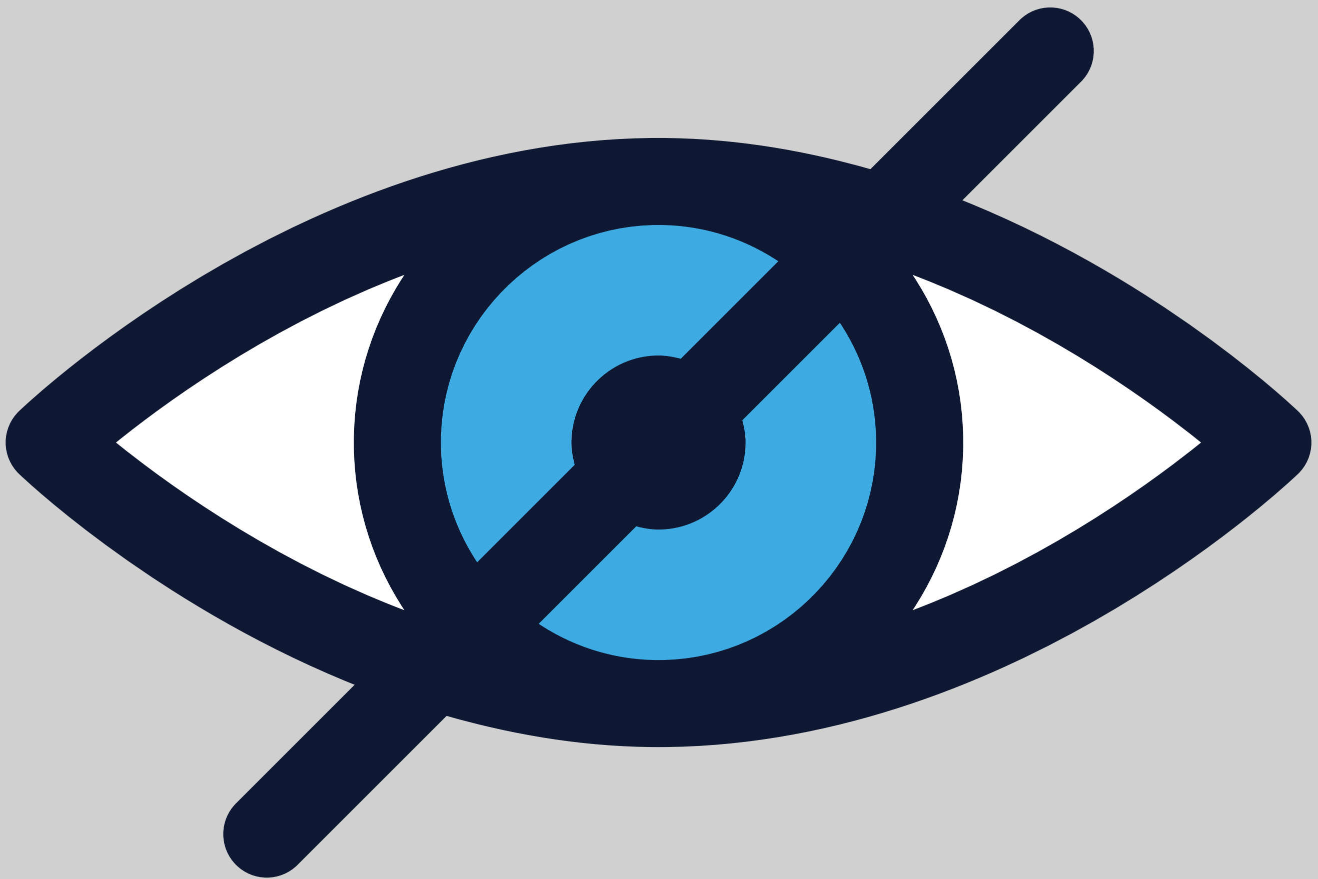
CONTENT REGULATION

JUNE 2022



Competition Policy International, a What's Next Media and Analytics Company

TechREG
EDITORIAL TEAM

Senior Managing Director
Elisa Ramundo

Editor in Chief
Samuel Sadden

Associate Editor
Andrew Leyden

TechREG
EDITORIAL BOARD

Editorial Board Chairman
David S. Evans – *GEG & University College London*

Martin Cave – *London School of Economics*

Avi Goldfarb – *University of Toronto*

Hanna Halaburda – *New York University*

Liyang Hou – *Shanghai Jiao Tong University*

Katharine Kemp – *University of New South Wales*

Kate Klonick – *St. John's University*

Mihir Kshirsagar – *Princeton University*

Philip Marsden – *Bank of England / College of Europe*

Saule Omarova – *Cornell University*

Eric Posner – *University of Chicago*

Xavier Vives – *IESE Business School*

LETTER FROM THE EDITOR

Dear Readers,

The issue of how to manage online content scarcely leaves the headlines. The key issues include questions of content moderation, algorithm transparency, and the potential for online platforms to be used to abuse or victimize certain individuals or groups. This can take the form of hate speech or personal harassment, among other forms of problematic speech.

Recent controversies have included the alleged use of social media platforms to spread disinformation surrounding recent elections, whether certain political figures should be excluded from key social media platforms, and the underlying question of whether given online platforms (and individuals) simply hold too much power over public discourse.

Perhaps inevitably, these questions have led to calls for regulation. But this opens a Pandora's Box. What entities should be regulated? What type of content should fall within the scope of such regulation? What criteria should be used to determine what type of content is acceptable and what is not? And who should make such determinations? Should the industry self-regulate or is greater government oversight required? And how is such regulation to be squared with the fundamental value of free speech (as understood under the First Amendment to the U.S. Constitution, Article 10 of the European Charter of Fundamental Rights, and countless human rights instruments and constitutions worldwide)?

The timely articles in this Chronicle address these and other issues in light of the latest developments the world over.

Natascha Just assesses the relevant provisions of the forthcoming European Digital Services Act ("DSA"). As the article notes, the DSA is one of many pieces of legislation that seek to negotiate a workable social contract, taking into account the interests of various participants

in terms of social roles and acceptable online behavior. In particular, the article assesses how the DSA addresses questions of platforms' responsibility for content moderation by creating an asymmetric system of due diligence obligations.

In turn, **Terry Flew** raises a home truth inherent to any attempt to regulate online content. The distinctive platform business models of the key players raise significant challenges for regulators. To keep up with the pace, the paper argues that regulators need to "think like Google": they need to be able to adopt holistic strategies that can apply across industry silos and different regulatory responsibilities.

Scott Babwah Brennen & Matt Perault turn to a more specific concern: the banning of former U.S. President Donald Trump from several online platforms. The paper argues that we still lack sufficient empirical analysis of the positive and negative consequences these bans have had on public discourse and extremism. The article makes an important contribution by setting out 13 metrics that analysts or researchers could consider as a means of evaluating the impacts of this (and other) bans of public figures from social media platforms.

Imanol Ramírez turns to the difficulty of reaching consensus over how online content should be moderated around the world. The result has been regulatory fragmentation, which has increased the cost of operating in digital global markets due to greater entry and expansion barriers. Nonetheless, policymakers and researchers could take advantage of these divergences to test and prove the effects of different regulations by using the rich dataset such fragmentation produces. Specifically, the rich dataset divergences produce can be used order to question and test the effectiveness of different approaches through empirical methods. This would allow policymakers to better understand how the different regimes shape the conduct of intermediaries and make policy decisions accordingly.

Gregory Day explores the social and societal costs of market concentration in digital platforms, specifically as they relate to mental health and perceptions of beauty. Image and video sharing markets are particularly concentrated, with only three major players accounting for the majority of usage worldwide. This fascinating piece outlines a problem known to aesthetics scholars, but which has evaded legal scholarship: the effect of tech programs on perceptions of beauty and attendant dangers. It then discusses the growing demand for regulations of certain types of apps, platforms, and tech companies in order to present potential ways that the law could ameliorate some of the alleged harms.

Finally, **Marta Cantero Gamito** notes that any policy choices made as regards the regulation of online content will impact freedom of expression, as the new rules promote a sort of "standardization" of content moderation procedures. Compliance with such regulations might be ensured by adopting recognized European and international standards. That said, there is the risk that a "one-size-fits-all" approach would run the risk of compromising constitutional pluralism and result in preventive censorship across platforms. As the article warns, critical political decisions should not be lost in seemingly technical discussions.

In sum, this set of articles provides a fascinating snapshot of how the content regulation debate is currently evolving. As many of the authors note, it is early days yet, and this is a debate that will doubtless rage on for years to come. As the first generation of content regulation makes itself felt, platforms, users, and public figures alike will have much to say as to how such regulation applies in practice.

As always, many thanks to our great panel of authors.

Sincerely,
CPI Team

TABLE OF CONTENTS

Letter from the Editor 04

Summaries 06

The Taming of Internet Platforms – A Look at the European Digital Services Act
by **Natascha Just** 08

Why Tech Regulators Need to Think Like Google
by **Terry Flew** 16

What we talk about when we talk about deplatforming Trump
by **Scott Babwah Brennen & Matt Perault** 24

Aesthetics, Technology, And Regulations
by **Gregory Day** 32

CPI COMPETITION POLICY™ INTERNATIONAL

CONTENT REGULATION

JUNE 2022

40

Randomized Experiments for Online Content Moderation Policy

by Imanol Ramírez

46

Platform Content Standardization

by Marta Cantero Gamito

54

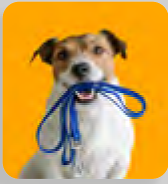
What's Next?

54

Announcements

© 2022 Competition Policy International All Rights Reserved

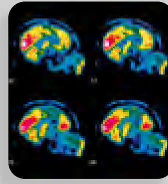
SUMMARIES



THE TAMING OF INTERNET PLATFORMS – A LOOK AT THE EUROPEAN DIGITAL SERVICES ACT

By Natascha Just

Various regulatory attempts at taming global Internet platforms have entered the stage worldwide. These set out to renegotiate the cornerstones of a workable social contract and the expectations of the various participants in terms of social roles, acceptable behavior and reasonable means. In this vein, the European Digital Services Act ("DSA") takes on questions of platforms' responsibility for content moderation with an asymmetric system of due diligence obligations. This comprises the assessment of systemic risks that may arise from platform services, which includes risks resulting from the dissemination of illegal content or those that negatively affect the exercise of fundamental rights such as the freedom of expression and information. The fact that the responsibility for this systemic risk assessment and the deployment of mitigation measures against these risks rest primarily with very large online platforms ("VLOPs") and their interpretive sovereignty raises various concerns. A major question is what cultural imprint this will inflict on fundamental rights in Europe and what normative values will eventually be accentuated.



WHY TECH REGULATORS NEED TO THINK LIKE GOOGLE

By Terry Flew

As many of the world's largest companies are platform-based technology companies, there has been a growing push worldwide to regulate these companies to address issues arising from economic, political and communications power. At the same time, their distinctive platform business models raise new challenges to regulators, such as what industries they are in, what problems connect to which regulatory authority, and who has jurisdictional authority and regulatory capacity. The paper argues that regulators increasingly need to "think like Google": they need to be able to adopt holistic strategies that can apply across industry silos and different regulatory responsibilities. There is also a need to empower the notion of regulation in the public interest, to challenge the ideational power of tech companies that they are superior stewards of public good to government agencies.

WHAT WE TALK ABOUT WHEN WE TALK ABOUT DEPLATFORMING TRUMP

By Scott Babwah Brennen & Matt Perault

Following his role in the January 6th attacks on the Capitol, several online platforms banned President Donald Trump from their platforms. While scholars and commentators have spent 18 months debating the value of removing Trump from major online platforms, we still lack sufficient empirical analysis of the positive and negative consequences the ban has had on public discourse, extremism, and Trump's power and influence. But before we can assess the costs and benefits of removal, we first must develop a common understanding of the metrics we might use to evaluate the impacts of the ban. When we talk about whether deplatforming worked or failed, what do we mean by "worked" and what do we mean by "failed"? This article provides 13 metrics that analysts or researchers could consider as a means of evaluating the impacts of the ban. While analyzing any of these will involve overcoming serious challenges, from data access, to resources, to how to attribute observed changes narrowly to Trump's removal, identifying what metrics we should consider is an important first step to understanding the impact of banning Trump.

AESTHETICS, TECHNOLOGY, AND REGULATIONS

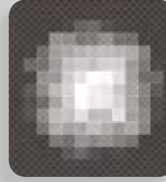
By Gregory Day

Technology companies have increasingly come under regulatory fire for impairing society, markets, competition, and free speech, among other things. The underlying belief is that consumers require protection in digital markets. Despite the spectrum of harms attributed to Big Tech, relatively sparse attention has been paid to society's relationship to aesthetics and image. In contrast to traditional forms of media where people passively view images of others, tech platforms allow people to manipulate their own photographs. By doing so, a belief is that unhealthy perceptions of beauty are supercharged compared to conventional mediums. This piece isn't necessarily claiming that tech's effects on aesthetic perceptions must come under greater regulatory scrutiny. Rather the goal is to discuss the nature and depth of a largely underspecified issue, which is related to many problems that have drawn the ire of commentators. It is indeed important to acknowledge how tech platforms influence perceptions of beauty and even views of self-worth in ways that were previously unknown — and whether this should implicate modern demands for tech regulation.

RANDOMIZED EXPERIMENTS FOR ONLINE CONTENT MODERATION POLICY

By Imanol Ramírez

The difficulty of achieving consensuses over the regulation of online content moderation has created a stringent and divergent regulatory framework around the world. This fragmentation increases the cost of operating in digital global markets due to greater entry and expansion barriers. Given that countries' legal standards over the regulation of content moderation remain too far apart from each other, international law does not seem to offer a solution in this respect. Nonetheless, policymakers and researchers could start taking advantage of the divergent legal environment and the data richness that characterizes the digital economy to test and prove the effects of the different regulations in place. Today, opinions and proposals are based to a great extent on intuitive assumptions or theoretical ideas. Thus, it is necessary to start questioning and testing those ideas through empirical methods. Applying the rationale used in randomized experiments to validate the intuitive assumptions that fill the debate, such as the alleged effects of intermediary liability, including the chilling-effect over speech, would allow policymakers to better understand how the different regimes shape the conduct of intermediaries and make policy decisions accordingly.



PLATFORM CONTENT STANDARDIZATION

By Marta Cantero Gamito

As the enactment of the DSA is coming closer, anticipations around the new rules for curbing Big Tech power are mounting. A revision of the eCommerce Directive has been long due, but its modernization will indeed bring significant and necessary changes to make the online space a safer one, establishing procedural guarantees that protect fundamental rights and democracy online. However, it cannot be forgotten that the proposal relies on the harmonization of the internal market as a legal basis, and that this means that the rules will have an intense market regulation flavor. This policy choice will impact freedom of expression, as the new rules promote a sort of "standardization" of content moderation procedures. For example, compliance with regulatory obligations can be ensured by adopting recognized European and international standards. While, in principle, establishing similar guarantees for all platforms is needed, a corseted one-size-fits-all approach to content moderation could run the risk of compromising constitutional pluralism and to result in preventive cancelations across platforms. To avoid this, attention should be paid to discussions within standard-setting organizations following DSA's adoption. Critical political decisions should not get lost into seemingly technical discussions.

THE TAMING OF INTERNET PLATFORMS – A LOOK AT THE EUROPEAN DIGITAL SERVICES ACT

BY
NATASCHA JUST

Professor of communication at the IKMZ – Department of Communication and Media Research of the University of Zurich. Chair of the Media & Internet Governance Division.

01 INTRODUCTION

When William Shakespeare wrote his comedy "The Taming of the Shrew" in the late 16th century the term and the social role of the shrew were apparently set, as were the means to achieve what was expected from them. Etymologically, the shrew – a small insectivorous, mouse-like mammal with at that time a supposedly venomous nature – first became associated with spiteful people more generally to then stand only for unpleasant, ill-tempered and maladjusted women. The latter's ideal

© 2022 Competition Policy International All Rights Reserved

social role was to be meek and submissive and the male means to achieve this thuggish. But Shakespeare's play is also about the problem of illusion and reality, about false and true identities. This theme is set in the framing introduction in which an illustrious hunting party gets a drunk tinker to wake up believing he is a lord. The wife they put beside him is a man in disguise and the traveling players who eventually stage the main play are instructed to ignore his odd behavior, which does not befit his supposed rank.

Definitions and identities, social roles and norms, behavior, expectations and means – these are all important but similarly contested and rapidly changing facets in the governance of globally active Internet platforms that are currently facing tough negotiations worldwide. Are they tech or media companies? Should they assume sovereign tasks? Are they the only shrews here? Barely a quarter century has passed since John Perry Barlow's "A Declaration of the Independence of Cyberspace" in 1996.[2] This called on governments to stay away from it, declaring that cyberspace problems would be solved by its own social contract, and anyone would be able to enter without privilege or prejudice and to express beliefs without fear of being coerced into silence.

> ***When William Shakespeare wrote his comedy "The Taming of the Shrew" in the late 16th century the term and the social role of the shrew were apparently set, as were the means to achieve what was expected from them***

In the meantime, a handful of mostly U.S. companies with a corresponding cultural imprint have structurally transformed our societal communications system. Alphabet, with its core moneymaker Google, Meta with Facebook and Instagram, or Twitter and the like are now taking over, complementing and changing social functions that were traditionally held by national media and communications companies. At the same time, they dwarf any of these traditional companies with their economic power and play a specific role in the constructions of our realities with their intermediary gatekeeping powers[3] and by employing sophisticated, opaque automated algorithmic-selection services.[4] The large market shares and business practices of these platforms have increasingly become a cause for concern, as have their strategic role and influence regarding access to and curation of content.

This rise to power and the specific roles of Internet platforms within and for our public communications system increasingly direct attention to some more dysfunctional elements that have become evident, such as the dissemination of illegal content, disinformation or hate speech as well as chilling effects or discrimination. Events such as the U.S. Capitol riots in early 2021, which were incited by then-President Donald Trump's tweets alleging vote fraud, mark a sad turning point, directing attention away from the Internet's democracy-enhancing potential to its potentially democracy-endangering force. The early libertarian euphoria has given way to a kind of disillusionment, and various regulatory attempts at taming have entered the stage worldwide. These set out to renegotiate the cornerstones of a workable social contract and the expectations placed on the various participants in terms of social roles to be filled, acceptable behavior and reasonable means.

One of the more recent European attempts at this is the Digital Services Act ("DSA"),[5] which was proposed by the European Commission alongside the Digital Markets Act ("DMA")[6] in mid-December 2020. The latter is an *ex ante* regulation to assure fair and contestable markets and specifically targets large gatekeepers of core platform services such as online search engines and online social networking services.[7] The former aims to contribute to a safe, predictable and trusted online environment, where fundamental rights are effectively protected.[8] Among other things, it takes on questions of platforms' responsibility for content moderation with a system that leaves a surprisingly broad

2 John P. Barlow, "A Declaration of the Independence of Cyberspace," 1996, https://www.eff.org/de/cyberspace-independence.

