"We Don't Know What We Want": The Tug between Rights and Public Health Online

Jonathan Zittrain

Follow this and additional works at: https://dsc.duq.edu/dlr

Part of the First Amendment Commons, Internet Law Commons, Law and Society Commons, and the Science and Technology Law Commons

Recommended Citation
Jonathan Zittrain, "We Don't Know What We Want": The Tug between Rights and Public Health Online, 61 Duq. L. Rev. 183 (2023).
Available at: https://dsc.duq.edu/dlr/vol61/iss2/2

This Article is brought to you for free and open access by the School of Law at Duquesne Scholarship Collection. It has been accepted for inclusion in Duquesne Law Review by an authorized editor of Duquesne Scholarship Collection. For more information, please contact beharyr@duq.edu.
"We Don’t Know What We Want":
The Tug Between Rights and Public Health Online

Jonathan Zittrain*

Twitter and Facebook boast billions of subscribers, many of whom are real people. The companies are also roundly hated, particularly by tech experts—at least those who follow them for something other than their stock performance.1 Objections to platforms’ behavior are commonly expressed as amazement that they could be so obviously and consistently wrong in failing to police awful content their users post. There is also amazement about unobjectionable posts and comments from users that they take down.2 That, in turn, has led to pressure for regulatory initiatives to push the companies into doing what they so clearly ought to be doing in the first place.

1. 64% of Americans say social media has a mostly negative effect on the way things are going in the country today. Only 10% said mostly positive effect; 25% said neither negative nor positive effect. Study also shows people think social media has a mostly negative effect (misinformation was the most cited reason, cited by 28% of the 64% who said there was a negative effect). Brooke Auxier, 64% of Americans say social media have a mostly negative effect on the way things are going in the U.S. today, PEW RSCH. CTR. (Oct. 15, 2020), https://www.pewresearch.org/fact-tank/2020/10/15/64-of-americans-say-social-media-have-a-mostly-negative-effect-on-the-way-things-are-going-in-the-u-s-today/ [https://perma.cc/UB7J-XMMNJ]. A 2019 WSJ/NBC poll showed US adults across different age groups and political ideologies held a negative view of the effects of social media (even though 70% continue to use these services at least once a day). “[T]hey regard services such as Facebook to be divisive and a threat to privacy . . . .” Survey results show a majority of US adults think social media does more to: divide us, waste our time, spread lies/falsehoods, and spread unfair attacks/rumors. Results also showed low trust in tech companies, particularly Facebook (FB) (trusted less to protect people’s personal information than Amazon, Google, and even the federal government!). John D. McKinnon and Danny Dougherty, Americans Hate Social Media but Can’t Give It Up, WSJ/NBC News Poll Finds, WALL ST. J. (Apr. 5, 2019), https://www.wsj.com/articles/americans-agree-social-media-is-divisive-but-we-keep-using-it-1155445660 [https://perma.cc/9MRY-MSLS].


* Jonathan Zittrain is the George Bemis Professor of International Law at Harvard Law School, Professor of Public Policy at the Harvard Kennedy School of Government, Professor of Computer Science at the Harvard School of Engineering and Applied Sciences, Director of the Harvard Law School Library, and Co-Founder of the Berkman Klein Center for Internet & Society. He is the recipient of the 2023 Duquesne Kline Law School John and Liz Murray Excellence in Scholarship Award. This article is adapted from a book-in-progress on the future of digital governance, with permission from Penguin Press.
If there is to be a shot at understanding what social media is doing for us—and to us—and what boundaries there should be on how they act, we need to more closely examine what we believe is so “obviously” wrong with them. This turns out to be more elusive than you would think. Not only is consensus about platform problems absent among us—perhaps not entirely surprising in an era in which basic facts are in deep dispute—it is also quite commonly absent within us.

For a pretty simple example of a lack of consensus, consider how you might feel upon learning that photos of animal abuse are circulating on Facebook. At first you or I might condemn the platform’s irresponsible abdication of responsibility by leaving them up, or worse, amplifying them. But the same photos shared for the purpose of highlighting the depravities of trained dogfighting can evoke a view on platform censorship and amplification opposite from that of the very same picture shared for the purpose of valorizing dogfighting—as can a photo for the purpose of placing a dog with a new adoptive home, versus selling that dog. For years, Facebook has had controversial policies on just this sort of thing that try to respond to context, along with reproof for failing to well enforce them.\(^3\)

(One account from 2019 claims 136,000 photos are uploaded to Facebook every minute; this could make such subtle policies difficult to apply consistently.\(^4\))

Or consider Clearview AI, a startup that scraped more than three billion images of people across the Internet, such as Facebook and LinkedIn profile pictures, along with names.\(^5\) The company then used a machine learning technique so that it could retrieve any matching photos (and, often, names) when presented with a new one. As I wrote in a 2020 *Washington Post* essay inveighing against Clearview’s activities and very existence:

The upshot? The fundamental comfort—and liberty—of being able to walk down a street or enter a supermarket or stadium without the authorities, or fellow strangers, immediately knowing who you are is about to evaporate.

---

without any public debate about whether that’s okay. It’s as if someone invented eyeglasses that could see through walls, sold them to a select few, and everyone else inexplicably shrugged.\(^6\)

I think what Clearview did was reprehensible and likely illegal. I am not alone; there is a reason its massive database was compiled by an unknown company rather than by a sedate stalwart like Microsoft or Oracle or even Palantir. I think Clearview’s database should be destroyed and the earth under it salted, with legal and technical measures taken not to permit its private or public recreation.

