Widening the Lens on Content Moderation

Jenna Ruddock
Technology, Law & Security Program, jruddock@wcl.american.edu

Justin Sherman
Technology, Law & Security Program

Follow this and additional works at: https://digitalcommons.wcl.american.edu/research

Part of the Business Organizations Law Commons, Communications Law Commons, Computer Law Commons, Consumer Protection Law Commons, Internet Law Commons, Marketing Law Commons, and the National Security Law Commons

Recommended Citation
https://digitalcommons.wcl.american.edu/research/69

This Article is brought to you for free and open access by the Program on Information Justice and Intellectual Property and Technology, Law, & Security Program at Digital Commons @ American University Washington College of Law. It has been accepted for inclusion in Joint PIJIP/TLS Research Paper Series by an authorized administrator of Digital Commons @ American University Washington College of Law. For more information, please contact DCRepository@wcl.american.edu.
Widening the Lens on Content Moderation

Mapping the Ecosystem of Online Content Dissemination

July 2021

AUTHORS:
Jenna Ruddock, Senior Researcher
Justin Sherman, Research Fellow

PROJECT DIRECTOR:
Gary Corn

Tech, Law & Security Program
American University Washington College of Law
Executive Summary

The Internet has enabled global communication and collaboration on an unprecedented scale. It has also become an incredibly effective means for distributing harmful content, engaging in harmful behaviors, and coordinating harmful acts. Some of these online harms exacerbate preexisting offline harms—like the distribution of child sexual abuse material—while others are relatively unique to the internet, like the ability to launch mass anonymous harassment campaigns, which disproportionately target women and people of color.¹

In the United States, targeted online disinformation campaigns during recent election cycles have super-charged calls for a comprehensive approach to addressing harms online.² The focus of these conversations has largely fallen on the small group of companies operating the most-visible content distribution platforms, namely Facebook, Twitter, and Google (primarily in its role as YouTube’s parent company).³ But Facebook is not the internet, and focusing almost exclusively on the companies behind these high-profile social media platforms means that the majority of the internet is often left out of the conversation about how to effectively combat harmful content online while also protecting fundamental rights and civil liberties.

This report engages in that larger conversation by discussing the full internet ecosystem’s role in generating, curating, and disseminating online content. We begin by introducing several case studies that illustrate the ways in which companies throughout the internet have taken action to deny services to specific actors or stem the tide of particular content, with varying degrees of success. We then propose a mapping of what we call the online information ecosystem: the full range of functions and services collectively enabling users to disseminate and consume content online. Finally, using this mapping as a framework, we examine the primary technical and contractual mechanisms available throughout the ecosystem to moderate online content.

Importantly, the goal of this paper is descriptive, not prescriptive. In order to have a nuanced and robust conversation about the roles, responsibilities, and decision-making of actors throughout the internet ecosystem in addressing harmful content online, it is necessary to first reach a common understanding about what is possible. This report is intended to provide a diverse range of audiences with a working understanding of the online information ecosystem’s collective role in the generation and dissemination of content online.

Acknowledgments: The authors would like to first and foremost thank Jennifer Daskal, without whom this project never would have been possible—during her time as faculty director of the Tech, Law & Security Program, she not only conceptualized this project but was the driving force behind it—as well as TLS
program director Gary Corn, whose support, insight, and guidance has been essential to its continuation. The authors would also like to thank the many individuals, who shall remain anonymous, who participated in multiple private roundtables on this issue set. Finally, the authors would like to thank Gary Corn, Corin Stone, Paul Rosenzweig, Bill Woodcock, Laura Draper, and Blake Reid for their comments on an earlier draft of this paper. Any errors are solely attributable to the authors.
Contents

Executive Summary i
Introduction 1
Moderation Beyond Platforms 3
8chan 3
Covid-19 4
Mapping the Ecosystem: How Content Moves Online 5
A Deeper Dive: Levers of Control 7
Conclusion: A Wider Lens 12
About the Project 14
Endnotes 15
Introduction

On January 6, 2021, the United States witnessed an unprecedented attack on the U.S. Capitol building during a joint session of Congress, resulting in several deaths. Parler—a social media platform where scores of participants in the Capitol riot had connected, organized, and posted live video footage of the attack—soon became the subject of intense public scrutiny. Within the following forty-eight hours, Google Play and the Apple App Store had both announced that they would be suspending downloads of the Parler mobile application. One day later, Amazon Web Services similarly confirmed that it would stop providing cloud hosting to Parler, temporarily forcing the platform’s operational back-end offline.