3 Natascha Just and Michael Latzer, "Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet," *Media, Culture & Society* 39, no. 2 (2017): 238–58, https://doi.org/10.1177/0163443716643157.it builds on (co-

4 Michael Latzer et al., "The Economics of Algorithmic Selection on the Internet," in *Handbook on the Economics of the Internet*, ed. Johannes M. Bauer and Michael Latzer (Cheltenham, Northhampton: Edward Elgar, 2016), 395–425.

5 European Commission, "Proposal for a Regulation of the European Parliament and the Council on a Single Market For Digital Services (Digital Services Act) and Amending Directive 2000/31/EC, COM(2020) 825 Final," 2020.

6 European Commission, "Proposal for a Regulation of the European Parliament and the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act), COM(2020) 842 Final," 2020.

7 For an assessment see Prabhat Agarwal, "The EU's Proposal for a Digital Markets Act – an Ex Ante Landmark," *TechREG Chronicle*, January (2022): 9–15.

8 DSA proposal, art. 1 para 2.

© 2022 Competition Policy International All Rights Reserved

scope for discretion to very large online platforms ("VLOPs") in what are socially particularly sensitive areas where fundamental rights such as free speech are at stake.

The relevant European institutions agreed on the DSA this late April, but the final text is not yet publicly available. The following comments are therefore based on the December 2020 proposal of the European Commission and the amendments agreed by the EU Parliament in its first reading in January 2022.[9]

02

THE EUROPEAN DIGITAL SERVICES ACT

Over the last twenty years, Internet platforms have been regulated in Europe by the Directive on electronic commerce of 2000 (thereafter the e-Commerce Directive),[10] which introduced liability privileges for content hosted by them. Accordingly, Internet platforms are not legally responsible for illegal content they host but required to remove or to disable access to it once they know of it. While the European liability provisions were originally inspired by the 1998 U.S. Digital Millennium Copyright Act and its rules relating to copyright infringements,[11] they also have parallels with Section 230 of the 1996 U.S. Communications Decency Act.[12] This protects providers from liability on the grounds that they are not to be treated as the publisher or speaker of any information provided by another content provider. In addition, they are also protected, provided they act in good faith and restrict access to or availability of content that the provider or users deem objectionable.

The European and the U.S. provisions were both introduced at a time when the Internet landscape was completely different. In the light of recent platform power and attendant dysfunctions, discussions about the reasonableness and fairness of these rules and the extent of relief for Internet platforms from liability have moved center stage. Reforms are being suggested,[13] have already been implemented (e.g. the contentious Art. 17 of the European Copyright Directive, which governs the use of protected content by online content-sharing service providers),[14] or have been agreed upon as in the case of the DSA.

The DSA is a Regulation, which, unlike the instrument of a Directive, which requires transposition into national law, will directly be applicable in all EU members when it enters into force. The choice of Regulation is deliberate and intended to counter legal fragmentation in the European internal market that may arise upon transposition. Further, it aims at curbing solo national efforts such as recent laws that tackle content moderation and complaints regarding illegal online content, for example the KoPl-G in Austria (in force since 2021)[15] and the NetzDG in Germany (in force since 2017).[16] The DSA applies to all providers irrespective of their place of establishment, providing they offer services to recipients in the European Union, and further amends the e-Commerce Directive, from which it transfers the liability regime for Internet platforms with some additions.

9 European Parliament, "Amendments Adopted by the European Parliament on 20 January 2022 on the 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 (COM(2020)0825 – C9-0418/2020 – 2020/0361(COD)), P9_TA(2022)0014," 2022.

10 European Parliament and Council of the European Union, "Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on Certain Legal Aspects of Information Society Services, in Particular Electronic Commerce, in the Internal Market ('Directive on Electronic Commerce')," 2000.

11 17 USC Section 512 (c).

12 47 U.S.C. § 230.

13 For debates on the reform of Section 230 see, for example Danielle Keats Citron and Mary Anne Franks, "The Internet as a Speech Machine and Other Myths Confounding Section 230 Reform," *The University of Chicago Legal Forum*, 2020, 45–76; Jeff Kosseff, "A User's Guide to Section 230, and a Legislator's Guide to Amending It (or Not)," 2021, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3905347.

14 European Parliament and Council of the European Union, "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," 2019.

15 "Bundesgesetz über Maßnahmen zum Schutz der Nutzer auf Kommunikationsplattformen (Kommunikationsplattformen-Gesetz – KoPl-G) (Federal Act on Measures for the Protection of Users on Communication Platforms (Communications Platforms Act - KoPl-G)) StF: BGBl. I Nr. 151/2020 (NR: GP XXVII RV 463 AB 509 S. 69. BR: 10457 AB 10486 S. 917.)," 2020.

16 "Gesetz Zur Verbesserung Der Rechtsdurchsetzung in Sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG) (Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act)) of September 1, 2017 (BGBl. I p. 3352), as Last Amended by Article 1 of the Act to Amend the Network Enforcement Act of June 3, 2021 (BGBl. I p. 1436)," 2017.

© 2022 Competition Policy International All Rights Reserved

A. Liability Regime

The liability regime (Chapter II) and the due diligence obligations (Chapter III) are the linchpin for dealing with illegal content and therefore an important cornerstone of the DSA. It principally maintains the liability exemptions of the eCommerce Directive, which, as further elaborated in case law of the Court of Justice of the European Union ("CJEU"), offer exemptions for passive or neutral providers of intermediary services.[17] There is no liability in cases of "mere conduit" or "caching" when the providers of such service assume no active role in the transmission of content. In cases of "hosting," they are excluded from liability if the content is not provided under their control or authority, they have no knowledge of the illegal content or activity and expeditiously remove or disable access to it once they become aware of it. Further, providers do not have any general monitoring or active fact-finding obligations.

Novel to the DSA is a Good Samaritan clause similar to Section 230. Accordingly, providers will not forfeit their liability exemptions if they voluntarily carry out activities to detect, identify and remove illegal content in a diligent manner and in good faith. In this case especially, the distinction between the active and passive role of the provider as elaborated by the CJEU may be put to test and prove difficult in practice. In addition, there are new rules that indicate how providers must react when they receive an order from a national judicial or administrative authority informing them about illegal content or requesting information on a specific user and what that order must contain. Among other things, providers have to explain how they have complied with the order and when. Further, the authorities' have to explain why a specific content is illegal or required, to indicate the exact URL or other information to enable its identification as well as the territorial scope of the order and redress options.

"Novel to the DSA is a Good Samaritan clause similar to Section 230

B. Due Diligence Obligations

The DSA introduces a four-layered asymmetric system of due diligence obligations for Internet platforms and includes, among other things, notice-and-action mechanisms for illegal content, the possibility to challenge a platforms' content-moderation decision, and the obligation to conduct assessments of systemic risk. The precise obligations depend on the role, size and impact of the provider and are cumulatively applied to intermediary services, hosting providers, online platforms, and VLOPs.[18] This new scheme can be visualized either as a four-layered pyramid, where the bottom layer (i.e. intermediary services) has the least obligations and the apex (i.e. VLOPs) the most, or as a concentric nested layered system, where the outer layer has the least and the innermost the most.

The following briefly summarizes some of the due diligence obligations related to content moderation and then focuses specifically on the risk assessments that VLOPs must conduct. These touch upon very sensitive areas of fundamental rights and thus raise the question of whether Internet platforms should be the ones in charge of assessing the systemic risks their services pose and of devising the appropriate mitigation measures themselves.

1. Intermediary Services

In their terms of service all providers are required to inform publicly and unambiguously about their content-moderation policies and procedures, including algorithmic decision-making and human review, and their activities have to respect the fundamental rights of the recipients as enshrined in the Charter of the European Union.[19] In addition, they all have to publish transparency reports on content moderation, e.g. information on their own-initiative content moderation, the number of complaints received through their internal complaint-handling system, together with the types of alleged illegal content and the time needed for taking decisions.

2. Hosting Services and Online Platforms

Providers of hosting services, including online platforms, have to further install notice-and-action mechanisms that permit easy notification of illegal content, including possibilities to submit all necessary information to identify the illegality of content (e.g. explanations of illegality, URL, name

17 According to the DSA, intermediary services include three categories: (1) mere conduit services, which transmit information of a recipient or provide access to a communication network, (2) caching services, which transmit information and involve its automatic or temporary storing for the sole purpose of more efficient onward transmission, and (3) hosting services, which store the information provided by and at the request of a recipient. The latter comprises online platforms and very large online platforms that store and disseminate information to the public at the request of a recipient of the service.

18 See also *supra* note 17.

19 European Union, *Charter of Fundamental Rights of the European Union*, *Official Journal of the European Union C83*, vol. 53 (Brussels: European Union, 2010).

© 2022 Competition Policy International All Rights Reserved

of submitter). Providers must confirm receipt of the notice and inform about their decision in a timely, diligent and objective manner. This also includes obligations to provide the user who provided the content with detailed statements of the reason for its removal (e.g. alleged illegality or incompatibility with terms of service, including reference to the contractual grounds and information whether the decision was reached by automated means), and the requirement to publish decisions and statements in a publicly available database.

3. Online Platforms

In addition to all the above, online platforms are required to install user-friendly, easy to access and free internal complaint-handling systems that allow for complaints in cases of content removal or suspension and termination of service. Further, users have the right to resort to impartial certified out-of-court bodies to settle disputes relating to platforms' content-moderation decisions. Online platforms must cooperate with these bodies and carry the cost of resolution. Moreover, they are obliged to process and decide on notices of certified trusted flaggers with priority and without delay. Trusted flaggers are impartial entities with proven experience in the realm of illegal content who represent a collective interest. The status is awarded upon application by the Digital Services Coordinators of establishment, which are the primary national authorities designated by the member states for the consistent application of the DSA. There is also a duty to notify suspicions of criminal offences and various protections against misuse. Online platforms, for example, are to suspend users who frequently provide illegal content, or pause the handling of complaints in cases of frequent unfounded notices. There are additional transparency-reporting obligations, among other things, on the number of disputes submitted to out-of-court bodies for settlement, the number of suspensions imposed and the use of automated content moderation.

4. Very Large Online Platforms

VLOPs, which serve on average at least 45 million monthly active users in the EU, are finally the category that is subjected to all of the above plus additional obligations due to their specific systemic role in facilitating public debate and economic transactions and the attendant highest level of risk to society that may stem from their activities. Accordingly, they are obligated to identify, analyze and assess significant systemic risks. These include those resulting from the dissemination of illegal content, those that negatively affect the exercise of fundamental rights such as the freedom of expression and information, or entail the intentional manipulation of services, also by inauthentic use or automated exploitation of the service, with actual or foreseeable effects, among other things on civic discourse, electoral processes and public security. For this assessment they are particularly required to consider how their content-moderation practices as well as their recommender and ad-display systems affect systemic risks.

The DSA introduces a four-layered asymmetric system of due diligence obligations for Internet platforms

The design of the risk-mitigation measures rests with VLOPs too, and may involve adaptations to their content-moderation or recommender systems, restrictions on advertising, cooperation with trusted flaggers or the establishment and adjustment of codes of conduct. In turn, the board – an independent advisory group of the Digital Services Coordinators – and the European Commission publish an annual report on the most prominent and recurring systemic risks as reported by the VLOPs, including best practices to mitigate such risks. In addition, at their own cost VLOPs are subject to annual independent audits to assess compliance with the DSA and may have to provide access to data to vetted researchers for investigations that contribute to the identification and understanding of systemic risks. There are further transparency obligations, for example regarding the mode of operation of recommender systems, including options for users to modify or influence the relevant parameters regarding the order of information presented or profiling. Moreover, they are required to set up a publicly available repository on the advertising, comprising information on its content, sponsor, reach and whether and to what extent specific targeting of users was involved. The European Commission further reserves an enhanced right to supervise, investigate, and monitor VLOPs and enforce the DSA, thus adding a further layer to an already scattered oversight and enforcement system.

The DSA exhibits a particularly conspicuous accentuation of fundamental rights, which was partly strengthened by the suggested amendments from the European Parliament in its first reading, expanding it by including further articles of the EU Charter and for example stressing the freedom of the media and pluralism, the protection of personal data, human dignity, or effects on democratic values more generally.[20]

Despite comprehensive measures to assure the compliance by VLOPs through independent audits, enhanced supervision by the European Commission or sanctions, the fact that the responsibility for systemic risk assessments and the deployment of mitigation measures against these risks

20 *Supra* note 9. As noted, the final agreed on version of the DSA is not public yet and the suggested amendments may or may not have been considered.

© 2022 Competition Policy International All Rights Reserved

rest primarily with VLOPs and their interpretive sovereignty raises various concerns. A major question is what cultural imprint this will inflict on fundamental rights in Europe and what normative values will eventually be accentuated.

03 WHAT VALUES FOR FUNDAMENTAL RIGHTS IN EUROPE?

Fundamental rights quite often appear to be set in stone in western democracies and the colloquial understanding of their substance is almost monolithic. But a closer look reveals considerable differences in the way they are interpreted by courts and people, and public ignorance of what is protected and who they protect against. This fact is particularly important when platforms socialized within an U.S. speech-protection environment are entrusted with risk assessments regarding these fundamental rights and other central democratic functions in other jurisdictions. Most recently, Elon Musk's deal to buy Twitter, which is on hold again, is a good example of conflicts that may arise in the interpretation of free speech. A 2018 survey by the Freedom Forum Institute on the state of the First Amendment found that 77 percent of U.S. citizens are supportive of it and the freedoms it protects but two-fifths of them (40 percent) could not name a single freedom it guaranteed and another third (36 percent) could only name one.[21]

While freedom of speech was the most-commonly recalled right (56 percent),[22] this freedom has also been identified as a constitutional challenge given the new ways of expressing opinions, constituting publics and public space or censoring and moderating speech online.[23] Communication rights, such as the freedom of speech and expression or the freedom to receive and disseminate information are rights where a cultural imprint is particularly accentuated and there are significant differences between the U.S. and Europe. In the U.S., free speech is an almost absolute right with only a few restraints, while in Europe – which may not be considered as a monolithic bloc either – freedom of expression does not trump all other rights, is not necessarily granted a preferred position and needs to be balanced against other competing rights, for example rights to privacy. As research shows, the reasons why and how platforms moderate content is also related to the underlying free-speech norms.[24] Of the various interpretations of the First Amendment by the U.S. Supreme Court, it is probably the marketplace-of-ideas theory that has most support.[25] This posits the discovery of truth through the competition of contrasting ideas, which also include ideas that may not be legal in European contexts, such as certain hate speech or holocaust denial.

In addition, the traditional function of fundamental rights is a defense against the state and its actions and powers inhibiting these rights. Thus they do not directly protect against private companies or community censorship – two instances that currently pose a great risk to freedom of expression, together with new techniques of speech control that lie in the design of the network infrastructure itself or the operations at work in its applications such as search engines or recommender systems.[26] Altogether, this calls for a rethinking of how to protect freedom of expression online and how and whether private entities should be directly or indirectly committed to these rights.

21 Freedom Forum Institute, "The 2018 State of the First Amendment," 2018, https://www.freedomforuminstitute.org/wp-content/uploads/2018/06/2018_FFI_SOFA_Report.pdf, at 3.

22 *Ibid*.

23 Gero Kellermann, "Die Meinungsfreiheit als verfassungspolitische Herausforderung (Freedom of Expression as a Constitutional Challenge)," *Datenschutz und Datensicherheit - DuD* 45, no. 6 (2021): 363–67.

24 Kate Klonick, "The New Governors: The People, Rules and Processes Governing Online Speech," *Harvard Law Review* 131 (2018): 1598–1670.

25 Clay Calvert, Dan V. Kozlowski, and Derigan Silver, *Mass Media Law*, 20th ed. (Mc Graw Hill, 2018).

26 Jack M. Balkin, "Old-School/New-School Speech Regulation," *Harvard Law Review* 127 (2014): 2296–2342; Wolfgang Hoffmann-Riem, *Recht Im Sog der Digitalen Transformation (Law in the Wake of Digital Transformation)* (Tübingen: Mohr Siebeck, 2022); Matthew P Hooker, "Censorship, Free Speech & Facebook: Applying the First Amendment to Social Media Platforms Via The Public Function Exception," *Washington Journal of Law, Technology and Arts* 1, no. 15 (2019): 36–73; Simon Jobst, "Konsequenzen einer unmittelbaren Grundrechtsbindung Privater (Consequences of a Direct Commitment of Private Parties to Fundamental Rights)," *NJW – Neue Juristische Wochenschrift*, 2020, 11–16; Lawrence Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999); Joel R. Reidenberg, "Lex Informatica: The Formulation of Information Policy Rules through Technology," *Texas Law Review* 76, no. 3 (1998): 553–93; Tim Wu, "Is the First Amendment Obsolete?," *Michigan Law Review* 3, no. 117 (2018): 547–81.\\uc0\\u8220{}Old-School/New-School Speech Regulation,\\uc0\\u8221{} {\\i{} Harvard Law Review} 127 (2014

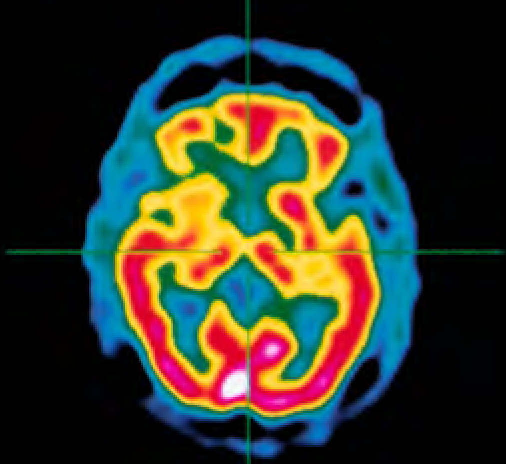
© 2022 Competition Policy International All Rights Reserved

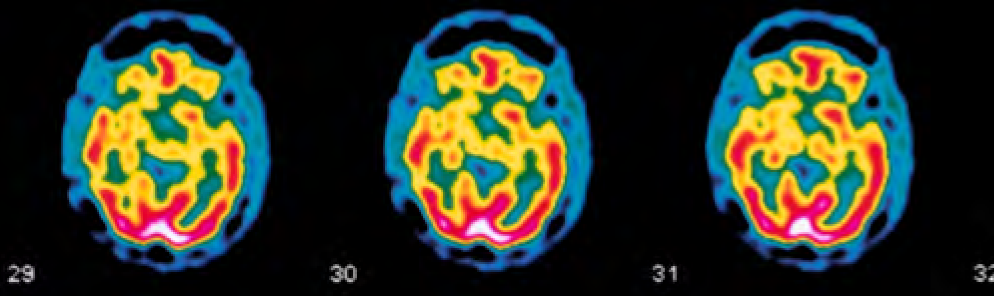
In addition, communication rights fulfil individual and societal functions that need to be balanced. The individual dimension is essentially aimed at personal self-realization regardless of benefits to society at large, while the social dimension sees communication rights as instruments for the protection of democracy. Communication rights thus guarantee further fundamental rights and aid in the control of political power, will formation and the free exercise of political rights. It is these social functions and values that need to be accentuated in an environment where platforms increasingly orient their content moderation to individuals with the aim of maximizing private economic advantages, and where the right of the speaker is considered more important than the disadvantage to those who have to listen and face the potentially negative consequences that may arise from it.

