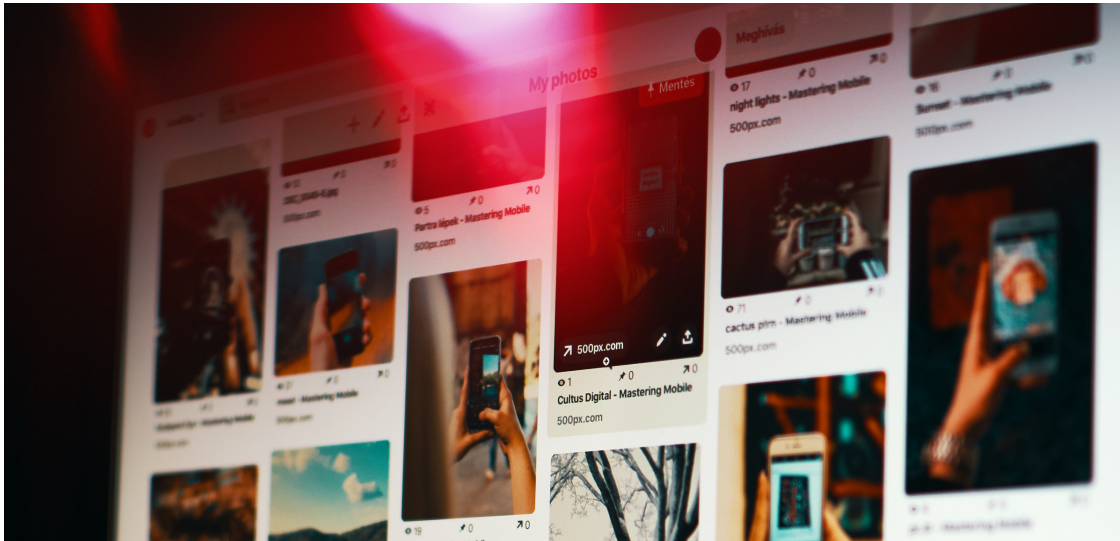


A PARTICIPATORY POLICY EXERCISE



CITIZEN PANEL: SOCIAL MEDIA & CONTENT MODERATION

AUGUST 2021

ORGANIZED BY
**UNIVERSITY OF MARYLAND COLLEGE
OF INFORMATION STUDIES
ETHICS AND VALUES IN DESIGN LAB**



Acknowledgements

This event was sponsored by the National Science Foundation under award CNS-1452854. UMD IRB 1682807-2.



COLLEGE OF
INFORMATION
STUDIES

EViD

Facilitation Team:

- Anna Lenhart, PhD Student, College of Information Studies, University of Maryland
- Dr. Sarah Gilbert, Postdoctoral Associate, College of Information Studies, University of Maryland
- Ari Perez Montes, Ronald E. McNair Fellow, B.A. University of Maryland

Advisor:

- Dr. Katie Shilton, Associate Professor, College of Information Studies, University of Maryland

Project Statement

In recent years, social media content moderation has become the subject of intense policy debate, with lawmakers from a range of ideological backgrounds calling for regulation. Much of the debate centers around Section 230 of the Communications Decency Act (“Section 230”). Section 230 shields interactive computer services (such as Facebook, Twitter, and YouTube) from liability for most content on their platform and provides these companies with the freedom to “restrict access to or availability of material that the provider or user considers to be obscene” (Communications Decency Act of 1996, Section 230 (C)(2A)).

While calls for reform typically come from lawmakers and academics, the protections offered by Section 230 are directly relevant to everyday users whose submissions make up the bulk of content found on these sites and who are directly impacted by social media policies and government protections. An individual’s relationship with social media intersects with their perspectives on commerce, sexuality, health, democracy, entertainment, crime and terrorism—values which will frame their perspectives on the types of policies government and social media employ as well as how those policy decisions will be carried out.

Citizen panels convene a small representative group of community members who go through a learning process on a technological subject, engage experts, and identify critical issues and values for policy makers to consider. To help social media users learn more about policy development at the governmental and company levels, the UMD Ethics and Values in Design (EViD) lab designed and tested a game in which participants debate, experience, and make decisions about platform governance, titled Content Moderation by Design (CMBD). Throughout the game, players come to understand how Section 230 impacts platform governance, and they build empathy for social media trust and safety teams and content moderators. By offering an experience with online governance, participants can begin to imagine, ideate and advocate for a safer internet.

On August 14th, 2021, 9 people from across the country convened virtually to answer the charge...



What are the values, in order of priority, that should be reflected in federal legislation overseeing content moderation on Social Media sites?

What action (if any) should federal legislators take?



Citizen Panel Process

The event took place virtually on August 14, 2021. A full academic case study will be published within the year.

