January 19, 2023

UNESCO
7 Pl. de Fontenoy
75007 Paris, France

RE: Comment Regarding Guidance for Regulating Digital Platforms

The undersigned are affiliated with the Institute for Technology, Law & Policy (ITLP) at the University of California, Los Angeles (UCLA). ITLP is a collaboration between the UCLA School of Law and the UCLA Samueli School of Engineering whose mission is to foster research and analysis to ensure that new technologies are developed, implemented, and regulated in socially beneficial, equitable, and accountable ways. We are grateful for the opportunity to submit a comment and invite UNESCO to consider this input.

I. Introduction

UNESCO's "Guidance for regulating digital platforms: a multistakeholder approach" framework ("the framework") addresses some of the most complex and challenging regulatory problems of our era: related to online speech and user generated content; hate speech, misinformation, and disinformation; and protection of freedom of expression on the internet. Interventions designed to curb problematic speech can easily be weaponized to silence free expression, particularly expression by dissidents and minority voices. Rules designed to safeguard free expression can lead to a proliferation of hate and disinformation, an erosion of accurate and reliable information, and ultimately can undermine trust in public institutions.

We commend UNESCO for its engagement in this vital area. However, we feel that there are important changes which will help the framework achieve positive change. Specifically, the framework should address the surveillance economy that underlies the platforms' content moderation decisions; the framework should account for alternate digital platform models, like nonprofit platforms or distributed (federated) social media; the framework's transparency requirements, for both governments and platforms, should be reworked to be more effective; and there should be a variety of changes with the framework's guidelines on content restrictions.

II. Freedom of Expression and the Surveillance Economy

First, the framework cannot be effective without addressing the surveillance economy that underlies the business models of nearly all of the largest and most widely used digital platforms. This business model is at the core of the platforms' baseline incentive to prioritize engagement above all else, not just to show users more ads, but also to collect more data from them.[1] This data is then used to refine ad targeting algorithms in a neverending iterative process. This dynamic, in turn, drives important content decisions such as promoting polarizing content above, say, enriching or educational content.[2] From Myanmar,[3] to the Philippines,[4] the surveillance economy, and its emphasis on targeted advertising, has played an enormous role in driving the development of our modern content paradigm.[5]

The draft framework leaves the surveillance economy entirely unaddressed.[6] Without considering the business model, and its accompanying incentive structure, which drives the online discourse, the framework is at best incomplete and, at worst, risks further entrenching the internet's "original sin".[7] While we understand a desire to limit the scope of the framework, in order to keep it from stretching too far from UNESCO's core areas of focus, freedom of expression and digital privacy have become inextricably linked in the digital age.

We respectfully ask that the framework be revised to address the challenges to freedom of expression flowing from the surveillance economy. This requires detailed and specific recommendations, including:

---

[1] Nathalie Maréchal & Ellery Roberts Biddle, *A Tale of Two Algorithms*, *in* It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge (last updated Mar. 17, 2020), https://www.newamerica.org/oti/reports/its-not-just-content-its-business-model/a-tale-of-two-algorithms.

[2] *See, e.g.*, Adam Greenwood, *Enrage to Engage: How Social Media's Algorithms Disseminate Radical Content to Maximise Screen Time*, LinkedIn (Feb. 25, 2021), https://www.linkedin.com/pulse/enrage-engage-how-social-medias-algorithms-radical-screen-greenwood.

[3] Paul Mozur, *A Genocide Incited on Facebook, With Posts From Myanmar's Military*, New York Times (Oct. 15, 2018), https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.

[4] Lauren Etter, What Happens When the Government Uses Facebook as a Weapon, BLOOMBERG (Dec. 7, 2017), https://www.bloomberg.com/news/features/2017-12-07/how-rodrigo-duterte-turned-facebook-into-a-weapon-with-a-little-help-from-facebook.