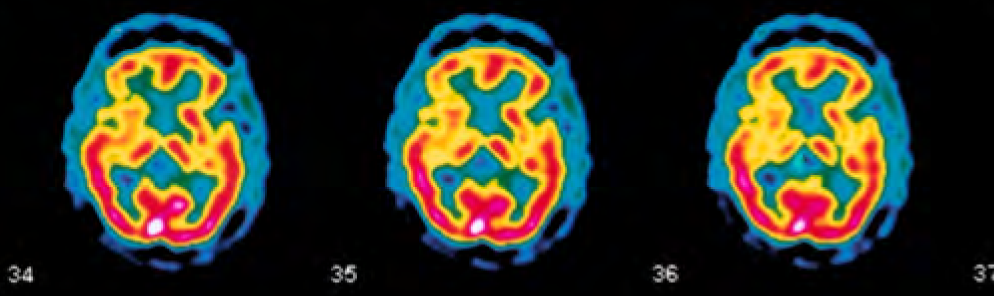
Altogether, the DSA is a further step in tackling definitions and identities, social roles, norms, behavior, expectations and means in the governance of Internet platforms. There are high hopes that it will remedy many of the visible dysfunctions, but in the end the extent to which it will indeed contribute to taming Internet platforms, and – indirectly – all Internet users' socially detrimental behavior, remains to be seen.

> ***Ensuring fair competition and a level playing field, including with technology firms, within an open data space is an essential aspect in this context***

© 2022 Competition Policy International All Rights Reserved

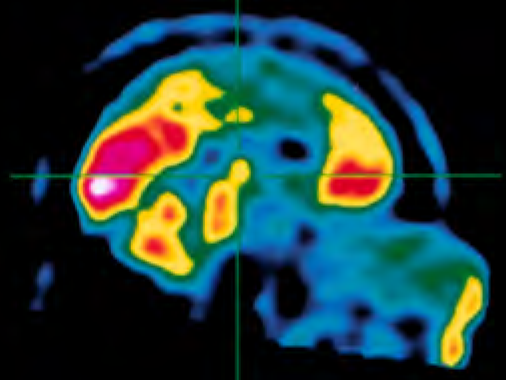


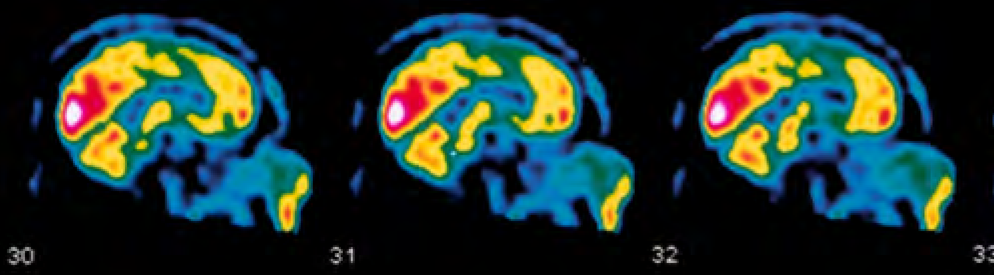


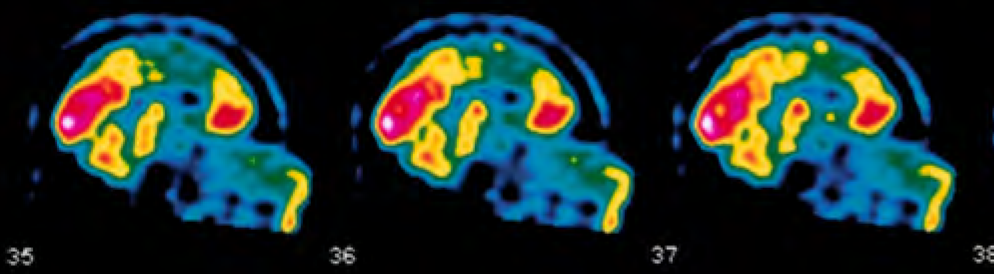


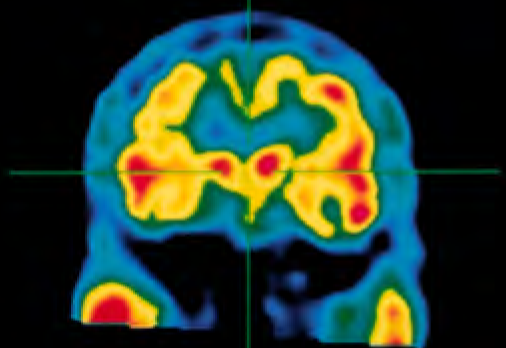
Feet to Head

Transversal



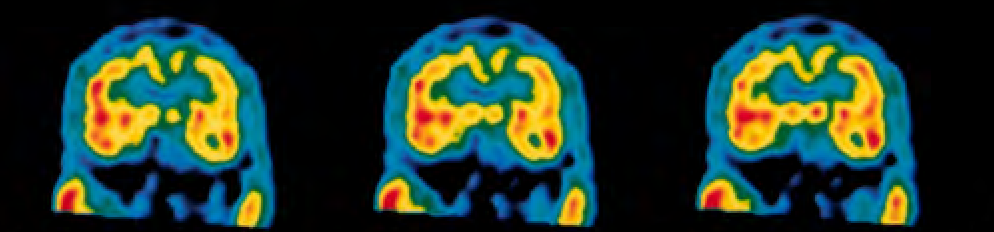


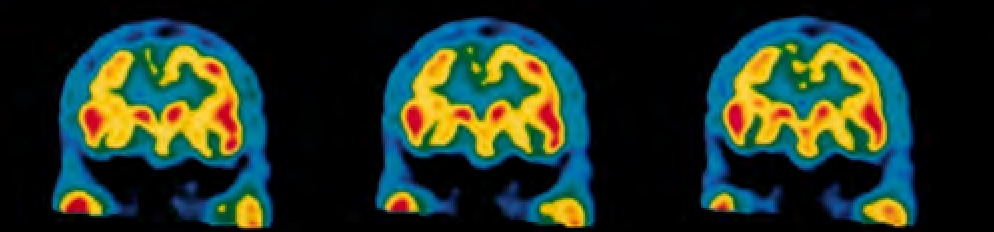




Right to Left

Sagittal





WHY TECH REGULATORS NEED TO THINK LIKE GOOGLE

BY

TERRY FLEW

Professor of Digital Communications and Culture, Faculty of Arts and Social Sciences, The University of Sydney.

01 THE CHALLENGE OF PLATFORM BUSINESS MODELS

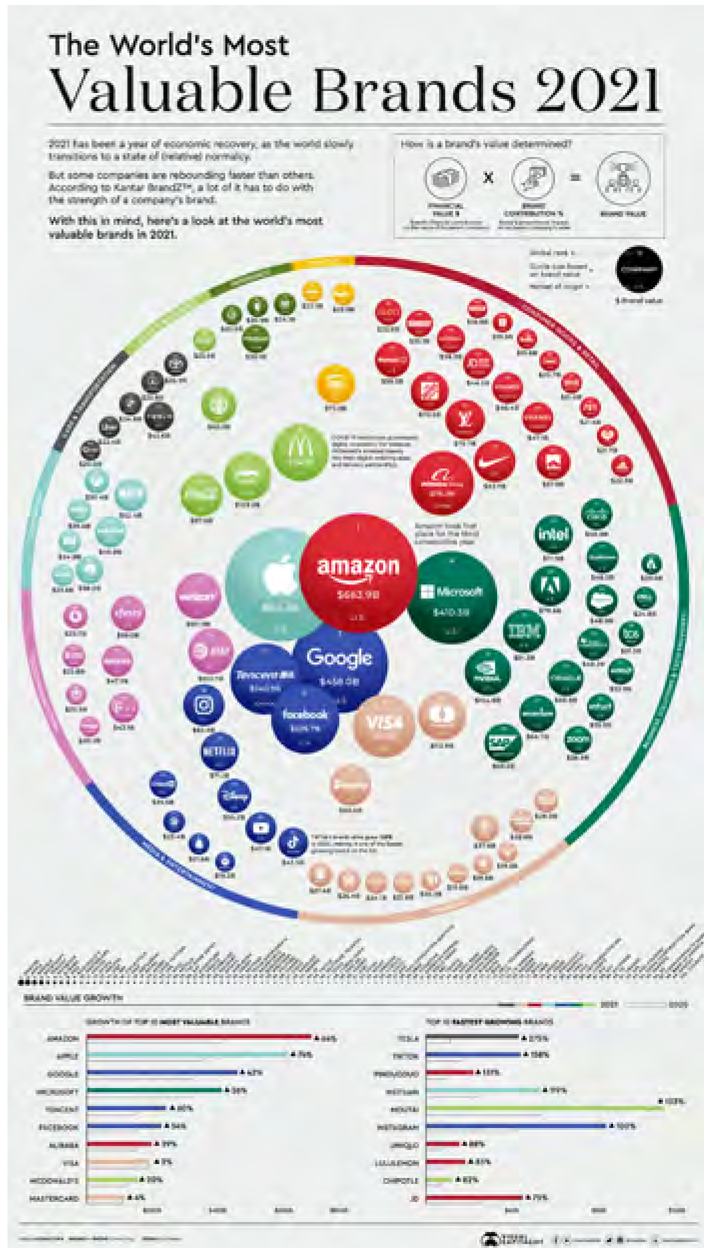
The 2020s have been a period where the policy spotlight has been thrown upon the power of Big Tech. The reasons are not surprising. Platform-based technology companies experienced phenomenal growth during the 2010s and early 2020s, to become the world's largest companies and most valuable brands.

According to Kantar BrandZ, the seven most valuable brands in 2021 were platform-based technology companies (in order): Amazon, Apple, Google, Microsoft, Tencent, Facebook, and Alibaba.[2] They sit alongside a slew of other tech-based companies, including those in me-

2 Carmen Ang, "The World's 100 Most Valuable Brands in 2021," *Visual Capitalist*, 6 October 2021.

© 2022 Competition Policy International All Rights Reserved

dia and entertainment (Netflix, Disney, TikTok), business-based technology solutions (IBM, Adobe, Intel, Cisco, SAP, Oracle and Zoom), telecommunications (AT&T, Verizon, T-Mobile, Vodafone) and technology hardware (Huawei, Xiaomi). Even traditional industries have been disrupted by the digital platform challengers, as seen with Amazon and Alibaba in retail, Tesla and Uber in cars and transportation, and PayPal in the financial sector.



The majority of these companies did not exist prior to 2000, and many have risen to their dominant position on the basis of a platform-based business model. The general features of the platform-based business model are well known, and include the brokering of interactions between buyers, sellers and third parties (e.g. advertisers) across multiple markets, the accumulation of data from multiple sources through online transactions, and the use of algorithms and machine learning for behavioral targeting of consumers based upon prior revealed preferences in order to better match them to products and advertising content. [3]

Critics of this model point to the rise of "platform capitalism" and "surveillance capitalism" whereby these companies access the innermost thoughts and preferences of people and harvest them as data to generate what Shoshana Zuboff termed a "behavioral surplus" for profit. [4] Even those who see the impact of digital platforms as broadly positive for society would nonetheless concede that it has opened up substantial risks for the abuse of power, through privacy loss and data breaches, the capacity to algorithmically amplify misinformation and unchecked hate speech, and the generation of filter bubbles and the distortion of public speech in the face of polarization and incivility in online discourse.

> *The majority of these companies did not exist prior to 2000, and many have risen to their dominant position on the basis of a platform-based business model*

There are also the economic challenges presented by the rise of digital platform companies. Most prominent among these is the "winner-takes-most" nature of many digital markets, which make it difficult to challenge an incumbent in key fields such as search and social media. Once a single provider reaches a certain scale, it becomes very difficult to challenge their market dominance, as direct and indirect network effects become mutually reinforcing, the superior quantity of data generates a superior quality of service (e.g. Google knows your pre-existing search habits better than any new provider can), and consumers find it increasingly costly to switch from one service to another

3 Geoffrey Parker, Marshall van Alstyne & Sangeet Paul Choudary, *Platform Revolution*, New York, W. W. Norton & Co, 2016. See also Hal Varian, "Computer-Enabled Transactions," *American Economic Review* 100(1), 1-10, 2010.

4 Shoshana Zuboff, *The Rise of Surveillance Capitalism*, New York, Public Affairs, 2019.

© 2022 Competition Policy International All Rights Reserved

(e.g. losing messaging histories if you leave Facebook).[5] As the U.S. House of Representatives Subcommittee on Antitrust, Commercial and Administrative Law concluded, “companies that once were scrappy, underdog startups that challenged the status quo have become the kinds of monopolies we last saw in the era of oil barons and railroad tycoons.”[6]

02
WHAT INDUSTRY ARE THEY IN?

While the digital platform giants may have the monopoly power of the oil barons and rail tycoons of yore, one important difference is about what industry they are in. Industries such as rail, food, transportation, even media, were once clearly delineated by the products or services they provided, and the production processes that went into making those products or providing those services. The legacy of these linear value chains – or what are now referred to as pipeline business models – is seen in the ways in which an industry is constructed for purposes as various as Standard Industry Classifications (“SIC”) data, industry support and industry lobby groups, and what constitutes a market for purposes of antitrust and competition policies.

By contrast, digital platform companies can operate in multiple industries simultaneously. Amazon is the paradigmatic example of this. The company started in online bookselling, less because its founder Jeff Bezos had a passion for books, which is the usual reason for getting into bookselling. Rather, the publishing industry had long developed detailed taxonomies for the types of books that were in its catalogues – romance novels, military histories, cookbooks, adult fiction, academic textbooks etc. – and this lent itself very well to a data-driven business model where recommendations to users based upon past purchases would be critical to repeat business. And of course the Internet had mitigated the need to invest heavily in physical retail infrastructure, which was an error made by Borders, which was Amazon’s principal competitor at the time.

Having pioneered core platform business innovations such as recommendation engines, the use of algorithms to direct consumers to products based in revealed preferences is now core to all platform-based businesses. But the other major innovation of Amazon has been to operate in disparate industries, generating large quantities of diverse sources of data that can be mutually reinforcing in the company’s overall business model. Amazon is in the video streaming business through Amazon Prime, but, unlike other streaming services such as Netflix, can run it as a “loss leader” as Amazon Prime is also a highly lucrative premium shipping service for its retail customers. When Amazon acquired the upmarket U.S. grocer Whole Foods in 2017 its value lay les in the markup that can be put on the price of an organic avocado, than in the additional data sources it provided on the shopping preferences of middle-class consumers, as well as the capacity to acquire physical distribution sites to complement Amazon’s wider e-commerce business.

Other examples can be given. Google’s free provision of search and email to users is cross-subsidized by its highly profitable cloud computing services; Amazon can also cross-subsidies other activities (e.g. cost of delivering products to homes) by its cloud computing activities. This is different to a multi-divisional corporate structure, or even to the conglomerate business model that emerged in the 19970s, and which has been a feature of giant media companies. Conglomerates are typically loosely interconnected by their engagement with a particular line of business, and the aim is to generate synergies between one line of business and another – think about the relationship that Disney theme parks have to Disney films and TV shows, or the capacity to franchise a myriad array of products from the characters in *Frozen*. While platform businesses can have integrated product lines with a common look and feel – think of Apple – the integration is actually hidden from view, and arises out of the capacity to share and repurpose data across multiple, and often disparate, online activities.

> ***By contrast, digital platform companies can operate in multiple industries simultaneously. Amazon is the paradigmatic example of this***

5 Patrick Barwise and Leo Watkins, “The Evolution of Digital Dominance: How and Why we Got to GAFA,” in Martin Moore and Damian Tambini (eds.), *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*, Oxford, Oxford University Press, 2018, pp. 21-49.

6 US House of Representatives, *Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations*, Washington, DC, 2020, p. 7.

© 2022 Competition Policy International All Rights Reserved

Among the many challenges that this inversion of conventional industry and market structures presents, one is the difficulty in identifying economic power. In conventional industry economics, and the competition policy and antitrust initiatives that stem from it, economic power is measured in terms of market dominance, and conduct which can follow from it, such as the ability to set prices above what would be the norm in more competitive markets, with limited risk of new competitors emerging due to economies of scale and scope.[7] Think for instance about the power to set prices for parking at or near an airport. But market power in digital markets dominated by platform businesses operates differently.

As Lina Khan observed "equating competition with 'consumer welfare', typically measured through short-term effects on price and output – failed to capture the architecture of market power in the 21st century marketplace."[8] Instead of focusing on price, Khan proposed that we needed to instead focus on questions such as whether a company's structure enables anticompetitive practices and generates conflicts of interest, whether it can cross-leverage market advantage across distinct lines of business, and whether the structure of the market incentivizes and permits predatory conduct towards potential competitors. A further measure, noted by the Stigler Center for the Study of Economy and the State at the University of Chicago, is whether such market structures increasingly constitute a barrier to innovation, promoting a mindset that seeks to buy out potential competitors rather than responding to competition by investing in new digital products, services, and processes.[9]

Among the many challenges that this inversion of conventional industry and market structures presents, one is the difficulty in identifying economic power

03

WHAT ARE THE PROBLEMS?

The challenges presented by digital platform power are thus a mix of traditional and new concerns. Traditional concerns include the capacity to engage in dominant conduct in digital markets and the use of their substantial wealth to lobby politicians for favorable legislative outcomes.[10] But these forms of economic and political power co-exist with newer challenges, which make the regulatory task more complex. Three in particular stand out:

1. Economic power that manifests itself, not in higher prices to consumers, but in the capacity to squeeze other stakeholders in digital platform markets through unequal *bargaining power*. This of course includes workers (campaigns to unionize tech workers have been a major feature of the last decade, particularly in the U.S.), but also includes those who provide the content that is distributed through digital platforms. A major flashpoint in a number of countries has proven to be news. News publishers have a "frenemy" relationship to digital platforms, in that they have been reliant upon them as distribution channels, yet lose advertising revenue to the platforms, who contribute very little to the costs of producing news.[11] This has led governments in countries such as Australia and Canada to establish "mandatory news bargaining codes," whereby major platforms such as Google and Facebook contribute to the costs of producing news through contracts with news publishers.

2. In addition to political power associated with corporate lobbying of politicians, donations to political parties etc., digital platform companies have come to have considerable *ideational power*, in that they ideas that they generate about how to do things come to shape public policy. One example is former Google CEO Eric Schmidt's observation that tech companies are better placed to address policy issues because they are more agile, and employ brighter people, than government bureaucracies. During the COVID-19 global

7 Economic power of this sort may indeed exist in digital markets. The Australian Competition and Consumer Commission (ACCC) identified this as being an issue with Google's dominant position in online advertising. See ACCC, *Digital Advertising Services Inquiry: Final Report*, 2021.

8 Lina Khan, "Amazon-An Infrastructure Service and its Challenge to Current Antitrust Law," in Martin Moore and Damian Tambini (eds.), *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*, Oxford, Oxford University Press, 2018, pp. 100.

9 Stigler Center for the Study of Economy and the State, *Stigler Committee on Digital Platforms: Final Report*, 2019.

10 Zephyr Teachout, *Break 'Em Up: Recovering Our Freedom from Big Ag, Big Tech, and Big Money*, New York, All Points Books, 2020.