Recruitment & Selection

Participants were recruited to apply through digital ads on Facebook and various listservs, Slack channels and social media posts affiliated with the facilitation team. Over 40 people applied to participate and the facilitation team selected 9 citizen panelists and 2 alternates. Selection was based on who could bring a diverse perspective to the conversation based on age, gender identity, race/ethnicity, experiences and attitude towards social media. Selected participants were NOT experts in technology policy or online content moderation. The table in Appendix A describes the background of the participants. Participants received a \$100 honorarium.

Event Agenda

- 9:00am-9:15am (EST): Introduction & Ground Rules
- 9:15am-12:00pm (EST): Play Content Moderation by Design Game
- Game End - 12:30pm (EST): Personal Reflection & Lunch
- 12:30pm-1:15pm (EST): Presentation of key concepts not included in the game
- 1:15pm - 1:20pm (EST): Break
- 1:20pm - 2:25pm (EST): Whiteboard session centered around the charge & Prepare for Expert Panel
- 2:25pm -2:30pm (EST): Break
- 2:30pm- 3:45pm (EST): Expert Panel
- 3:45pm -4:00pm (EST): Afternoon snack & water break
- 4:00pm - 4:45pm (EST): Drafting Official Statement
- 4:45pm- 5:00pm (EST): Close & take aways



Content Moderation by Design Game (CMbD)

To play CMbD, the nine panelists were split into three groups of three and dispersed into separate virtual rooms. Each virtual room had a facilitator. Players were told they work for a brand new social media startup--Contentr--a company that is determined to avoid the same mistakes as its predecessors.

In the beginning of the game, Section 230's liability shield was described, along with other important platform policies such as the Digital Millennium Copyright Act and Fight Online Sex Trafficking Act (FOSTA)/ Allow States and Victims to Fight Online Sex Trafficking Act (SESTA).

The game had three cycles, each representing a category related to types of content: cycle one, sexually explicit content and illegal activity, cycle two, self harm and graphic content, and cycle three, harassment and hate speech and quality contributions. Within each cycle there are two rounds. In the first round players work together to develop content moderation policies (i.e., determine which content policy cards should be allowed or banned), in the second round, players switch into the role of a content moderator where they use their policy to make moderation decisions based on real life examples. Finally, players see how their decisions play out through event cards that change their ad revenue and expenses.

“I was surprised at how well, in the game, my fellow players and I were able to come to an agreeable consensus on many issues, but disagreements became opportunities to explore issues that were high-priority for other people. My experience in this game has changed the way I [think] about content moderation...I have a greater appreciation of the subtlety of the problem”

~Citizen Panelist



Expert Panelists

The following three experts agreed to answer questions that the citizen panelists had following the CMbD game. Their participation is in no way an endorsement for the ideas presented in the *Official Statement*.

- Dean Jackson, Project Manager of the Influence Operations Researchers' Guild, a component of the Partnership for Countering Influence Operations at the Carnegie Endowment for International Peace
- Tejas N. Narechania, the Robert and Nanci Corson Assistant Professor of Law at the University of California, Berkeley, School of Law
- Kyle Pittman, Community Moderator for Reddit.com's r/AskHistorians and r/IndianCountry

Drafting the Official Statement

After the expert panel, participants convened to draft their *Official Statement* which included an executive summary and their response to the charge. Because of time constraints the group was not able to finalize their responses during the Citizen Panel event. As a group, the facilitators and participants felt it would be best if the facilitators listened to the recordings from the whiteboard sessions, gleaned all the ideas and sent the list of ideas around for votes the following week. Over the week of Aug 15-22, 2021, each participant selected the values and actions they believed were most important and any ideas they wanted included in the executive summary. The votes were used to rank the lists on the following page and the executive summary is a light rewording of the key ideas submitted. *This Official Statement is not a consensus report but rather represents a range of ideas.*



Official Statement

“Social platforms are social goods...this is not a battlefield, we don’t need a sword and a shield*, we need bread and water, we need to be feeding ourselves”

~ Citizen Panelist

*the “shield and sword” metaphor is frequently used to describe Section 230 and was used in the CMbD game.

Executive Summary

Most federal legislation was written when media spread through newspapers and television. It’s necessary to update laws to reflect the way we express ourselves today. We hope that lawmakers engage a wide range of independent groups, academics and social media companies to craft new regulations. New laws should not only discourage harmful practices but work to encourage a thriving, value-driven media ecosystem.

Values

What are the values, in order of priority, that should be reflected in federal legislation overseeing content moderation on social media platforms?