This was not the first—nor will it be the last—time that internet companies beyond the major social media platforms have taken voluntary action to limit the accessibility of specific content online. In October 2020, a website hosting service stopped providing services to the all-male, self-styled “Western chauvinists” known as Proud Boys, following pressure from Google. Just one year earlier, on August 3, 2019, a user posted “a manifesto decrying a ‘Hispanic invasion of Texas’” on 8chan, a website and messaging service widely known as a place where white extremists congregate online. Hours later, a gunman presumed to be the author of the manifesto opened fire in a crowded shopping mall in El Paso, killing 22 unarmed men, women, and children, and injuring another 24. 8chan had already been tied to other hate-motivated killings, including the March 2019 Christchurch massacre targeting two New Zealand mosques and the fatal San Diego synagogue attack that same spring. Still, it wasn’t until the aftermath of the El Paso shooting that Cloudflare, a company that sells content delivery services, terminated its services to 8chan. Four years prior, Cloudflare—along with content hosting company Digital Ocean, Google, and GoDaddy, one of the web’s largest domain name registrars—had also stopped providing services to the Daily Stormer, one of the then-most popular neo-Nazi websites on the internet, following intense public pressure. Like 8chan, the Daily Stormer struggled to find and retain new online service providers, bouncing between the dark and clear web—described by journalist Talia Lavin as “wander[ing] the digital wilderness.” In 2010, a similar series of termination decisions by web hosting providers and payment processors temporarily forced WikiLeaks offline in the wake of U.S. government pressure over WikiLeaks’ publication of classified documents, courtesy of U.S. Army Specialist Chelsea Manning.

These actions highlight the under-discussed roles that companies across the internet ecosystem play in delivering content to users and, by extension, their ability to moderate that process. While debates rage about the roles and responsibilities of the major social media platforms in addressing the distribution of harmful content online, there is far less attention paid to the roles
and responsibilities of other companies across the internet ecosystem that are also essential to the generation and distribution of online content.¹⁷

We seek to widen the lens beyond the content dissemination functions that have thus far dominated these discussions. Who is responsible for addressing concerns arising from the vast universe of content that exists online beyond the social media platforms—the proverbial last mile of the content distribution ecosystem? What happens when social media companies fail to respond to—or themselves explicitly promulgate—violence, fraud, and a range of other harms? At what point is it the responsibility of the numerous other companies that comprise the online information ecosystem to take action? Using what methods? According to what substantive and procedural standards? And what are the potential collateral costs—to freedom of speech, to security, and to privacy—of doing so?

This report is part one of a larger project seeking to answer these questions—issues that are critical to the ongoing efforts to respond to an array of harms being perpetuated online. What types of content and harms merit a response is, of course, a source of constant debate.¹⁸ For the purposes of this particular paper, we are leaving that debate to the side. Our focus here is on describing the mechanisms used by companies across the online information ecosystem to address whatever falls into the bucket of agreed-upon, need-to-be addressed harms.

Examining the full range of companies that comprise the internet’s content generation, storage, and delivery ecosystem—and the various ways they engage with and manage that content—is necessary to ensure truly robust public debate on a diverse set of urgent issues. As noted above, this report is intended to provide a range of audiences with a working understanding of the online information ecosystem’s collective role in the generation and dissemination of content online. By necessity, we do not capture every nuance of internet architecture here, though we have done our best to not sacrifice accuracy for accessibility. We also hope our endnotes highlight ample additional resources.
Moderation Beyond Platforms: 8chan

From 2013 to 2019, 8chan gained popularity as an online gathering place for white supremacists, conspiracy theorists, and self-styled trolls. While 8chan’s terms of use nominally restricted posting, requesting, or linking to “any content that is illegal in the United States of America,” violent and extremist content went largely unmoderated. Like most websites, 8chan relies upon a web of online service providers to remain accessible to users.

In March 2019, a mass shooting at two mosques in Christchurch, New Zealand, was livestreamed on Facebook. Facebook did not remove the livestream until after it had ended, at which point the footage at already made the jump over to independently hosted platforms, including 8chan. In response, major Internet Service Providers (ISPs) in New Zealand and Australia took the dramatic step of blocking country-wide access to 8chan, as well as to several other sites that were hosting the footage.

In August 2019, yet another mass shooter took to 8chan – this time posting a racist manifesto before opening fire in a Walmart in El Paso, Texas, killing 22 people. The next day, content distribution network provider Cloudflare ceased providing services to 8chan. The site temporarily found a new CDN provider, until that service provider’s parent company pulled its support as well. Tucows, the world’s second-largest domain registrar, also cut ties with 8chan, disrupting traffic to the 8chan.com domain. 8chan ultimately survived, rebranding as 8kun.top, though the reach of its content and its accessibility to potential new users have been restricted by mainstream service providers’ refusal to associate with the site.