11 ACCC, *Digital Platforms Inquiry: Final Report*, Commonwealth of Australia, Canberra, 2019.

© 2022 Competition Policy International All Rights Reserved

pandemic, Microsoft CEO Satya Nadella called for a "partnership" between tech companies and governments, on the basis that tech companies are better at "doing thing" than governments are, and that the role of governments is to set overarching policy goals.[12] A particularly striking recent example has been Elon Musk's use of the Twitter platform to generate support for his proposed takeover of the company, pointing to faults with current Twitter management on Twitter, and conducting plebiscites of his followers about how to "fix" the platform.

3. The rise of digital platforms as distributors, moderators, amplifiers, and arbiters of public speech has placed *communications power* issues on the political agenda in new ways. Whereas the classic debates about free speech and censorship was around the relationship between governments and individual citizens, authors such as the legal theorist Jack Balkin now refer to a "free speech triangle," where private companies increasingly mediate the flows of public discourse in direct and largely unaccountable ways. [13] How to establish accountability and transparency in the decisions being made by digital platform companies about speech rights and content moderation has proven to be a lively political question, from across the political spectrum.

The rise of digital platforms as distributors, moderators, amplifiers, and arbiters of public speech has placed communications power issues on the political agenda in new ways

In noting these issues, it becomes apparent that the agencies which could or should have regulatory oversight vary depending upon the nature of the concern. If the primary concern is content rules and speech rights, the focus is on communications agencies, and the capacity to extend existing media and communications laws and policies into the digital realm. If the primary concern is lack of market competition, then the locus of responsibility shifts to economic agencies. Some argue that the new nature of the problems requires the creation of new mega-regulators purpose built for the platform environment, while others foresee the rise of "neo-regulation," whereby existing laws and regulations are stitched together through cooperation among existing government agencies across Ministerial portfolios. [14]

04
WHO HAS RESPONSIBILITY, AND WHO SHOULD HAVE IT?

Overarching questions around what should be regulated, and how, is the issue of who is best placed to undertake such regulation. In an age of the global Internet, it is apparent that this issue emerges differently in some jurisdictions compared to others. In China, to take an obvious example, discourses of cyber-sovereignty dictate that it is the nation-state that remains the primary locus of regulatory oversight, and there are limits to the capacity of digital platform companies to self-regulate their activities.

In the liberal democracies, by contrast, the Internet arose at a time when there was a strongly "hands off" view of the role of government. While the platformization of the Internet has clearly challenged the idealism of the early Internet, concerns that nation-state governments may act to restrict or suppress speech, or that they lack the regulatory capacity or competence to oversee digital environments, remains significant. [15]

Others feel that civil society groups and NGOs need to take a more active role in platform governance, so as to safeguard against nation-state overreach. In *Speech Police: The Global Struggle to Govern the Internet*, David Kaye argues that, while it is important to have mechanisms to keep in check unaccountable corporate power over digital technologies, regulatory measures "risk capture by ill-intentioned governments," and favors a stronger role for

12 The Eric Schmidt and Satya Nadella statements are cited in Terry Flew, *Regulating Platforms*, Cambridge, Polity, pp. 123, 131.

13 Jack Balkin, "Free Speech is a Triangle," *Columbia Law Review*, 117(8), pp. 2011-56, 2018.

14 Philip Schlesinger, *The Neo-Regulation of Internet Platforms in the UK*, CREATe Working Paper 2021/11, 2021.

15 See e.g. Jeff Jarvis, "Statement to the Judiciary subcommittee on antitrust," *Medium*, February 13, 2021.

© 2022 Competition Policy International All Rights Reserved

civil society organizations, operating in some instances in direct collaboration with the digital platform companies.[16] Tim Berners-Lee's "Contract for the Web" prioritizes civil society in a similar way, favoring a largely hands-off role for governments beyond commitments to invest in digital infrastructure and promote digital literacy. [17]

The problem with these civil society-led initiatives, and associated measures such as promoting corporate social responsibility, shareholder activism or "social media councils," is that they do not adequately address the issues of legitimacy and accountability. The desire to develop binding global standards fails to take account of the diverse preferences and circumstances that exist within and between nations, while it is far from clear than NGOs possess more legitimacy than democratically elected nation-state governments.

In the absence of some form of transnational government, democratic elections continue to provide the "least worst" means of addressing questions of legitimacy and accountability, in a context where nation-state remain the primary rule-makers in geographically defined territories. As the late British democratic socialist Tony Benn put it, a question that we always have to ask about power is "How do we get rid of you?" Models for third party quasi-self-governance, such as the Facebook Oversight Board and NGO-led initiatives, cannot adequately address the need for nation-states to underpin forms of counterpower to digital platform power that have legitimacy, accountability, and legal effectiveness.

"The problem with these civil society-led initiatives, and associated measures such as promoting corporate social responsibility, shareholder activism or "social media councils," is that they do not adequately address the issues of legitimacy and accountability

05

REGULATORS LEARNING FROM GOOGLE

Effective regulation of digital platforms requires the articulation of a new public interest vision for the Internet. As the legal theorist Julie Cohen has observed, it requires a preparedness to see regulatory institutions as sites of policy innovation, rather than as slow-moving bureaucracies from the pre-digital era, a vision long promoted by thought leaders in the tech sector. [18] This does mean learning lessons from the tech sector about how to act effectively, without ceding responsibility for regulation to the tech sector.

Notably, it requires a preparedness to think laterally. Just as Google, Amazon and other digital platform giants can operate in multiple businesses and markets simultaneously and apply lessons from one area of operations to others. In this respect, a key challenge for nation-state policy agencies is the capacity for coordinated action based on shared information. What was described as neo-regulation is likely to become the new normal in addressing the multifaceted challenges associated with digital platform power. This is not a call for a new "super-regulator" or even necessarily for new laws: it is a call for adapting well-established principles to the changed environment for content, competition and markets associated with the rise of digital platforms.

It also requires a new idealism on behalf of the public interest. For much of the last 30 years, idealism has been associated with the open Internet, and those seeking greater regulation in the public interest are too often assumed to have base cynical motivations. There is no shortage of statements of what ethically sound and socially just principles are with regards to media and communications policy.[19] What is needed are advocacy skills for the kinds of practical measures that can move digital societies closer to such ideals.

Finally, regulators can learn from Google the capacity to be clear about what they are seeking. If the primary public interest concern is the lack of market competition, then measures can be put in place around that question. If the primary concerns are around hate speech and potentially

16 David Kaye, *Speech Police: The Global Struggle to Govern the Internet*, New York, Columbia University Press, 2019, p. 88.

17 Tim Berners-Lee, 30 years on, what's next #for the web? *Web Foundation*, March 12, 2019.

18 Julie Cohen, *Between Truth and Power: The Legal Constructions of informational Capitalism*, Oxford, Oxford University Press.

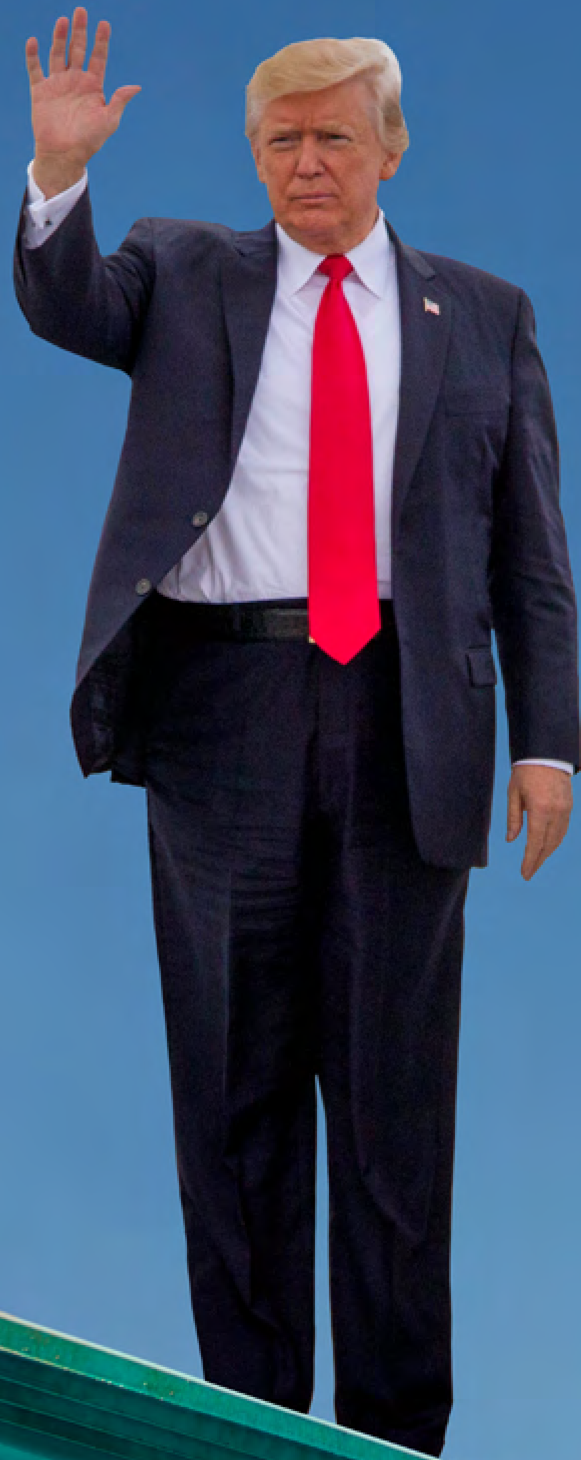
19 See e.g. Robert Picard and Victor Pickard, *Essential Principles for Contemporary Media and Communications Policymaking*, Reuters Institute for the Study of Journalism, Oxford, 2017.

© 2022 Competition Policy International All Rights Reserved

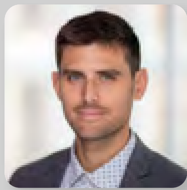
objectionable content in the public sphere, then that requires its own mechanisms for addressing. The need will be for advocates who can speak with clarity about such issues, and who can identify realistic and achievable pathways to addressing these concerns that draw upon the legitimacy provided by nation-state governments.

> *Effective regulation of digital platforms requires the articulation of a new public interest vision for the Internet*

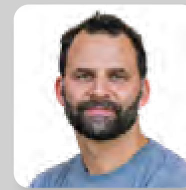
© 2022 Competition Policy International All Rights Reserved



WHAT WE TALK ABOUT **WHEN WE TALK ABOUT DEPLATFORMING TRUMP**



BY
SCOTT BABWAH BRENNEN



&
MATT PERAULT

Scott Babwah Brennen is head of online expression policy at the Center on Technology Policy at UNC-Chapel Hill. Matt Perault is the director of the Center on Technology Policy at UNC-Chapel Hill, a professor of the practice at UNC's School of Information and Library Science, and a consultant on technology policy issues at Open Water Strategies.

01 INTRODUCTION

Following his role in the January 6th attacks on the Capitol, a series of online platforms, including Twitter, YouTube, Facebook, and Instagram, banned President Donald Trump from their platforms. Eighteen months later, as the midterms and the next presidential campaign approach, Twitter might have new ownership, and Facebook's ban on Trump is set to expire in January 2023, there is once again a national discussion about the value of removing Trump from major online platforms.

In a post explaining Facebook's decision to remove Trump in early 2021, Mark Zuckerberg wrote, "[w]e removed these statements yesterday because we judged that their effect – and likely their intent – would be to provoke further

© 2022 Competition Policy International All Rights Reserved

violence."[2] Rather than arguing that Trump expressly violated community standards, Zuckerberg justified the ban by appealing to the likely effect of allowing Trump to remain on the platform.

But what were the "effects" of removing Trump from the platforms? Did it decrease the amount or reach of problematic content? Did it encourage followers to seek out more radical communities on alternative platforms? Did it buttress or undermine Trump's political power?

Empirical understanding of the effects of policy decisions can and should be used to design and target policy interventions more efficiently and effectively. Any consideration of whether to reinstitute Trump to major platforms should be based, at least in part, on a rigorous examination of the good and bad consequences the ban has had on public discourse, on extremism, and on Trump's power and influence. The value of this exercise extends beyond a decision on Trump's account. When platforms face decisions in the future on whether or not to remove users or sitting government officials, those decisions should be informed by an assessment of similar interventions in the past.

But before we can assess the costs and benefits of removal, we first must develop a common understanding of the metrics we might use to evaluate the impacts of the ban. When we talk about whether deplatforming worked or failed, what do we mean by "worked" and what do we mean by "failed"?

> *Empirical understanding of the effects of policy decisions can and should be used to design and target policy interventions more efficiently and effectively*

Toward this end, we provide 13 metrics that analysts or researchers could consider as a means of evaluating the impacts of the ban. While analyzing any of these will involve overcoming serious challenges, from data access, to resources, to how to attribute observed changes narrowly to Trump's removal, identifying what metrics we should consider is an important first step to understanding the impact of banning Trump.

02
WHAT WE KNOW ABOUT DEPLATFORMING

Recently, scholars have offered some empirical insight into the effects of removing individuals or communities from major platforms. Broadly speaking, analysis suggests that deplatforming reduces the amount of prohibited and/or problematic content on the platform from which the user or community is banned. Analyses of the removal of prominent influencers on Twitter (Jhaver et al., 2021)[3] and of communities on Reddit (Chandrasekharan et al., 2017)[4] both observed declines in related conversations, the activity of followers, and the toxicity of content. In contrast, observing conspiracy communities on Facebook, Innes & Innes (2021) found that "minion accounts" worked to "replatform" and share the content that had once been shared by now-banned accounts and groups, meaning that much of the problematic content continued to circulate.[5]

There is also evidence that, in many cases, followers of banned individuals or communities moved to alternative platforms. Notably, when users move to alternative platforms, many become more active, posting more content (Ali et al., 2021)[6] — at least initially (Rauchfleisch & Kaiser, 2021).[7] In some cases, users' content on alternative plat-

2 https://about.fb.com/news/2021/01/responding-to-the-violence-in-washington-dc/.

3 Jhaver, S., Boylston, C., Yang, D., & Bruckman, A. (2021). Evaluating the effectiveness of deplatforming as a moderation strategy on Twitter. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–30.

4 Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., & Gilbert, E. (2017). You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 31:1-31:22. https://doi.org/10.1145/3134666.

5 Innes, H., & Innes, M. (2021). De-platforming disinformation: Conspiracy theories and their control. *Information, Communication & Society*, 1–19.

6 Ali, S., Saeed, M. H., Aldreabi, E., Blackburn, J., De Cristofaro, E., Zannettou, S., & Stringhini, G. (2021). Understanding the effect of deplatforming on social networks. *13th ACM Web Science Conference 2021*, 187–195.

7 Rauchfleisch, A., & Kaiser, J. (2021). Deplatforming the Far-right: An Analysis of YouTube and BitChute (SSRN Scholarly Paper No. 3867818). Social Science Research Network. https://doi.org/10.2139/ssrn.3867818.

© 2022 Competition Policy International All Rights Reserved

forms is more toxic than on primary platforms (Ali et al, 2021; Horta Ribeiro, 2021).[8] However, the reach of this more toxic and more abundant content is significantly lower (Rauchfleisch & Kaiser, 2021; Ali et al., 2021; Horta Ribeiro, 2021) — the audience on alternative platforms simply cannot replace that which was lost on mainstream sites. Importantly, however, existing analyses only examine the movement of users to single alternative platforms, they do not capture activity spread across multiple platforms.

Existing analyses provide less specific insight into the impacts of deplatforming Trump from major platforms. Most studies focus narrowly on how the ban impacted problematic content. An analysis by the for-profit Zignal Labs covered by the Washington Post (Dwoskin & Timberg, 2021)[9] found that misinformation about the 2020 presidential election on Twitter reduced by 73 percent after Trump was banned from the platform. However, a *New York Times* analysis (Alba et al., 2021)[10] found that after Trump was banned from major platforms, a handful of his statements still eventually received as many likes or shares as his posts before the ban. Similar to Innes & Innes (2021), the article observes that "Mr. Trump's most ardent supporters continue to spread his message – doing the work that he had been unable to do himself."

Given the narrowness of both the literature on the impact of deplatforming in general and of deplatforming Trump specifically, we need more, broader analyses of the range of potential impacts of the ban across platforms and across media. But in order to produce those analyses, we first need a shared understanding of what we mean when we talk about "impact."

> *There is also evidence that, in many cases, followers of banned individuals or communities moved to alternative platforms*

03

METRICS TO UNDERSTAND THE IMPACT OF DEPLATFORMING

We identify a series of outcomes or metrics that could be used to better assess the complexity of the impact of removing Trump from major social media platforms. We've grouped these 13 outcomes into 5 categories.

Some of the metrics we propose will be easier to study (e.g. amount of problematic content) than others (change in radicalized beliefs), but all will involve substantial challenges. Attributing any observed change in outcomes to Trump's deplatforming will be extraordinarily difficult. Importantly, our focus here is on providing a set of criteria that could guide our assessments of impact, rather than offering a plan for that assessment.

A. Content

The amount and prominence of hate speech, misinformation, and/or illegal content has been a major concern both of academic analysis of platforms (e.g. Grinberg et al., 2018)[11] and of industry transparency reports.[12] As discussed above, it also has anchored many of the existing efforts to assess the impact of Trump's removal.

We offer four specific metrics related to content that can provide insight into the impact of removing Trump from major platforms. For each of the four, it is important that analysis considers the impact of the ban on content across *users, platforms, and media.* Analysis should consider not only how the bans impacted content from Trump, but also content produced and shared by other users. Analysis needs to examine both the impact of the ban on content on the platform that removed Trump, as well as the impact on other platforms that may have seen increases in Trump supporters (Sanderson et

8 Horta Ribeiro, M., Jhaver, S., Zannettou, S., Blackburn, J., Stringhini, G., De Cristofaro, E., & West, R. (2021). Do Platform Migrations Compromise Content Moderation? Evidence from r/The_Donald and r/Incels. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 316:1-316:24. https://doi.org/10.1145/3476057.

9 https://www.washingtonpost.com/technology/2021/01/16/misinformation-trump-twitter/.

10 Alba, D., Koeze, E., & Silver, J. (2021, June 7). What Happened When Trump Was Banned on Social Media. *The New York Times*. https://www.nytimes.com/interactive/2021/06/07/technology/trump-social-media-ban.html.

11 Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374–378.

12 https://transparency.fb.com/data/community-standards-enforcement/.

© 2022 Competition Policy International All Rights Reserved

al., 2021).[13] Finally, given the interconnections between social media and other media types, analysis should examine how the ban impacted content on other media, including TV, radio, podcasts, and political ads (Benkler et al., 2018).[14]

While we have our own personal views on the question of whether decreased volume, distribution, and engagement of Trump's content is positive for our society, our intent here is to avoid those political judgments. Some people may view such decreases as evidence deplatforming worked and others might view them as evidence deplatforming failed. Those value-based determinations are not our focus. Independent of whether these impacts are seen as positive or negative, they are important metrics for assessing the impact of the intervention.

> ***We offer four specific metrics related to content that can provide insight into the impact of removing Trump from major platforms***

1. Amount of Problematic Content

As described above, examining whether banning Trump increased or decreased the amount of problematic online content is one of the few metrics existing analysis employs (Dwoskin & Timberg, 2021).[15] The landscape of such content has been seen as indicative of the type of environment that platforms provide to users.