But that does not mean they are not useful, whether for something as straightforward as solving a crime, or for more novel applications. For the latter, Clearview is said to have been used to help identify Russian soldiers killed in Ukraine whose relatives are unaware of their fates.\(^7\) For the former, in January 2021, a mob of Trump supporters, some peaceful and many not, pushed and beat its way through light Capitol Police cordons to break into the Capitol itself, some in search of members of Congress to attack. Law enforcement reinforcements were slow enough in coming, and the situation volatile enough, that efforts to regain control of the Capitol complex entailed simply gently ushering people out rather than arresting them, even as Congressional demonstrators in earlier contexts had been arrested for as little as laughing during a hearing.\(^8\)

After the insurrection, authorities were left with the daunting puzzle of identifying hundreds of rioters who had converged in Washington from around the country, largely on the basis of photos and video taken during the riots—most by the rioters themselves. Among the tools used: Clearview AI.\(^9\) In those cases in which Clearview AI might be the only way to make an identification, and instantly at that, is it a good thing that my and others’ calls against Clearview were not heeded? For our purposes here, the practical case study is susceptible to a Tolkienesque challenge: those who

---

6. Id.
push technology to its ethical and legal limits forge artifacts of unprecedented power, and experience a temptation to use them as deep as their potential for corruption and misuse.

With respect to content moderation, the core confusion within a shell of certitude can be explained by a number of factors: the distinct (and now absent) context of American beliefs about the primacy of free speech in the 1990s, during which the Internet went mainstream; our tendency to generalize too much from our own singular experiences; the ways in which social media has newly and vastly empowered intermediaries in ways both users and platforms have not dealt with before; an interregnum of about twenty years in which American law largely took a pass on considering, much less making, new demands of platforms; and most important, the fact that ethically speaking, there is good reason to be tugged in competing directions when thinking through what a healthy public sphere looks like.

Among those many directions in which to be tugged, two stand out. The first is a rights framework—specifically around the right to free speech as developed in late twentieth century American law and culture. The second framework is that of public health, both by metaphor and, more recently in the wake of a pandemic, quite literally.

"FREE SPEECH MUST BE DEMANDED FOR ALL":
THE RIGHTS ERA, 1995-2010

The rights framework dominated the conventional discourse—and corresponding law—about internet responsibilities from the 1990s to at least 2010. It has innate appeal, pushing for individual rights and free speech online above all else, irrespective of whether it appears to some or most observers not to have much benefit for anyone. To see it in a recent snapshot, consider the case of Lasse Gustavson. Forty years ago, he was a newly minted firefighter in Sweden. During his first week on duty, Gustavson responded to the scene of a gas leak and was severely injured when twenty tons of propane exploded. After spending a few months in a coma, he recovered, but lost his ears, hair, and fingers, and had extensive burns and scarring. Wanting to share his journey, he became a motivational speaker. 10 A few years ago, on the occasion of his sixtieth

birthday, a friend posted a photo of Gustavson on Facebook.¹¹ Within an hour Facebook removed the photo, without explanation. The friend posted it again, and it was removed again. The friend began publicly agitating with a new post sans photo—on Facebook, of course—over the removals, while lodging an appeal with Facebook by using the “report a problem” box on the site. This post garnered more than thirty thousand shares, and the friend then received a note from Facebook that a “member of our team accidentally removed something you posted on Facebook. This was a mistake, and we sincerely apologize for this error. We’ve since restored the content, and you should now be able to see it.”¹²

Many burn survivors had similar experiences with removal of their photos; apparently Facebook’s actions were company policy.¹³ “Apparently” because the rules for acceptable posting were at the time expressed only at a general level. More detailed policies existed—after all, Facebook’s own content moderators needed concrete guidance on what to allow and what to take down—but they were not publicly released until after some internal slides documenting them leaked in 2017.¹⁴

It’s hard to imagine a persuasive defense of Facebook’s decision to—and apparent policy of—removing photos of people who happen to be scarred. Certainly, Facebook never offered one; instead, once there was some measure of outcry, it claimed an accidental removal and apologized. More broadly, many might ask what business Facebook has at all in judging the speech of its users. To remove these photos was not, by these lights, a misapplication of an otherwise-noble content restriction policy. Rather, short of moderating illegal content, which offers its own set of puzzles, the thought goes, Facebook should not be judging speech at all. We would be horrified if a telephone company were to enforce an “acceptable content policy” on telephone calls—including conference calls—nor would we want to hold the telephone company responsible for paper posters that agitators might staple to its physical telephone poles.

¹². Id.
To invite that kind of moderation is to ask for a suffocating form of censorship, even if it is “merely” censorship by a private company rather than the government. I described it in 2020 this way:

Although we don’t have consensus about what we want, no one would ask for what we currently have: a world in which two unelected entrepreneurs are in a position to monitor billions of expressions a day, serve as arbiters of truth, and decide what messages are amplified or demoted. This is the power that Twitter’s Jack Dorsey and Facebook’s Mark Zuckerberg have.  

Facebook’s business model has generally reflected a distaste for content moderation. It is simpler to simply emphasize connecting users to one another, ricocheting their posts around in proportion to how engaging others would find them, and exposing everyone to tailored advertising while on and off the site. Facebook found itself following this philosophy in the heart of the online rights era, when people were still learning how to share content from family events to homemade meals, and it was yet unclear how that content might be used for harm. But abstention also has a philosophical basis in individual speech rights that is in turn reinforced by American law. Before turning, then, to the public health framework for regulating online activities, I want to offer an account of the spread of free-speech-oriented thinking from its pre-Internet mainstream legal incarnations to broader popular culture, through to the architecture and law of online discourse.

