Before the

Office of the United States International Trade Commission

Washington, D.C.

In re

Foreign Censorship Part 1: Policies and Practices Affecting U.S. Businesses

Investigation No. 332-585

Written Comments of the
Center for Democracy & Technology

The Center for Democracy & Technology (CDT) welcomes the opportunity to provide comments to the United States International Trade Commission, in response to its Investigation No. 332-585, *Foreign Censorship Part 1: Policies and Practices Affecting U.S. Businesses*. CDT is a non-profit public interest advocacy organization that works to defend internet users' human rights and civil liberties in the U.S., the E.U., and around the world. Through our research and public policy advocacy, we have identified a number of trends in how governments around the world exert control over online speech and information flows.

"Digital censorship" can involve direct or indirect state action that seeks to prevent or suppress online communication, or to punish online speakers, through laws, policies, or practices that are inconsistent with states' international human rights obligations. In some cases, this censorship is definite and overt, such as the Chinese government's decades-long project to build a "Great Firewall," within the bounds of which it exerts direct control over speech.[1] But digital censorship can come from any quarter; it is not the province, alone, of strictly authoritarian regimes with aggressive censorship policies. Rather, the suppression of expression and information, through methods contrary to the rule of law and the protection of fundamental

[1] Geremie R. Barme and Sang Ye, *The Great Firewall of China*, Wired (June 1, 1997) https://www.wired.com/1997/06/china-3/; Yaqiu Wang, *In China, the 'Great Firewall' Is Changing a Generation*, POLITICO (Sept. 1, 2020) https://www.politico.com/news/magazine/2020/09/01/china-great-firewall-generation-405385.

rights, is being enacted or contemplated by many countries around the world today, including key U.S. diplomatic allies and trading partners, and by policymakers in the U.S. itself.

In some case, this trend towards greater state control over online speech and speakers may be borne of legitimate aims, including concerns over privacy and data protection, efforts to fight hate and discrimination targeted at vulnerable populations, or attempts to stymie the corrosive effects of disinformation and online voter suppression on our democracies. But many of the tactics and techniques that democratic governments are using to pursue these legitimate aims are vulnerable to abuse, including by more authoritarian governments that seek to suppress journalism, human rights advocacy, and dissent of all kinds.

Online intermediaries are the primary focal point for state efforts to control online expression. These online service providers host, transmit, amplify, index, link to, and otherwise facilitate user-generated content and are essential to enabling communication among individuals online. They are also obvious targets for state actors who aim to censor online speech, as these intermediaries are easier to identify and control than individual speakers.

Moreover, many of the most prominent online content hosts are U.S.-headquartered businesses operating services that are potentially available worldwide. These companies make strategic decisions about whether and how extensively to invest in specific overseas markets based in part on considerations of the legal and regulatory operating environment. A country with harsh intermediary liability laws and strict limitations on freedom of speech will create substantial legal risk for the online service provider and may be a major impediment to investment and offering of services. At the same time, we have seen an acceleration in efforts by governments to obtain jurisdiction over these U.S. companies in order to force them to comply with national laws or extralegal orders or requests restricting speech, or to give them strong incentives to do so.

In the comments below, CDT describes four major trends emerging worldwide in the governance of online intermediaries that enable digital censorship:

- Requirements of private companies to make determinations about the legality of speech;
- Requirements or government pressure on intermediaries to implement automated content filtering;
- Government officials' manipulation of private content moderation processes; and

- Mandates to locate data and personnel in-country to increase the government's leverage over a private company.

Each of these trends imperils freedom of expression and access to information for individuals and creates substantial barriers to trade and investment by U.S. companies. In each section below, we discuss recommendations the Commission can make for how to reduce these barriers to trade and improve respect for human rights worldwide.

I. Non-judicial determinations about the legality of speech

Intermediary liability laws are essential components of the regulatory frameworks that shape the environment for freedom of expression online. In general, these frameworks lay out the legal risk that online intermediaries bear for hosting, transmitting, or otherwise enabling access to user-generated content. They may take the form of broad, unconditional shields from liability, such as 47 U.S.C. § 230 in the U.S., or conditional notice-and-action regimes, such as the U.S.'s Digital Millennium Copyright Act or the E.U.'s E-Commerce Directive, which specify requirements intermediaries must meet upon being notified of illegal content, in order to maintain their statutory safe harbor from liability. While there is currently considerable debate about the optimal contours of intermediary liability frameworks in the U.S. and many other countries around the world, there remains broad understanding of the significant risk to freedom of expression if these frameworks permit governments or other third parties to leverage intermediaries' services for censorship.[2]

Central to these intermediary liability concerns is the question: Is a specific item of user-generated content illegal? Under the rule of law and international human rights standards, only independent courts have the authority to adjudicate this question. However, many governments around the world have recently adopted or proposed regulations that would put private companies in the position of making determinations of the legality of users' speech. These laws may require service providers to evaluate whether content is illegal after receiving a

[2] Frank La Rue, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, Human Rights Council of the United Nations (May 16, 2011) https://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf; David Kaye, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, Human Rights Council of the United Nations (May 11, 2016) https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/32/38.

notification from an average user, or require providers to remove content pursuant to an order from a non-judicial government agency—or risk facing liability for the content themselves. In either case, the result is that content is taken down as allegedly illegal, without the input of an independent judicial authority.

Authorizing non-judicial entities to determine what constitutes illegal speech is a threat to the freedom of expression. A 2018 report by the UN Special Rapporteur on the freedom of expression found that governments should only restrict access to or remove content pursuant to the order of an impartial judicial body in order to remain consistent with international human rights principles.[3] Regulations that impose heavy fines or other sanctions on platforms for failure to remove content that users or government agencies have flagged can have a chilling effect on speech. Such regulations lead platforms to err on the side of over-policing content, without adequate regard for users' due process rights.[4]

This issue pervades intermediary liability frameworks around the world, creating a global regulatory environment that increasingly pushes U.S. companies to play the roles of judge and jury over individuals' speech rights. One of the most notorious examples of this is the Chinese model, in which intermediaries are provided with extensive lists of prohibited content and required to actively police their services for it.[5] But the issue arises in democratic countries, as well. For example, the Indian government has recently enacted rules that require online services to take action on the basis of non-judicial determinations about the legality of content. Under the 2021 Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules (the Indian Intermediary Rules), online services are required to remove illegal content within thirty-six hours of receiving an order from a government agency.[6] Platforms are also required to remove certain categories of content—including sexually explicit material—within 24 hours of

---

[3] David Kaye, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression,* Human Rights Council of the United Nations 19 (Aug. 17, 2018) https://digitallibrary.un.org/record/1304394.