[5] *See, e.g.*, Diego Naranjo & Jan Penfrat, *Surveillance-based Advertising: An Industry Broken by Design and by Default*, European Digital Rights (Mar. 9, 2021), https://edri.org/our-work/surveillance-based-advertising-an-industry-broken-by-design-and-by-default; Bennett Cyphers & Adam Schwartz, *Ban Online Behavioral Advertising*, Electronic Frontier Foundation (Mar. 21, 2022), https://www.eff.org/deeplinks/2022/03/ban-online-behavioral-advertising; Sarah Myers West, *Data Capitalism: Redefining the Logics of Surveillance and Privacy*, *Business & Society (2019)*, 58(1), 20–41, 10.1177/0007650317718185; Shoshanna Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Public Affairs (2019); Paul Lewis, 'Fiction is Outperforming Reality': How YouTube's Algorithm Distorts Truth, The Guardian (Feb. 2, 2018), https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth; Safiya Noble, *How Search Engines Amplify Hate — in Parkland and Beyond*, Time (Mar. 9, 2018), https://time.com/5193937/nikolas-cruz-dylann-roof-online-white-supremacy.

[6] The framework has cursory mention of privacy as a human right, but only as a counterbalance to requirements that the framework sets forth. *See* framework at para. 28, 29, 38, 55. The framework does not directly address privacy or corporate surveillance.

[7] Ethan Zuckerman, *The Internet's Original Sin*, The Atlantic (Aug 14, 2014), https://www.theatlantic.com/technology/archive/2014/08/advertising-is-the-internets-original-sin/376041/.

- Requiring digital platforms to have data minimization and purpose limitation for the data they collect from users.[8]
- Give users clear control over collection, processing, and sharing of any data that digital platforms are allowed to collect from them.[9]
- Restrict how digital platforms are able to target users with content.[10] For instance, digital platforms should not be allowed to target users with content in a way that amounts to discrimination based on protected characteristics.[11]
- Support the passage of robust and enforceable national privacy and data protection frameworks.

Please note that these points are meant to be illustrative, not exhaustive as addressing an issue of this scale will require significant additional review and consultation.[12]

III. Accounting for Alternate Digital Platform Models

The introduction to the framework states that the guidance is developed for platforms whose services have the largest size and reach, while recommending that minimum safety requirements should be applied to all platform service companies regardless of size. While this is a sensible distinction, it fails to account for factors other than size and scale that are important differentiating factors for digital platforms. These include for-profit/nonprofit status, as well as the emergent alternatives to centrally managed platforms, such as federated or distributed platforms. These operate very differently from the predominant platforms who have deep pockets and are centralized in terms of control and supervision over content, necessitating a nuanced, differentiated approach to regulation.

For instance, despite having millions of registered users[13] and billions of visitors[14] from around the world, the non-profit online encyclopedia Wikipedia relies on user donations and volunteer

---

[8] Nathalie Maréchal et al., *Key Recommendations for Policymakers*, *in* Getting to the Source of Infodemics: It's the Business Model (last updated May 27, 2020), https://www.newamerica.org/oti/reports/getting-to-the-source-of-infodemics-its-the-business-model/key-recommendations-for-policymakers/.
[9] *Id.*
[10] *Id.*
[11] *Id. See also* Chandler Nicholle Spinks, *Contemporary housing discrimination: Facebook, targeted advertising, and the Fair Housing Act*, 57 Hous. L. Rev. 925 (2019-2020).
. 2019.. Hous. L. Rev. 57 (2019), 925.
[12] *See also* Article 19, *ARTICLE 19 Submission to the Second UNESCO Consultation on a "Model Regulatory Framework for the Digital Content Platforms to Secure Information as a Public Good"*, 6 (Dec. 2022) ("We note that any regulation of platforms that does not take into account competition and data protection aspects will prove ineffective in limiting the harmful effects of these platforms' business model.").
[13] English Wikipedia currently has 44,782,008 registered users. *Wikipedia:Wikipedians*, Wikipedia (last accessed Jan. 8, 2023), https://en.wikipedia.org/wiki/Wikipedia:Wikipedians.
[14] In May 2022, Wikipedia had 5.1 billion unique global visitors. *Worldwide visits to Wikipedia.org from December 2021 to May 2022,* Statista (last accessed Jan. 8, 2023), https://www.statista.com/statistics/1259907/wikipedia-website-traffic/.