THE SKOKIE MARCH

Play word association with Skokie, Illinois, with many Americans over 50 and the reaction is instantaneous: Nazis. For those under 50, a blank stare. So, a quick review of the Skokie-Nazi episode is a fitting introduction to the Rights Era, as it near-perfectly encapsulates the free speech marinade that had been so thoroughly absorbed by the early Internet. It provides a bookend to new questions around the censorship of neo-Nazi activity in the wake of a “Unite the Right” march in Charlottesville, Virginia, in 2017, and the attack by Trump supporters on the Capitol in January of 2021.


In 1977, a group of neo-Nazis announced plans to picket on the sidewalks in front of the town hall of Skokie, Illinois, wearing World War II-era German Nazi storm trooper uniforms with swastika armbands. More than half of Skokie’s residents at the time were Jewish, and several thousand were survivors of the Holocaust. On the other side, the American Civil Liberties Union (ACLU) represented the Nazis, arguing that the First Amendment compelled the government to allow the march.

The trial court sided with the town, sparking appeals that worked their way to the United States Supreme Court and then back to the Illinois Supreme Court. There, the ACLU’s free speech arguments prevailed, although after winning the right to protest, the Nazis ultimately decided to make their march in nearby downtown Chicago, rather than Skokie.

The ACLU’s national director, Aryeh Neier, was himself a Jewish refugee of Nazi Germany. His defense of free speech rings across the decades:

Did Jehovah’s Witnesses or birth control advocates have a right to pass out leaflets in Catholic neighborhoods? . . . Did Martin Luther King Jr. have a right to march in Selma, Alabama, or in Cicero, Illinois? To all of these questions, the A.C.L.U.’s answer is ‘Yes.”

As a New York Times editorial put it at the time, echoing Neier’s arguments, “if his organization is not faithful to the principle that

19. Id. at 350–51.
free speech must be demanded for all, then it does not deserve the words 'civil liberties' in its name."25

The Illinois Supreme Court, in ultimately supporting the Nazis’ right to march, rounded up high-minded quotations from a number of U.S. Supreme Court decisions in First Amendment cases:

It is firmly settled that under our Constitution the public expression of ideas may not be prohibited merely because the ideas are themselves offensive to some of their hearers.

The constitutional right of free expression is powerful medicine in a society as diverse and populous as ours. It is designed and intended to remove governmental restraints from the arena of public discussion, putting the decision as to what views shall be voiced largely into the hands of each of us, in the hope that use of such freedom will ultimately produce a more capable citizenry and more perfect polity and in the belief that no other approach would comport with the premise of individual dignity and choice upon which our political system rests.

To many, the immediate consequence of this freedom may often appear to be only verbal tumult, discord, and even offensive utterance. These are, however, within established limits, in truth necessary side effects of the broader enduring values which the process of open debate permits us to achieve. That the air may at times seem filled with verbal cacophony is, in this sense not a sign of weakness but of strength. We cannot lose sight of the fact that, in what otherwise might seem a trifling and annoying instance of individual distasteful abuse of a privilege, these fundamental societal values are truly implicated. “[S]o long as the means are peaceful, the communication need not meet standards of acceptability.”26

There were plenty of people who condemned this defense of free speech. Tens of thousands of ACLU members resigned over its decision to take up the Skokie case, and for years after, the observation that someone was a “card-carrying member of the ACLU” was intended as an insult—a claim of radicalism. George H.W. Bush

used it in his stump speech when running for President against Michael Dukakis. But Dukakis had already proudly described himself that way, and he did not back down. (To be sure, he lost his Presidential race.) The ACLU seized on the issue to successfully raise money, producing a public service announcement including famed actor Burt Lancaster. "You know the kind of people who support the ACLU," Lancaster says: "Radicals’ like Douglas MacArthur, Dwight Eisenhower, Harry Truman." 27

The Skokie case was a near-perfect encapsulation of mainstream late twentieth century characterization of the right to free speech in America. The courts and the ACLU couched this case, and others like it, as a grand principle of free speech versus the parochial Puritanism of those who might be scandalized, perhaps enough to themselves be incited to violence, and allow that offense to translate into an inappropriate use of state power against citizens on the basis of their (concededly awful) views. In the Skokie case, no one argued that the handful of Nazi marchers could turn violent in ways that the police could not handle. Nor was Nazism deemed in danger of spreading as a result of publicity from a peaceful march. New York Times columnist William Safire drew from the argument of the town of Skokie itself when he identified any threat of violence as instead coming from among the thousands of counter-protesters who would appear: "The American Nazis’ object was, and is, to trigger a violent counterdemonstration, thereby making themselves martyrs at the hands of Jews shouting, 'Never again!'" 28

And that threat of violence could not be the basis for banning the original march. The heckler’s veto should not triumph. As the Illinois Supreme Court put it: "We accordingly, albeit reluctantly, conclude that the display of the swastika cannot be enjoined under the fighting-words exception to free speech, nor can anticipation of a hostile audience justify the prior restraint." 29

Instead, suggested Safire, the citizens of Skokie should see the wisdom of allowing the Nazis to march. He thought it would demonstrate a commitment to civil liberties that would, ironically for the marchers, undermine the Nazi cause and others like it. "Grumbling all the way, I have to agree: There can be no greater affirmation of freedom than ostentatiously to respect the rights of those who

would destroy that freedom.” As the ACLU’s Neier put it in a symposium ten years later:

> It was a twist of fate that placed me in a spot where I was engaged at one stage in my life as the defender of civil liberties for Nazis. But I do not believe that I was deceiving myself when I asserted then, as I would assert today, that the defense of rights for all, even Nazis, is just what is needed to ensure that Nazism never again prevails.