[4] Jacob Mchangama and Joelle Fiss, *The Digital Berlin Wall: How Germany (Accidentally) Created a Prototype for Global Online Censorship*, Justitia 5 (Nov. 2019) https://justitia-int.org/wp-content/uploads/2019/11/Analyse_The-Digital-Berlin-Wall-How-Germany-Accidentally-Created-a-Prototype-for-Global-Online-Censorship.pdf.

[5] *Freedom on the Net 2020: China*, Freedom House B2 (Oct. 2020) https://freedomhouse.org/country/china/freedom-net/2020.

[6] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Rule 3(1)(d) https://egazette.nic.in/WriteReadData/2021/225464.pdf [hereinafter "2021 Indian Intermediary Rules"].

receiving a complaint from any user about the material.[7] The new intermediary liability rules are stringent and could lead to jail time for employees of platforms who fail to comply with requests to take down illegal content (see section IV below).[8]

Germany's Network Enforcement Act (NetzDG) also obliges intermediaries to determine the legality of speech, and has become something of a model regulation for other nations.[9] Enacted in 2017, NetzDG specifies that online platforms can be subjected to fines of up to $50 million for failing to remove "manifestly illegal" speech within 24 hours of receiving a complaint.[10] Users and non-governmental organizations (NGOs) may make allegations that user-generated content is illegal, and online service providers must consider these allegations to provide actual knowledge of illegal content on their services. In 2020, the German government extended NetzDG to require platforms to report certain kinds of allegedly illegal content directly to the Federal Criminal Police Office.[11] In 2021, rules designed to make it easier for users to submit complaints about illegal content and to appeal content moderation decisions went into effect. The 2021 amendment also imposed new transparency reporting requirements and expanded the supervisory powers of the German Federal Office of Justice.[12] While the transparency and user-redress concepts in NetzDG are important safeguards for individuals' free expression on these services, NetzDG has become a template for governments seeking to censor online content throughout the world, as similar proposals have been adopted in Kenya, the Philippines, Malaysia, Vietnam, India, Singapore, Venezuela, Honduras, France, the UK, Russia,

---

[7] 2021 Indian Intermediary Rules, Rule 3(2)(b).

[8] Namrata Maheshwari and Emma Llansó, *Part 1: New Intermediary Rules in India Imperil Free Expression, Privacy and Security*, Center for Democracy and Technology (May 25, 2021) https://cdt.org/insights/part-1-new-intermediary-rules-in-india-imperil-free-expression-privacy-and-security/; *see also* Global Network Initiative, *GNI Analysis: Information Technology Rules Put Rights at Risk in India* (Mar. 30, 2021) https://globalnetworkinitiative.org/india-it-rules-2021/ (explaining that failure to comply with the new intermediary rules can lead to a loss of safe harbor protections under the IT Act and ultimately result in prison terms of up to seven years for platform employees based in India).

[9] *See* Gesetz zur Verbesserung der Rechtdurchsetzung in sozialen Netzwerken [Netzwerkdurchsetzungsgesetz] [NetzDG] [Act to Improve Enforcement of the Law in Social Networks], Sept. 1, 2017, BUNDESGESETZBLATT, Teil I [BGBL 1] at 3352 (Ger.) https://germanlawarchive.iuscomp.org/?p=1245 [hereinafter "NetzDG"].

[10] *Overview of the NetzDG Network Enforcement Law*, Center for Democracy & Technology (July 17, 2017) https://cdt.org/insights/overview-of-the-netzdg-network-enforcement-law/.

[11] Natasha Lomas, *Germany tightens online hate speech rules to make platforms send reports straight to the feds*, Tech Crunch (June 19, 2020) https://techcrunch.com/2020/06/19/germany-tightens-online-hate-speech-rules-to-make-platforms-send-reports-straight-to-the-feds/.

[12] *Germany: Network Enforcement Act Amended to Better Fight Online Hate Speech*, Library of Congress (2021) https://www.loc.gov/item/global-legal-monitor/2021-07-06/germany-network-enforcement-act-amended-to-better-fight-online-hate-speech/.

and Australia.[13]

These same issues are under active debate at the E.U. level, as the European Parliament considers the draft Digital Services Act, released by the European Commission in late 2020. The DSA is poised to make significant changes to the regulatory framework for online speech, and online service providers, in the E.U. Concerningly, Article 14 of the DSA would require online service providers to receive notifications of allegedly illegal content from average users, NGOs, and law enforcement officials, in addition to court orders.[14] Article 5 of the DSA specifies that when these providers receive these notifications, they have actual knowledge sufficient to give rise to liability for the flagged content.[15] This type of notice-and-action framework creates strong incentives for intermediaries to treat all notifications of allegedly illegal content as legitimate, and to remove that content rather than expose themselves to legal risk. This creates an environment ripe for abuse by state actors and private citizens alike who seek to silence speech and opinion with which they disagree or find offensive. Zealous compliance with such notices can also have a significant extraterritorial impact, if companies broadly restrict access to content on the grounds that it may be illegal in a particular jurisdiction.

In addition to these changes to the core liability frameworks that govern the hosting of user-generated content, governments are also turning to more indirect requirements that non-judicial actors make determinations about unlawful speech. In 2016, the European Commission entered into a voluntary agreement, known as the E.U. Code of Conduct on countering hate speech online, with four major technology companies, Facebook, Microsoft, Twitter, and YouTube, in order to improve the companies' response times to content flagged as illegal hate speech.[16] The companies agreed to review the majority of requests to take down content within 24 hours. The European Commission has been evaluating the code periodically over the last five years, and has found dramatic increases in the speed at which platforms remove flagged content.[17] Free speech advocates have warned that the Code has contributed to a global

---

[13] Mchangama & Fiss, *supra* note 4, at 6-16.
[14] Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC, at 51, COM (2020) 825 final (Dec. 15, 2020) [hereinafter "Digital Services Act"].
[15] *Id.* at 47.
[16] Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 Notre Dame L. Rev. 1036, 1052 (2018), https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=4772&context=ndlr.
[17] Elizabeth Schulze, *EU says Facebook, Google, and Twitter are getting faster at removing hate speech online*, CNBC (Feb. 4, 2019)

censorship creep that affects lawful speech.[18] For example, internet law scholar Danielle Keats Citron has argued that the Code's mandates could lead to censorship of legitimate critiques and discussion, including criticism of the Catholic church for its handling of child abuse scandals, challenges to anti-feminist ideas in Islamic fundamentalism, and stories shared by users who had been the targets of racist or hateful language.[19] As Facebook ramped up its efforts to police hate speech on its platform in 2017, it came under fire for precisely this kind of censorship, when users found that their posts simply discussing their experiences as the victims of racism were being deleted.[20] For now, the E.U. Code of Conduct remains voluntary, but lawmakers have called to adopt legally binding content moderation requirements as part of the Digital Services Act package.[21]