administrators to run.[15] As such, despite its size and scale, Wikipedia would not be able to implement the kind of moderation and administration structures that one might expect from Facebook or Google.[16]

The framework should explicitly differentiate between for-profit and nonprofit platforms, with non-profit platforms being afforded greater flexibility in responding to regulatory challenges and being subject to less severe compliance requirements than their for-profit counterparts. Minimum safety standards should be differentiated, accounting for the differences in measures that are practicable for for-profit and nonprofit platforms.

In addition to for-profit status, the framework should also account for alternate technological models, particularly federated or distributed platforms. Unlike most popular social media platforms, such Facebook, TikTok, Google and WeChat, which own the proprietary code their services run on, and the servers on which profiles and content are hosted, distributed and federated social media platforms use open source code that users can download and use to maintain social media profiles, allowing them to set up their own servers for their personal networks and set their own content guidelines.[17] A site or server participating in such social media networks is interoperable with other participating sites or servers. While federated networks have multiple centers,[18] decentralized networks have no center at all.[19] Popular federated social media platforms include Mastodon, a crowdfunded platform where users can engage through various nodes with different policies. This allows users to pick the nodes with the policies they prefer, while allowing for interoperability between different nodes. These models create digital platforms that offer greater user flexibility, diversity of services and innovation, while doing away with a central authority that exercises control over content.[20]

The framework should account for federated and distributed digital platforms, differentiating these from traditional for-profit platforms that exercise greater control over the code, policies and content moderation practices of their services. The impact of the standards created under the framework on federated and distributed digital platforms should be considered.

[15] *Support Wikipedia*, Wikimedia Foundation (last accessed Jan. 8, 2023), https://wikimediafoundation.org/support/.
[16] Meta, the company that runs Facebook and Instagram, reported $22.05 billion operating expenses in Q3 of 2022, *see Facebook | Meta | FB - Operating Expenses*, Trading Economics (last accessed Jan. 8, 2023), https://tradingeconomics.com/fb:us:operating-expenses, while the Wikimedia Foundation that runs Wikipedia reported $140 million operating expenses for the whole of 2022, *see Wikimedia Foundation*, Wikipedia (last accessed Jan. 8, 2023), https://en.wikipedia.org/wiki/Wikimedia_Foundation#Expenses.
[17] Richard Esguerra, *An Introduction to the Federated Social Network*, Electronic Frontier Foundation (Mar. 21, 2011), https://www.eff.org/deeplinks/2011/03/introduction-distributed-social-network.
[18] These have been described as "distributed network[s] of centralized networks." *Beyond Distributed and Decentralized: What is a Federated Network?*, Institute of Network Cultures (last accessed Jan. 8, 2023), https://networkcultures.org/unlikeus/resources/articles/what-is-a-federated-network/.
[19] *Id.*
[20] *Supra note* 17.

IV. Transparency

Robust transparency is vital to an effective and accountable regulatory framework, both in terms of decision-making by governments and by online platforms. While the Draft's mention of increased transparency as a primary goal for the proposed regulatory framework is appreciable,[21] its current language leaves some pressing issues unaddressed.

Transparency of Government requests

With governments increasingly interested in digital platforms as sites of public discourse and information sharing, platforms have reported a record number of requests from governments to take down or restrict access to content.[22] While takedown requests may be a legitimate tool against illegal content, they are prone to abuse from governments, who may weaponize them to suppress critical voices by blocking content that is legal, but which is viewed as politically problematic.[23] This makes it important to create safeguards around government takedown requests to ensure transparency and accountability in their use.

