eBay. And she believes that there are better examples out there than the ones people tend to reach for.

"I do think that your adoption of certain paths – that you're getting into bed with certain investors, that your decisions regarding the growth of your company, of how fast and how and under what circumstances – do affect your ability to guide things by moral orientation. Those decisions do affect you, who you hire, and why. But blaming it on rapid growth I do not think is actually a reasonable excuse, because we are all growing like crazy."

She points to companies she has experience of that have structured themselves precisely to better align with their users. Flickr focused on subscription, not on advertising. Etsy was structured as a B-Corp, which allowed it to focus on staying a sustainable, socially-responsible business (new management changed it to a more traditional status in 2017 as they sought to become "more profit-driven.")

And Kickstarter, the crowdfunding site in which Fake was an early investor, grew rapidly on the back of the same sort of network effects that excite social platforms. But it was deliberately structured to

stay in line with its user base's needs. Kickstarter became a Public Benefit Corporation in 2015.[64]

"It was designed into Kickstarter from the get go," she says. "[Co-founder] Perry Chen had specifically resisted taking investment from most of the people who were interested. He chose his business model, he chose his investors, he chose his team members, he chose his board members extremely carefully and deliberately so that they aligned with him and the values that he held."

For a long time, technology companies hewed to the letter of the law, rather than its spirit. Protecting themselves from liability while carefully manipulating attention was not problematic, as long as the Maginot Line remained intact. But a misunderstanding of bad actors, combined with a lack of understanding among consumers of data and damage, meant that decisions which seemed simple and entirely legal now feel like swampy ethical territory.

Talking about moral values may seem like an odd place to end up. After all, despite the criticism, Silicon Valley is not without them—the John Perry Barlows and Andrew Macgillivrays and the Mark Zuckerbergs have all blazed a trail for certain sets of values. But perhaps ultimately, tackling the dangerous influence of private platforms is about a different kind of value set entirely. Not one that involves hewing to a profit motive, or kneeling prostrate to the public good, but adopting a dedication to seeing all the potential problems, staring them in the eyes... and managing to find a way through anyway.

"Technology, and in particular social networks, have not been built with a realistic view of human nature," says the Center for Humane Technology's Sandy Parakilas.

"That can mean both that there is technology that is intentionally meant to exploit people's vulnerabilities, and there is technology that is built with good intention but because it doesn't understand human vulnerabilities people overuse it, or misuse it, or things happen that aren't intended. What we want to do is to move the entire industry towards a better, clearer, more accurate view of human nature so that those human vulnerabilities that are baked into us over millennia of evolution are being protected and not exploited."

---

[64] https://www.kickstarter.com/blog/kickstarter-is-now-a-benefit-corporation

Exhibit 1: YouTube Community Guidelines

Exhibit 2: CDA 230

Title II of the Communications Act of 1934 (47 U.S.C. 201 et
seq.) is amended by adding at the end the following new section:
`SEC. 230. PROTECTION FOR PRIVATE BLOCKING AND SCREENING OF
OFFENSIVE MATERIAL.

`(a) FINDINGS- The Congress finds the following:

`(1) The rapidly developing array of Internet and other
interactive computer services available to individual Americans
represent an extraordinary advance in the availability of
educational and informational resources to our citizens.

`(2) These services offer users a great degree of control
over the information that they receive, as well as the
potential for even greater control in the future as technology
develops.

`(3) The Internet and other interactive computer services
offer a forum for a true diversity of political discourse,
unique opportunities for cultural development, and myriad
avenues for intellectual activity.

`(4) The Internet and other interactive computer services
have flourished, to the benefit of all Americans, with a
minimum of government regulation.

`(5) Increasingly Americans are relying on interactive media
for a variety of political, educational, cultural, and
entertainment services.

`(b) POLICY- It is the policy of the United States--

`(1) to promote the continued development of the Internet and
other interactive computer services and other interactive media;

`(2) to preserve the vibrant and competitive free market that
presently exists for the Internet and other interactive
computer services, unfettered by Federal or State regulation;

`(3) to encourage the development of technologies which
maximize user control over what information is received by
individuals, families, and schools who use the Internet and
other interactive computer services;

`(4) to remove disincentives for the development and
utilization of blocking and filtering technologies that empower
parents to restrict their children's access to objectionable or
inappropriate online material; and

`(5) to ensure vigorous enforcement of Federal criminal laws
to deter and punish trafficking in obscenity, stalking, and
harassment by means of computer.