Both direct and indirect requirements that companies remove content pursuant to non-judicial notices threaten freedom of expression and the ability of U.S.-based technology companies to conduct business abroad. CDT recommends:

- The Commission should emphasize the importance of U.S. trade negotiators seeking clear and stable intermediary liability frameworks in trade agreements. The U.S. should support the position that actual knowledge of the illegality of content can only come from an independent judicial authority and that non-judicial government actors should not be able to directly or indirectly force intermediaries to remove or suppress speech.
- The Commission also should encourage U.S. policymakers to advocate that intermediaries adopt fair and transparent content moderation practices consistent with the

https://www.cnbc.com/2019/02/04/facebook-google-and-twitter-are-getting-faster-at-removing-hate-speech-online-eu-finds--.html.

[18] *EU: European Commission's Code of Conduct for Countering Illegal Hate Speech Online and the Framework Decision*, ARTICLE 19 https://www.article19.org/data/files/medialibrary/38430/EU-Code-of-conduct-analysis-FINAL.pdf.

[19] Citron, *supra* note 16

[20] Tracy Jan and Elizabeth Dwoskin, *A white man called her kids the n-word. Facebook stopped her from sharing it*, Washington Post (July 31, 2017) https://www.washingtonpost.com/business/economy/for-facebook-erasing-hate-speech-proves-a-daunting-challenge/2017/07/31/922d9bc6-6e3b-11e7-9c15-177740635e83_story.html.

[21] Natasha Lomas, *On illegal hate speech, EU lawmakers eye legally binding transparency for platforms*, Tech Crunch (June 23, 2020) https://techcrunch.com/2020/06/23/on-illegal-hate-speech-eu-lawmakers-eye-binding-transparency-for-platforms/.

Santa Clara Principles in order to respect and support users' online speech rights.[22]

II. Pressure to use automated content filtering techniques

A second concerning trend in digital censorship worldwide is the growing pressure on online platforms to use automated content filters in content moderation. Given the significant constraints facing even state-of-the-art content analysis techniques, mandated reliance on these tools can lead to the suppression of lawful speech.

The pressure to use content filtering tools is significant. Massive scale is a defining feature of internet-enabled communications, where even the smallest online services can host enormous amounts of user-generated content, far beyond what any service provider could hope to review before publishing. Providers have relied on forms of automated filtering to handle spam and malware from the earliest days of the commercial web. Increasingly, however, the pressure comes from policymakers who see content filters, and in particular the promise of artificial intelligence, as a magic tool for fighting everything from terrorist abuse of the internet to COVID-19 misinformation.

Despite recent advances in machine learning and artificial intelligence, however, automated content analysis techniques have significant limitations that create risks to human rights, and filtering should not be mandated by law. Automated tools continue to struggle to analyze new, previously unseen types of multimedia.[23] Even minor distortions in the way an image is presented can fool an automated content classifier.[24] These tools also perform poorly when required to account for contextual information that human decision-makers might consider obvious.[25]

Because of these limitations, when platforms rely exclusively on these tools in their content removal decisions, they are more likely to inadvertently remove content that is lawful and consistent with their Terms of Service. If biases exist in the data used to train an automated

---

[22] The Santa Clara Principles are a set of best practices in online content moderation developed by digital rights advocates during a conference in Santa Clara, CA, on February 2nd, 2018. The principles are available at https://santaclaraprinciples.org/.

[23] Carey Shenkman, Dhanaraj Thakur, Emma Llansó, *Do You See What I See? Capabilities and Limits of Automated Multimedia Content Analysis,* Center for Democracy & Technology 22-24 (2021) https://cdt.org/insights/do-you-see-what-i-see-capabilities-and-limits-of-automated-multimedia-content-analysis/.

[24] *Id.*

[25] *Id.* at 29-31.

classifier, the classifier will reproduce those biases and disproportionately censor content posted by members of marginalized groups.[26] The tools' propensity for bias is compounded by technical limitations that often make it difficult or impossible to explain to users why a tool flagged or removed their content.[27]

Requirements that companies use automated content moderation tools may come in the form of direct mandates to use filtering tools, or indirectly, as an inescapable consequence of another regulation. For example, Article 17 of the E.U. Copyright Directive makes platforms liable for user-generated copyright infringing content.[28] Under Paragraph 4(b) of Article 17, if the providers cannot obtain permission from the copyright holder, the providers must make their best efforts, "in accordance with high industry standards of professional diligence" to eliminate unauthorized reproductions of copyrighted works.[29] Critics have noted that it is impossible to comply with this mandate without the use of automated content filters that cross-reference all content uploaded by users with a database of copyrighted works. These filters are prohibitively expensive for all but the biggest technology companies and are likely to lead to the over-removal of legal content.[30] An early version of Article 17 directly required platforms to use "copyright filters" but lawmakers eventually deleted the mention of filters from the final version of the directive.[31]

Indirect regulations to use automated content filters also take the form of requirements that companies remove content on sharply abbreviated timelines. For example, Germany's NetzDG imposes a 24-hour timeline on platforms to remove "manifestly" unlawful content.[32] Illegal content that is not manifestly unlawful must be removed within seven days.[33] Similarly, the 2021 Indian Intermediary Rules require that platforms remove content within 36 hours after

---

[26] Natasha Duarte, Emma Llansó, and Anna Loup, *Mixed Messages? The Limits of Automated Social Media Content Analysis,* Center for Democracy & Technology 14 (Nov. 2017) https://cdt.org/wp-content/uploads/2017/11/Mixed-Messages-Paper.pdf.
[27] *Id.* at 33-35.
[28] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, O.J. (L 130) https://eur-lex.europa.eu/eli/dir/2019/790/oj [hereinafter "EU Copyright Directive"].
[29] EU Copyright Directive Title IV, Chapter 2, Article 17(4)(b).
[30] Cory Doctorow, *The European Copyright Directive: What Is It, and Why Has It Drawn More Controversy Than Any Other Directive in EU History?*, Electronic Frontier Foundation (Mar. 19, 2019) https://www.eff.org/deeplinks/2019/03/european-copyright-directive-what-it-and-why-has-it-drawn-more-controversy-any.
[31] *Id.*
[32] NetzDG, Article 1§3(2)2.
[33] NetzDG, Article 1§3(2)3.

receiving a government order.[34] Platforms are required to remove certain other categories of content within 24 hours.[35] These laws do not directly require the use of automated content moderation, but that is the inevitable effect because it is infeasible for platforms to comply with these mandates with a human workforce alone. Directly or indirectly forcing overreliance on automated tools to determine whether content should be blocked leads to errors in the content moderation process that infringe on users' free speech rights. CDT recommends:

- The Commission should recommend that U.S. trade negotiators reject legal obligations for companies to adopt automated content filtering in trade agreements.
- To the extent that intermediary liability frameworks are the subject of negotiations and agreements, the Commission should advocate evaluation of proposals for the risk that they will indirectly impose filtering obligations on U.S. companies.