While the framework states that governments should be open, clear and specific about the type and volume of takedown requests they make to digital platforms, it goes on to envisage situations of 'sensitivities' around publicizing some requests, listing national security related concerns and prevention of serious crime as examples.[24] This is at odds with the endorsement of transparency as a key value of the framework, and risks legitimizing 'secret' blocking through tacit approval of such government conduct. This also potentially conflicts with recommendations in 23.4 and 23.5 of the framework, requiring transparency and due process of law around any government takedown requests. Exempting categories such as national security and prevention of serious crimes from disclosure requirements opens the door to abuse from governments, especially given the lack of guardrails around what types of content could fall into such categories. National security exemptions in particular risk becoming a free pass for unaccountable government discretion - as

[21] The framework mentions platforms' responsibility to be transparent about their operations and policies as well as that of governments to be transparent about takedown requests and other requirements from platforms.
[22] Elizabeth Culliford, *Twitter Sees Record Number of Govt Demands to Remove Content*, Reuters (Jan. 25, 2022), https://www.reuters.com/technology/exclusive-twitter-sees-record-number-govt-demands-remove-content-japan-russia-2022-01-25.
[23] Shreya Tewari, *Stark Increase in Government Takedown Requests in Lumen*, Lumen (July 19, 2022), https://www.lumendatabase.org/blog_entries/stark-increase-in-government-takedown-requests-in-lumen.
[24] *See* framework at para. 18.

has been demonstrated in the context of takedown requests in multiple contexts,[25] as well as in other areas of law, such as trade and tariff laws.[26]

The framework should clearly state that all government takedown requests should be subject to public disclosure in keeping with the governments' responsibility to remain transparent and accountable.

In addition to encouraging public disclosure of government takedown requests, the framework should address the prevalent practice of government Internet Referral Units (IRUs) which submit non-binding requests to content platforms, asking them to voluntarily remove legal content from their platforms that is (in the IRU's interpretation) in violation of the platform's policies. This has been criticized as a new system of informal governance which allows governments to remove content that they could not remove through legal takedown requests, effectively sidestepping legal or constitutional limits on the government's ability to target speech.[27] Such informal governance poses multiple risks - an intrusion on free speech and public law norms, company adoption of a government's interpretation of their terms of service, and impeding human rights documentation efforts through IRU's broad interpretation of terms such as 'terroristic content'.[28] In 2019, the Internet Archive was on the receiving end of hundreds of false takedown requests from a European IRU, with falsely identified 'terrorist content' URLs including out-of-copyright literature and legal, public domain content.[29]

The framework should address the growing use of IRUs and other informal measures, making it clear that specific content or accounts should only be removed at the behest of governments through a formal takedown request.

Transparency for Platforms

---

[25] See, e.g., Ernesto Van der Sar, *Google Reveals Surge in Questionable Removal Requests From Russian Government*, TorrentFreak (Dec. 10, 2022), https://torrentfreak.com/google-reveals-surge-in-questionable-removal-requests-from-russian-government-221209/; Damien Sharkof, *What Russia Does Not Want You to See: Kremlin Tops List for Flagging Google Content*, Newsweek (July 21, 2017), https://www.newsweek.com/what-russia-does-not-want-you-see-kremlin-tops-list-flagging-google-content-640114.

[26] Scott Lincicome & Inu Manak, *Protectionism or National Security? The Use and Abuse of Section 232*, Cato Institute (Mar. 9, 2021), https://www.cato.org/policy-analysis/protectionism-or-national-security-use-abuse-section-232.

[27] Rabea Eghbariah & Amre Metwally, Informal Governance: Internet Referral Units and the Rise of State Interpretation of Terms of Service, 23 Y. J.L. and Tech. 542 (Spring 2021), https://yjolt.org/sites/default/files/23_yale_j.l._tech._542_informal_governance.pdf.

[28] *Id.*

[29] Gareth Halfacree,
Archive.org Hit by False Terrorist Takedown Notices, bit-tech (Apr. 11, 2019), https://bit-tech.net/news/tech/software/archiveorg-hit-by-false-terrorist-takedown-notices/1/.