The lifecycle of 8chan illustrates not only the potential impact that companies throughout the internet ecosystem can have on the accessibility of harmful content—it also highlights the overwhelmingly ad-hoc and reactionary bases upon which these decisions have largely been made. While ISPs took action against 8chan in New Zealand and Australia, those same companies took no action against much larger platforms like Facebook, where hundreds of thousands of duplicate or slightly-altered versions of the footage were uploaded in the hours following the attack. ISPs cited the large platforms’ ongoing removal efforts as justification, yet blocked several smaller websites found to be hosting the footage that were also actively working to remove the content. And although Cloudflare and Tucows ultimately also took action against 8chan, both companies expressed extreme reluctance both before and after doing so, with Cloudflare in particular citing the lack of clear independent standards to guide their decision-making as a cause for concern.
Moderation Beyond Platforms: Covid-19

False information about coronavirus treatments has been a persistent problem during the Covid-19 pandemic. Authoritarian governments have used the pandemic as a reason—in many cases illegitimately—to crack down on the spread of “false information” within their borders. Another key issue has been the spread of false coronavirus treatment information through traditional media, such as then-U.S. President Donald Trump’s dangerous suggestion in an April 2020 press briefing that individuals inject disinfectants to treat the virus. Yet the spread of false information about Covid-19 treatments online has been a persistent problem far beyond traditional media and high-profile social media platforms, and companies across the online information ecosystem have taken action to limit its dissemination—including within the Domain Name System (DNS).

In April 2020, the United Kingdom’s domain name registry Nominet (the official registry for .uk domain names) began screening coronavirus-related websites by default when they were registered with the company; only once the website was deemed legitimate would the domain name be usable. Nominet said in April 2020 that this had already led to the suspension of over 600 coronavirus-related domain names under its umbrella. “We don’t want to prevent legitimate registration from getting through, but I think the current situation warrants further checks at the point of registration,” Nominet’s managing director of registry services told ZDNet.

Also in April 2020, the U.S. Department of Justice (DOJ) announced a disruption of hundreds of online scams related to the coronavirus “without requiring legal processes.” The DOJ said it worked directly with domain registries and registrars, passing along complaints about domains that hosted false information or were even selling fake vaccines and fake cures. It also passed along the names of domains that were hosting malware or that operated fraudulent charities purporting to help address the coronavirus. Many of these domain registries and registrars, the DOJ said, then reviewed the domains, found them to be in violation of their policies, and took down the domains voluntarily. The DOJ itself received these leads from industry and was also using new industry tools to detect new domains that hosted false Covid-19 information and other scams. It did not specify how the takedowns were technically carried out, and it did not name any specific domain registrars or registries involved. Quad9, a public recursive DNS resolver, independently monitored malicious Covid-specific domain sets and blocked access attempts.

Companies across the DNS have a variety of levers at their disposal to limit access to online content, described in greater detail below. These actions taken to limit the spread of Covid-19 false information help illustrate how these levers
can be deployed relatively selectively, or relatively broadly, to limit the accessibility of online content and services to internet users globally.\textsuperscript{40}

**Mapping the Ecosystem: How Content Moves Online**

When using the internet in our daily lives, we rarely think about all of the steps that it takes for content to travel from another device to our own screens, and all of the parties involved in making that exchange of information possible. When we post on Facebook, for example, the only obvious parties to that transaction are our own devices (and by extension, the companies that produce them), the company providing our internet connection, and Facebook itself. But even sharing such a simple piece of content online requires a number of other critical actors whose role in the distribution of content throughout the internet ecosystem is oftentimes overlooked. Without these many parties—and in the age of a privatized internet, these organizations are in large part private companies—the poster would never make it online, the post would not get to the right place, and it certainly would not be accessible to others, whether half a mile away or halfway around the world.

Echoing but diverging slightly from traditional, more abstracted models of the internet protocol “stack”,\textsuperscript{41} the mapping we propose is intended to provide a framework for discussing how content is distributed throughout the internet ecosystem by identifiably involved actors. As with any modeling exercise, this mapping seeks to provide workable categories, thus eliding the full set of complexities. It is not, as a result, exhaustive. It also is not the only way to categorize or describe the practical functions carried out,\textsuperscript{42} and it is critical to note that:

- certain functions often overlap;
- certain companies fall into multiple service provider categories; and,
- increasingly, related functions are often bundled together and presented as a single service.

What this mapping offers is a model that sheds light on the broad spectrum of functional roles played by companies throughout the online information ecosystem, all of which, to varying degrees, enable users to share or access content on their devices:

**Accessing:** This is the part of the online information ecosystem that connects devices (like computers and phones) to the online world, via a combination of wires, cables, servers, routers, and more. **Internet service providers** (ISPs) like AT&T, Comcast, Verizon, and Vodaphone all operate in this space. It also includes **Virtual Private Networks** (VPNs) like Tor, which enable users to access the internet in ways that mask location and are thus often utilized to evade local content
controls.

**Delivering**: This part of the ecosystem routes internet traffic from users to sought-after content (such as websites, streaming videos, and apps) and back again. It includes the **registries**, **registrars**, and **Domain Name System operators** that effectively create and operate a phone book for the internet, the range of **content delivery networks** like Cloudflare and Akamai that allow content publishers to scale to meet user demand, and an array of other private and quasi-governmental actors that help deliver traffic along routes from point A to point B.

**Hosting and securing**: Hosting services provide a place for content to sit and business logic to be executed. This includes **cloud providers** like Microsoft, Google, and Amazon and **website hosting platforms** like Bluehost. **DDoS mitigation services** protect hosting services, and the servers that such hosting services use, by filtering attacks on the system. Cloudflare, Verisign, and Amazon (AWS Shield) are among the companies providing these services. Increasingly, hosting and DDoS mitigation services are offered by the same providers, though key companies operating in this part of the ecosystem, such as Cloudflare, do not offer web hosting services.43

**Browsing**: **Web browsers**, like Google Chrome, Safari, and Firefox, provide the digital gateway to accessing public-facing content on the public internet as well as on the “dark web.”