III. Unaccountable censorship through government manipulation of private content moderation processes

Many governments also pursue the removal of online speech through means other than statutory provisions or court orders. Increasingly, governments rely on service providers' own content policies to obtain removal of online content or accounts. Rather than challenging content in court as a violation of law, the government flags and reports it to the provider for removal on the basis that the content violates the provider's Terms of Service.

Under this tactic, a government actor leverages providers' Terms of Service to indirectly censor online speech of which it disapproves. Because providers' Terms of Service may prohibit a variety of types of speech, the government can selectively target speech within those categories to censor based on viewpoint or content.[36] In addition, government referrals can be coercive.

---

[34] 2021 Indian Intermediary Rules, Rule 3(1)(d).
[35] 2021 Indian Intermediary Rules, Rule 3(2)(b).
[36] In some instances, governments have pressured companies to amend their Terms of Service to prohibit certain disfavored speech that the government cannot legally prohibit itself. For example, as part of its strategy to counter online extremism, the UK government called on providers "to strengthen their terms and conditions," in order to "ensure fewer pieces of extremist material appear online." Scott Craig & Emma Llansó, *Pressuring Platforms to Censor Content is Wrong Approach to Combatting Terrorism*, Center for Democracy & Technology (Nov. 5, 2015), https://cdt.org/insights/pressuring-platforms-to-censor-content-is-wrong-approach-to-combatting-terrorism. However, the UK broadly defines "extremism" as "vocal or active opposition to fundamental British values, including democracy, the rule of law, individual liberty and mutual respect and tolerance of different faiths and beliefs." *Id.* If incorporated into a provider's Terms of Service, this definition raises the risk of abuse and the removal of speech far beyond what governments may permissibly restrict. By pressuring companies to alter their

While a provider's removal of the speech may purportedly be done voluntarily under its own Terms of Service, government notification may exert significant pressure on the provider to comply with removal requests. And, in some countries, providers face mandatory regulations for refusing to comply with government removal requests[37] or can be stripped of liability protection for user-generated content based on a government notification.[38]

Reliance on a provider's Terms of Service also allows governments to obtain extraterritorial removal of online speech, because providers' Terms of Service typically apply worldwide.[39] Thus, when a government actor flags content that violates a provider's Terms of Service for removal, the provider will remove it everywhere around the world, not just in the country that originally flagged it.

As a result, even if a government limits itself to flagging only content for removal under a providers' Terms of Service that is also illegal under its laws—which, as explained below, is not always the case—the content may also be removed in countries where it is legal. For example, European law prohibiting "illegal hate speech" is interpreted differently across Member States[40] and bars at least some speech protected by the First Amendment.[41] Yet if the government of one Member State flags for removal particular hate speech that is illegal under its interpretation of the law, the provider will remove it around the globe, including in other Member States and the U.S. where it may be legal, if it violates the provider's Terms of Service.

Terms of Service referrals also allow governments to have even content that is legal in the referring country selectively removed. Service providers may—and often do—prohibit far

---

Terms of Service and then notifying them of speech that violates those altered Terms, governments can effectively use providers to remove online speech that they themselves cannot censor legally.

[37] Tomer Shadmy & Yuval Shany, *Protection Gaps in Public Law Governing Cyberspace: Israel's High Court's Decision on Government-Initiated Takedown Requests*, Lawfare (Apr. 23, 2021), https://www.lawfareblog.com/protection-gaps-public-law-governing-cyberspace-israels-high-courts-decision-government-initiated (stating that the Israeli Cyber Unit "has the power to subject the online platforms to mandatory regulations should they systematically refuse to comply with its takedown requests").

[38] Jim Killock, *Informal Internet Censorship: The Counter Terrorism Internet Referral Unit (CTIRU)*, Open Rights Group (Mar. 5, 2019), https://www.openrightsgroup.org/blog/informal-internet-censorship-the-counter-terrorism-internet-referral-unit/ (describing the impact of detailed notification under the E-Commerce Directive on platforms' "actual knowledge" of criminal content and subsequent potential liability for that content).

[39] Citron, *supra* note 16, at 1055.

[40] *Letter to European Commissioner on Code of Conduct for "Illegal" Hate Speech Online*, Center for Democracy & Technology (June 3, 2016), https://cdt.org/insights/letter-to-european-commissioner-on-code-of-conduct-for-illegal-hate-speech-online/.

[41] *See, e.g.*, *R. A. V. v. St. Paul*, 505 U.S. 377 (1992).

more speech on their services than that which is prohibited by law. Even when providers' Terms of Service ban content that may, at first blush, appear to coincide with legal restrictions on speech, such as threats,[42] incitement to violence,[43] or harassment,[44] the providers' definitions may not align entirely with legal definitions. If a government flags for removal content that is legal in its country, such as the documentation of human rights violations, the provider will remove it so long as it violates the provider's Terms of Service.

Moreover, governments may refer for removal, pursuant to a provider's Terms of Service, even content that *cannot* be made illegal consistent with international human rights standards. Indeed, providers' privately developed Terms of Service are typically more restrictive—and often much more restrictive—than what governments may permissibly restrict under law. Providers consider a wide variety of factors when developing their Terms of Service, often based on the type of user they wish to attract and the subject matter of the site. For example, a comedy website may ban politics and require all posts to be humorous[45]—a reasonable response for a website operator, but something the government could not do. In addition, even the most clear and narrow content policies will tend to go well beyond what governments may permissibly censor. In seeking content removal through Terms of Service referrals, the government relies on providers' definitions of terms such as "threatening or promoting terrorism," "supporting or praising leaders of dangerous organizations," and "hate, harassment and abuse,"[46] rather than definitions that are supported by international human rights law.