The framework's approach to transparency for platforms requires additional clarification to adequately account for business and technological realities. In its current form, it calls for "cooperation between the companies providing services and the regulatory systems while being effective and implementable and providing real accountability."[30] While a multistakeholder approach may be useful in dealing with regulatory challenges around digital platforms, in endorsing such an approach, the framework should explicitly address concerns around consolidation of power in unaccountable cooperation structures.[31] This is especially relevant as a handful of companies control the dominant digital platforms, exercising enormous power and influence through network effects and economies of scope and scale.[32] As scholars[33] and governments[34] globally grow increasingly concerned about consolidation in the digital space and its antitrust implications,[35] the framework should go beyond merely calling for accountability.

In order to create meaningful accountability, cooperation structures should include civil society organizations and digital rights watchdog groups as participants on an equal footing with companies and governments, and issue regular, publicly accessible reports detailing their deliberations and decisions. Further, the framework should explicitly consider the impacts of such cooperation structures on competition, and the potential problems related to harmonizing content policies across the dominant platforms.

Additionally, while the framework approaches transparency for platforms primarily through the lens of explainability of policies[36], this is not, by itself, enough. To be an effective driver of accountability, the 'right to explanation', must be accompanied by other transparency requirements that work in conjunction with the right to create systematic transparency around algorithmic decisions.[37] While the right to explanation is an individualized right that exists *post facto* (i.e. once an algorithm has already been created and used to make decisions) the General Data Protection Regulation (GDPR) also offers a number of ways to implement algorithmic accountability at various levels. First, at the government level, the GDPR empowers regulators to get access to

---

[30] *See* framework at para. 20.

[31] https://www.justsecurity.org/72603/gifct-possibly-the-most-important-acronym-youve-never-heard-of/

[32] Charlotte Slaiman, *Why Dominant Digital Platforms Need More Competition*, Centre for International Governance Innovation (Apr. 13, 2020), https://www.cigionline.org/articles/why-dominant-digital-platforms-need-more-competition/.

[33] Steven C. Salop, *Dominant Digital Platforms: Is Antitrust Up to the Task?*, 130 Y.L.J. Forum (Jan. 18, 2021), https://www.yalelawjournal.org/forum/dominant-digital-platforms.

[34] *Abuse of Dominance in Digital Markets*, Organisation for Economic Co-operation and Development (last accessed Jan. 8, 2023), https://www.oecd.org/daf/competition/abuse-of-dominance-in-digital-markets.htm.

[35] *FTC Alleges Facebook Resorted to Illegal Buy-or-Bury Scheme to Crush Competition After String of Failed Attempts to Innovate*, Federal Trade Commission (Aug. 19, 2021), https://www.ftc.gov/news-events/news/press-releases/2021/08/ftc-alleges-facebook-resorted-illegal-buy-or-bury-scheme-crush-competition-after-string-failed.

[36] *See* framework at para 22.2.

[37] Kaminski, Margot E., The Right to Explanation, Explained (June 15, 2018). U of Colorado Law Legal Studies Research Paper No. 18-24, Berkeley Technology Law Journal, Vol. 34, No. 1, 2019, Available at SSRN: https://ssrn.com/abstract=3196985 or http://dx.doi.org/10.2139/ssrn.3196985.

information about algorithms,[38] and envisions general data protection audits carried out by regulators.[39] Second, at the firm level, companies using algorithms to make decisions are required to set up internal accountability and disclosure regimes, perform data protection impact assessments[40] and provide information to an internal, independent data protection officer. Third, when companies use algorithms to make decisions with a "high impact on individuals", guidelines suggest they should use independent third party auditing, and provide the auditor with "all necessary information about how the algorithm or machine learning system works".[41]

The framework should include measures to create systematic accountability, including internal, third party/expert and regulatory oversight over the development of such algorithms from their inception. This should include consideration of what type of information is datified, and how that data is collected, aggregated, and incorporated into machine learning and AI.