Cases like Skokie were perfect vehicles for civil libertarians to extol higher values over knee-jerk instincts, in part because no one would reasonably think their support of Nazis’ rights to march arose because they shared the Nazis’ views. Hence the Illinois court’s “reluctant” conclusion, and Safire’s “grumbling” agreement with it.

The Skokie episode tells us a lot about the origins of digital governance, because the ethos of American civil libertarianism that reached apogee in the 1960s and 70s political and legal establishments in turn infused the mainstream use of the Internet.

AN ENSHRINEMENT OF ONLINE SPEECH RIGHTS

Indeed, the free speech values of the Skokie case were echoed by many of those promoting and celebrating the widespread adoption of the Internet beginning in the 1990s. John Perry Barlow might have been sent by Central Casting for this purpose. Barlow was a sometime Wyoming cattle rancher, a lyricist for the Grateful Dead, and a co-founder of the Electronic Frontier Foundation (EFF), a digital rights non-profit organization on whose board I’ve served.

In 1996, a time of deep interest around what kind of revolution the Internet portended, a slightly tipsy Barlow penned “A Declaration of the Independence of Cyberspace.” It is well known among

---


34. Cindy Cohn, *Inventing the Future: Barlow and Beyond*, 18 DUKE L. & TECH. REV. 69–77 (2019). (“In talking about the Declaration at Electronic Frontier Foundation (EFF) many years later, Barlow admitted that when he stepped out of a party at Davos to write it, he was both a little drunk and trying desperately to channel Thomas Jefferson.”) *Id.* at 70.
Internet history aficionados, not so much with the public at large. It says:

Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.

We have no elected government, nor are we likely to have one, so I address you with no greater authority than that with which liberty itself always speaks. I declare the global social space we are building to be naturally independent of the tyrannies you seek to impose on us. You have no moral right to rule us nor do you possess any methods of enforcement we have true reason to fear . . . .

We are creating a world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity . . . . In our world, all the sentiments and expressions of humanity, from the debasing to the angelic, are parts of a seamless whole, the global conversation of bits. We cannot separate the air that chokes from the air upon which wings beat . . . .

We will create a civilization of the Mind in Cyberspace. May it be more humane and fair than the world your governments have made before.35

Barlow’s argument’s power in part relied on aptly characterizing the digital space as being one of speech rather than action—host to a “global conversation of bits” and a home of “Mind”—just as Skokie rightly characterized marches as expressive activities offering a range of ideas, none of which should be judged by the governments in a position to license them. Indeed, if Skokie’s expression was to be allowed even at the risk of violence by those physically counterprotesting, how could there be objection to speech lobbed into one another’s screens at an even safer distance?

By this reckoning, the Internet stood to allow for a vast expansion of available sidewalk space, a realization of a vision that discussion and expression should not be confined to—or by—the powerful, nor,

as much as possible, by the physical constraints of who happens to be nearby to experience it. Thanks to the Internet, people could choose to see the equivalent of a lone zealot’s sandwich board as readily as they could tune in to Walter Cronkite’s national evening newscast. The distasteful or debasing was simply a (perhaps reluctant and grumbling) price to be paid for this freedom, inextricable from its exercise.

Barlow was no outlier in praising this arrangement; the Supreme Court itself championed a First Amendment rights discourse for online activities. On the very day in 1996 that Barlow published his Declaration of the Independence of Cyberspace, a new Federal law called the Communications Decency Act (CDA), part of a broader Telecommunications Reform Act, went into effect. The CDA primarily concerned itself with the unhappy fact that kids could newly gain access to pornographic content online that would have previously been much more difficult for them to get.

That the first grand legal battle over online speech rights would concern pornography might not be surprising. A 1994 *New York Times* article entitled “Porn, the Low-Slung Engine of Progress,” declared that “In the history of communications technology, sex seems to be the most enduring killer app . . . . Sometimes the erotic has been a force driving technological innovation; virtually always, from Stone Age sculpture to computer bulletin boards, it has been one of the first uses for a new medium.” That online free speech would first be prominently fought on those grounds reinforced the narrative of prim provincialists trying to fan moral panic over the natural progression of loosening societal norms—and failing that, trying to hold everyone else to their blinkered views. Uptightness and taking offense were weak counters to the soaring ideals of free speech, and they would fail as surely as the attempts of archetypical parents of a 1950’s Pleasantville to stop their kids’ embrace of rock and roll did.

But the CDA would try. Its most powerful intervention was to criminalize the posting of such content, even when legal for adults to consume, if the poster failed to ensure that kids could not get to it—as represented by checking to see if an Internet user had a functioning credit card before granting access. (There was no need to charge anything to the card, but merely to validate that it worked,
on the then-credible hypothesis that most kids under 18 did not have credit cards.) The CDA's core provisions seemed entirely at odds with Barlow's Declaration. After all, the CDA was a form of censorship by a "Government of the Industrial World." Barlow certainly saw it that way, and his contemporaneous release of the Declaration was no coincidence. He wrote:

> In the United States, you have today created a law, the Telecommunications Reform Act, which repudiates your own Constitution and insults the dreams of Jefferson, Washington, Mill, Madison, De Tocqueville, and Brandeis. These dreams must now be born anew in us.