Government use of referrals of content to providers for removal under their Terms of Service in order to target critics, rivals, or activists is well documented.[47] For example, Amnesty

[42] *See, e.g.*, *Violent Threats Policy*, Twitter (Mar. 2019), https://help.twitter.com/en/rules-and-policies/violent-threats-glorification.
[43] *See, e.g.*, *Violence and Incitement*, Facebook, https://www.facebook.com/communitystandards/credible_violence (last visited July 8, 2021).
[44] *See, e.g.*, *Hateful Conduct and Harassment*, Twitch (Apr. 7, 2021), https://www.twitch.tv/p/en/legal/community-guidelines/harassment/.
[45] *See, e.g.*, *r/funny*, Reddit, www.reddit.com/r/funny (last visited July 10, 2021) (stating that "[a]ll posts must make an attempt at humor" and prohibiting "politics or political figures"). Similarly, a website may ban political speech in support of only one candidate, political figure, or party, *see, e.g.*, *Aja Romano, "Everyone uses Ravelry": why a popular knitting website's anti-Trump stance is so significant*, Vox (June 27, 2019), https://www.vox.com/2019/6/27/18744347/ravelry-trump-ban-backlash-community-reaction, a restriction that would be entirely inappropriate and inconsistent with international human rights law if imposed by a government.
[46] *See* David Kaye, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression,* Human Rights Council of the United Nations, A/HRC/38/35 at para. 26–27 (Apr. 6, 2018), https://undocs.org/en/A/HRC/38/35.
[47] *See, e.g.*, *1st Referral Action Day Against Right-Wing Terrorist Online Propaganda*, Europol (May 28, 2021), https://www.europol.europa.eu/newsroom/news/1st-referral-action-day-against-right-wing-terrorist-online-propagan

International has reported that the Vietnamese government engages in "mass reporting campaigns" in which it relies on social media sites community reporting functions to have "large numbers of users . . . simultaneously 'report' a particular account or specific content with the aim of having it deleted or suspended by social media companies on the basis of it violating community standards."[48] According to news reports, the Vietnamese government has used the mass reporting technique to target journalists and human rights activists on Facebook."[49] In another example, in Ecuador, the Office of the President reported a news site supportive of a rival politician to the site's hosting provider for alleged copyright violations, resulting in its temporary takedown.[50] The Ecuadorian government is also suspected of initiating the suspensions of several Twitter accounts of its critics.[51]

In some countries, governments have formalized Terms of Service referrals using Internet Referral Units (IRUs). IRUs are government entities formed to flag user-generated content directly to the service provider that hosts it, often using the provider's own content-flagging mechanisms, so the provider will remove the content under its Terms of Service.[52] IRUs often focus on the removal of online terrorist and violent extremist content, and they are increasingly common. Among other countries, the United Kingdom,[53] Belgium, France, Germany, Italy, Israel, and the Netherlands all have IRUs or an equivalent government office.[54] The European Police Office (Europol) also operates an IRU.[55] Some IRUs, like the UK's Counter-Terrorism Internet

da (describing a "Referral Action Day" to refer rightwing terrorist and extremist content to platforms for removal, in which 28 international partners, including Australia, Austria, Croatia, Czech Republic, Denmark, Georgia, Greece, Hungary, Ireland, Latvia, Lithuania, Luxembourg, Moldova, Montenegro, North Macedonia, Norway, Portugal, Romania, Serbia, Slovakia, Spain, and Sweden participated).

[48] *'Let Us Breathe!': Censorship and Criminalization of Online Expression in Viet Nam*, Amnesty International 53 (2020), https://www.amnesty.org/download/Documents/ASA4132432020ENGLISH.pdf.

[49] Russel Brandom, *Facebook's Report Abuse button has become a tool of global oppression*, Verge (Sept. 2, 2014), https://www.theverge.com/2014/9/2/6083647/facebook-s-report-abuse-button-has-become-a-tool-of-global-oppression.

[50] *Freedom on the Net 2019: Ecuador*, Freedom House (Nov. 2019), https://freedomhouse.org/country/ecuador/freedom-net/2019.

[51] *Id.*

[52] Jason Pielemeier & Chris Sheehy, *Understanding The Human Rights Risks Associated With Internet Referral Units*, VOX-Pol (Mar. 26, 2020), https://www.voxpol.eu/understanding-the-human-rights-risks-associated-with-internet-referral-units/; Craig & Llansó, *supra* note 36.

[53] Pielemeier & Sheehy, *supra* note 52.

[54] *Id.*; Jen Patja Howell, *The Lawfare Podcast: Israel's 'Cyber Unit' and Extra-legal Content Take-downs*, Lawfare (Apr. 29, 2021), https://www.lawfareblog.com/lawfare-podcast-israels-cyber-unit-and-extra-legal-content-take-downs.

[55] *Europol: Non-transparency cooperation with IT companies*, EDRi (May 18, 2016), https://edri.org/our-work/europol-non-transparent-cooperation-with-it-companies/; Pielemeier & Sheehy, *supra* note 52.

Referral Unit, refer content to providers for removal only if they deem it to violate national law;[56] importantly, however, the determination of illegality is typically made by law enforcement officers—not a court. Other IRUs, like Europol's EIRU, can refer even legal content to providers for removal.[57]

There is little publicly available information about IRUs, and "[e]ven those that have issued transparency reports . . . tend to mostly include cumulative statistics on referrals or 'content removed' in ways that make verification and accountability a major challenge."[58] This lack of transparency results in critical due process violations. Providers and governments are not required to notify users when their content is removed because it was flagged by an IRU or other government program.[59] As a result, affected users may not realize their speech was subject to government review and flagging, and, even if they suspect that it was, they may lack a legal route to challenge the removal. For example, the Israeli Supreme Court recently rejected a constitutional challenge by two nongovernmental organizations to the Israeli Cyber Unit's practice of requesting that online platforms remove unlawful content.[60] The Court held the NGOs failed to provide evidence that their content was removed as a result of a government request, and therefore "failed to establish a connection between the practice of the Cyber Unit and the infringement of any specific constitutional rights."[61]

In short, government referrals of online speech to service providers for removal under providers' Terms of Service, whether through IRUs or other mechanisms, allow governments to use private companies to pursue expedited, privatized removal of content worldwide that they otherwise could not legally censor. Accordingly, CDT recommends:

---

[56] Counter-terrorism: Question for Home Office, UK Parliament (Mar. 14, 2016), https://questions-statements.parliament.uk/written-questions/detail/2016-03-14/30893.
[57] EDRi, *supra* note 55.
[58] Pielemeier & Sheehy, *supra* note 52. Indeed, while some internet companies also make data available about government requests for removal based on Terms of Service available, Emma Llansó, *Twitter Transparency Report Shines a Light on Variety of Ways Governments Seek to Restrict Speech Online*, Center for Democracy & Technology (May 4, 2017), https://cdt.org/insights/twitter-transparency-report-shines-a-light-on-variety-of-ways-governments-seek-to-restrict-speech-online/, the statistics released by at least some IRUs have been criticized as "inconsistent with transparency reports at major platforms," Killock, *supra* note 38.
[59] Killock, *supra* note 38.
[60] Shadmy & Shany, *supra* note 37.
[61] *Id*.