While dealing with platforms' content management policies, the framework envisions situations where there may be a tension between national laws and global human rights standards, and states that platforms will be expected to report on how it responds to requests to remove content that is illegal under national law in violation of international human rights law.[42] While we do not disagree with this, we believe that platforms should universally report their responses to all government requests, regardless of jurisdiction. The current language of the framework may create confusion as to the universal nature of this requirement.

The framework should modify this language and make it clear that such reporting is required universally, and not only in select jurisdictions or cases where there is a tension between national laws and global human rights standards.

On the issue of election integrity and political advertising, the framework recommends that platforms retain political advertisements and relevant information on funding in a publically accessible online library. However, given the difficulty of determining whether an advertisement is political or not and the benefits of having a library of all targeted advertising, this section should be expanded to include requirements that ad libraries for all targeted advertising meeting a certain threshold be included .

---

[38] Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1 at art. 58(1)(e) (authorizing the authority to carry out data protection audits, and "obtain, from the controller and the processor, access to all personal data and to all information necessary for the performance of its tasks").
[39] *Id.* at art. 58(1)(b)
[40] *Id.* at art. 35(3)(a) (requiring a data protection impact assessment "in a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person")[hereinafter GDPR]; GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, at 29–30 (explaining that this requirement "will apply in the case of decision-making including profiling).
[41] *Id.* at arts. 22, 13, 14, 15.
[42] *See* framework at para 27.2.

We agree with this recommendation, and further suggest that such information be recorded in open formats and harmonized styles between platforms, ideally in a common library of all ads, shared by all relevant platforms.[43]

V. Content Restrictions

We take issue with several aspects of the framework's approach to content moderation:

- **Paragraph 27.1:** *"Platforms should, in policy and practice, through adequately trained and staffed personnel, ensure that, at a minimum, there is quick and decisive action against child sexual abuse materials, promotion of terrorism, promotion of genocide, clear threats of violence, gender-based violence and incitement to hatred based on protected characteristics."* There are two issues here. First, "promotion of terrorism" is not defined, which opens the door to undue pressure by governments who often seek to define "terrorism" in politically advantageous ways.[44] Dissidents and anti-government protesters are often censored by risk-averse platforms that do not want to leave up protest speech that some government calls "promotion of terrorism."[45] Second, Paragraph 27.1 demands the same "quick and decisive action" against both child sexual abuse materials (CSAM) and hate speech or threats of violence. These kinds of content should not have the same enforcement protocols. CSAM, as a general category of proscribed content, is vastly easier to identify through automated means since it is generally not contextually dependent, and the risk of "by-catch" related to filtering mechanisms which target CSAM is much lower than hate speech or threats of violence, which require a more nuanced and contextual assessment.[46] By requiring the same standard for context-dependent categories of expression like hate speech, the framework encourages trigger-happy censorship that risks shutting down legitimate expression.

---

[43] Radsch, Courtney. "Transparency Reporting: Good Practices and Lessons from Global Assessment Frameworks." GIFCT Transparency Working Group. Global Internet Forum to Counter Terrorism (GIFCT), July 2022. https://gifct.org/wp-content/uploads/2022/07/GIFCT-22WG-ResearchAgendaScopingPaper-1.1.pdf.

[44] Courtney Radsch, *Taking Down Terrorism Online While Preserving Free Expression*, Medium (May 15, 2019), https://courtneyr.medium.com/taking-down-terrorism-online-while-preserve-our-free-expression-89249fb260f9.

[45] Note that dissidents and anti-government protesters already experience high levels of censorship on digital platforms. *See, e.g.*, Courtney Radsch, *On Christchurch Call Anniversary, a Step Closer to Eradicating Terrorism Online?*, Just Security (May 21, 2021), https://www.justsecurity.org/76607/on-christchurch-call-anniversary-a-step-closer-to-eradicating-terrorism-online/; Simon Speakman Cordall, *Facebook Deactivates Accounts of Tunisian Political Bloggers and Activists*, The Guardian (June 4, 2020), https://www.theguardian.com/global-development/2020/jun/04/facebook-deactivates-accounts-of-tunisian-political-bloggers-and-activists.