**Content-curating**: Content curation is the user-interfacing part of the ecosystem that has as its core the aggregation and curation of content online. We use the term “curation” here broadly to include both algorithmic and human curation that ranks, sorts, prioritizes, suggests, or otherwise filters content. This comes in a variety of forms, including platforms, search engines, app stores, and a range of other entities that enable or support user-generated content.

**Platforms** — **Social platforms** including Facebook, Twitter, YouTube, Instagram, Reddit, and TikTok and **online marketplaces** like Amazon, Airbnb, and Alibaba both directly curate content and also enable user curation.

**Search** — **Search engines** are a unique and critically distinct part of the online ecosystem, influencing the ways in which their users encounter content by algorithmically curating search results for specific queries.
**App stores** – App stores, like the Apple iOS App Store and the Google Play Store, serve as both gateways to communities of internet users as well as content-curators in their own right. In addition to operating as file-hosting services that store mobile application software files, which users download onto their devices to then access content hosted elsewhere, app stores create both technical and content-oriented rules that govern how applications are posted to, maintained through, and removed from their platforms.

**Financially Facilitating:** These are the plugins and services that enable the processing of payments, an essential component of e-commerce and other fee-for-service interactions online. Well-known examples include Alipay, M-Pesa, PayPal, Amazon Pay, Stripe, Square, Visa, and MasterCard.

How these various subcomponents of the online information ecosystem interact at any given time to produce and disseminate content is not static, and importantly, these categories are not discrete. Single companies often act in a variety of functional roles and thus have potential control over a variety of these functions. Google, for example, operates across each of these areas. It builds fiber-optic internet cables (access role); operates a registry, a registrar, and a recursive resolver (delivery role); provides scaled cloud services (hosting and securing role); runs a digital wallet and online payment system (financially facilitating role); offers an internet browser (browsing role); manages an app store (app store role); and runs the dominant search engine across the United States and Europe (content-curating role); all in addition to its core business functions of advertising and data monetization.

**A Deeper Dive: Levers of Control**

Building on the basic mapping and case studies introduced above, the following section explores in greater detail the functionality provided by each part of the ecosystem, the corresponding mechanisms of content mediation available, and further illustration of how these levers can be—and have been—utilized.

**Accessing**

The precursor to communicating online is accessing the internet. This requires access to both the infrastructure layer—the actual cables, cellular system, routers, switches, and other hardware via which the 0s and 1s that make up our digital communications travel—and the logical layer—meaning the internet protocols that computers use to route information in the digital world. As a practical matter, this access is generally provided by Internet Service Providers (or “ISPs”). ISPs connect individual users to the internet, enabling them to access and share content.
Just as ISPs can provide access, they can also turn off access—something that typically happens in response to government mandates, though ISPs can act independently as well, as they did in the 8chan case. These actions can be regional, or platform- and service-specific. ISP can cut access to services altogether, cut access to services for a particular geographic area, or block particular content through techniques like deep packet inspection (a form of opening up traffic and filtering for certain key words or images) or blocking ranges of IP addresses. These firms can also take steps to “throttle” access to content, or slow down its speed of transmission, which raises the costs (temporally, sometimes also financially) for users to access internet content. Virtual private networks (VPNs) can provide an alternative way for users to connect—enabling users to mask their location and access internet content blocked by local content controls imposed on their local ISPs through laws and regulations, particularly intellectual property regimes. However, VPN companies can themselves make decisions about what kinds of traffic to allow or disallow through their encrypted connections—even if most allow users to access whatever content they choose—and they can also be regulated by governments to prevent access to certain kinds of content.

**Browsing**

Once users are connected to the internet on a technical level—with the requisite hardware installed, and electronic signals flowing—users need to navigate the internet in words and languages that they understand. Browsers are software applications, installed on a user’s device, that enable users to initiate the process of requesting and retrieving web content by typing a user-friendly domain name (URL) into an address field—for example, www.american.edu. Certain devices might have specific browsers pre-installed (e.g., Safari on Apple devices), but users are generally able to install and use a variety of browsers to navigate to content.

Though often overlooked, browsers can and do determine what information is made available to users online. Browsers can block access to particular webpages or websites, both for technical reasons (such as the destination site not having HTTPS enabled) as well as content-oriented ones. They can redirect users to different websites, and they can also issue interstitial notices that appear before access is granted. For example, when a user on Google Chrome attempts to visit a site known to host malware, a warning is typically delivered while the browser intercepts and stops the webpage from loading. These kinds of interstitial notices can operate as warning signals that users have to click through; they can require entry and certification of particular information, such as self-verification of age or parental approval; or they can block access to the website altogether.