You are terrified of your own children, since they are natives in a world where you will always be immigrants. Because you fear them, you entrust your bureaucracies with the parental responsibilities you are too cowardly to confront yourselves.\(^\text{39}\)

While Barlow lyricized free speech, the ACLU, the EFF, the American Library Association, Human Rights Watch, and others put his lyrics to lawyerly music in filing suit to block the CDA.\(^\text{40}\) Their arguments echoed those of the Skokie case and Barlow's Declaration, and focused on the fact that the online pornography criminalized by the CDA had already been deemed by the Supreme Court to be protected by the First Amendment with respect to adults consuming it. It was only with respect to kids that the authorities could seek to ban it, and then only in ways not unduly burdensome to adults.\(^\text{41}\)

To be sure, that kind of harmful-to-minors pornography in question is not to be confused with entirely obscene materials, which the Supreme Court had found to be not protected at all by the First Amendment.\(^\text{42}\) Obscene content, the Court had said, was a narrow category of work which depicts sexual conduct in a "patently offensive way" that appeals to the "prurient interest" and that "lacks serious literary, artistic, political, or scientific value."\(^\text{43}\) (This

\(^\text{39. Barlow, supra note 35.}\)
\(^\text{42. Miller v. California, 413 U.S. 15, 36 (1973).}\)
\(^\text{43. Id. at 24–25.}\)
standard eclipsed Justice Potter Stewart’s concurring formulation that “I know it when I see it.”44 So Playboy or Penthouse magazines could be banned from carrying obscene content, and adults could be punished in many instances for possessing it, but those magazines still had lots of legal pornography they could offer, so long as there were efforts made by distributors to keep it away from kids. Stores could sell Playboy magazines, and states could insist that they be placed on a high rack in the store and the shopkeepers verify that someone was an adult before allowing them to purchase it.45

The CDA’s censorship was viewed skeptically by the special three-judge lower court that first weighed in on it. “Cutting through the acronyms and argot that littered the hearing testimony, the Internet may fairly be regarded as a never-ending worldwide conversation,” wrote one concurring judge hearing the case along its way to the Supreme Court.46 “[J]ust as the strength of the Internet is chaos, so the strength of our liberty depends upon the chaos and cacophony of the unfettered speech the First Amendment protects,” he added.47 “For these reasons, I without hesitation hold that the CDA is unconstitutional on its face.”48

The Supreme Court also agreed with the civil liberties organizations that the CDA’s core provisions, those criminalizing harmful-to-minors pornography in the absence of something like a credit card check, were too restrictive.49 Under those provisions, too much content would be swept behind gates online or eliminated entirely, whether through overly-cautious responses to the law—for example, restricting access to materials about breast cancer—or through withholding of legal pornography itself. “As a matter of constitutional tradition, in the absence of evidence to the contrary, we presume that governmental regulation of the content of speech is more likely to interfere with the free exchange of ideas than to encourage it,” wrote Justice John Paul Stevens for the Court.50 “The interest in encouraging freedom of expression in a democratic society outweighs any theoretical but unproven benefit of censorship.”51

45. Ginsberg, 390 U.S. at 629.
47. Id.
50. Id. at 889.
51. Id.
Court struck down the CDA’s core provisions before they came into effect.  

Team Cyberspace Free Speech spiked the football. “The case is likely to be the Brown v. Board of Education of cyberspace[,”] said the ACLU’s Barry Steinhardt, grandiosely invoking the unanimous landmark 1954 Court decision striking down legally-mandated racial segregation in schools.

The Supreme Court set an extraordinarily high barrier against further attempts to limit online expression in the United States. The decision has had an impact around the world. I have had the opportunity to discuss it with jurists from Bulgaria to South Africa and they view it as an important precedent for their own countries.

Realspace revelers turned out in San Francisco shortly after the decision came down. “Let today be the first day of a new American Revolution—a Digital American Revolution!” declared EFF attorney Mike Godwin.

Thus, as the Internet found mass public adoption—or, perhaps in more recent parlance, solidified its grip upon the public’s consciousness—the values and rhetoric of rights exemplified in Skokie, then further articulated by Barlow and his contemporaries, and finally embraced in Reno v. ACLU, became the standard frame in which to argue about Internet freedom and governance.

Whether or not private companies providing Internet access or other web services should themselves adopt this kind of Rights attitude, that is declining to moderate the speech they facilitated, was another question that could become muddled, since a private party’s voluntary act of censorship of others could itself be defended as a choice that should be free from government interference. But the overall ethos was for intermediaries to butt out, apart from specifically family-friendly and therefore minor corners of the Internet. Should a private company enforce moderation, it was fine; there were innumerable alternatives online. And many were at the time all too happy to abdicate, as their business models thrived on volume of bits moved rather than on careful curation, whether that of

52. Id.
54. Id.
an ISP offering broadband service, of search engines rapidly indexing the Web for search rather than presenting carefully taxonomized top-down directories, or of the early days of social media like Twitter and Facebook. Twitter’s general counsel in 2014 described the company as the “free speech wing of the free speech party.” And there was not just a convenience to that, but an allure of the sort a private university might offer when it chooses to allow free speech under its umbrella as expansively as the First Amendment does for a citizen at a public one: the values of Skokie upheld.

So, the Rights era thrived—for a while.

FROM THE RIGHTS ERA TO THE PUBLIC HEALTH ERA

Of course, just as many citizens of Skokie were profoundly emotionally disturbed at the prospect of neo-Nazis marching through their neighborhoods in 1977, so too were many people disturbed at what they were exposed to from the earliest days of the Internet. Hate, doxxing, harassment—all were present from the start.