1. Human autonomy (over data and content) (7 votes)
2. Truth rooted in the scientific process (5)
3. Mental, emotional and physical health (5)
4. Diversity of thought (5)
5. Flexibility and adaptability (4)
6. Empowerment of every stakeholder (3)
7. Truth rooted in the scientific process and lived experiences (3)
8. Justice (enabling enforcement of laws) (3)
9. Democracy and distributions of power (1)
10. Inclusiveness (1)
11. Growth (as in “growth mindset”) (1)
12. Utilitarianism (harm to the last number of individuals) (1)



Actions

What action (if any) should federal legislators take to address challenges posed by social media?

1. **Pass comprehensive legislation that provides internet users with data rights** such as privacy, anonymity, data access and data deletion. (7 votes)
2. **Improve the process for public-facing oversight of platforms** (beyond public Congressional hearings). (4)
3. **Improve access to voting** (votes are a way for people to choose the values they want reflected in the nation's leadership and therefore the media). (3)
4. **Create a diverse panel** to provide guidance on what content should be considered accurate/inaccurate. (3)
5. **Incentivize K-12 education systems to prioritize** technical skill development, critical thinking, information science and statistics. (3)
6. **Incentivize the creation and adoption of standards for digital forms of consent** (i.e. pgp, encryption keys, etc) for images containing nudity. (3)
7. **Enable mechanisms for investigations** into platforms to better understand harms and the spread of disinformation. May include policies that mandate transparency into how platforms categorize content, moderation practices and the platform's algorithms. (2)
8. **Encourage distributed online communities** with community-based content moderation practices. (2)
9. **Fund efforts to improve science communication** through research grants and dedicated staff within agencies. (2)
10. **Fund public institutions to disseminate information** in the public interest through social media platforms. (2)



Actions Cont.

11. **Mandate mental health support** for content moderators. (2)
12. **Increase lawmaker's knowledge** and understanding of social media platforms. (2)
13. **Hold platforms accountable for content** that harms/victimizes users, allow the courts to set precedent. (2)
14. **Mandate interoperability and portability** between social media platforms to facilitate the emergence of a more distributed approach to content moderation. (2)
15. **Designate social media as a utility (common good)** and insist social media platforms promote the best of human values as opposed to a focus on advertising revenue. (2)
16. **Collaborate with international lawmakers.** (1)
17. **Create a new government agency or department** focused on the attention economy that is tasked with incentivizing--through grants, tax-breaks, etc--responsible platform design. (1)
18. **Incentivize social media platforms to engage a range of stakeholders**, community members and employees in the creation of their content moderation practices.



APPENDIX

It was the facilitators goal to recruit a group of nine participants that represent the views and background of the country. The final group skewed towards high levels of educational attainment, and we struggled to recruit older and rural Americans. Based on comments during the application process, we believe the low honorarium and broadband requirements contributed to recruitment challenges.

Demographic	United States	Ideal # of Participants	Actual # of Participants
Gender			
Female	50% (a)	4	5
Prefer to Self-Describe	-	1	1
Male	-	4	3
Race/Ethnicity			
Caucasian/White	60% (b)	5	4
Persons of Color/Multiracial	40%	4	5
Age (c)			
18-34	30%	3	5
35-54	32%	3	3
55 & Over	38%	3	1
Urban/Rural (d)			
Urban	61%	5	1
Suburban	25%	2	8
Rural	14%	2	0
Educational Attainment			
Less than High School	12%	1	0
High School, GED or Some College	56%	5	1
College Degree	32% (e)	3	8
Views About Social Media Regulation			
Strong	unknown	3	4
Neutral	unknown	3	2
Weak	unknown	3	3
Political Leaning (f)			
Right Leaning	44%	4	2
Independent	7%	1	1
Left Leaning	49%	4	6

a) "Female persons, percent" (<https://www.census.gov/quickfacts/fact/table/US/RHI825219>)

b) "White alone, not Hispanic or Latino, percent"(<https://www.census.gov/quickfacts/fact/table/US/RHI825219>)

c) Single Year of Age and Sex Population Estimates: April 1, 2010 to July 1, 2019 - CIVILIAN (SC-EST2019-AGESEX-CIV), percentage based on adult population over 18

d) Data by Pew Research. Urban defined as "urban core" + "smaller metropolitan" <https://www.pewresearch.org/social-trends/2020/07/29/prior-to-covid-19-urban-core-counties-in-the-u-s-were-gaining-vitality-on-key-measures/>

e) "bachelor's degree or higher, percentage of persons age 25years +, 2015-2019" (<https://www.census.gov/quickfacts/fact/table/US/RHI825219>)

f) Data by Pew Research. <https://www.pewresearch.org/fact-tank/2020/10/26/what-the-2020-electorate-looks-like-by-party-race-and-ethnicity-age-education-and-religion/>