



THE
FUTURE
OF
FREE
SPEECH



Global Freedom of Expression
COLUMBIA UNIVERSITY



JUSTITIA

*“Regulating the Online
Public Sphere:
From Decentralized Networks
to Public Regulation”,
a retrospect*

OCTOBER 3 & 4, 2022

“Regulating the Online Public Sphere: From
Decentralized Networks to Public Regulation”,
a retrospect.

“Regulating the Online Public Sphere: From Decentralized Networks to Public Regulation”, a retrospect.

On October 3 & 4, 2022, Columbia Global Freedom of Expression hosted the conference titled: **“REGULATING THE ONLINE PUBLIC SPHERE: From Decentralized Networks to Public Regulation”**. In this event, a diverse group of panelists, comprised of leading academics, industry experts, and notable activists, gathered to discuss the new challenges that the regulation of the global public sphere pose. The event was held in person with over 50 participants in attendance each day and live-streamed on YouTube, drawing over 2,000 views. The conference webpage, including the agenda, speaker bios and the archived video is available [here](#).

On the first day, the speakers discussed new models of decentralized networks and their importance in the current social media ecosystem. The panelists had the opportunity to debate the advantages of new network models for freedom of expression and reflected on central issues regarding this novel approach. For instance, how would content moderation work in a decentralized network? What can be done differently and how does it affect user experience? Should new regulatory frameworks be different when it comes to decentralized networks?


On the second day, panelists debated and analyzed the different regional approaches to public and private regulation of content moderation on the Internet. Speakers focused mainly on discussing current regulatory frameworks such as the Digital Services Act (DSA) in Europe or the work of Meta’s Oversight Board, to explain and understand the impact of content moderation on public discourse and how it shapes the internet’s architecture.

A diverse group of panelists, comprised of leading academics, industry experts, and notable activists, gathered to discuss the new challenges that the regulation of the global public sphere pose.

Feedback on the conference was overwhelmingly positive from a broad range of industry experts.¹ Daphne Keller, Director of Program on Platform Regulation, Stanford Cyber Policy Center [Tweeted](#), “This was the best discussion of interoperability, “middleware,” or “protocols not platforms” I’ve ever been in, hands down. These ideas are *really* important. They might be the best path forward for law makers responding to today’s platform speech pathologies.”

David Kaye, Clinical Professor of Law at the University of California, Irvine, and the former United Nations Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2014-2020) [Tweeted](#), “Another excellent panel w @Klonick @article19law @McAndrew @ZoeDarme. And from [Rachel Wolbers](#), Head of Global Engagement at Oversight Board Administration, “Thanks GFoE for putting together such a great conference!”

This conference was made possible thanks to [The Future of Free Speech Project](#), a collaboration between [Justitia](#), [Columbia University’s Global Freedom of Expression](#), and [Aarhus University’s Department of Political Science](#)— with support from the [John Templeton Foundation](#).

 **Daphne Keller**  [@daphnehk](#) ...

This was the best discussion of interoperability, “middleware,” or “protocols not platforms” I’ve ever been in, hands down. These ideas are *really* important. They might be the best path forward for lawmakers responding to today’s platform speech pathologies.

[Traducir Tweet](#)

 **jose_moreno** [@jose_moreno](#) · 4 oct.

This panel yesterday, moderated by @mmasnick, with @daphnehk, among others, was excellent. Worth seeing > DAY 1 | Regulating the Online Public Sphere: From Decentralized Networks... [youtu.be/OYsnGkFG7o8](#) via @YouTube

10:14 a. m. · 4 oct. 2022

 **David Kaye**  [@davidakaye](#) ...

another excellent panel w @Klonick @article19law @McAndrew @ZoeDarme. and as a bonus, @doctorow's zoom studio looks exactly what you'd think cory's zoom studio would look like.

[Traducir Tweet](#)

 **Columbia Global Freedom of Expression** [@ColumbiaGFoE](#) · 3 oct.

● LIVE NOW! Day 1 of the #OnlinePublicSphere2022 is kicking off with welcome remarks by Hawley Johnson, Associate Director of GFoE.

Watch the livestream HERE: [youtu.be/OYsnGkFG7o8](#)

11:26 a. m. · 3 oct. 2022

¹ Highlights from Twitter are available [here](#) and a full list of relevant social media posts are [here](#).