Whitney Phillips, a scholar of rhetoric, reflects on those times as ones that perhaps today elicit too much nostalgia in the mainstream.

I used to believe that the internet used to be fun. Obviously the internet isn’t fun now. Now, keywords in internet studies—certainly, keywords in my own internet studies—include far-right extremism, media manipulation, information pollution, deep state conspiracy theorizing, and a range of vexations stemming from the ethics of amplification . . . .

The more jagged, trollish edges of “internet culture” may have been sanded off for family-friendly consumption, but the overall category and its distinctive aesthetic—one that hinges on irony, remix, and absurd juxtaposition—has in many ways fused with mainstream popular culture . . . .

It was hard to take Nazi memes all that seriously when they were sandwiched between sassy cats and golf course enforcement bears, and so, fun and ugly, ugly and fun, all were flattened into morally equivalent images in a flip book. Others selectively ignored the most upsetting

images, or at least found ways to cordon them off as being “just” a joke, or more frequently, “just” trolling, on “just” the internet.

For those whose identities were targeted and corroded by all that ironic, arm’s length laughter, or whose personal and professional lives were under constant threat, often for the sin of not being a white man in public, the ugliness of the forest wasn’t so easily obscured by the fun of the trees. People tend to see the things that have the potential to harm them. 57

The Rights framework acknowledges those phenomena as readily as it did the discomfort of Skokie and offered the same small solace: that’s the price of free speech. What dimmed the star of Rights—or at least produced a competing one—was the prospect that the real world could be affected, even infected, by what happened online.

THE RIGHTS ERA SHARES SPACE WITH A PUBLIC HEALTH ONE

Consider the comparatively recent case of Earthley, a natural health products affiliate marketing company whose founding story entails a claim of doctors misprescribing antibiotics for an infant—inspiring the search to discover, and market, natural remedies. 58

As Caroline Haskins reported in BuzzFeed in early 2020, the company ran targeted ads on Facebook offering a free guide for pertussis, also known as whooping cough. 59 Pertussis is a highly contagious respiratory tract infection that can in rare cases be fatal, especially in infants. 60 Haskins writes about one of the Earthley store’s online pamphlets:

The document falsely claims that the whooping cough vaccine contains levels of the element aluminum that could cause neurological damage, and it offers Earthley

products—like elderberry elixir, vitamin C powder, and a mixture of herbs—as an alternative.  

Haskins goes on to quote a pediatrician and senior scholar of tropical medicine who describes this marketing—discouraging vaccines and selling unproven remedies in their stead—as “dangerous.”

Facebook’s policies at the time did not prohibit user-posted misinformation about vaccines, though its policies did ban advertisements containing vaccine misinformation as judged by “global health organizations including the World Health Organization.” When asked whether the pointer to the quoted pamphlet in an Earthley advertisement violated that policy, the company said it did not, since the misinformation was in the linked brochure rather than in the advertisement itself. (For its part, the Earthley co-founder told BuzzFeed that it “is up to Facebook to properly review advertisements and choose to reject the ones it does not want on its platform.”)

By March of 2020 the worldwide COVID pandemic was in full swing, and by that December Facebook announced tightened policies against misinformation specifically around the COVID vaccines that had been approved a few weeks earlier. Misinformation, said Facebook, would not be allowed, whether in advertisements or posted organically by users or on companies’ Facebook pages. Facebook said that it would “start removing false claims about these vaccines that have been debunked by public health experts on Facebook and Instagram.”

The new policy notwithstanding, Earthley posted a message on its Facebook home page that simply substituted a new word (well, non-word) for “vaccine” to tout its remedies, explicitly noted as a way to avoid detection by Facebook (who, the company argues, is trying to “shadow ban” its account): “It’s v@ccyne Injury Awareness Month. (We spelled it that way to beat the censors...they’re trying to hide this from you!” (alternative detection-avoidance ways of spelling include “va))ines” and “v@66ine”).

---

61. Haskins, supra note 59.
62. Id.
63. Id.
64. Id.
It’s entirely possible that your assessment of Facebook’s moral responsibility with respect to vaccine denialism is in exactly the same category as that of Facebook’s allowing portraits of people who happen to have scars: the company should not be in the business of judging otherwise-legal content. If so, the Rights framework runs strong in you.

But beginning in the 2010s and accelerating in the 2020s, both before and during the Covid pandemic of 2020, most commentators weighing in on vaccine denialism excoriated Facebook for not doing more to rein it in, particularly that which was grounded in manifestly incorrect claims. Facebook itself did not attempt to defend inaction as it might have years ago, say by borrowing from Twitter’s earlier identification as the free speech wing of the free speech party. Instead, the company simply first parried by parsing its existing policies in a narrow way to continue to allow links to misinformation rather than the outright repeating of it. And by this time Twitter itself had repudiated its free speech stance. In 2013 it introduced a “report abuse” button. In 2015 Twitter made reference to the need to “keep Twitter safe.” As Sarah Jeong wrote in 2016 after another revision of rules to tighten what was permitted there:

The new Rules are radical because they rewrite Twitter’s story of what it is and what it stands for. The old Twitter fetishized anti-censorship; the new Twitter puts user safety first . . .