**Delivering**

For online traffic to get routed from one point to another, it needs to have a digital address and a means of finding that address—a function the Domain Name System
DNS) performs. This is a multi-step process: First, the plain language web address (e.g., facebook.com, worldwildlife.org, or whitehouse.gov) is translated into a unique digital address (known as an IP address) by a recursive DNS nameserver, which retrieves the relevant information and then sends it back to the initial requester.\textsuperscript{50} Notably, the DNS nameservers do not create or maintain this digital directory—they merely serve as publishers, distributing it to the public. A range of private- and public-sector actors perform the role of creating and maintaining the directory. Registry operators like Verisign (.COM and .NET), Nominet (UK), and Afnic (France) maintain the top-level domains—the last part of the website address. Registrars like GoDaddy, Squarespace, Bluehost, and Google Domains sell subsidiary domain names—what goes to the left of the .com, .org, or .edu.\textsuperscript{51} There are also country-specific top level domain names, called “ccTLDs,” such as .cn for Chinese websites, .de for German websites, and .ir for Iranian websites.

The DNS is fertile ground for controlling what is and is not accessible online.\textsuperscript{52} As discussed above, Tucows, one of the world’s biggest domain registrars, suspended its agreement with the parent company of 8chan in the wake of the 2019 El Paso shooting.\textsuperscript{53} Without a registrar servicing a domain (in that case, 8chan.com), content retrieval requests from users go unanswered, making the website inaccessible. Similarly, GoDaddy and then Google stopped servicing the domain for the Daily Stormer in the wake of its support for and disturbing commentary on the 2017 Charlottesville Unite the Right rally and terrorist attack.\textsuperscript{54} This past year, in an effort to combat harmful content relating to the Covid-19 pandemic, a stakeholder group representing major registrars around the world published a guide for “Registrar approaches to the Covid-19 Crisis.”\textsuperscript{55}

Content delivery network (CDN) services also play a critical role in scaling the delivery of websites and other content to users. Generally speaking, CDNs consist of a geographically distributed set of servers that accelerate the delivery of internet content by holding cached content of a website (content stored locally so that new requests can be filled faster). This protects against long delays in accessing requested content; regionally staged CDNs can respond to local requests with relative speed.

CDNs can be a source of control independent of the security services that CDNs also provide (described below). Before the Chinese government banned Google from operating in China, the government blocked the Google cache.\textsuperscript{56} In 2020, Vietnam’s state-owned telecommunications companies forced Facebook’s local Vietnamese cache servers offline for a period of seven weeks.\textsuperscript{57} The connection was so slow as to make Facebook effectively unavailable to many users. It was only after Facebook negotiated with the government and agreed to a range of content take-down requests that the domestic cache servers were allowed back online.\textsuperscript{58}
Hosting and securing

Content needs a home, because otherwise there will be nothing to route and deliver upon request. This is where hosting services come in. Web hosting is enabled by servers, or computers running programs designed to listen for and process incoming internet requests.\(^{59}\) The most familiar website hosting is public. Subscribers pay for website hosting via companies like DigitalOcean or one of the many cloud providers like Amazon Web Services, Google Cloud, or Microsoft Azure that publish websites publicly for any internet user to access (though full access may then be conditioned on subscriptions, membership, etc.).\(^{60}\)

While not a distinct part of the technical architecture, DDoS mitigation services play a key functional role in this part of the ecosystem by supporting hosting and delivery functions. They ensure availability in the face of distributed denial of service (DDoS) attacks on target servers, thereby ensuring the continued operation of websites, messaging services, and digitally connected devices, among other parts of the internet.\(^{61}\) Related service providers—like web hosts or CDNs—often bundle DDoS mitigation services with their core services.

Content controls in this part of the internet ecosystem can take several forms. Website hosts can block requesting IP addresses (or support plugins that do the same). Hosts can also modify the content displayed to users depending on their geographic location. For instance, a user could go to a federal website to learn about a policy and then have the same website load state-specific information, tied to the user’s location, as determined by the user’s IP address. Similarly, the host of a streaming website, such as Amazon Web Services (which hosts Netflix), can ensure that only regionally licensed videos are available, or otherwise geographically segmented services are provided.\(^{62}\)

DDoS mitigation services are equally critical to determining what is and is not available online. The same fundamental infrastructure that enables content to be hosted, routed, and delivered to users in response to legitimate requests for access is vulnerable to exploitation by malicious actors looking to deny users access to content. This is what DDoS attacks do—flood a target device with traffic in order to overwhelm that device and render it inaccessible by other users. Protections against such attacks, therefore, also enable content availability on the internet. 8chan, for example, was forced offline once Cloudflare stopped providing DDoS protections.\(^{63}\) Absent DDoS mitigation services, website hosts and others are at risk of being flooded by attackers, or by normal use (the “slashdot effect”), ultimately rendering them inaccessible or forcing them to shut down.