DAY 1 – MONDAY, OCTOBER 3, 2022

Introduction to the First Day - Jacob Mchangama, Founder and Executive Director of Justitia and The Future of Free Speech Project

Jacob Mchangama, Founder and Executive Director of [Justitia](#) and [The Future of Free Speech Project](#), provided introductory remarks. In order to frame the discussions to come, Mchangama provided some historical context about the underlying philosophy of the internet in relation to freedom of expression and how it has varied over time. Mchangama recalled Tim Berners Lee’s — best known as the inventor of the World Wide Web—original vision of the internet in 1999 which “encompass[ed] the decentralized, organic growth of ideas, technology and society” without imposed filters, and juxtaposed it with how it is seen today. As an example of the radical shift in perception of the internet and social media’s role in society and the political process, he shared a quote from former US President Barack Obama from 2006 where Obama described the internet as a space that was “neutral,” “without corporate media middlemen,” and where he could speak without censorship to energize young voters. Yet by 2020 Obama asserted that online disinformation was “the single biggest threat” to democracy. According to Mchagama, the driver of the narrative shift away from the early techno-optimism, is the platformization of the internet where it is now dominated by centralized tech companies, creating “choke-points” for freedom of expression.

“Even if we can learn from the past, I think it’s clear there’s a dire need to find solutions for today and our rapidly developing future. One of the big issues is whether the solution could be a way forward where we could benefit from the huge advantages of free and equal online speech while limiting the harms and costs.”

Despite the exponential growth of available information online in recent years, Mchangama argues we are experiencing a “free speech recession.” According to Freedom House’s Freedom On The Net Report, over the last 11 consecutive years online Freedom has been in the decline, and while it is most notable in authoritari-

an and illiberal states, it is also showing in democratic states which are advocating for increased restrictions on freedom of expression to manage online harms. Mchangama observed that some of this backlash is normal as there are historical parallels with previous periods of dramatic change heralded by technological innovation. Mchangama noted that —contrary to the popular belief that challenges to freedom of expression posed by technological developments are a novel problem of the digital age— throughout history, the introduction of new technologies, such as the printing press, the telegraph, or the radio, similarly introduced complex challenges to freedom of expression, and elicited staunch criticism extolling the dangers of new technological developments. In doing so, Mchangama underscored the possibility of learning from the past when it comes to confronting current challenges. However, today there is a “dire need to find solutions and chart a more healthy culture of free speech globally.” To that end, the conference aimed to explore if and how a return to decentralized models could offer solutions.

Mike Masnick, CEO at [The Copia Institute](#), moderated

Session I: Mapping the Decentralized Ecosystem.

the first panel of the day. In his introductory comments, Masnick stated that the purpose of this first session was to understand the impact of novel decentralized networks on freedom of expression and the influence of regulations in the technological ecosystem. To frame the discussion Masnick said that over time attitudes towards the internet have radically shifted from unreserved optimism to consternation as to how online platforms have been abused. Masnick seconded Mchangama's observations about the excitement over the promise of the internet in the early days of the 1990s and early 2000s when it was truly decentralized. The power of the internet was seen then

as the freedom to communicate without gatekeepers, to build your own services and move around a variety of services without being beholden to them. In recent years, Masnick noted, "I started to look at that and realize how do we get back to that world, how do we move from a world that was becoming more and more cen-



tralized and more in control of a few platforms or a few companies to one that was more decentralized, where the powers were more pushed out towards the ends and that could maybe lead to what we had been promised early on. " Even though the situation is so different today, perhaps the spirit of that early promise can shape the way forward. One of the main concerns today is that the policy discussions are about how to manage the 5-6 big tech companies and how to regulate them, rather than how to "empower more users to be able to have more power themselves and not necessarily hand that off to the large companies."

Daphne Keller, Director of the [Program on Platform Regula-](#)

[tion at Stanford Cyber Policy Center](#), responded to Masnick's request that she give a short summary of the primary concerns people have about freedom of speech and the way the internet is dominated by a group of large platforms.

Keller opened by observing that every human rights compliant democracy is facing the same problem of how to deal with the profusion of “lawful but awful” speech that offends social norms. Even in the United States which has the strongest protections for speech under the First Amendment, there is still pressure for platforms such as Facebook to remove offensive or allegedly harmful speech. The result is a “spiraling dynamic” where the platforms are removing more and more content and there are legal proposals to further restrict speech that used to be legal, such as in the UK Online

Safety Bill. As that direction is problematic for freedom of speech, and breaking up the platforms also comes with network effects in the long-term, the question then becomes what other ways can we reduce the power of the platforms over discourse, and give users more decision making power over what content they see. One answer is what Mike Mas-



nick has coined “protocols not platforms” and what Keller calls “Magic APIs.” This approach is somewhat of a hybrid model that allows users to house their data and content in a safe place, but then allow third-parties to provide content moderation services, and curate incoming content. While this model too has its challenges, namely economic incentives for those businesses, it is worth exploring an alternative in which content moderation is a service provided to centralized platforms and their users by third parties. For her, “diversifying the sources of control over the feeds that we see reduces the problem of having Mark Zuckerberg in charge of everything and also allows people to choose the discursive universe they want to operate in.”