Defining problematic content provides serious challenges. Different platforms prohibit or action different types of content. Misinformation, hate speech, and illegal content, may all be broadly described as problematic, but are all different. It is possible that Trump's banning had different impacts on different types of (problematic) content.

2. Reach of Problematic Content

Beyond assessing the quantity of problematic content, analysis should examine how banning Trump impacted the reach of problematic content. The gross amount of problematic content on a platform is in some ways less relevant than the number of people who *saw* that content. For example, if Trump followers moved to alternative platforms, they may have continued to post objectionable content, yet that content may have reached far fewer users.

However, users view huge quantities of content on social media each day, little of which is meaningfully considered. The relationship between viewership and impact is complex, and deeply contextual.

3. Engagement with Problematic Content

Assessing how removing Trump did or did not impact the engagement with problematic content — including likes, shares, or comments — across platforms could provide a better sense of how many people actively considered that problematic content, especially if considered across platforms.

As with the amount and viewership of content, a great deal of scholarship has complicated the relationship between engagement with content and the impact of that content (Bennett & Igengar, 2008).[16] Most importantly, people may like, share, or comment on a piece of content for a number of different reasons (Marwick, 2018;[17] Pennycook & Rand, 2021).[18] Distinguishing between types of engagement may provide a more granular account of the impact of removing Trump.

4. Substance of Problematic Content

The three metrics discussed above lend themselves to quantitative analysis: studying how the bans impacted amounts of content, reach, or engagement. Yet, we also need a sense of how the bans impacted the substance of that content: the narratives, themes, and arguments made. At the same time, while it is essential that we understand

13 Sanderson, Z., Brown, M., Bonneau, R., Nagler, J., & Tucker, J. (2021). Twitter flagged Donald Trump's tweets with election misinformation: They continued to spread both on and off the platform | HKS Misinformation Review. *Misinformation Review*, *2*(4). https://misinforeview.hks.harvard.edu/article/twitter-flagged-donald-trumps-tweets-with-election-misinformation-they-continued-to-spread-both-on-and-off-the-platform/.

14 Yochai Benkler, Robert Faris & Hal Roberts , *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*, Oxford University Press, October 2018.

15 Dwoskin, E., & Timberg, C. (2021, January 16). Misinformation dropped dramatically the week after Twitter banned Trump and some allies. *Washington Post*. https://www.washingtonpost.com/technology/2021/01/16/misinformation-trump-twitter/.

16 Bennett, W. L., & Iyengar, S. (2008). A New Era of Minimal Effects? The Changing Foundations of Political Communication. *Journal of Communication*, *58*(4), 707–731. https://doi.org/10.1111/j.1460-2466.2008.00410.x.

17 Marwick, A. (2018). Why Do People Share Fake News? A Sociotechnical Model of Media Effects. *Georgetown Law Technology Review*. https://www.georgetownlawtechreview.org/why-do-people-share-fake-news-a-sociotechnical-model-of-media-effects/GLTR-07-2018/.

18 Pennycook, G., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007.

© 2022 Competition Policy International All Rights Reserved

how the bans may have impacted problematic content, we should also understand how it has impacted acceptable content as well.

B. Networks

Beyond its potential impact on social media content, Trump's removal may have affected Trump's networks of influence and support on and off social platforms.

1. Distribution of Followers and Content

One analysis found that while the bans decreased Trump's reach on major platforms, his supporters on the mainstream platforms increased their activity to help share and distribute Trump's (off-platform) statements (Alba et al., 2021).[19] Further analysis could better examine the impact of these shifts in the distribution patterns of Trump's content.

For example: irrespective of its impact on the total amount of views/engagements with content, by leading to an increase in the number of people willing to directly support Trump on platforms, the bans may have mediated Trump's influence. Recently, social movement scholars have emphasized the importance of networked organizing to supplement "easy" online activism (Tufekci, 2017).[20] We need to better understand if and how the bans resulted in a larger or more diverse network willing to support Trump and share content, and what impact this might have had on Trump's influence and on public discussion.

2. Cross Media influence

Many have observed that Trump has benefited from "free" media coverage, as news outlets across the political spectrum follow him closely (Wells et al., 2016;[21] Lawrence & Boydstun, 2016).[22] Notably, journalists regularly covered Trump's posts on Twitter and Facebook, granting Trump significant influence over the news agenda.[23] It is important that we understand better how being removed from major platforms influenced both Trump's coverage in mainstream news as well as his ability to shape the topics discussed across outlets.

However, users view huge quantities of content on social media each day, little of which is meaningfully considered

C. Beliefs

While studying the impact of Trump's removal on content may help us understand how the ban might have impacted the social media landscape, we should also consider how that content may or may not have affected users. First, we consider measures that speak to the impact on users' *beliefs*.

For nearly a century, media effects research has complicated the relationship between viewing content and being impacted by that content (Bennett & Igengar, 2008). While media content can shape the issues or topics we care about (McCombs & Shaw, 1972),[24] impacts on opinions or actions are much harder to tease out. This does not, however, mean that (lack of) access to Trump's content had no impact on followers—only that we need research that can identify and detangle the complex impact from the constellation of forces shaping users' beliefs.

1. Radicalization or Extreme Partisanship

How did banning Trump impact the number of users holding extremely partisan or radical beliefs or opinions? While radicalization has taken on many different meanings, here we follow Marwick et al (2022),[25] defining radicalization as "the process whereby individuals come to adopt an 'extremist' mindset or, more directly, escalate from nonviolent to violent political action over time." Importantly, there is

19 Alba, D., Koeze, E., & Silver, J. (2021, June 7). What Happened When Trump Was Banned on Social Media. The New York Times. https://www.nytimes.com/interactive/2021/06/07/technology/trump-social-media-ban.html.

20 Tufekci, Z. (2017). *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. Yale University Press. http://gen.lib.rus.ec/book/index.php?md5=b7f5b30b96ae4b3de38fe32ccfa5ac8b.

21 Wells, C., Shah, D. V., Pevehouse, J. C., Yang, J., Pelled, A., Boehm, F., Lukito, J., Ghosh, S., & Schmidt, J. L. (2016). How Trump Drove Coverage to the Nomination: Hybrid Media Campaigning. *Political Communication*, *33*(4), 669–676. https://doi.org/10.1080/10584609.2016.1224416.

22 Lawrence, R. G., & Boydstun, A. E. (2017). What We Should Really Be Asking About Media Attention to Trump. *Political Communication*, *34*(1), 150–153. https://doi.org/10.1080/10584609.2016.1262700.

23 https://www.businessinsider.com/elon-musk-twitter-donald-trump-ban-amplified-right-wing-experts-2022-5?r=US&IR=T.

24 McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, *36*(2), 176–187.

25 Marwick, A., Clancy, B., & Furl, K. (2022). Far-Right Online Radicalization: A Review of the Literature. *The Bulletin of Technology & Public Life*. https://citap.pubpub.org/pub/jq7l6jny/release/1.

© 2022 Competition Policy International All Rights Reserved

little evidence that online content "causes" radicalization in any direct sense (Marwick et al., 2022). Research on online radicalism has consistently found a weak relationship between viewing online content and becoming radicalized (Gil et al., 2015)—yet online platforms can play catalyzing roles in radicalization, including by normalizing extreme content (Munn, 2019)[26] and by aiding community formation and identity development (Markwick et al., 2022).

2. Substance of Extreme Views

It is important that analysis not only captures the change in intensity of beliefs, but also the *qualitative* difference in the content of radical beliefs. How did it impact the narratives circulating in radical communities across platforms?

3. Trust in or Views of Social Media

Analysis should also consider how the ban impacted users' trust in social media platforms. It could examine how a change in trust has impacted platform use and how it shapes and is shaped by a broader decline in trust across institutions.

Scholars and opinion polls have traced a notable decline in trust in nearly all institutions across several decades (e.g. Gallup, 2022);[27] today, major platforms see low levels of public trust (Kelly & Guskin, 2021).[28] Within this context, we need to better understand what impact the bans may have had on the broader rejection of platforms.

Scholars and opinion polls have traced a notable decline in trust in nearly all institutions across several decades

D. Actions

It is also important that we understand how Trump's removal impacted the actions of followers on and off social media. While data about online or offline violence could be useful, qualitative analysis could help us better understand how followers understand and narrate the impact of Trump's removal.

1. Online Violence

Has Trump's removal impacted the number of examples of online violence, including harassment, stalking, or doxing? Has it impacted the forms or severity of online harassment?

2. Offline Violence

How has Trump's removal impacted both the frequency and the nature of offline violence, including hate crimes or politically motivated violence? While attributing observed differences in amounts or types of hate crimes to the ban on Trump is unlikely, we need to understand better how banning Trump, in conjunction with changes in other forces that can radicalize users together may have played a role in increasing or decreasing offline violence.[29] Interviews with perpetrators of such crimes could help us better understand the broad constellations of forces that combined to facilitate violence, and what role, if any, Trump may have played in it.

E. Politics

Finally, analysis should consider if removing Trump from the major platforms had explicit political impact. Without some account of the influence of deplatforming on Trump's political power and on the greater landscape of politics, we will be missing an important part of the story.

1. Trump's Political Influence

While it is difficult to operationalize political impact, analysis could explore if the bans impacted Trump's ability to raise money for himself and allies, the power of his endorsements, or his ability to influence the Republican party platform.

2. Likelihood of Increased Regulation of the Tech Sector

Congress, state governments, and governments in other countries are considering a wide array of reform proposals. In the wake of the decision to deplatform Trump, a number of government officials – not only Republicans in the United

26 Munn, L. (2019). Alt-right pipeline: Individual journeys to extremism online. *First Monday*. Available online at https://firstmonday.org/ojs/index.php/fm/article/view/10108.

27 https://news.gallup.com/poll/1597/confidence-institutions.aspx.

28 Kelly, H., & Guskin, E. (2021, December 22). Americans widely distrust Facebook, TikTok and Instagram with their data, poll finds. *Washington Post*. https://www.washingtonpost.com/technology/2021/12/22/tech-trust-survey/.

29 As a result of the 1990 Hate Crime Statistics Act, the Justice Department must collect and report annual data on hate crimes. Unfortunately, the latest data released is from 2019.

© 2022 Competition Policy International All Rights Reserved

States,[30] but also leaders[31] of other countries – pointed to the decision as evidence of platform power and the need for stronger government regulation.

Deplatforming Trump could affect the likelihood that technology reform is passed. For instance, Republicans have historically been skeptical of increasing government intervention in private industry, but have been more inclined[32] to support regulation of the tech sector because of the perception that it is biased against conservatives. A number of antitrust and content regulation proposals included both Republican and Democrat co-sponsors.[33] Deplatforming Trump therefore may impact the likelihood that regulation is passed, which could in turn influence the product experience for users and competitive dynamics in the industry.

04 CONCLUSION

Touching off 18 months of debate, the decision to ban the sitting President of the United States was an unprecedented move by online platforms. As politicians and commentators on the left and right argue over the merits of deplatforming Trump, discussion has been hamstrung by not having a shared understanding of how we might assess either the impacts or the effectiveness of the bans. In this paper we aim to start the conversation by offering a set of metrics that could anchor deeper and more granular analyses of the impact of Trump's ban. Using these metrics to empirically assess the bans could result in a deeper understanding of how the ban has impacted online content, networks, and politics.

Although we recommend researchers conduct empirical analysis to better understand the impacts of deplatforming, we recognize that no matter how exhaustive the data, this issue is unlikely to be resolved definitively. Conflicting value judgments and political perspectives means this issue will continue to be hotly contested. Nevertheless, judgements by policymakers and platforms should be informed by the full range of impacts deplatforming has had across public discussion and public life. Those assessments must begin with a shared understanding of what "impact" might mean...

Touching off 18 months of debate, the decision to ban the sitting President of the United States was an unprecedented move by online platforms

30 https://www.latimes.com/politics/story/2021-05-06/how-big-tech-pushed-the-gop-into-the-corner-of-bernie-sanders.

31 https://www.reuters.com/article/usa-trump-germany-twitter/germany-has-reservations-about-trump-twitter-ban-merkel-spokesman-says-idUSL8N2JM4ES.

32 https://www.vox.com/recode/2019/10/29/20932064/senator-josh-hawley-tech-facebook-google-mark-zuckerberg-missouri.

33 https://slate.com/technology/2021/03/section-230-reform-legislative-tracker.html.

© 2022 Competition Policy International All Rights Reserved

AESTHETICS, TECHNOLOGY, AND REGULATIONS

BY
GREGORY DAY

Gregory Day is an Assistant Professor of Legal Studies at the University of Georgia Terry College of Business and holds a courtesy appointment in the School of Law. He is also a Visiting Fellow at Yale Law School's Information Society Project.

01 INTRODUCTION

Technology companies have increasingly come under regulatory fire for impairing society, markets, competition, and free speech, among other things. Whether these proposals to reign in "Big Tech" include antitrust enforcement, new legislation, or executive orders, the underlying belief is that consumers and "consumer welfare" require protection in digital markets. This movement has even united left and right wings with leaders such as Elizabeth

© 2022 Competition Policy International All Rights Reserved

Warren and Ted Cruz in general agreement about the need to regulate platforms and Big Tech.[2]

> *Technology companies have increasingly come under regulatory fire for impairing society, markets, competition, and free speech, among other things*

Despite the spectrum of harms attributed to Big Tech, relatively sparse attention has been paid to society's relationship to aesthetics and image. The effects of media on perceptions of beauty and self-value are far from new, as scholars have long sought to understand how magazines, television, and other mediums influence individuals. The conventional belief is that people develop unrealistic views of beauty from witnessing altered images (e.g. photoshopped or airbrushed) of idealized models, but tech platforms might present even greater or specialized types of harms.

For instance, research has uncovered the dangers of when individuals use filters on Instagram (and other platforms) to "smoothen," "enhance," or even "fix" one's face to match certain ideals. In contrast to traditional media where people passively view images of *others*, platforms enable individuals to manipulate their own photographs. By allowing persons to do so, a belief is that unhealthy perceptions of beauty are supercharged compared to conventional mediums.

A related issue concerns the impact of market concentration on societal perceptions of beauty. In current times, people use only about four platforms to share images and videos: Instagram, Facebook, TikTok, and Snapchat. And Facebook owns Instagram, leaving only three unique companies. The problem concerns whether the lack of options regarding filters and editing programs leaves people to augment images using a select few programs. This has allegedly standardized views of beauty; almost anyone who uses Instagram — 1 billion people at the moment — may be influenced by a singular program.

This piece isn't necessarily claiming that tech's effects on aesthetics must come under greater regulatory scrutiny. Rather the goal is to discuss the nature and depth of an underspecified issue, which is related to problems drawing the ire of Big Tech's critics. It is indeed important to acknowledge how tech platforms influence perceptions of beauty and even self-worth — and whether it should implicate demands for tech regulation.

This Piece proceed in two parts. The first Part introduces a problem known to aesthetics scholars but has evaded legal scholarship: the effect of tech programs on perceptions of beauty and attendant dangers. Then the Second Part discusses the growing demand for regulations of apps, platforms, and tech companies in order to present potential ways that the law could ameliorate some of the alleged harms.

02

INSTAGRAM FACE ETC.

Social media and similar technologies have altered views of beauty on societal and individual levels. While media has long influenced aesthetic perceptions, the unique and even heightened effects of modern technology is explained in this Part.

On a simpler level, technology has increased the amount of time that people focus on themselves.[3] For example, when Zoom emerged during the pandemic, concern for one's appearance mounted as users could watch themselves on video — "One of the strangest things about zoom is you're looking at yourself, usually we don't look at ourselves when

2 Jessica Guynn, *Ted Cruz Threatens to Regulate Facebook, Google, and Twitter over Charges of Anti-conservative Bias*, USA Today (Apr. 10, 2019, 3:41 PM), https://www.usatoday.com/story/news/2019/04/10/ted-cruz-threatens-regulate-facebooktwitter-over-alleged-bias/3423095002/; Cristiano Lima, *Facebook Backtracks After Removing Warren Ads Calling for Facebook Breakup*, Politico (Mar. 11, 2019, 6:32 PM), https://www.politico.com/story/2019/03/11/facebook-removes-elizabeth-warren-ads-1216757.

3 See generally Liraz Margalit, *The Rise of "Instagram Face,"* Psychology Today (May 5, 2021), https://www.psychologytoday.com/us/blog/behind-online-behavior/202105/the-rise-instagram-face.

© 2022 Competition Policy International All Rights Reserved

we meet with other people."[4] This has, as we'll see, driven reliance on filters as well as inspired users to seek out plastic-surgery.[5]

Far from isolated to Zoom, usages of filters prevail on each of the platforms. For instance, FaceTune cures "imperfections" and edits faces on about *one million images* exported to third parties *daily*.[6] One observer estimated that 95 percent of the most followed individuals rely on FaceTune.[7]

In fact, the manner in which platforms enable users to edit pictures of themselves may do more to create unhealthy perceptions of beauty than conventional media. As an observer described, "what is taking it to the next level with these filters is it's not just seeing an image of a celebrity who is unrealistic and measuring yourself against that person, it's measuring your real self against a pretend imagine of yourself."[8] This phenomenon is exacerbated when edited versions receive likes and comments, generating a positive feedback loop. To this end, observers have coined the term "Snapchat Dysmorphia" after witnessing individuals seek out plastic surgeons[9] — per the American Academy of Facial Surgery, a majority of plastic surgeons have noted performing a procedure to conform a person to their snapchat images.[10]

On a societal level, the popularity of only a few filters has driven a narrow view of beauty. For instance, commentators have discussed "Instagram Face," which is an aesthetic ideal prominent among celebrities.[11] While many descriptions of Instagram Face exist — e.g. "It's a young face, of course, with poreless skin and plump, high cheekbones… It looks at you coyly but blankly, as if its owner has taken half a Klonopin and is considering asking you for a private-jet ride to Coachella"[12] — a consensus has generally emerged.[13] By providing 1 billion users with a selection of filters, Instagram has not only created a community spanning the world but also an "extremely specific aesthetic."[14]

> ***Social media and similar technologies have altered views of beauty on societal and individual levels***

4 Anna Haines, *From "Instagram Face" to "Snapchat Dysmorphia": How Beauty Filters Are Changing the Way We See Ourselves*, Forbes (Apr. 27, 2021; 1:19PM), https://www.forbes.com/sites/annahaines/2021/04/27/from-instagram-face-to-snapchat-dysmorphia-how-beauty-filters-are-changing-the-way-we-see-ourselves/?sh=38c477d54eff.

5 *Id.*

6 *Id.*

7 Jia Tolentino, *The Age of Instagram Face*, The New Yorker (Dec. 12, 2019), https://www.newyorker.com/culture/decade-in-review/the-age-of-instagram-face.

8 *Id.*

9 Kamleshun Ramphul & Stephanie G. Mejias, *Is "Snapchat Dysmorphia" a Real Issue?*, 10 Cureus 1,1 (2018).

10 Margalit, *supra* note 3.

11 Catherine Wright, *What is "Instagram Face" and Which Celebrities Have It?*, Celebrity CheatSheet (Jun., 22, 2020), https://www.cheatsheet.com/entertainment/what-is-instagram-face-and-which-celebrities-have-it.html/.