Content-curating

At the edge of this ecosystem are the open web and—likely the most discussed and analyzed part of the ecosystem—the content curators. Content curators include
Search engine operators like Google and Microsoft (Bing); social platforms like Facebook, Twitter, YouTube, Instagram, Reddit, and TikTok; mobile app stores, such as the Apple iOS App Store and Google’s Play Store; and a range of other entities that support, rank, sort, and otherwise filter user-generated content. Some companies in this space—such as Airbnb, Alibaba, Uber, and eBay!—curate content to connect people in order to provide an offline service. Still others, such as online dating apps and Eventbrite, connect people to people. A range of content disseminators serve closed ecosystems of users; think, for example, of Dropbox. All of these actors are user-interfacing, in that they directly interact with users online, employing different models to aggregate or curate content for users to consume. They each play, as a result, a key role in determining what content is and is not available to users online.

The tools available within this part of the online ecosystem are diverse, though there are several key methods of moderation. Platforms that host user-generated content—for example, social media platforms or online newspaper comment sections—are typically able to target individual pieces of content for removal. These removals are largely informed by “community standards” or similar terms of service.

More discreetly, many content curators regulate the degree of visibility of content rather than removing it from a platform (or in the case of search engines, from the internet). Google’s search algorithm, for example, exercises a significant amount of control over which websites and thus which information each user sees in response to a particular query. 64 Similarly, platforms like Facebook and YouTube use algorithmic curation to suggest content that users might not otherwise encounter.65 Twitter’s introduction of flags and filters for misinformation relating to U.S. election integrity and the Covid-19 pandemic previewed another set of tools that stop short of actually removing potentially harmful content.66

Mobile app stores, such as the Apple iOS App Store and Google’s Play Store, play a unique role in the ecosystem due to the widespread use of mobile applications, the dominance of the two major app stores, and the fact that users’ choice of app store is largely predetermined by their choice of mobile device. App stores serve as gateways to online communities of users (e.g., of iPhones or Android devices) that rely on those app store platforms to download approved applications to their devices. Once the software files are downloaded from an app store, users run the applications on their device to access content hosted (in most cases) by parties other than the app store (e.g., the TikTok app connects to servers for tiktok.com which are hosted on cloud services separate from Apple and Google).

App stores have levers of control over which applications are posted on their stores, how applications are maintained through their stores (e.g., content policies, security updates), and how applications are removed from their stores. These controls are both technical and content-oriented; for example, while Parler’s return to Apple’s App Store following its takedown in the wake of the January 6th attack on the Capitol
seems to be contingent upon stricter moderation of hate speech, Parler will not filter the same content for users accessing the platform by other means, such as browsers.67

Financially Facilitating

Underpinning many websites, mobile applications, and other key internet services are financial services plugins that enable payment processing. These payment plugins add features to websites and can be used across multiple different services. They either manage payments directly within a website or redirect purchasers to their respective third-party website (e.g., logging into PayPal in a second browser window in order to pay for something in the first). Apple, Alipay, PayPal, Amazon (Amazon Pay), Stripe, Square, Visa, and MasterCard all provide these payment services. They protect website owners from having to worry about the technical and regulatory difficulties of creating a secure online monetary transaction system.

Just as financial interactions are often required to access, publish or distribute information in hard copy, financial interactions online can govern access to information—as with subscription-based platforms—as well as facilitate or prevent content consumers from supporting content creators on- or offline. Payment processors can refuse—and at various times have refused—to support sellers that are deemed objectionable. In 2010, for example, WikiLeaks was effectively “starved” of 95% of its revenue after the U.S. government threatened reprisal against Julian Assange and his website and pressured both U.S. and U.K. payment processors, including Visa, MasterCard, and Moneybookers, to stop providing their services to WikiLeaks.68 Similarly, copyright owners have long turned to payment processors to “choke the flow of money to ‘pirate sites,’” eventually forcing the websites to shut down.69 In 2015, an ambitious sheriff in Cook County, Illinois, sought to choke Backpage.com—an online classified ad forum that was widely known as a place to hire sex workers—by demanding that credit card companies refuse payment for ads on Backpage.70 Sex workers are frequently targeted in this way, having their accounts suspended and funds frozen.71

Conclusion: A Wider Lens

From screen to screen, every piece of online content relies on the majority of the internet ecosystem mapped above—a diverse network of companies providing related, but distinct, functions that enable us to tweet, to blog, to comment on the news, to view others’ content or to share our own. This same ecosystem also empowers harmful actors to launch harassment campaigns, spread disinformation, and exchange illegal and abusive material.

Each of the functions described above presents unique opportunities to regulate this content as it moves through the online information ecosystem. Companies exercising these various functions not only can but also, as illustrated, already do take action to target specific content. These actions are not cost free. They often involve attendant
risks of collateral and disproportionate impacts. The consequences of allowing—or even requiring—companies controlling core internet infrastructure (such as Internet Service Providers) to moderate certain online content are vastly different from allowing or requiring those on the edge of the ecosystem, such as content curators, to do the same.