- The Commission should recommend that U.S. trade negotiators make clear that governments should not pursue IRUs and other forms of unaccountable censorship through enforcement of providers' Terms of Service.
- If governments continue to refer online speech to providers for removal under their Terms of Service or content policies, including through IRUs, U.S. policymakers should advocate that governments produce transparency reports documenting these referrals with sufficient detail to allow the public to understand how many referrals are made and the legal basis for each referral.

IV. Data and personnel localization requirements chill speech and increase government power to demand content removals.

Legal requirements that internet companies locate data or personnel in a particular country, known respectively as "data localization" and "personnel localization," also create risks to human rights and can be mechanisms through which states exert control over online speech. In particular, data localization increases government surveillance capabilities, which in turn can chill free expression. Personnel localization requirements, sometimes referred to as "hostage provisions,"[62] make it more difficult for intermediaries hosting user-generated content to resist abusive government demands to remove content, including content that reveals wrongdoing by the government or embarrasses or criticizes government officials, because of the threat that failure to comply with those demands will result in punishment, including imprisonment, of the local personnel.

The censorial impact of data localization and personnel localization mandates is felt worldwide. When an internet user in one country self-censors her speech due to increased fear of government surveillance and persecution as a result of data localization laws, audiences outside that country do not receive that speech. And when an intermediary removes user-generated content because of threats a country makes to its employees present in the country as a result of personnel localization laws, it may be required or choose to take down the content not just in the threatening country, but around the globe.

[62] *GNI Submission to European Commission Consultation on the Digital Services Act*, Global Network Initiative (Apr. 1, 2021), https://globalnetworkinitiative.org/dsa-submission-mar-21/ [hereinafter "GNI, Digital Services Act Submission"].

Data localization and personnel localization requirements are increasingly popular around the world in both nondemocratic countries such as China and democracies such as India and in Europe. While the risks that data and personnel localization requirements pose to free expression may be highest in nondemocratic, authoritarian regimes, they are problematic in democratic countries as well.

A. Data localization

Data localization mandates are on the rise around the world.[63] According to a recent report by the Information Technology & Innovation Foundation (ITIF), "the number of data-localization measures in force around the world has more than doubled in four years," from 2017 to present.[64] In recent years, China, Vietnam, Pakistan, Russia, Turkey, Indonesia, India, and Brazil have all enacted or proposed mandated data localization provisions.[65] ITIF identifies China as "the most data-restrictive country in the world, followed by Indonesia, Russia, and South Africa."[66]

While specific data localization laws vary from country to country, they can include "rules preventing information from being sent outside the country, rules requiring prior consent of the data subject before information is transmitted across national borders, rules requiring copies of information to be stored domestically, and even a tax on the export of data."[67] Data localization mandates may apply to only particular types of data, such as telecommunications metadata, health data, or personal data.[68] Increasingly, countries are requiring localization of data in "broad and vague categories involving data deemed 'sensitive,' 'important,' 'core,' or related to national security."[69]

Data localization laws pose significant risks to privacy, freedom of expression and belief, and due process among other human rights.[70] Foreign governments often justify these laws on

[63] Nigel Cory & Luke Dascoli, *How Barriers to Cross-Border Data Flows Are Spreading Globally, What they Cost, and How to Address Them,* Information Technology & Innovation Foundation (July 2021), https://itif.org/sites/default/files/2021-data-localization.pdf; Adrian Shahbaz, Allie Funk, & Andrea Hackl, *Special Report 2020: User Privacy or Cyber Sovereignty?*, Freedom House (July 2020), https://freedomhouse.org/sites/default/files/2020-07/FINAL_Data_Localization_human_rights_07232020.pdf.
[64] Cory & Dascoli, *supra* note 63, at 3.
[65] Shahbaz et al., *supra* note 63.
[66] Cory & Dascoli, *supra* note 63, at 1.
[67] Anupam Chander & Uyen P. Le, *Data Nationalism*, 64 Emory L.J. 677, 680 (2015).
[68] Shahbaz et al., *supra* note 63.
[69] Cory & Dascoli, *supra* note 63, at 1, 4.
[70] Shahbaz et al., *supra* note 63.

the grounds of national security and other rationales for "keep[ing] personal or financial transaction data in-country where they are subject to access and local regulation."[71] However, in reality, these justifications are often a smokescreen for governmental efforts to enhance their surveillance capabilities, exercise greater control over the internet, or other similar motivations.[72]

Data localization laws are not necessary to allow governments to obtain data for legitimate law enforcement investigations. Rather, governments can obtain data for such investigations through existing options, and, when necessary, seek to reform and improve those options. For example, U.S. communications service providers may make voluntary disclosures to foreign governments of non-content information, because such disclosures are unregulated under U.S. law even when the data is in the U.S.[73] In addition, many foreign governments can obtain data through mutual legal assistance treaties (MLATs), which operate by subjecting the foreign country's demand to U.S. law requirements.[74] Over time, foreign governments may also be able to obtain data through CLOUD Act agreements, through which the foreign demand is made under the laws of the foreign country.[75]

Whatever the ostensible rationale for data localization laws, increased government surveillance capabilities from data localization chills speech. As Freedom House has explained, data localization mandates provide government authorities with access to a more extensive dataset of their population's speech and other information that can reveal their locations and associations.[76] This access "opens the door for arbitrary, disproportionate, and discriminatory surveillance" and prosecutions.[77] For example, ITIF reports that Pakistan's data localization requirement requires companies "to retain information, including traffic data linked to blocked content, and decrypted information about subscribers and their activity" and "allows the Pakistan Telecommunication Authority to avoid existing data access and privacy safeguards, and to