`(c) PROTECTION FOR `GOOD SAMARITAN' BLOCKING AND SCREENING OF
OFFENSIVE MATERIAL-

`(1) TREATMENT OF PUBLISHER OR SPEAKER- No provider or user

of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

　`(2) CIVIL LIABILITY- No provider or user of an interactive computer service shall be held liable on account of--

　　`(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

　　`(B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).

`(d) EFFECT ON OTHER LAWS-

　`(1) NO EFFECT ON CRIMINAL LAW- Nothing in this section shall be construed to impair the enforcement of section 223 of this Act, chapter 71 (relating to obscenity) or 110 (relating to sexual exploitation of children) of title 18, United States Code, or any other Federal criminal statute.

　`(2) NO EFFECT ON INTELLECTUAL PROPERTY LAW- Nothing in this section shall be construed to limit or expand any law pertaining to intellectual property.

　`(3) STATE LAW- Nothing in this section shall be construed to prevent any State from enforcing any State law that is consistent with this section. No cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section.

　`(4) NO EFFECT ON COMMUNICATIONS PRIVACY LAW- Nothing in this section shall be construed to limit the application of the Electronic Communications Privacy Act of 1986 or any of the amendments made by such Act, or any similar State law.

`(e) DEFINITIONS- As used in this section:

　`(1) INTERNET- The term `Internet' means the international computer network of both Federal and non-Federal interoperable packet switched data networks.

　`(2) INTERACTIVE COMPUTER SERVICE- The term `interactive computer service' means any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions.

　`(3) INFORMATION CONTENT PROVIDER- The term `information content provider' means any person or entity that is

responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.

  `(4) ACCESS SOFTWARE PROVIDER- The term `access software provider' means a provider of software (including client or server software), or enabling tools that do any one or more of the following:

    `(A) filter, screen, allow, or disallow content;

    `(B) pick, choose, analyze, or digest content; or

    `(C) transmit, receive, display, forward, cache, search, subset, organize, reorganize, or translate content.'.

12. Hate Speech

Policy Rationale

We do not allow hate speech on Facebook because it creates an environment of intimidation and exclusion and in some cases may promote real-world violence.

We define hate speech as a direct attack on people based on what we call protected characteristics — race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, gender identity, and serious disability or disease. We also provide some protections for immigration status. We define attack as violent or dehumanizing speech, statements of inferiority, or calls for exclusion or segregation. We separate attacks into three tiers of severity, as described below.

Sometimes people share content containing someone else's hate speech for the purpose of raising awareness or educating others. Similarly, in some cases, words or terms that might otherwise violate our standards are used self-referentially or in an empowering way. When this is the case, we allow the content, but we expect people to clearly indicate their intent, which helps us better understand why they shared it. Where the intention is unclear, we may remove the content.

We allow humor and social commentary related to these topics. In addition, we believe that people are more responsible when they share this kind of commentary using their authentic identity.

Do not post:

Tier 1 attacks, which target a person or group of people who share one of the above-listed characteristics or immigration status (including all subsets except those described as having carried out violent crimes or sexual offenses), where attack is defined as

• Any violent speech or support for death/disease/harm;

• Dehumanizing speech including (but not limited to): Reference or comparison to filth, bacteria, disease, or feces; Reference or comparison to animals that are culturally perceived as intellectually or physically inferior; Reference or comparison to subhumanity; Mocking the concept, events or victims of hate crimes even if no real person is depicted in an image; Designated dehumanizing comparisons in both written and visual form

Tier 2 attacks, which target a person or group of people who share any of the above-listed characteristics, where attack is defined as

• Statements of inferiority implying a person's or a group's physical, mental, or moral deficiency; Physical (including but not limited to "deformed," "undeveloped," "hideous," "ugly"); Mental (including but not limited to "retarded," "cretin," "low IQ," "stupid," "idiot"); Moral (including but not limited to "slutty," "fraud," "cheap," "free riders"); Expressions of contempt, including (but not limited to)"I hate""I don't like""X are the worst"; Expressions of disgust, including (but not limited to)"Gross""Vile""Disgusting"Cursing at a person or group of people who share protected characteristics

Tier 3 attacks, which are calls to exclude or segregate a person or group of people based on the above-listed characteristics. We do allow criticism of immigration policies and arguments for restricting those policies.

• Content that describes or negatively targets people with slurs, where slurs are defined as words commonly used as insulting labels for the above-listed characteristics.