However, due at least in part to the lack of widespread attention paid to the functional roles these companies play in disseminating content online, when these companies have acted there have been few guiding norms or standards to reference. Content, users, and even entire platforms have been denied services on a largely ad-hoc basis, presenting significant speech- and security-related concerns while simultaneously reducing the likelihood that these actions will have a meaningful long-term impact in terms of effectively disrupting the dissemination of harmful content online. Widening the lens of content moderation conversations to account for the full online information ecosystem allows us to engage in more nuanced, comprehensive, and thus hopefully more effective conversations about when, where, and how to most effectively address the dissemination of harmful content online given the potential collateral costs. We look forward to continuing these conversations in follow-on reports.
About the Project

The Technology, Law & Security Program (TLS) is a new initiative at American University Washington College of Law that tackles the challenges and opportunities posed by emerging technology—offering innovative solutions, engaging our students, and training the leaders of tomorrow.

This is the first report from TLS’s Addressing Harmful Content Online project. The project was launched to examine how actors throughout the Internet ecosystem – beyond the major social media platforms – engage in content moderation and to help craft relevant, rights-protective regulations and standards.

This project has received funding from the following organizations:

- Knight Foundation
- Craig Newmark Philanthropies
- Cloudflare
Endnotes

1 Viktorya Vilk, What to Do if You’re the Target of Online Harassment, Slate (June 3, 2020), https://slate.com/technology/2020/06/what-to-do-online-harassment.html (noting that “[p]eople of color and LGBTQ+ people are disproportionately targeted, and women are twice as likely as men to experience sexual harassment online.”).


6 Id.

7 Id. (“The decision, which went into force on Sunday at 11:59 p.m. Pacific time, will shut down Parler’s website and app until it can find a new hosting provider.”). Parler finally secured a new hosting provider, a Los Angeles-based company named SkySilk, though early reports indicated that users were having trouble connecting to and browsing the site. See Bobby Allyn & Rachel Treisman, After Weeks Of Being Offline, Parler Finds A New Web Host, NPR (Feb. 15, 2021), https://www.npr.org/2021/02/15/968116346/after-weeks-of-being-off-line-parler-finds-a-new-web-host.

8 Catalin Cimpanu, Proud Boys website kicked off web host, Google Cloud, ZDNet (Oct. 10, 2020), https://www.zdnet.com/article/proud-boys-websites-kicked-off-google-cloud. Google did not provide services directly to the websites in question but provided services to a hosting provider that in turn provided services to the Proud Boys. Proud Boys was, however, able to find a new hosting provider within days. Id.


13 Matthew Prince, Terminating Service for 8chan, Cloudflare (Aug. 4, 2019), https://blog.cloudflare.com/terminating-service-for-8chan; see also McLaughlin, supra note 11.


23 *Id.*

24 8chan was also connected to the tragic Poway synagogue shooting in April 2019, during which a gunman opened fire in a synagogue on the last day of Passover, wounding three and killing Lori Gilbert-Kaye as she protected the congregation’s rabbi. A racist and antisemitic manifesto purportedly authored by the shooter was posted to 8chan shortly before the attack. However, none of 8chan’s online service providers took action following this shooting. *See* EJ Dickson, *After

25 Danner, supra note 11.

26 Prince, supra note 13.


31 Paul Smith, Telecos caught in social media crackdown, Australian Financial Review (Mar. 20, 2019), https://www.afr.com/companies/telecommunications/pm-drags-isps-into-postchristchurch-content-control-meeting-as-sites-get-blocked-20190320-h1cl2q [noting that “o]n Wednesday afternoon a Telstra spokesman said it had started to unblock sites that reached out to it and were proactively removing related content from their websites. He conceded some websites had been blocked despite the fact they had already been removing the footage, and said those sites had been unblocked.”].

32 See Prince, supra note 24.


37 Id.


39 It is also unknown if the individual domain registrars and registries that voluntarily took action disclosed their decisions to the affected individuals—e.g., the domain owners—but it is reasonably safe to assume that some form of notification was given.

40 In March 2020, the U.S. Department of Justice also took legal action against a website engaging in a Covid-19-related wire fraud scheme. After the Department became aware of the website, it requested an injunction from a U.S. district judge. The judge granted the request and issued a temporary restraining order requiring the domain’s registrar to immediately take down the website while a DOJ investigation continued. In this case, the DOJ publicly disclosed the details of the takedown action
after it occurred, and it even provided the name of the specific domain in its press release, “coronavirusmedicalkit.com.” The DOJ release did not name the specific registrar, but it was discoverable through public sources using the domain name; TechCrunch subsequently reported that the domain registrar was NameCheap, and one of its spokespeople confirmed to the publication that it had responded to the injunction and suspended the website, rendering it inaccessible. See Zack Whittaker, Justice Dept files its first coronavirus takedown: a bogus vaccine website, TechCrunch, March 22, 2020, https://techcrunch.com/2020/03/22/justice-department-coronavirus.