---

[71] *Internet Way of Networking Use Case: Data Localization*, Internet Society (Sept. 30, 2020), https://www.internetsociety.org/wp-content/uploads/2020/09/IWN-Use-Case-Data-Localization-EN.pdf.
[72] Cory & Dascoli, *supra* note 63, at 4.
[73] Although the Electronic Communications Privacy Act (ECPA) bars U.S. service providers from voluntarily disclosing metadata to "governmental entities," 18 U.S.C. § 2702(c)(6), ECPA defines governmental entity to include only U.S. federal, state and local government departments or agencies, 18 U.S.C. § 2711(4). This definition does not include foreign governments. Therefore, U.S. communication service providers are permitted to voluntarily disclose user metadata—be it of a U.S. or non-U.S. person—to other governments.
[74] *See generally* U.S. Department of State, 7 Foreign Affairs Manual § 962.1, https://fam.state.gov/FAM/07FAM/07FAM0960.html#M962_1.
[75] *See* Clarifying Lawful Overseas Use of Data Act (CLOUD Act), H.R. 4943, 115th Cong. (2018).
[76] Shahbaz et al., *supra* note 63.
[77] *Id.*

intervene on behalf of law enforcement agencies to ask social media companies to provide user data."[78] By increasing the risk of government surveillance, data localization laws chill speech.

B. Personnel localization

In addition to data, more countries are requiring internet companies to locate personnel within their borders, sometimes coupled with a mandate that the in-country personnel be high ranking within the company. Personnel localization requirements give a country greater leverage over intermediaries, because the government can subject their in-country employees to personal civil or criminal liability if they refuse to comply with government demands to censor user-generated content or reveal users' data or information. In 2016, for example, Brazil arrested a Facebook executive after the company refused to reveal information about users of its subsidiary WhatsApp—which the company said it did not have—in response to a court order.[79] The executive was released 24 hours later after a higher court overturned his arrest.[80]

In some cases, however, a company may have no personnel in a country or may purposefully remove personnel in response to government threats.[81] In response, some countries have begun to require internet companies to locate personnel within their jurisdiction. One of the earliest examples is Germany's NetzDG, which, among other things, requires social media platforms with more than 2 million users in Germany to "appoint a local German representative."[82] Since its enactment, other countries have passed or considered personnel localization requirements modeled on NetzDG.[83]

---

[78] Cory & Dascoli, *supra* note 63, at 8.
[79] Jonathan Watts, *Brazilian police arrest Facebook's Latin America vice-president*, Guardian (Mar. 1, 2016), https://www.theguardian.com/technology/2016/mar/01/brazil-police-arrest-facebook-latin-america-vice-president-diego-dzodan.
[80] Brad Haynes, *Facebook executive released from jail in Brazil*, Reuters (Mar. 2, 2016), https://www.reuters.com/article/us-facebook-brazil/facebook-executive-released-from-jail-in-brazil-idUSKCN0W4188.
[81] *See* Paul Sonne & Sam Schechner, *Google to Shut Down Engineering Office in Russia*, Wall Street Journal (Dec. 12, 2014), https://www.wsj.com/articles/google-to-shut-engineering-office-in-russia-1418401852 (reporting that Google closed its engineering office in Russia in 2014 amid a government crackdown on internet freedom and growing concerns about "the safety of its staff in Russia should Google run afoul of the new Russian laws"). In January 2021, President Putin instructed the government "to create a set of additional rules for foreign tech companies operating in Russia, including a requirement to open branch offices in the country." *Russia: Social Media Pressured to Censor Posts*, Human Rights Watch (Feb. 5, 2021), https://www.hrw.org/news/2021/02/05/russia-social-media-pressured-censor-posts.
[82] Vittoria Elliott, *New laws requiring social media platforms to hire local staff could endanger employees*, Rest of World (May 14, 2021), https://restofworld.org/2021/social-media-laws-twitter-facebook/.
[83] Svea Windwehr & Jillian C. York, *Turkey's New Internet Law Is the Worst Version of Germany's NetzDG Yet*, Electronic Frontier Foundation (July 30, 2020), https://www.eff.org/deeplinks/2020/07/turkeys-new-internet-law-worst-version-germanys-netzdg-yet (describing

In 2020, for example, Turkey passed a law that, among other things, requires social media platforms to both store users' data locally and to appoint a local representative in Turkey to respond to government requests to remove or block access to content.[84] The local representative must be a Turkish national[85] and can be held personally liable for failing to comply with court-ordered content removals within the statutory time period.[86] After some initial resistance, many major social media sites have now appointed local representatives in Turkey.[87]

Though Turkish lawmakers justified the law as a response to online hate speech and harassment,[88] according to Human Rights Watch, the Turkish government makes an "enormous number" of content removal requests that are "in violation of the right to freedom of expression and information."[89] Tellingly, the law enacting the personnel localization requirement was passed "after a series of allegedly insulting tweets aimed at President Erdogan's daughter and son-in-law."[90] Human rights activists have argued that the law enables more extensive government censorship by making it more difficult for companies to resist politically-motivated or otherwise unjustified removal requests and increases Turkish surveillance powers by putting users' data within easy reach of the Turkish government.[91]

The recently enacted Indian Intermediary Rules also require certain social media companies to have at least three responsible company employees resident in India, including a Chief Compliance Officer (CCO).[92] The CCO must be a "key managerial personnel from the company"[93] and is personally liable for the company's failure to comply with the rule's

---

laws passed in Turkey, Russia, Malaysia, the Philippines, Venezuela, and Singapore inspired by NetzDG); *see also* Mchangama & Fiss, *supra* note 4.

[84] Windwehr & York, *supra* note 83.

[85] Deniz Yuksel, *Turkey's Government Wants Silicon Valley to Do Its Dirty Work*, Lawfare (Dec. 9, 2020), https://www.lawfareblog.com/turkeys-government-wants-silicon-valley-do-its-dirty-work.

[86] Ayla Jean Yackley, *Turkey's social media law: A cautionary tale*, POLITICO (Mar. 29, 2021), https://www.politico.eu/article/turkeys-social-media-law-a-cautionary-tale/.

[87] *Id*.

[88] Windwehr & York, *supra* note 83.

[89] *Turkey: YouTube Precedent Threatens Free Expression*, Human Rights Watch (Dec. 18, 2020), https://www.hrw.org/news/2020/12/19/turkey-youtube-precedent-threatens-free-expression [hereinafter "*Turkey: YouTube Precedent*"].

[90] Windwehr & York, *supra* note 83.

[91] *Id.*; Yackley, *supra* note 86; HRW, *supra* note 89; *Freedom on the Net 2020: Turkey*, Freedom House (Oct. 2020), https://freedomhouse.org/country/turkey/freedom-net/2020.