12 *Id.*

13 *Id.* ("When you look at Kim, Megan Fox, Lucy Liu, Halle Berry, you'll find elements in common," a Beverly Hills plastic surgeon told Tolentino in the New Yorker: "the high contoured cheekbones, the strong projected chin, the flat platform underneath the chin that makes a ninety-degree angle."); https://www.michigandaily.com/michigan-in-color/the-instagram-face-and-its-implications/ ("1. A youthful, heart-shaped face 2. A small button nose with an upturned tip 3. Full lips with a defined philtrum 4. Full, but well-groomed brows. 5. Upturned, cat-like eyes 6. A defined, forward-pointing chin and a chiseled jawline to match 7. High cheekbones 8. Defined lashes sometimes achieved through extensions 9. Tan, dewy skin 10. The length of the nose perfectly trisects the rest of the face 11. Distance between the eyes being equal the width of one eye 12. Natural-looking makeup 13. Voluptuous bust and buttocks 14. A tiny waist with defined abdominals 15. Long, shiny hair 16. Never repeating an outfit and always trendy.").

14 Poorva Misra-Miller, *You Look Familiar — "Instagram Face" and the De-racialization of Beauty*, Swaay (Oct. 27, 2020), https://swaay.com/instagram-face-and-the-de-racialization-of-beauty.

© 2022 Competition Policy International All Rights Reserved

Exacerbating this issue is artificial intelligence and machine learning. When users interact with platforms, apps, and filters, the programs receive input, incorporate it, and then improve the interface based upon this feedback. It creates a snowball effect whereby users seek out a certain aesthetic and then apps and platforms evolve whereby they promote the ideal back to users.[15]

Also notable are the effects levied on adolescents who average over 5.5 hours per day online as well as young women. It was found by one researcher that "52% of girls use filters every day and 80% have used an app to change their appearance before the age of 13."

In sum, a new form of body dysmorphia has seemingly emerged, turbo charged from prior iterations. Instead of motivated from exogenous sources (i.e. a picture of someone else), a primary catalyst comes from a person's ability to edit themselves; one surgeon "noticed that if in the past patients came to him and brought pictures of celebrities they wanted to look like, today they come for with filtered pictures of themselves."[16] So should aesthetics and body dysmorphia demand digital regulations?

03

BIG TECH AND REGULATION

Big Tech has become a relentless target of regulators, though scholars and legislators have rarely cited unhealthy views of beauty as a reason. This raises questions of what the way forward should resemble. Part A examines the demands to regulate apps, platforms, and tech companies and then Part B discusses whether current or proposed forms of regulation can or should be applied to the effects of tech on dysmorphia and perceptions of beauty.

A. Regulation of Platforms in General

A frequent source of regulatory anxiety is that platforms and apps alter or even manipulate behaviors. Rather than passive players, tech companies collect and analyze data gathered from users — typically in conjunction with artificial intelligence and machine learning — to constantly improve their platforms.

But the term "improve" is loaded. While this could constitute enhancing a user's experience, it is notable that tech companies generate revenue by increasing the amount of time spent and engagement on their apps (e.g. clicks, swipes, scrolls, etc.). This is because greater interactions allow firms to advertise, collect insights, target products, and build value. So the concept of "improving" can refer to generating usage or even addiction — even if users do not find an app to be materially "better." Platforms have thus come under regulatory scrutiny for allegedly designing manipulative and exploitative techniques.[17]

Big Tech has become a relentless target of regulators, though scholars and legislators have rarely cited unhealthy views of beauty as a reason

For instance, Facebook has incurred volleys of criticisms for its strategies meant to allegedly increase a user's attention. One method has involved promoting posts on feeds when it garners a greater number of angry emojis versus happy ones.[18] The intended effect was supposedly to increase the amount of time spent by users reading, engaging, and debating "angry posts" but it has also, as scholars allege, fostered societal polarization, misinformation, and anxiety.[19] While Facebook had unlikely wanted to polarize America, it seems like a foreseeable result.

15 Liraz Margalit, *The Rise of the "Instagram Face,"* CMS Wire (May 3, 2021), https://www.cmswire.com/digital-experience/the-rise-of-the-instagram-face/.

16 *Id.*

17 Gregory Day, Attention, *Antitrust, and the Mental Health Crisis*, 106 Minn. L. Rev. __, __ (forthcoming in 2022).

18 Jeremy B. Merrill & Will Oremus, *Five Points for Anger, One for a 'Like': How Facebook's Formula Fostered Rage and Misinformation*, Wash. Post (Oct. 26, 2021), https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji -algorithm.

19 Adrienne LaFrance, *History Will Not Judge Us Kindly*, Atlantic (Oct. 25, 2021), https://www.theatlantic.com/ideas/archive/2021/10/facebook-papers-democracy-election-zuckerberg/620478.

© 2022 Competition Policy International All Rights Reserved

Similar examples of manipulation have surfaced. For instance, scholars have discussed Snapchat's "streaks" which are linked with "Snapchat depression"[20] or YouTube's algorithm designed to select videos, which is said to curate extremist content.[21] It can indeed be difficult to differentiate whether an app is producing an intended effect or negative externality, or a mix of the two.

This has drawn a significant response in favor of regulating Big Tech. Proposals have included increasing or altering antitrust enforcement to govern digital markets.[22] Commentators have also proposed new types of regulations intended to promote privacy,[23] free speech,[24] competition,[25] mental health,[26] and/or data integrity.[27] Whether these proposals are supposed to remedy a specific injury or consumers in general, the belief is that apps, platforms, and tech companies levy too much harm for them to operate with free rein. So what about aesthetics?

This has drawn a significant response in favor of regulating Big Tech. Proposals have included increasing or altering antitrust enforcement to govern digital markets

B. Regulating Tech's Effects on (Mis)perceptions of Beauty?

Notably absent in this discourse has been tech's effects on aesthetics, body imagine, and beauty. The inference is that tech firms understand the deleterious consequences yet promote filters anyway.[28] In 2021, the *Wall Street Journal* published an exposé entitled "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Shows."[29] Facebook determined, as the article revealed, that "Thirty-two percent of teen girls said that when they felt bad about their bodies, Instagram made them feel worse" and, per a slide in an internal Facebook presentation, "We make body image issues worse for one in three teen girls."[30] Nevertheless, the manner in which platforms lead to injuries such as Instagram Face and Snapchat

20 Nassir Ghaemi, *Snapchat Depression*, Tufts Now (Apr. 17, 2018), https://www.now.tufts.edu/articles/snapchat-depression.

21 Casey Newton, *How Extremism Came to Thrive on YouTube*, Verge (Apr. 3, 2019), https://www.theverge.com/.interface/2019/4/3/18293293/youtube-extremism-criticism-bloomberg.

22 See e.g. Gregory Day & Abbey Stemler, Infracompetitive Privacy, 105 Iowa L. Rev. 61 (2019).

23 *Id*.

24 See e.g. Marco Rubio, *Rubio Introduces Sec 230 Legislation to Crack Down on Big Tech Algorithms and Protect Free Speech*, Marco Rubio U.S. Senator for Florida (Jun. 24, 2021), https://www.rubio.senate.gov/public/index.cfm/2021/6/rubio-introduces-sec-230-legislation-to-crack-down-on-big-tech-algorithms-and-protect-free-speech.

25 Cecilia Kang, *Lawmakers, Taking Aim at Big Tech, Push Sweeping Overhaul of Antitrust*, N.Y. Times (Jun. 29, 2021), https://www.nytimes.com/2021/06/11/technology/big-tech-antitrust-bills.html.

26 Day, *supra* note 17..

27 Day & Stemler, *supra* note 22.

28 Elle Hunt, *Faking It: How Selfie Dysmorphia Is Driving People to Seek Surgery*, The Guardian (Jan. 23, 2019; 1:00PM), https://www.theguardian.com/lifeandstyle/2019/jan/23/faking-it-how-selfie-dysmorphia-is-driving-people-to-seek-surgery ("She liked the sense of having a platform, she says, with the average selfie getting 300 replies. "It was like, 'Oh my God, I'm so popular – I've gotta show my face.'" But the filters were also part of the appeal. The Londoner had long been insecure about the slight bump in her nose. Snapchat's fun effects, which let you embellish your selfies with dog ears, flower crowns and the like, would also erase the bump entirely. "I'd think, 'I'd like to look how I look with this filter that makes my nose look slimmer.'").

29 Georgia Wells, Jeff Horwitz, & Deepa Seetharaman, *Facebook Knows Instagram Is Toxic for Teen Girls, Company Document Shows*, Wall St. J. (Sept. 14, 2021; 7:59AM), https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739.

30 *Id*.

© 2022 Competition Policy International All Rights Reserved

Dysmorphia have received little or no regulatory scrutiny.

So what could be accomplished if anything? A factor imploring regulatory *restraint* is that popular media has always harmed bodies and fostered dysmorphia. Given the costs and unintended effects of regulation, the question is whether it's even worth attempting to regulate tech's effect on aesthetics.

But if we should regulate Big Tech, an initial suggestion comes from antitrust law, which observers have frequently cited when endeavoring to regulate digital markets. The inkling is that tech's harms are embellished by concentrated markets; if more competitors existed, it would ostensibly introduce more and better products valuing privacy, mental health, and other virtues.[31] To this end, a potential solution would be to enhance antitrust's presence in digital markets, which would theoretically encourage entrants to design products meant to limit dysmorphia. It could also add variety, limiting the greater effects of Instagram Face where the dominance of one filter or app renders societal consequences. Whether antitrust enforcement may actually achieve this goal, however, is doubtful.

But if we should regulate Big Tech, an initial suggestion comes from antitrust law, which observers have frequently cited when endeavoring to regulate digital markets

There is also potential in enacting agency rules or new legislations. For instance, the UK's Advertising Standards Authority announced in 2021 that social media influencers must cease using "misleading" filters, though the regulation was primarily intended to ban influencers from deceptively filtering their faces to appear like makeup — the harm was false advertising.[32] But commentators have proposed similar rules in hopes of tempering unrealistic perceptions of beauty.[33] These sentiments have been expressed by celebrities such as Jameela Jamil.[34]

Perhaps a more viable way of limiting Snapchat Dysmorphia concerns raising awareness rather than enacting regulations. As discussions of tech's effects on aesthetics emerge, numerous companies have released media campaigns pledging not to use filters or editing programs to promote more realistic versions of beauty and bodies — e.g. Dove's "Reserve Selfie" campaign.[35] Firms have also pledged to hire models representing a greater and more realistic scope of bodies and faces. Other possibilities include installing parental controls on editing programs. If these techniques prove effective, it would reflect a more organic manner of neutralizing Big Tech's relationship with body and face dysmorphia.

Again, this Piece isn't meant to advocate for greater regulations or a specific law. It is intended to highlight research from the field of aesthetics that has largely gone unseen in legal scholarship. Given the volumes of criticisms levied at Big Tech, it seems relevant to understand Snapchat Dysmorphia, Instagram Face, and similar issues. Perhaps there is no legal answer, or maybe greater attention might spark a solution.

31 Day & Stemler, *supra* note 22.

32 Chloe Laws & Laura Hampson, *Influencers Can No Longer Use "Misleading" Filters on Beauty Ads, ASA Rules*, Glamour Mag. (Feb. 3, 2021), https://www.glamourmagazine.co.uk/article/enhanced-photos-social-media-law.

33 Ella Bennett, *London Influencer Calls for Social Media Platforms to Ban Filters in Bid to Tackle Unrealistic Beauty Expectations*, My London (Apr. 24, 2021), https://www.mylondon.news/lifestyle/london-influencer-calls-social-media-20429487.

34 *Id.*

35 Amy Houston, Ad of the Day: Dove Deepfakes Highlight Toxic Beauty Advice on Social Media, The Drum (Apr. 27, 2022), https://www.thedrum.com/news/2022/04/27/ad-the-day-dove-deepfakes-highlight-toxic-beauty-advice-social-media.

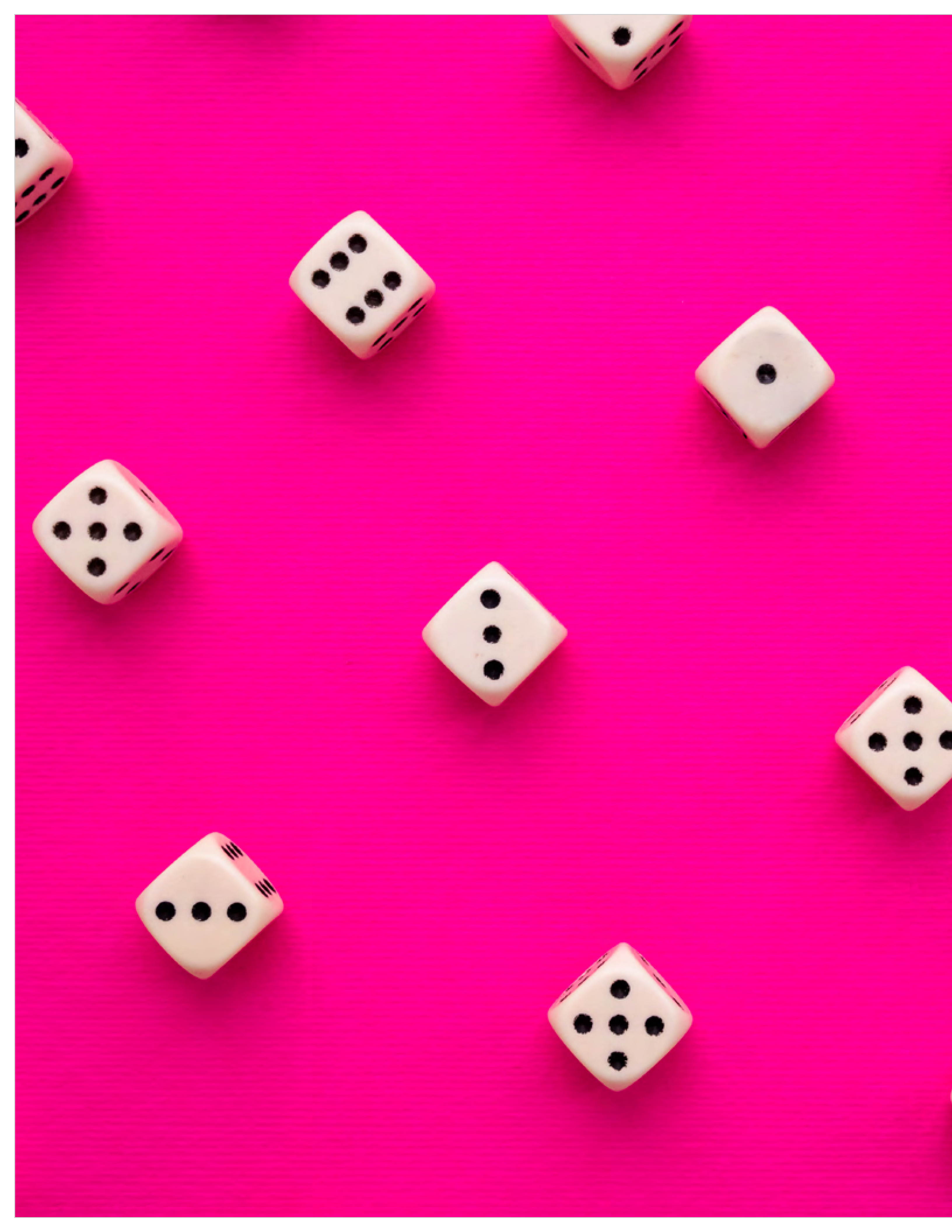
© 2022 Competition Policy International All Rights Reserved

04
CONCLUSION

Perhaps because unhealthy perceptions of faces and bodies have long posed problems, the effects of platforms and apps have largely gone unnoticed by regulators. Or maybe the problem is that regulating the issue might be too difficult or even impossible. Whether this issue is a matter for regulators or not, this Piece's goal is certainly to increase attention.

"Perhaps because unhealthy perceptions of faces and bodies have long posed problems, the effects of platforms and apps have largely gone unnoticed by regulators

© 2022 Competition Policy International All Rights Reserved



RANDOMIZED EXPERIMENTS FOR ONLINE CONTENT MODERATION POLICY

BY
IMANOL RAMÍREZ

Senior Associate at Dentons López-Velarde (Mexico City). Lawyer (Universidad Autónoma del Estado de Hidalgo), LL.M. (Harvard Law School). The article represents only the author's views. Contact: iramirez@llm20.law.harvard.edu and https://www.linkedin.com/in/imanol-ramirez/.

01 INTRODUCTION

Much ink has been spilled on how to balance the different policy objectives of online content moderation regulation. These policy goals include preventing online harms, promoting free speech, encouraging technical innovation, and guaranteeing competition, among others.[2] Nonetheless, optimal solutions or consensuses have proven difficult to achieve and today, as a result, there is a fragmented legal and regulatory landscape over online content moderation all over the world.

2 See Joris van Hoboken and Daphne Keller, Design Principles for Intermediary Liability Laws, Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression (Oct. 8, 2019) at 2-3. Available at https://www.ivir.nl/publicaties/download/Intermediary_liability_Oct_2019.pdf.

© 2022 Competition Policy International All Rights Reserved

Although this increasingly stringent and divergent legal environment raises various concerns, especially for innovation and competition policy due to increased entry and expansion barriers to markets and data, the way forward in terms of harmonization and cohesion does not look promising. International law, which could be the right avenue to address issues with global dimensions, does not seem at hand since countries' legal standards remain too far from each other.

For instance, in the debate of intermediary liability/immunity, the United States and now its trading partners under the United States-Mexico-Canada Agreement ("USMCA") maintain broad protections for intermediaries who, with certain exceptions, are not liable for the effects of user-generated content posted in their platforms and of the moderation decisions taken with respect to it. On the other hand, the European Union and its member countries aim to establish liability-based regimes for intermediaries, including significant fines for failure to comply with regulations, as demonstrated by the recent approval of the Digital Services Act ("DSA").

Moreover, regulation proposals tend to be based on intuitive assumptions or even anecdotal evidence, most likely reflecting commentators' values or implicit guesses about the possible effects of regulation. At a theoretical or logical level, there are plenty of compelling arguments supporting many of the views out there, such as the pros and cons of intermediary liability/immunity, which proponents from both sides of the debate have extensively put forward. However, we may be failing to look at the facts and evidence available around these discussions in order to test our theories and ideas through empirical methods.

Legislators and policymakers could flip the script and take advantage of the fragmented regulation across jurisdictions to assess and determine the way forward. Empirical evaluation methods such as randomized experiments have been successfully applied to health and economic development policies, significantly advancing our understanding of how regulation could be designed. Digital markets offer unprecedented and valuable data to that effect which could allow policymakers to understand how different regimes shape the conduct of intermediaries and make policy decisions accordingly.

The following paragraphs elaborate on these ideas as follows: (i) the fragmentation of online content moderation regulation; (ii) moving beyond assumptions and theoretical debates; and (iii) randomized experiments for the design of content moderation policies.