---


“Freedom of expression means little as our underlying philosophy if we allow voices to be silenced because they are afraid to speak up,” said a Twitter spokesperson in response to a request for comment... Striking the right balance will inevitably create tension, but user safety is critical to our mission at Twitter and our unwavering support for freedom of expression.70

The argument against abdication and for more intervention, Rights rhetoric notwithstanding, is powerful. The way social networks had developed since the resolution of the CDA case was not towards deep and elaborate debate—Barlow’s aspiration for a Civilization of Mind—but rather, perhaps, something less humane and fair than what might be encountered in an exchange of letters to the editor in a Twentieth Century hometown newspaper. When online activities result in persistent harassment, that amplifies the true costs of speech—at least for some—in the Rights framework. And material harms can be documented when this discourse results in viral disinformation tied directly to physical health. Many people were going online, absorbing propaganda of unknown provenance from uncredentialed sources confidently holding themselves out to be in the know, and taking that information to heart, informing decisions about their own lives and those of their loved ones, including their children.

The Rights framework might have some subtleties intended to account for that—the First Amendment does not protect outright false advertising, for example—but the very presumption of a free flow of speech, subject to specific exceptions, itself less and less fits the circumstances of modern social media. For one thing, modern social media was already shaping people’s news feeds and recommendations in ways that differed from just providing the sidewalks on which anyone could digitally march. And for another, it simply turned out that people shared misinformation very, very readily, adding a community momentum to disinformation that far exceeded the analog world’s old rumor mill.

Whether the topic is vaccine denialism or violent extremism, the precision with which information’s movement can be tracked, alongside people’s coming to believe it, is also new. In 1977 there was no easy way short of prohibitively expensive (and dubiously accurate) polling to see how many people might be persuaded by a march of neo-Nazis to the Nazi cause. So, a Supreme Court Justice,

70. Id.
or *New York Times* columnist, could breezily declare that no one would be persuaded by nonsense—or that the display of nonsense would in fact actively repel people. Roughly forty years later, Facebook can trace exactly the flow of links and memes across its services, and the resulting spike in membership of neo-Nazi private groups—even as the company would have no incentive to, say, issue a press release with its findings. (To this day, it has not—and academics and journalists must make guesses about such cause and effect.)

Not long after the Skokie decision was handed down, alumni of Saturday Night Live starred in a feature film build-out of one of the show’s sketches. *The Blues Brothers* were musicians embarking on a romp around Chicago as they pulled together a benefit to save an orphanage. The comedy included a scene in which the Brothers encounter a neo-Nazi rally clearly modeled on the proposed march of the well-known Skokie case. A friendly cop leans into the heroes’ car, stopped in a line of traffic as the Nazis are gathered on a small bridge ahead: “Those bums won their court case, so they’re marching today.”

As police hold back hundreds of counter-protesters, the heroes rev the engine, drive around the line, and make straight for the bridge. The Nazis jump off and into the stream below just in time to avoid being hit, wet, and humiliated but unharmed. It was a scene that came and went from pop cultural consciousness, a comeuppance fantasy in a fictional movie where all the variables and consequences could be consummately controlled.

Forty years later, and with the aid of online organizing, neo-Nazis from across the country converged in Charlottesville, Virginia—an open carry jurisdiction—for a rally nominally in defense of Confederate statues. Police appeared reluctant to referee fist fights that broke out, lest they spark shooting. A neo-Nazi drove his car into a crowd of counter-protesters, killing one of them. This grim inversion of the *Blues Brothers* scene gained its own macabre online momentum after protests for racial justice sprang up in American cities and towns in the summer of 2020. Persistent calls remain online for attacks against protesters blocking streets or highways, and

---


At the Capitol incursion three years later, the fear of contagious domestic radical extremism hit a peak. Incitement in its pre-Internet First Amendment form was about a specific speech leading in direct and immediate sequence to mob violence by those inspired by the speech. With materials of radicalization available a click away—or offered up as recommendations with no click at all—incitement could take place over a period of months, leading to “stochastic terrorism” in which individuals might be moved to violence in unpredictable places and at unpredictable times.

These circumstances have made a Public Health framework a meaningful counterpart and competitor to the Rights framework. Each responds to a distinct set of problems. The Rights framework anticipates a threat from the government. The familiar argument is that states have a monopoly on the use of force; they are distinctly powerful, and the potential for abuse of that power requires setting limits and fixing them beyond easy revision—including, in a democracy, beyond mere majority preferences. “Congress shall make no law . . . abridging the freedom of speech” is subject only to the near-impossible hurdle of Constitutional amendment, and of narrowing interpretations by the judiciary, which has often staked out deeply unpopular free-speech positions, such as not permitting the criminalization of private American flag burning.\footnote{73}{Texas v. Johnson, 491 U.S. 397, 418 (1989).} The Rights framework defends not only against abuse by those with power, but offers support for the idea that they should not be entrusted, even in good faith, to make many speech-related tough judgments. Only when there is a truly compelling interest, and no less-restrictive alternatives, are government impingements on speech for its content or viewpoint to be entertained. The spirit of the Rights framework, if not its First Amendment letter, flowed naturally towards an unwillingness to have large corporate intermediaries, including social media platforms, make decisions around speech, once scarcity of speech distribution was no longer a constraint.

The Public Health framework, by contrast, sees threats from many more sources. For some of these threats, the state can be helpful, even necessary, in marshaling a common defense against risky behavior by fellow citizens. Literal public health strategies

\footnote{73}{Texas v. Johnson, 491 U.S. 397, 418 (1989).}
emphasize education of the public, collection and use of personal information by authorities to track and deploy resources against health threats, and at times large-scale environmental modifications to deal with an infectious disease. When used as a metaphor, then, a public health-style framework can lend itself to diagnosing the overall health of, say, an online community, and thinking through interventions to make it better—a vastly different focus than the idea of hypothesizing some kind of neutral, level playing field on which people act and interact and then otherwise getting out of the way.