[92] Maheshwari & Llansó, *supra* note 8; *Letter to MeitY and GNI Analysis of the IT Rules*, Global Network Initiative (Mar. 30, 2021), https://globalnetworkinitiative.org/wp-content/uploads/2021/03/GNI-Letter-Analysis-IT-Rules30March21.pdf [hereinafter "GNI, *Letter to MeitY*"].

[93] GNI, *Letter to MeitY*, *supra* note 92.

requirements regarding content removals, facing penalties of up to seven years in prison and significant fines for noncompliance.[94] This is not an idle threat; even before the effective date of the new rules, India threatened to jail employees of Twitter, Facebook, and WhatsApp for their failure to comply with takedown requests related to protests by Indian farmers against the government.[95]

The United Nations Special Rapporteurs for Freedom of Expression, Association, and Privacy have raised concerns that the Indian personnel localization requirement "incentivizes the restriction of content and is likely to cause a chilling effect on freedom of expression."[96] Numerous human rights and other internet freedom organizations have also criticized the requirement.[97] As CDT has explained, the Indian personnel localization rule appears "tailor-made to give the government greater leverage over [certain social media companies] by effectively taking a CCO hostage in order to force the removal of content the government does not like."[98] In addition, as Software Freedom Law Center, India has explained, the personnel localization requirement poses significant financial and operational barriers to smaller companies operating within India and may mean that smaller or nonprofit companies, like encrypted messaging service Signal, cannot offer their services in India.[99]

Personnel localization requirements are also poised to spread across all of Europe. Article 11 of the DSA would require intermediaries that operate in the Europe Union to designate a legal or natural person in one of the member states as their legal representative.[100] The designated legal representative must have sufficient power to cooperate with government authorities and can be

---

[94] Maheshwari & Llansó, *supra* note 8; GNI, *Letter to MeitY*, *supra* note 92.
[95] *See* Jeff Horowitz & Newley Purnell, *India Threatens Jail for Facebook, WhatsApp and Twitter Employees*, Wall Street Journal (Mar. 5, 2021), https://www.wsj.com/articles/india-threatens-jail-for-facebook-whatsapp-and-twitter-employees-11614964542.
[96] Letter from Irene Khan, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Clement Nyaletsossi Voule, Special Rapporteur on the rights to freedom of peaceful assembly and of association, & Joseph Cannataci, Special Rapporteur on the right to privacy to the Government of India (June 11, 2021), https://spcommreports.ohchr.org/TMResultsBase/DownLoadPublicCommunicationFile?gId=26385.
[97] *See, e.g.*, Maheshwari & Llansó, *supra* note 8; GNI, *Letter to MeitY*, *supra* note 92; Katitza Rodriguez, Sasha Mathew, & Christoph Schmon, *India's Strict Rules For Online Intermediaries Undermine Freedom of Expression*, Electronic Frontier Foundation (Apr. 7, 2021), https://www.eff.org/deeplinks/2021/04/indias-strict-rules-online-intermediaries-undermine-freedom-expression.
[98] Maheshwari & Llansó, *supra* note 8.
[99] *Analysis of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code Rules, 2021*, Software Freedom Law Center, India (Feb. 27, 2021), https://sflc.in/analysis-information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021.
[100] Digital Services Act, Art. 11.

held personally liable for a company's non-compliance with requirements under the DSA.[101] As with other personnel localization requirements, this provision risks encouraging intermediaries to engage in overly-aggressive removal of user-generated content in response to government demands. Moreover, as the Global Network Initiative has noted, it "sets a troubling precedent as non-democratic governments insert 'hostage provisions' in their content regulations in order to increase their leverage over intermediaries."[102]

Data and personnel localization requirements chill speech by giving governments greater control over and access to internet user's data and leverage over intermediaries in the form of employees that can be jailed or fined if intermediaries resist government demands to censor user-generated content. However, CDT recognizes that governments have a legitimate interest in obtaining user data using legal processes for valid criminal investigations and in enforcing lawful content removal demands that comply with international human rights standards. Accordingly, CDT recommends:

- The Commission should recommend that U.S. trade negotiators seek to bar trade counterparts from enacting data localization and personnel localization requirements and/or seek commitments that personnel from U.S. companies will not be imprisoned or punished for content decisions made by those companies.
- The Commission should recommend that U.S. trade negotiators advocate that, in negotiating new bi-lateral and multilateral agreements to facilitate cross border data demands outside the MLAT process, governments should include protections for fundamental human rights and ensure that any expanded ability to access data is bounded by strong privacy and procedural protections.[103] With respect to CLOUD Act agreements, the Commission should recommend that the U.S. maintain its commitment to not enter

[101] Digital Services Act, Art. 11.
[102] GNI, Digital Services Act Submission, *supra* note 62.
[103] *See, e.g., Joint Statement Encouraging EU Legislators to Fight for Fundamental Rights Protections in e-Evidence Legislation*, Center for Democracy & Technology (Jan. 6, 2020), https://cdt.org/insights/joint-statement-encouraging-eu-legislators-to-fight-for-fundamental-rights-protections-in-e-evidence-legislation/; *CDT Comments, Cybercrime Convention Committee (T-CY), Preparation of a 2nd Additional Protocol to the Budapest Convention on Cybercrime, Octopus Conference, 11-13 July 2018 Council of Europe, Strasbourg, France*, Center for Democracy & Technology (June 25, 2018), https://cdt.org/insights/cybercrime-convention-committee-2nd-additional-protocol-to-the-budapest-convention-on-cybercrime-discussion-guide/.

into such agreements with countries "that do not respect the rule of law and fundamental human rights."[104]

V. Conclusion

Digital censorship practices are spreading across the globe in not only authoritarian countries, but also U.S. trade partners and allies. Such censorship creates legal risks for U.S. companies seeking to operate overseas, inhibits trade and cross-border flows of speech and information, and jeopardizes the free expression and privacy rights of the users of online services, both within the U.S. and without. When providers of online services decline to invest in foreign markets, they often have less incentive to tailor their services and policies to the needs of users in those markets. This can mean less investment in translation of policies and support documentation, few or no staff with language and cultural competency to make content moderation decisions, and an overall lack of awareness of how a service is affecting individuals and communities in the country. The Commission should recommend that U.S. trade policy seek to promote trade agreements, laws and policies that prohibit or limit government digital censorship practices and help ensure that both governments and companies will protect and respect individuals' free expression and other human rights.

July 22, 2021

Emma Llansó
Caitlin Vogus
Maura Carey

Center for Democracy & Technology

[104] Cory & Dascoli, *supra* note 63.