41 While our mapping does not use the terminology of “layers” or the “stack,” this is primarily to accommodate categories of actors—such as financial facilitators—that do not neatly fit into a more technical hierarchy of infrastructure service providers. The principles articulated by scholarship on layer-conscious Internet regulation, however, remain highly relevant to this discussion. See e.g. Bridy, supra note 18 (noting that “different layers of modern digital networks have very different economic and technological attributes” and citing to Blevins’ The New Scarcity); John Blevins, The New Scarcity: A First Amendment Framework for Regulating Access to Digital Media Platforms, 79 Tenn. L. Rev. 353, 359 (2012) (detailing the distinctions between “network-layer platform regulations and application-layer platform regulations”). See also Mark A. Lemley & Lawrence Lessig, The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era, 48 UCLA L. Rev. 925, 930.

42 Donovan, supra note 17.


44 See Centre for International Governance Innovation & Chatham House, The Royal Institute of International Affairs, Global Commission on Internet Governance 3-4 (2016).


47 Brodkin, supra note 23.

48 In 2019 alone, some 122 “major” blackouts – defined as blackouts that were either regional or country-wide – were recorded in a range of different countries including Chad, Iran, Myanmar, Sudan; the longest and most wide-spread shutdowns were ordered by India. Samuel Woodhams & Simon Migliano, Top10VPN (Jan. 3, 2021), https://www.top10vpn.com/cost-of-internet-shutdowns (noting that “India imposes internet restrictions more often than any other country”): Feliz Solomon, Internet Shutdowns Become a Favorite Tool of Governments: ‘It’s Like We Suddenly Went Blind’, Wall Street Journal (Feb. 25, 2020), https://www.wsj.com/articles/internet-shutdowns-become-a-favorite-tool-of-governments-its-like-we-suddenly-went-blind-11582648765. ISPs can also restrict certain kinds of services and platforms. As just some examples of many, ISPs have been ordered to cut off streaming services in the precursor to the 2018 Bangladeshi elections to prevent the exchange of pictures and videos. See Reuters Staff, Bangladesh slows down mobile internet speeds ahead of election: Daily Star, Reuters (Dec. 27, 2018), https://www.reuters.com/article/us-bangladesh-election-internet-idUSKCN10R06E.


50 Chris Frost, What Is the Difference Between Authoritative and Recursive DNS Nameservers?, Cisco (July 16, 2014), https://umbrella.cisco.com/blog/2014/07/16/difference-authoritative-recursive-dns-nameservers. There are four categories of domain nameservers. The root servers operate as telephone books for their respective roots (such as .com, .org, and .gov) of which there are 13. Each is managed
by its own TLD nameserver, which manages information for each domain name with the same root—in turn backed up by an authoritative nameserver.


52 In China, for example, a government body operates and administers the registry for China’s top-level domain.cn and thus allows the government a lever of control over who is able to register domains to host content. See “Domain name registration in China,” Government of Canada, n.d. (Last accessed November 23, 2020), https://www.tradecommissioner.gc.ca/china-chine/market-facts-faits-sur-le-marche/76359.aspx?lang=eng.


58 Id.


60 Hosting can also be private, rendering websites accessible only to authorized personnel connected to an internal network; for example, a university may provide a database accessible only to certain researchers logged into the university network system, while a business might provide an internal website for Human Resources accessible only by employees using company workstations.

61 Cloudflare, Verisign, and Amazon (AWS Shield) are among the companies providing this suite of services.


68 Annemarie Bridy, Internet Payment Blockades, 67 FLA. L. REV. 1523, 1524-25 (2015) (clarifying that the payment processors terminated their services based on their own respective business judgments, not pursuant to any court order).

69 Annemarie Bridy, Copyright’s Digital Deputies: DMCA-Plus Enforcement by Internet Intermediaries, in RESEARCH HANDBOOK ON ELECTRONIC COMMERCE LAW 185, 201 (John A. Rothchild ed., 2016); see also Bridy, supra note 68 (describing how, after SOPA and PIPA failed in 2012, corporate IP owners entered into an agreement with payment processors, pursuant to which IP owners notify the payment processors of infring ing sales and then they payment processors in turn notify the online merchants, demanding that they desist and terminating services if they fail to do so.) As Professor Bridy describes it, payment processors provide a particularly powerful source of online control: “Whereas it is trivially easy for the operator of a seized or blacklisted domain name to relocate objectionable content to another domain, it is much more difficult for a website operator to replace a canceled banking relationship.” Annemarie Bridy, Internet Payment Blockades, 67 FLA. L. REV. 1523, 1525 (2015). This is due to market power: approximately 80% of online transactions use a credit or debit card as a method of payment, and most of those transactions go through one of two payment systems: MasterCard or Visa. In the words of Mark MacCarthy, “One conclusion that can be drawn from these examples is that payment intermediary action has been effective.” Mark MacCarthy, What Payment Intermediaries Are Doing About Online Identity and Why It Matters, 25 Berkeley Tech. Law J. 1037, 1059 (2010), https://bit.ly/data/articles2015/vol25/25_2/25-berkeley-tech-l-j-1037-1120.pdf.

70 In this case, the Seventh Circuit called the attempted “suffocation” a form of censorship that violated the First Amendment. Backpage.com v. Dart, 807 F.3d 229, 231, 235 (7th Cir. 2015).