02

FRAGMENTATION OF REGULATION OVER ONLINE CONTENT MODERATION

I have written separately about online content regulation and competition policy, arguing that the push for regulation over content moderation around the world is creating a stringent and divergent legal environment with increased liability for firms operating in the digital landscape, including the mandatory use of technology, the establishment of substantial fines and even criminal responsibility.[3]

We have seen regulations establishing increasingly strict intermediary liability in several countries, including Germany, the United Kingdom, India, Thailand, and Australia. More recently, the European Union reached an agreement on the DSA establishing, among others, the possibility for users to challenge, either judicially or through an out-of-court mechanism, content moderation decisions taken by intermediaries and providing for significant fines in case of noncompliance.[4]

The immediate consequence of the fragmentation of the regulatory landscape is that the cost of operating in the market increases, as companies have to deploy the technology and resources to comply with legal requirements and avoid liability, which may be very costly in some cases.[5] This is particularly relevant considering the economic characteristics of some digital platform markets where legal barriers may hinder an entrant's ability to access large data sets, obtain scale and generate its own positive network effects that are necessary to challenge large incumbents.[6]

3 See Imanol Ramírez, *Online Content Regulation and Competition Policy*, Harvard Law School Antitrust Association, Cambridge, MA, December 3, 2020. Available at https://orgs.law.harvard.edu/antitrust/files/2020/12/Imanol-Ramirez-Online-Content-Regulation-and-Competition-Policy-HLSAntitrustBlog-2020.pdf.

4 European Commission, *Digital Services Act: Commission welcomes political agreement on rules ensuring a safe and accountable online environment*, April 23, 2022. Available at https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545.

5 See Imanol Ramírez, *Online Content Regulation and Competition Policy*, Harvard Law School Antitrust Association, Cambridge, MA, December 3, 2020. Available at https://orgs.law.harvard.edu/antitrust/files/2020/12/Imanol-Ramirez-Online-Content-Regulation-and-Competition-Policy-HLSAntitrustBlog-2020.pdf.

6 For a brief description of the economics of digital platform markets, see Imanol Ramírez, Merger Thresholds in the Digital Economy, Delaware Journal of Corporate Law,

© 2022 Competition Policy International All Rights Reserved

In this way, policymakers worldwide face a collective action problem whereby increased liability imposed in each jurisdiction may strengthen the market position of large incumbents operating globally by raising the costs of entry and expansion into digital markets. Although its pending implementation, the European Union's DSA attempts to address this issue by establishing obligations depending on companies' size, role, and impact.[7] Government interventions on content moderation need to use a mix of strategies to take advantage of both market forces and state regulation to effectively tackle online harms, promote free speech, and guarantee competition, among others.

Evidently, international law could set a better framework to face this collective action problem. Nonetheless, countries are not close enough with respect to the legal standards applicable to content moderation, including intermediary liability/immunity and the actual rules governing users' content, i.e. what should be considered illegal content and what should not.

On one side of the spectrum, the United States' Section 230 of the Communications Decency Act of 1996 ("CDA") is generally viewed as landmark legislation granting broad protections for intermediaries with respect to harms arising from user-generated content and the moderation decisions taken by intermediaries to this effect. This more laissez-faire oriented approach to intermediary liability is also reflected in Article 19.7 of the USMCA, which establishes a similar (although not identical)[8] provision to Section 230 containing broad protections for intermediaries.

On the other hand, the European Union and its member countries have been adamant about their attempt to impose a responsibilities-based regime on intermediaries. In Germany, the Network Enforcement Law establishes fines up to €5 million for internet companies with at least 2 million users that fail to remove manifestly unlawful speech within 24 hours and all illegal content within seven days of receiving a complaint, as well as reporting obligations on how complaints are handled.[9] More recently, in April 2022, the DSA approved by the European Union 27-member countries established landmark legislation for intermediary liability, including the possibility for users and civil society to challenge moderation decisions and seek redress, as well as transparency obligations, with fines up to 6% of the companies' global turnover for the failure of compliance.[10]

Just as with historical attempts to create an international antitrust regime,[11] the fact that the United States and the European Union remain too far from each other concerning the applicable standards to regulate online content moderation most likely puts them in a deadlock where the two blocks perceive that the benefits from international laws on content moderation would not exceed its costs.

In this context, it is unlikely that larger agreements at a multilateral or international level will be achieved, and even when international law may be the right avenue to create a coherent regime for issues with global dimensions, international law would not be available for content moderation at least in the short term. Moreover, a universal answer to online content moderation is unlikely, considering that views on freedom of speech vary across the spectrum, including for each individual, the leadership within each tech company, and societies as a whole.

03

BEYOND INTUITIVE ASSUMPTIONS AND THEORETICAL DEBATES

Today, there is a lengthy debate around the regulation of online platforms, including moderation of user-generated content. Despite the extent of the discussion, opinions tend to be based on intuitive assumptions or even anecdotal evidence, with little or no support from empirical evidence.

For instance, one of the main ideas behind the rationale of CDA Section 230 that academics, courts, and organizations have persistently put forward argues that increased liability

7 European Commission, *Digital Services Act: Commission welcomes political agreement on rules ensuring a safe and accountable online environment*, April 23, 2022. Available at https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545.

8 See Vivek Krishnamurthy and Jessica Fjeld, *CDA 230 Goes North American? Examining the Impacts of the USMCA's Intermediary Liability Provisions in Canada and the United States*, Harvard University - Berkman Klein Center for Internet & Society, July 7, 2020. Available at http://dx.doi.org/10.2139/ssrn.3645462.

9 Network Enforcement Act (Netzdurchsetzunggesetz, NetzDG), (Federal Law Gazette I, p. 3352 ff. Valid as from 1 October 2017) https://germanlawarchive.iuscomp.org/?p=1245.

10 European Commission, *Digital Services Act: Commission welcomes political agreement on rules ensuring a safe and accountable online environment*, April 23, 2022. Available at https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545.

11 See Anu Bradford, *International Antitrust Negotiations and the False Hope of the WTO*, 48 Harvard International Law Journal, 383, 2007. Available at https://scholarship.law.columbia.edu/faculty_scholarship/561.

© 2022 Competition Policy International All Rights Reserved

over online intermediaries would generate an over-removal or chilling effect on speech. The argument goes that when facing moderation decisions of third-party content in a stringent liability regime, intermediaries will most likely err on the side of caution to avoid liability and thus systematically censor content that eventually will hurt free speech.

Even when this is a strong logical argument, there is little empirical analysis of the validity of this idea that allows us to really understand the effects of increased regulation on speech. Most studies that have attempted to test this idea have focused on surveys without really measuring how people behave after actual changes in law since they are limited to people's claims on how hypothetical regulatory changes will affect them.[12]

At a theoretical or logical level, we can further develop more arguments about the validity or invalidity of the chilling effect principle. For example, it can be argued that the chilling effect idea fails to recognize the economic incentives that the industry has to maintain content online, which could offset or at least be in direct tension with the possibility of over-removal of speech.

Considering that user-generated content is at the core of some digital businesses and that it is central to profits in many cases, one could argue that it is possible to introduce monetary sanctions for the failure to remove harmful content without having a significant chilling effect on desirable speech. In this way, the underdeterrence concern of existing harmful speech can be weighed in the balance.

A rational company whose objective is to maximize shareholders' profits would balance the potential economic loss derived from a fine against the loss from removing profitable content, which is particularly important in industries with network effects, and will choose the smaller loss. Although other elements should be considered, such as reputational damages and the companies' need for good content, it can be argued that this is one reason why the "chilling-effect" argument, while powerful, is incomplete. The economic incentives of an intermediary will at least counter its incentives for over-removal of speech and could create a mix of incentives that results in better content moderation tools and efforts. Therefore, when facing a decision of potential removal of content, it should not be taken as a given that intermediaries will err on the side of caution to avoid fines.

This is even more compelling in cases where it is easy for intermediaries to identify harmful content and there is not much discussion about its adverse effects. Establishing fines for failing to eliminate the easy cases could suppress any incentives of companies to leave harmful content online that is highly profitable. It is also likely that a positive correlation exists between the degree of harm and the obviousness of harmful content. Thus, monetary penalties may be more effective in preventing and eliminating harmful content in these cases.

But at this point of the debate, shouldn't we move beyond these intuitive assumptions and, if possible, question the reality and strength of these ideas and other principles at an empirical level? Although it may be complex, data is at the heart of digital markets, and business models and platforms collect and store unprecedented information that could be used to measure the effects of the different policies in place.[13] Nonetheless, numerous studies show that the data collected is never used, signaling that incentives to gather it are there but not the incentives to analyze it.[14] To the extent possible, the ideas and assumptions over content moderation regulation should be accompanied by empirical methods to increasingly employ data generated by firms and be able to test and confirm the intended effects of regulation.

04 RANDOMIZED EXPERIMENTS

Given these circumstances, how can we make something good out of the increasingly fragmented regulatory framework for online content moderation and the data richness that characterizes the industry? Due to the somewhat accidental disagreement among the international community on the regulation of content moderation, policymakers could imagine themselves in a position where only one global content moderation program exists, but some of its features are varied across jurisdictions, in a situation similar to randomized experiments.

Following the logic of randomized experiments or randomized controlled trials, which have been successfully

12 See Suneal Bedi, *The Myth of the Chilling Effect*, Harvard Journal of Law & Technology, Volume 35, Number 1 Fall 2021. Available at https://jolt.law.harvard.edu/assets/articlePDFs/v35/Bedi-The-Myth-of-the-Chilling-Effect.pdf.

13 Daniel Björkegren and Chiara Farronato, *To Regulate Network-Based Platforms, Look at Their Data*, Harvard Business Review, October 18, 2021. Available at https://hbr.org/2021/10/network-based-platforms-must-be-regulated-but-how.

14 Viktor Mayer-Schönberger and Thomas Ramge, *The Data Boom Is Here — It's Just Not Evenly Distributed*, the MIT Sloan Management Review, Spring 2022, February 9, 2002. Available at https://sloanreview.mit.edu/article/the-data-boom-is-here-its-just-not-evenly-distributed/.

© 2022 Competition Policy International All Rights Reserved

employed for development economics and evaluation of public policies,[15] researchers could take advantage of the divergent regulatory framework to implement large-scale experiments that test the theories underlying the different legal regimes. Divergent regulation across jurisdictions will have the effect of assigning online intermediaries that operate internationally to different potential solutions. Thus, the effects in the conduct of an intermediary could be attributed to the differences in the regulation. The goal of these studies should be to understand how to better design content moderation regulation.

For instance, these empirical studies could shed some light on the intermediary liability/immunity debate. It may be possible to understand how moderation decisions of online intermediaries have been shaped by regulations that establish fines for failure to comply with the removal of harmful content, and how those decisions compare to other regimes that provide for a more laissez-faire approach. Moreover, comparisons among different intermediaries with divergent moderation policies but acting under the same legal framework could also allow policymakers to understand the effects of the rules governing users' content (i.e. companies' internal moderation policies). To achieve the preceding, regulators will require information such as the amount and type of material being taken down in each of the jurisdictions being studied, the reasons why a certain type of content is being censored or not, the results of flagged content or complaints, the differences in technology being deployed and investments, including in human resources and programs, among others.

To this effect, it should not be expected that online intermediaries provide their sensitive data voluntarily, especially proprietary information and trade secrets arising from their own data analysis. Simultaneously to the ongoing efforts, governments should consider increased transparency obligations and data access mandates for online intermediaries. Regulators would have to meet very high standards of confidentiality and protection of the information to guarantee the safety of the platform and that its competitiveness is not jeopardized, as well as to invest in teams that can analyze raw data provided by companies.

The randomization-like approach would not offer definitive answers or ultimate solutions to content moderation problems as every study group (in this case, the same companies acting in different jurisdictions) possesses intrinsic elements that would shape moderation decisions. However, no study is likely to demonstrate causality on its own. One of the main features of randomization is that it sets aside the observed and unobserved characteristics of the different study groups allowing attribution of any differences in the outcome to the intervention and thus measuring to some extent its effectiveness.[16]

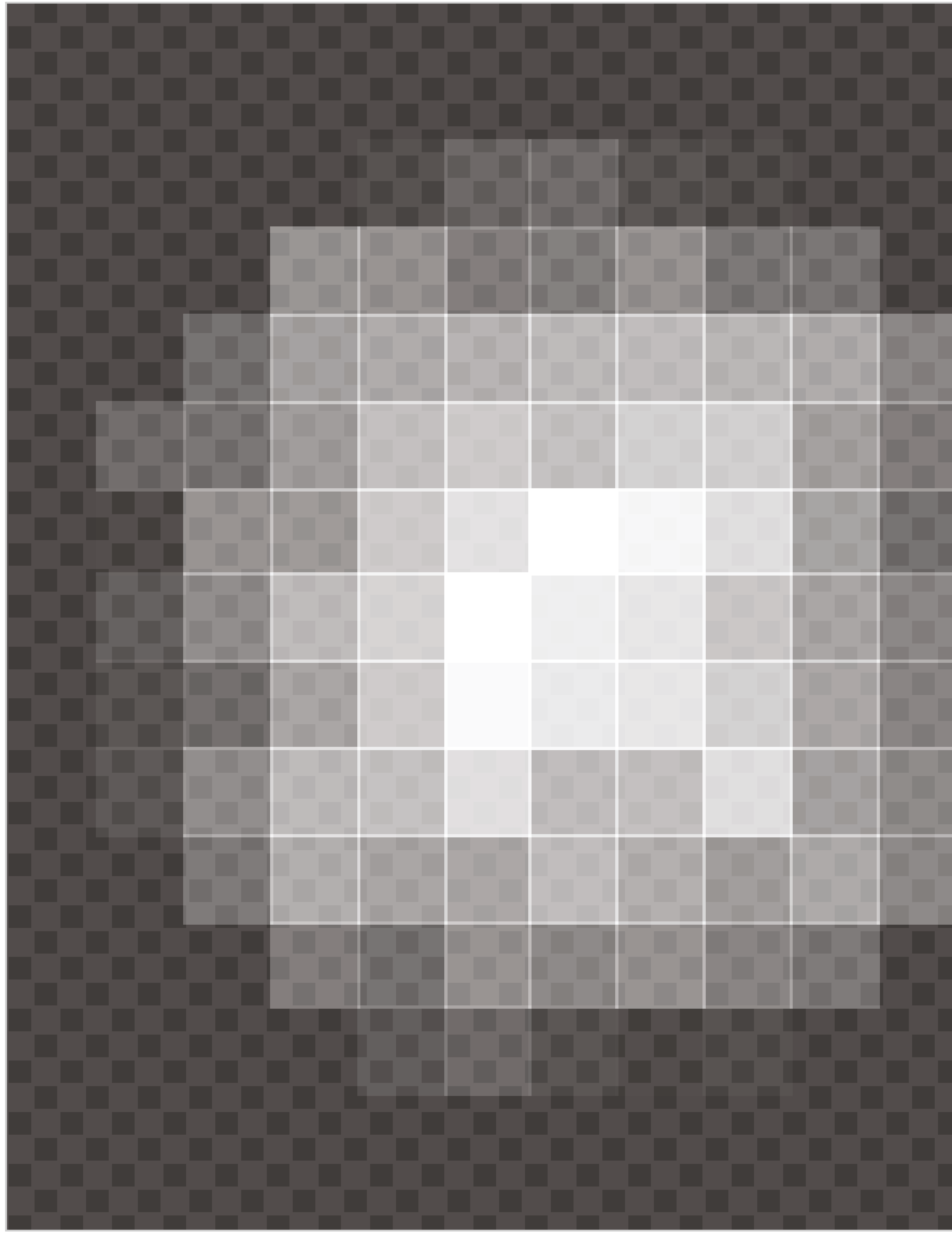
Applying the rationale behind randomized experiments will help to understand the effects of the different policy ideas and theories in place without solely relying on intuitive assumptions of alleged impacts of regulation. This may bring countries' standards closer and thus real possibilities of larger consensuses that further promote innovation and competition in the digital economy.

> ***Given these circumstances, how can we make something good out of the increasingly fragmented regulatory framework for online content moderation and the data richness that characterizes the industry?***

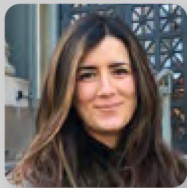
15 See Arthur Jatteau, *The Success of Randomized Controlled Trials: A Sociographical Study of the Rise of J-PAL to Scientific Excellence and Influence*, Historical Social Research, 43, no. 3, 165, 2018: 94–119. Available at https://www.jstor.org/stable/26491530.

16 Eduardo Hariton and Joseph Locascio, *Randomised controlled trials - the gold standard for effectiveness research*, Study design: randomised controlled trials. BJOG: an international journal of obstetrics and gynaecology, 125(13), 1716, December 2018. Available at https://doi.org/10.1111/1471-0528.15199.

© 2022 Competition Policy International All Rights Reserved



PLATFORM CONTENT STANDARDIZATION



BY
MARTA CANTERO GAMITO

Associate Professor of IT Law, University of Tartu. Research Fellow, School of Transnational Governance, European University Institute.

01 INTRODUCTION

Can (and should) moderation be standardized? Standardization can be defined as the development of consensus-based, often technical, non-binding guidelines to be followed by all the processes related to producing a product or performing a service. As part of market regulation, the legislator often delegates the definition of specific technical details to standardization. When that happens, voluntary compliance with the standards presumes conformity with the essential requirements es-

© 2022 Competition Policy International All Rights Reserved

tablished by legislation to place a product or to provide a service in the market.

By *platform content standardization* I refer here to the setting of standards that, not being necessarily technical, prescribe the specifications and procedural requirements to comply with regulatory obligations concerning content moderation. The emphasis on procedural steps in the forthcoming Digital Services Act ("DSA") can also be seen as an example of standardization; a platform can provide services in the EU internal market provide that it abides by specific procedural guidelines. In this regard, it can be argued that standards are generally non-binding while the DSA establishes *mandatory* regulatory obligations. Underlying the claims to regulate content moderation there is a justified distrust of platform's moderation policies and processes. However, by regulating procedures and not outcomes, the DSA opens the door to institutional variations, blurring the line between law and standards.

Administering speech is no easy task. But when faced with a high volume of cases, there is a trade-off between adjudicating in a rapid manner and compromising procedural guarantees in the protection of fundamental rights. While forthcoming due process provisions are welcome, there are already voices signaling the potential perils of the industrialization of content moderation.[2] Building on Keller's insightful analysis, this article joins the critical voices on the DSA by offering a view on the shift from law to standards (and codes). In this short article I shall explain how the forthcoming EU rules promote and fuel platform content standardization, and why this might lead to suboptimal outcomes that counteract the (seemingly) desired policy and regulatory goals of the DSA. Moreover, I reflect on how platform content standardization is leading a process of change in the way we perceive law and legal authority.

02

FROM LAW IS CODE TO CODE IS PROCEDURE

A platform in 2022 is something far removed from the lawmaker's concept of intermediaries when the safe harbors by the eCommerce Directive or Section 230 were drafted. Today's platforms connect, entertain, employ, make richer, give voice, as well as censor, cancel, and generally govern users. As a result, the emphasis has been put on the private rules, practices and procedures through which platforms exercise their power in an attempt to understand how they regulate us.[3] Platform (private) ordering plays an ever-increasing role in governing the conditions for freedom of expression and access to information. This is largely due to, first, the capacity and necessity of platforms to automate the administration of content and, second, the growing reliance of public regulation on private ordering and automated enforcement.