[46] Indeed, CSAM is filtered by many digital platforms using artificial intelligence. *See, e.g.*, Courtney Radsch, *Artificial Intelligence and Disinformation: State Aligned Information Operations and the Distortion of the Public Sphere*, Organization for Security and Co-Operation in Europe, 5 (July 2022), https://www.osce.org/files/f/documents/e/b/522166.pdf ("Social media and other internet platforms depend on AI systems to enforce their Terms of Service and rules governing acceptable speech and behavior on their platforms. For example, automated detecting, filtering and blocking of child sexual abuse material … .").

The framework should recognize that an appropriate standard for targeting CSAM may differ from appropriate protocols for responding to other categories of harmful content, that require a higher standard of human intervention due to their contextual nature.

- **Paragraph 33.4:** *"Content removal or de-platforming of users should be considered only when the intensity, and severity of content that has the intention to harm a group or individual occurs."* Platforms have their own freedom of expression, which includes the freedom to define their own standards of unacceptable speech.[47] If a platform decides that it will only host sports-related content, and all other content will be deleted, such a policy would be perfectly legitimate. Allowing the platforms this level of discretion is important for creating an online environment free of harassment and discrimination, which is necessary to fully realize freedom of expression for many users.

  The recommendation that content removal or de-platforming should only be contemplated in particularly severe cases should be removed.

- **Paragraph 35.3:** *"Political advertisements which refer to issues rather than parties or candidates should be scrutinised to ensure they are consistent with the overarching policies of the platform in relation to hate speech or speech targeting people with protected characteristics."* This language seems to establish a higher standard of review – and thus a greater likelihood of removal – for political expression. This runs counter to a broader recognition that, in a robust democracy, political speech should be afforded the highest level of protection.[48] It also risks chilling the speech of civil society organizations, activists, and journalists, as "political" is a broadly relevant term and advertisements may refer to any paid promotion of specific content, such as awareness raising around issues of public policy.

  The recommendation that political advertising should be more carefully scrutinized should be removed.

VI. Conclusion

[47] Indeed, in the United States, platforms are explicitly allowed by statute to censor content. 47 U.S.C. § 230(c) ("No provider or user of an interactive computer service shall be held liable on account of— (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.").

[48] *See*, *e.g.*, Michael Karanicolas, Subverting Democracy to Save Democracy: Canada's Extra-Constitutional Approaches to Battling 'Fake News', 7(2) Canadian Journal of Law and Technology 201, 206 (2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=342309 ("In a democratic society, political speech, and in particular speech around elections, cuts to the core of freedom of expression guarantees.arantees.").

While we appreciate the substantive contributions of the framework to the global discourse around online speech, we believe that there are areas of improvement, particularly through an added focus on expressive challenges related to the surveillance economy, and the role of nonprofit and federated/distributed digital platforms in the social media landscape. We also call for the transparency guidance to be expanded and clarified, and for additional changes to the guidance on content restrictions. Without these changes, the framework risks legitimating bad policies which are counter to global freedom of expression values.

Safeguarding freedom of expression online is a phenomenally complex challenge. Government regulations and platform policies intended to solve problems can easily cause harm. However, there are some clear paths forward. We appreciate the opportunity to contribute to this process, and look forward to further discussions in Paris and beyond.

Respectfully submitted,

Nicholas Wilson
Research Assistant
UCLA Institute for Technology, Law & Policy

Akshat Agarwal
Research Assistant
UCLA Institute for Technology, Law & Policy

Courtney Radsch
Resident Fellow
UCLA Institute for Technology, Law & Policy

Michael Karanicolas
Executive Director
UCLA Institute for Technology, Law & Policy