13. Graphic Violence

Policy Rationale

We remove content that glorifies violence or celebrates the suffering or humiliation of others because it may create an environment that discourages participation. We allow graphic content (with some limitations) to help people raise awareness about issues. We know that people value the ability to discuss important issues like human rights abuses or acts of terrorism. We also know that people have different sensitivities with regard to graphic and violent content. For that reason, we add a warning label to especially graphic or violent content so that it is not available to people under the age of eighteen and so that people are aware of the graphic or violent nature before they click to see it.

Do not post:

• Imagery of violence committed against real people or animals with comments or captions by the poster that contain: Enjoyment of suffering; Enjoyment of humiliation; Erotic response to suffering; Remarks that speak positively of the violence; or, Remarks indicating the poster is sharing footage for sensational viewing pleasure

• Videos of dying, wounded, or dead people if they contain: Dismemberment unless in a medical setting; Visible internal organs; Charred or burning people; Victims of cannibalism

For the following content, we include a warning screen so that people are aware the content may be disturbing. We also limit the ability to view the content to adults, ages eighteen and older: Imagery featuring mutilated people if it contains the following in a medical setting

• Dismemberment

• Visible internal organs

• Charred or burning people

• Victims of cannibalism

• Throat-slitting

• Videos of self-immolation when that action is a form of political speech or newsworthy

• Photos of wounded or dead people if they show: Dismemberment; Visible internal organs; Charred or burning peopleVictims of cannibalismThroat-slitting

• Imagery featuring animals that shows: Dismemberment; Visible internal organs; Charred or burning animals

• Videos of animal abuse, defined as: Repeated kicking or beating of a live animal by a personActs of torture by a person committed against animalsRepeated biting of a live animal by a person for abusive purposesRepeated animal-to-animal biting in staged fightsVideos of animals being killed in a hunting, manufacturing, or food preparation/processing context

• Videos that show child abuse, defined as: Repeated kicking, beating, slapping, or stepping on by an adult or animal; Strangling or suffocating by an adult or animal; Drowning by an adult or animal; Biting through skin by an adult or animal; Poisoning by an adult; Forcible restraint by an adultInflicting of burn or cut wounds by an adultForcible smokingTossing, rotating, or shaking of an infant (too young to stand) by their wrists/ankles, arms/legs, or neck

• Videos that show the violent death of a person or people by accident or murder when they depict the deceased body

• Still images depicting the violent death of a person or people

• Videos that show acts of torture committed against a person or people

• Videos of physical bullying or violence against minors in a fight context shared with a condemning caption

14. Adult Nudity and Sexual Activity

Policy Rationale

We restrict the display of nudity or sexual activity because some people in our community may be sensitive to this type of content. Additionally, we default to removing sexual imagery to prevent the sharing of non-consensual or underage content. Restrictions on the display of sexual activity also apply to digitally created content unless it is posted for educational, humorous, or satirical purposes.

Our nudity policies have become more nuanced over time. We understand that nudity can be shared for a variety of reasons, including as a form of protest, to raise awareness about a cause, or for educational or medical reasons. Where such intent is clear, we make allowances for the content. For example, while we restrict some images of female breasts that include the nipple, we allow other images, including those depicting acts of protest, women actively engaged in breast-feeding, and photos of post-mastectomy scarring. We also allow photographs of paintings, sculptures, and other art that depicts nude figures.

Do not post:

Images of

• Real nude adults, where nudity is defined as: Visible genitalia; Visible anus and/or fully nude close-ups of buttocks unless photoshopped on a public figure; Uncovered female nipples except in the context of breastfeeding, birth giving and after-birth moments, health (for example, post-mastectomy, breast cancer awareness, or gender confirmation surgery), or an act of protest

• Sexual activity, including: Sexual intercourse; Explicit sexual intercourse, defined as mouth or genitals entering or in contact with another person's genitals or anus, where at least one person's genitals are nude; Implied sexual intercourse, defined as mouth or genitals entering or in contact with another person's genitals or anus, even when the contact is not directly visible, except in cases of a sexual health context, advertisements, and recognized fictional images or with indicators of fiction

• Other sexual activities including (but not limited to): Erections; Presence of by-products of sexual activity; Stimulating genitals or anus, even if above or under clothing; Use of sex toys, even if above or under clothing; Stimulation of naked human nipples; Squeezing naked female breast except in breastfeeding context; Fetish content that involves; Acts that are likely to lead to the death of a person or animal; Dismemberment; Cannibalism; Feces, urine, spit, menstruation, or vomit