> ***A platform in 2022 is something far removed from the lawmaker's concept of intermediaries when the safe harbors by the eCommerce Directive or Section 230 were drafted***

Moderation is intrinsic to platforms' value propositions.[4] Among the 2.5 quintillion bytes of content created every day there are millions of photos, videos, posts, and comments uploaded to the internet. In order to sustain their business model, largely based on advertising, platforms must organize and categorize information. This is true not only for the purposes of structuring of participation but also for preventing abuses. Assuming that each of the videos, photos or posts may be potentially illegal, harmful or infringe platforms' Terms of Service, it seems almost inescapable that *effective* content moderation requires automation, or at least a certain level of it. Besides, the scale of moderation needed when platforms users' pools expand can only be approached with the use of computerization techniques.

2 Daphne Keller, *The DSA's Industrial Model for Content Moderation*, VERFBLOG (February 22, 2022), https://verfassungsblog.de/dsa-industrial-model/.

3 Kate Klonick, *The new governors: The people, rules, and processes governing online speech* 131 *HARV. L. REV.*, 1598-1670 (2017).

4 See TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA (2018); James Grimmelmann, *The virtues of moderation*. *YALE JL & TECH*., 17, 42 (2015).

© 2022 Competition Policy International All Rights Reserved

Automated curation, also referred to as algorithmic commercial content moderation,[5] allows not only a more efficient identification of inappropriate content, but also the possibility of taking immediate action consisting of removing or downgrading content and/or shadow banning users. By using digital hash technology or matching, filtering, and prediction (including the use of natural language processing or "NLP"), and image recognition tools, the detection of copyright-infringing, hate speech, extremist, and other types of unlawful content can be done in a matter of (micro) seconds.

Content moderation has thus become an ideal use case of AI for law enforcement purposes in digital environments. Aware of this, the legislator has been gradually supporting the use of algorithmic tools for regulatory compliance, evidencing the power of the Code and its capacity to constrain (online) behavior.[6] In Europe, there are rules that require platforms to monitor and enforce rules against any behaviour infringing child rights (Audiovisual Media Services Directive and the Child Sexual Abuse and Exploitation Directive), copyright (Copyright Directive), exhibiting terrorist content (Counter-Terrorism Directive), or racist and xenophobic hate speech (Counter-Racism Framework Decision). There are also some soft law initiatives on disinformation, hate speech and illegal content online such as the Code of Practice on Disinformation or the EU Code of conduct on countering illegal hate speech online. The use of algorithmic tools for copyright enforcement was highly debated, especially Article 17 of the EU Copyright Directive, which requires the use of automatic recognition and filtering tools, which is at odds with the prohibition of general monitoring obligations contained in the eCommerce Directive (Article 15).[7]

Yet, despite criticism, the EU legislator has continued supporting the role of online intermediaries to voluntarily take the necessary measures to comply with the requirements of EU law, recognising the role of platforms as regulatory intermediaries.[8] Now, due to legitimacy concerns regarding the adjudication of fundamental rights by private actors, the DSA will set a procedural safety net to ensure accountability and respect for fundamental rights. The threats over democracy escalate in highly concentrated markets where only a few dominant platforms are in charge of channeling public discourses, and the DSA stands as the promise for regulating platforms and the way they moderate.

Content moderation has thus become an ideal use case of AI for law enforcement purposes in digital environments

As the EU's commissioner Thierry Breton has warned Elon Musk over the new direction in Twitter's content moderation policies,[9] a more stringent approach towards content moderation by platforms seems to be not only a requirement for conditional immunity over user-generated and user-shared content, but also an entry condition to the EU's internal market. We should reflect upon the implications of governing fundamental rights with an (internal) market narrative. Seen from the perspective of market regulation, I argue that the legislator is standardizing content moderation by leaving the definition of the technical details to standard-setting organizations.

03 CAN CONTENT MODERATION BE STANDARDIZED?

Standards contribute to remove market barriers and to decrease compliance costs. Seen this way, there are good reasons why content moderation can be standardized.
First, platform moderation operates in a one-to-many

5 Robert Gorwa, Reuben Binns & Christian Katzenbach, *Algorithmic content moderation: Technical and political challenges in the automation of platform governance*, 7 *BIG DATA & SOCIETY* 1-15 (2020).

6 LAWRENCE LESSIG, *CODE: AND OTHER LAWS OF CYBERSPACE* (1999).

7 Giancarlo Frosio, *To Filter, or Not to Filter-That Is the Question in EU Copyright Reform,* 36 *CARDOZO ARTS & ENT. LJ*, 331 (2018); João Quintais, *The new copyright in the digital single market directive: A critical look*, 1 *EUROPEAN INTELLECTUAL PROPERTY REVIEW* 28; Martin Senftleben & Christina Angelopoulos, *The Odyssey of the Prohibition on General Monitoring Obligations on the Way to the Digital Services Act: Between Article 15 of the E-Commerce Directive and Article 17 of the Directive on Copyright in the Digital Single Market* (2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3717022.

8 Christoph Busch, *Self-Regulation and Regulatory Intermediation in the Platform Economy* in THE ROLE OF THE EU IN TRANSNATIONAL LEGAL ORDERING. STANDARDS, CONTRACTS AND CODES (Marta Cantero Gamito & Hans W. Micklitz, eds. 2020).

9 https://www.ft.com/content/22f66209-f5b2-4476-8cdb-de4befffebe5.

© 2022 Competition Policy International All Rights Reserved

environment. To date, every platform is governed by its own house rules, which determine the terms of use and access to information, including moderation policies and procedures. Therefore, standardizing platforms' moderation can contribute to preventing a situation whereby, within the internal market, what is allowed on one platform could be simultaneously prohibited on another. In this regard the DSA promotes the role of standards (and technological means) to facilitate the effective and consistent fulfillment of regulatory obligations.[10] For example, in the proposed draft, the European Commission encourages the development of industry standards covering technical procedures for the submission of notices regarding violating content or for interoperable advertising repositories, among others. One of the most criticized aspects of the DSA is Article 14, dealing with notice and action mechanisms, as it empowers platforms to make decisions about the (il)legality of content. To assess content's legality, the DSA commends the value of industry standards for helping to "distinguish between different types of illegal content or different types of intermediary services, as appropriate."[11] Different standards to identify illegal content can also help to prevent fragmentation. Existing legislative initiatives, such as the German Network Enforcement Act ("NetzDG") or the French Law on Countering Online Hatred ("Avia Law"), set different guidelines on content moderation, and this was considered to pose a barrier to the free movement of services and prevent interoperability. In this regard, standardization would contribute to harmonize the internal market – let's not forget that the DSA is based on the internal market harmonization legal basis.[12]

Standards contribute to remove market barriers and to decrease compliance costs

Secondly, standardization can contribute to reduce compliance costs. One of the downsides of the DSA is that, although not all, it establishes a set of procedures to be complied by all platforms regardless of size, significantly or even disproportionately increasing costs for small and medium online intermediaries and reinforcing the competitive disadvantage of these companies *vis-à-vis* incumbents. Standardization can indeed minimize the required efforts to monitor online activity by setting more efficient content review systems. But if the EU wants to protect European-born innovation, it needs to be more flexible with those smaller intermediaries trying to reach a significant user base. This could perhaps explain why there are other areas for which the DSA opens the door to standardization, in addition to submission of notices, such as submission of notices by trusted flaggers, interfaces to comply with regulatory obligations (including APIs), standards for auditing, data transmissions, or concerning the interoperability of advertisement repositories.[13] While this approach is a step towards levelling the playing field for big and small platforms, we should consider whether a one size fits all approach is suitable for content moderation.

04
SHOULD IT?

There are important problems associated with entrusting content moderation to standardization. Continuing with the example of Article 14 DSA, precise and adequately substantiated notices would constitute actual knowledge for the purposes of hosting liability. This means that standards are expected to define what content is illegal and what content is not. Provided that the submission of these notices would take away immunity, it is realistic to believe that platforms would likely remove notified content even when this is lawful.

This is problematic from the perspective of the assessment of content's legality. The impact and scale of platforms' action and power over the way we use our fundamental rights requires putting in place procedural safeguards but they can also reveal a darker side with undesired effects since content standardization can lead to automated and generalized cancelation. For example, while the DSA establishes due process obligations by regulating dispute resolution procedures in Articles 17 and 18, it does not provide any guidance on how platforms make decisions that generally affect the fundamental right to freedom of expression. This can be seen as a missed opportunity for the legislator to codify the existing (and abundant) case law concerning the limitation of fundamental rights and proportionality assessments. Any limitation to EU funda-

10 Article 34 and Recital 66.

11 Recital 66.

12 Article 114 of the Treaty on the Functioning of the European Union.

13 See Article 34 DSA.

© 2022 Competition Policy International All Rights Reserved

mental rights must be provided for by law.[14] Therefore, an assessment shall be made concerning whether potential limitations to freedom of expression made on the basis of standards' understanding of content legality would pass a proportionality test.

Resulting limitations to free speech do not result from any legal general obligations imposed on intermediaries to remove illegal content. The DSA does not require platforms to delete content but to take a decision with regard to the information considered to be illegal.[15] Instead, potential limitations would come for instance from either platforms' policies concerning users' claims of illegality or by the threshold to be set in the standard that contains the technical specifications governing notice and action mechanisms – or both. The European Court of Human Rights has already recognized that the requirement that limitations to fundamental rights must be provided for by law can be interpreted expansively and in a flexible manner so as to avoid excessive rigidity, which would allow the law to keep pace with changing circumstances.[16] From this perspective, standardization of content moderation seems suitable and compatible with the approach of the EU legislator supporting soft law and private regulatory initiatives to fight against hate speech or disinformation. However, this should not lead to intermediaries taking measures, or advocating for standards, that would affect the essence of the freedom of expression of users who share lawful content.[17]

From this perspective, platform content standardization would require the exploration of important questions. What does standardized and algorithmic moderation mean for the democratic control of the adjudication of fundamental rights? How does the regulatory reliance on standards and codes interplay with notions of authority? What does the recognized administrative power of platforms' private contracts and algorithms mean for our understanding of law, the legal system and, ultimately, the (digital) constitutional state? Can automated content moderation fix societal problems or is it the reason for societal problems to emerge? While these are questions not to be answered in this short article, some preliminary observations can be made.

Platform content standardization can affect fundamental rights on two levels: institutional (including procedural) and normative. Institutionally, "mandated" private ordering blurs the distinction between "standards" and "law." On the one hand, standardization can be seen as a tool to depoliticize content moderation. However, standardization is ultimately about normative choices. Standards are inherently political, as they involve adherence to a particular decisional approach or understanding, representing critical value choices. This has been widely discussed with regard to internet governance from the perspective of the internet's infrastructure. Despite that, we are witnessing a growing reliance on multistakeholder standard-setting in other areas, such as artificial intelligence. On the other hand, the regulative power of standardization is often contested.[18] However, inclusive standardization would allow the incorporation of multistakeholders' preferences in the regulatory process. The European Commission is currently trying to reinforce the democratic credentials of European standardization organizations ("ESOs") by improving their decision-making processes and requesting them to "modernise their governance to fully represent the public interest."[19] Efforts need also to be made to prevent strategic participation and geopolitical opportunism. Therefore, it is important to keep an eye on how standardization following the DSA will occur, as it can be reasonably expected that small platforms would not have the means or the reach to be part of the standard-setting process, which would allow big platforms to hold the pen.

Platform content standardization can affect fundamental rights on two levels: institutional (including procedural) and normative

At the normative level, platform content standardization is based on a non-traditional philosophical foundation of law and law enforcement. Automated moderation brings

14 Article 52(1) Charter of Fundamental Rights of the European Union.

15 Article 14(6) DSA.

16 European Court of Human Rights, 16 June 2015, *Delfi AS v. Estonia* (64569/09), para. 121.

17 Cf. Court of Justice of the European Union, Judgment of the Court (Grand Chamber) 26 April 2022, Case C-401/19 *Poland v. Parliament and Council*. ECLI:EU:C:2022:297.

18 MARIOLINA ELIANTONIO & CAROLINE CAUFFMAN (EDS.). *THE LEGITIMACY OF STANDARDISATION AS A REGULATORY TECHNIQUE: A CROSS-DISCIPLINARY AND MULTI-LEVEL ANALYSIS* (2020).

19 EU Strategy on Standardization, EU Commission Communication "An EU Strategy on Standardisation. Setting global standards in support of a resilient, green and digital EU single market," COM(2022) 31 final.

© 2022 Competition Policy International All Rights Reserved

the technology to the forefront of law enforcement, with algorithms massively filtering and removing content. Governing mandated content moderation practices by non-legislative actors raises a whole set of issues related to the incorporation of value choices into privately designed algorithms with enforcement capabilities, the interplay between the code and the legal systems, and their inevitable impact on the ethos of fundamental rights adjudication. The protection of fundamental rights online is replacing the law-based enforcement discourse with a new set of code-based enforcement values consisting of access (to platform services), neutrality, and algorithmic transparency.

From this perspective, standardized moderation can solve some problems (e.g. procedural fairness) while it exacerbates others such as lack of pluralism. What constitute harmful content can vary and mean different things depending on the geographical, cultural, substantive, and subjective context. Content moderation is not neutral, and every moderation model involves a trade-off between competing interests and values. For example, moderation obligations under DSA may empower platforms to use public regulation to pick winners and losers. There are indeed important concerns related to the ambitions of protecting democracy through standards and algorithms. The most important one is that standardizing content moderation has a fundamental design flaw. Can we rely on AI and standards to oversee public interests? With more than 500 million tweets per day, there are just so many and mutually inconsistent moderation demands to be met that renders impossible to please everyone. The solution to the "wrong speech" should not be "standardized speech." However, this triggers the following question: does the definition of values belong to the law?

> ***At the normative level, platform content standardization is based on a non-traditional philosophical foundation of law and law enforcement***

Lastly, technosolutionism is leading a process of change from states based on the rule of law to states increasingly centered on the rule of code and system-level bureaucracy.[20] This raises questions as to, first, whether attempts to constitutionalize the cyberspace are compatible with the EU's constitutional pluralism and, second, whether technical solutions improve or instead worsen citizens' fundamental rights.

05 CONCLUSION

The internet once favored a remarkable development towards greater democratization. It facilitated a mechanism for anonymous users to voice and amplify their opinions. Today's sentiment is the opposite; internet's governance has the power to manipulate and shape public opinion, eroding democracy.

The broad recognition of the role of platforms in the functioning of democracy prompted lawmakers to investigate platforms' activities more closely. As a result, there is much hope in the initiatives by the European Commission to curb Big Tech power. From this perspective, the DSA can be seen as an *in-progress* political building process for internet governance. Yet, in my opinion, it would be incorrect to think that the DSA would solve all the existing problems and that business models based on users' data extraction would not find alternative ways to sustain their revenue models. Most importantly, the DSA should not be a victim of its own success. The public expectation of increased responsibility taken by platforms is overshadowing an underlying process of institutional innovation and the use of alternative regulatory techniques, which includes an excessive reliance on less accountable codes and standards.

This article has outlined the pros and cons of legitimizing the regulative power of non-legislative regulatory tools. Not much consideration has been given regarding the protection fundamental rights with (internal) market regulation narratives, although this can be seen as a legacy problem in the historical constitutional configuration of the EU. In this regard, it is argued that, while improvements can be made, an EU-level approach towards platform regulation is indeed welcome and necessary for effectively and consistently protecting fundamental rights online.[21] However, discussions should not terminate with the final approval

20 Mark Bovens & Stavros Zouridis, *From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control*, 62 *PUBLIC ADMINISTRATION REVIEW* 2, 174-184 (2002).

21 The proposal for a Regulation laying down harmonized rules on artificial intelligence ("AI Act") and the forthcoming initiative for protecting media freedom, the European Media Freedom Act ("EMFA") are also using legal harmonization in the internal market (Article 114 TFEU) as their legal basis.

© 2022 Competition Policy International All Rights Reserved

of the DSA. Instead, attention should be paid to standardization, where actual value choices affecting fundamental rights are to be made. Moreover, it is important to monitor that standardization is not resulting in a level of substantial harmonization that would compromise constitutional pluralism. The focus should be on advocating for a process of greater institutionalization and accountability of standard-setting that aims at reproducing the more participatory decision-making structures of the early days of the internet to counterbalance the existing centralization of platform power.

> *The internet once favored a remarkable development towards greater democratization*

© 2022 Competition Policy International All Rights Reserved

WHAT'S **NEXT**

For July 2022, we will feature a TechREG Chronicle focused on issues related to **Gig Economy**.

ANNOUNCEMENTS

CPI TechREG CHRONICLES August & September 2022

For August 2022, we will feature a TechREG Chronicle focused on issues related to the **Editorial Advisory Board**. And in September we will cover **Connected Healthcare**.

Contributions to the TechREG Chronicle are about 2,500 – 4,000 words long. They should be lightly cited and not be written as long law-review articles with many in-depth footnotes. As with all CPI publications, articles for the CPI TechREG Chronicle should be written clearly and with the reader always in mind.

Interested authors should send their contributions to Sam Sadden (**ssadden@competitionpolicyinternational.com**) with the subject line "TechREG Chronicle," a short bio and picture(s) of the author(s).

The CPI Editorial Team will evaluate all submissions and will publish the best papers. Authors can submit papers in any topic related to competition and regulation, however, priority will be given to articles addressing the abovementioned topics. Co-authors are always welcome.

© 2022 Competition Policy International All Rights Reserved

ABOUT **US**

Since 2006, **Competition Policy International** ("CPI") has provided comprehensive resources and continuing education for the global antitrust and competition policy community. Created and managed by leaders in the competition policy community, CPI and CPI TV deliver timely commentary and analysis on antitrust and global competition policy matters through a variety of events, media, and applications.

As of October 2021, CPI forms part of **What's Next Media & Analytics Company** and has teamed up with **PYMNTS**, a global leader for data, news, and insights on innovation in payments and the platforms powering the connected economy.

This partnership will reinforce both CPI's and PYMNTS' coverage of technology regulation, as jurisdictions worldwide tackle the regulation of digital businesses across the connected economy, including questions pertaining to BigTech, FinTech, crypto, healthcare, social media, AI, privacy, and more.

Our partnership is timely. The antitrust world is evolving, and new, specific rules are being developed to regulate the so-called "digital economy." A new wave of regulation will increasingly displace traditional antitrust laws insofar as they apply to certain classes of businesses, including payments, online commerce, and the management of social media and search.

This insight is reflected in the launch of the **TechREG** Chronicle, which brings all these aspects together — combining the strengths and expertise of both CPI and PYMNTS.

Continue reading CPI as we expand the scope of analysis and discussions beyond antitrust-related issues to include Tech Reg news and information, and we are excited for you, our readers, to join us on this journey.

Scan to Stay Connected!

Scan here to subscribe to CPI's **FREE** daily newsletter.

© 2022 Competition Policy International All Rights Reserved

CPI
SUBSCRIPTIONS

CPI reaches more than **35,000 readers** in over **150 countries** every day. Our online library houses over **23,000 papers**, articles and interviews.

Visit **competitionpolicyinternational.com** today to see our available plans and join CPI's global community of antitrust experts.

CPI COMPETITION POLICY™ INTERNATIONAL

© 2022 Competition Policy International All Rights Reserved