The Public Health framework, focusing on group rather than individual behaviors and movements, thus entails balancing the costs and benefits for given interventions, without particularly privileging free speech independent of its impact. Its arrival outside the digital world can be tracked from declarations of rights at the United Nations (UN). In 1948, the UN adopted the Universal Declaration of Human Rights, a statement of guidance to states worldwide. Article 19 of the declaration is a sibling to the American First Amendment, and even seems to anticipate the Supreme Court's thinking in the CDA case:

Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.

Compare that with Article 19 of the International Covenant on Civil and Political Rights, adopted by the UN in 1966 and entering into force as a formal treaty in 1976:

1. Everyone shall have the right to hold opinions without interference.

2. Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.


75. Id.
3. The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:

(a) For respect of the rights or reputations of others;
(b) For the protection of national security or of public order (order public), or of public health or morals.\(^{76}\)

As the third-numbered point above indicates, by the time the aspirational rights of 1948 had been translated into a more binding form, broad exceptions had been elevated as well, balancing the unambiguous statement of rights in the first two clauses with a decidedly nebulous set of special-case exceptions for national security, public health, or morals.

**DO WE KNOW WHAT WE WANT?**

These exceptions emphasize our central problems with digital governance: we don't know what we want. Or put another way, we want it all: free speech for all (rights), along with restrictions against dangerous toxicities (public health). Governments have had a hard enough time figuring out the right balance over the years; private companies, lacking public servants' claim to legitimacy, are even harder put to know what to do.

Our confusion, whether as a group or within ourselves, is broad. It extends to whether to draw a line on speech at all, and when it is drawn, where exceptions should lie. It includes who should enforce the line, at a time when more and more options for enforcement present themselves. And it includes how clear private parties need to be about what they are doing, and what avenues of appeal to offer, when they enforce the rules they choose to set.

It may be that with enough thought, and some kind of Manhattan Project of subsidy to philosophers and political scientists, we will experience a collective breakthrough about how best to strike these difficult balances. I am all for trying. But we cannot wait on that. While we try to sort out the fundamental trade-offs in values and in real-world effects, the fact remains that digital intermediaries are becoming more and more powerful. They are more powerful in the contingent sense that they appear to have consolidated into just

---

a few of them. That consolidation will be tested as they apply
tighter and tighter rules around content, driving some users to
other services to better express what has been forbidden in the
mainstream. But they are also more powerful in their ability to
identify trends in speech and points of intervention even amidst bil-
lions of comments and other speech acts each day.

This is how we get to the point of describing companies like Fa-
cebook and Twitter both as far too powerful and far too hands off.
Clearview AI shouldn't exist—and we should also use it to find the
bad guys.

There are real and painful consequences for failing to resolve the
discord between rights and public health. Those who incline to pu-
rity around individual freedom of expression must contend with the
fact that neo-Nazis and their counterparts are real, that their ide-
ology is growing, facilitated, and cultivated by technologies that
didn't exist in the 1970s, and that the linkages between words and
violence—and more specifically between easy, private communica-
tion and violence—are plausible. (There is a reason that the law
specifically recognizes conspiracy—an agreement among two or
more people to commit a crime in which steps have been taken to
do it—as a crime unto itself. People are stronger together.)

And those who see First Amendment legalisms and culture as a
vestige of privilege for the already-strong, behind which obviously
bad acts can hide and thrive, should grapple with the dilemma of
how anyone but the already-strong could be empowered to judge
speech and censor it from the public sphere. (There is a reason that
censorship is so often associated with hollow orthodoxies and au-
thoritarian impulses.)

To reconcile rights and public health, and indeed to see how much
to ask intermediaries to act in the public interest rather than
simply to pursue profit, we must explore the growth of companies' power—and their newly ascendent desire to strategically shed that
power in order to avoid being forced to confront some of these hard
questions. Then we can put on the table ways to channel and use
that power that might accrue more legitimacy than simply version-
ing up a new terms of service every few months.

My account thus far describes a Rights-centric era from 1995–
2010, followed by a Public Health era from 2010 through today.
While emphasizing distinctly different, at times incommensurable,
values, each has compelling elements. What lies ahead may be a
third era—one of Process, or legitimacy-seeking—particularly for
companies at the center of, and whether they like it or not, active
participants in the construction and operation of our public spheres.
“Process” helps a polity answer questions where consensus around what we want substantively is elusive. It helps us observe, for example, that we might disagree with a new law, while understanding and respecting at some level that, so long as it was passed in accordance with the right procedures, including ones that incorporate elements of democratic governance, it was “right” for it to be enacted.

We are seeing new experiments in the online environment to help the “forum fit the fuss,” in the memorable words of scholars of alternative dispute resolution. Shaping speech and discourse has an impact and gravity that goes beyond a mere customer-service issue, both for those speaking and for those potentially listening and reacting. Given the scale of online discourse, not every utterance or platform’s moderation of same can be litigated in a federal case. An external review board such as the one Facebook has stood up to review its content moderation decisions against its stated community guidelines, represent a process innovation in this area—indeed, a novel form of binding arbitration on the company. And elsewhere I’ve suggested novel forms of incorporating outside judgments—such as those of sortitioned high school students—that might prove more acceptable as means of judging truth in political advertisements than having the company that hosts the advertising decide.

It may prove less vital to get things right when we truly lack even rough society-wide consensus on these issues, than it is to get the process by which we—not just platforms, and not simply governments—examine these issues, come to closure, and revisit and re-examine on a regular and sufficiently structured basis.