• Digital content that meets our definition of sexual activity unless any of the following conditions exist: Content was posted in a satirical or humorous context; Content was posted in an educational or scientific context; Imagery is not sufficiently detailed and only body shapes or contours are visible

• Sexually explicit language, defined as description that goes beyond mere mention of: A state of sexual arousal; An act of sexual intercourse unless posted in an attempt at humor or satire, or if educational in nature

15. Cruel and Insensitive

Policy Rationale

We believe that people share and connect more freely when they do not feel targeted based on their vulnerabilities. As such, we have higher expectations for content that we call cruel and insensitive, which we define as content that targets victims of serious physical or emotional harm.

Do not post:

• Content that depicts real people and mocks their implied or actual serious physical injuries, disease, or disability, non-consensual sexual touching, or premature death

Exhibit 4: Speech by Neelie Kroes, European Commissioner for Competition Policy

On the decision to impose € 899 million penalty on Microsoft for non-compliance

Brussels, 27th February 2008

Today the European Commission has imposed a substantial fine – to be precise € 899 million – on Microsoft for its non-compliance up until 22 October 2007 with its obligations under the Commission's March 2004 Decision to provide interoperability information on reasonable terms.

Today's decision finds that, prior to 22 October 2007 Microsoft charged unreasonable prices for access to indispensable interoperability information.

In plain English this is to say that Microsoft continued to abuse its powerful market position after the Commission's March 2004 decision requiring it to change its practices. Microsoft continued to stifle innovation by charging other companies prohibitive royalty rates for the essential information they needed to offer software products to computer users around the world. Charging such an unreasonable price effectively rendered the offer of the information pointless.

Microsoft's behaviour did not just harm a few individuals or a handful of big companies ... directly and indirectly this had negative effects on millions of offices in companies and governments around the world.

It gives me no pleasure to be here again announcing such a fine; nor to see that a company flouted the law when it had many opportunities to choose other courses of action. But given the nature of Microsoft's non-compliance – considering both its effects and its duration, today's fine is proportionate and necessary.

There are lessons that I hope Microsoft, and any other company contemplating similar illegal action, will learn.

Talk is cheap; flouting the rules is expensive.

We don't want talk and promises - we want compliance.

If you flout the rules you will be caught, and it will cost you dear.

This is the third time in four years that we have had to impose fines or penalty payments on Microsoft.

The first fine was for the abuse itself – everything up until March 2004.

Then the Commission had to impose a penalty payment in July 2006 for non-delivery of complete and accurate interoperability information. That was the first time that a penalty payment for non-compliance with a Commission decision had been necessary in fifty years of EC Competition Law.

And finally today's penalty payment for unreasonable royalty rates up until October 2007. That the Commission has been forced to levy these three fines reflects a clear disregard by Microsoft of its legal obligations.

The Commission's fine is a reasonable response to a series of quite unreasonable actions.

Microsoft has gone from insisting on charging a 2.98% royalty for non-patented interoperability information to charging a flat fee of €10 000 for the same information.

Finally, after three years of illegal behaviour, it appears that Microsoft has come into compliance with the 2004 decision.

Today's decision and penalty payment are about the 2004 Decision only, and are not about any of Microsoft's other business practices.

I must stress that distinction in the clearest possible terms.

You will, of course, be aware that the Commission opened two antitrust inquiries in January looking at several of Microsoft's other practices. One concerns interoperability information across a broad range of products, including Microsoft's Office suite and a number of its server products. The other concerns tying of Internet Exporer and other separate software products with the Microsoft PC operating system.

These investigations are separate and continuing – I can offer no comment until the investigation is complete.

As always, we will take into account any changes that Microsoft makes to its business practices that are relevant to those investigations.

We take positive changes to business practices very seriously.

But again, I stress that a press release, such as that issued by Microsoft last week on interoperability principles does not necessarily equal a change in a business practice.

If change is needed – and as I say, I have reached no conclusion on that – then the change will need to be on the market. Not in the rhetoric.

To finish I think it is important to share a little insight into our consistent approach to illegal behaviour.

It's not my job to tell people or companies what to buy … but it is my job to ensure they have a choice about what they can buy.

Competition policy is about making sure that people and companies have the right to choose.

And when people and companies are given that right to choose, then markets deliver so much more – this is why I am passionate about making markets work better.

Today's decision sends a strong signal that consumers can expect choice in Europe, and can expect the European Commission to fight for their right to choose.

Thank you.