Golda Velez, Co-Founder of [Cooperation.org](https://cooperation.org) and Community builder for [BlueSky](https://blue.sky), described how to go about creating a decentralized social media system and some of the ones she has worked on. She expressed concern for what she described as two types of attacks against freedom of expression on social platforms. The first type of attack comes from those who try to suppress content and the second comes from those who try to push content that benefits specific undemocratic agendas through disinformation. The extent of the attacks, Velez noted, is correlated to the reach of the platforms.

BlueSky, for example, was established by Twitter's Jack Dorsey as an independent organization to create a "durable protocol" that could be employed in different use cases. Velez described the decentralized ecosystem as a community of "walled gardens," some similar to Matrix rooms, including a few "big citadels like Facebook and Twitter." In this environment there are also, "attackers" who are seeking to invade privacy, suppress content or flood the gardens with content. Those who manage the walled gardens protect them in a variety of ways. Matrix, for instance, has blacklists against certain people coming into the rooms. On Mastodon each node has to do their own moderation, and other systems are even more closed and you can only talk to your friends. Velez pointed out that in these environments there are two important variables to consider. One is freedom of speech and the other freedom of reach. Communities in those walled gardens protect freedom of speech and are in charge of their own moderation based on the norms established by that community.



Some gardens create communications channels to communi-

cate with other gardens, which provides greater reach. However, those connections also open them up to attackers. New types of federated systems and tools are being created to manage the attacks and make communications safer between communities. Velez also pointed out that it is important to develop tools that stake one's own credibility online, especially when considering that anonymity, as currently fostered by the internet, incentivizes problematic behaviors.

Alex Feerst, CEO of [Murmuration Labs](#) – Law, Policy, Trust & Safety, responded to Masnik’s question about the extent to which the need for content moderation, trust and safety goes away in the protocol/interoperable or distributed environments? According to Feerst, centralized content moderation eventually finds itself limited by its own structural limits. Feerst described content moderation as an umbrella term that encompasses different processes, that range from amplification of content up to deletion of content. As such, content moderation, as it has been conceptualized, assumes that humans are going to act in a predetermined way while ignoring that humans also act “along with the specific incentives and structures and mazes we make for them”.

Feerst replied that there are inherent structural constraints in doing content moderation at scale. Those constraints built into the technology, in turn create incentives which shape online behaviours, both good and bad. If people respond to incentives that are built into the design of the technology, to change the behavior, you need to change the design of new technologies with an understanding of the incentives to get a different outcome. No matter the system, there will always be some need for intervention, either through regulation, law enforcement, or privately through platforms. Most interventions take place outside of court rooms due to the scale of human communication, either through tooling, filtering or norm development. Feerst believes “the goal is to do it in a way where the power allocates to a larger number of hands and each of those larger number of hands has less power.” To him, the notion of non-intervention is fundamentally fallacious if you really think about the issues. Ultimately, the promise of decentralization is the ability to do the allocation differently so that the notion of hardcore enforcement of content moderation bleeds into a much



more organic set of human behavior, such as fostering normative behavioral changes by putting pressure on people through mentorship and encouragement and curation as it is done offline. It is important to recognize that societies intervene into other people's expression off line all the time in subtle ways, and now we need to have tools that biomimic those human mechanisms for nudging and steering towards more civil communications.

Alan Z. Rozenshtein, Associate Professor of Law at the [University of Minnesota Law School](#), discussed his research work on Mastodon, which is a federated system and is a slightly different approach from a purely decentralized, purely distributed system. Masnick started by asking about his thoughts on federated systems and how they differ from the fully decentralized distributed systems and what are the pros and cons of one versus the other.

Rozenshtein started by explaining that Mastodon is decentralized based on its architecture, rather than as a matter of policy. This is different from Reddit's federation, for instance, which is decentralized in so far as it has subreddits that act with some degree of autonomy. However, its policies are also federated which allows for Reddit HQ to intervene at any moment. In other words, policy federation or policy decentralization only gets you so far.

To illustrate how this all works, Rozenshtein discussed Gab, the far-right social media platform. Gab began using Mastodon's open protocol system which created a problem for Mastodon because it has a norm, or policy, against Neo Nazi content. There was not much Mastodon could do about Gab's content because the protocol was designed so that it could not be interfered with. Hence, the broader network of Mastodon took action and began isolating Gab. The major federated instances started cutting ties with Gab, and then cutting ties with anyone who would not cut ties with Gab, and then in retaliation Gab decided to cut ties with the rest of the federation. In effect, Gab is now an island in the federated system. Rozenshtein believes this was the best outcome as the "Gabbers" still get to gab freely amongst themselves, and the rest of the federation does not need to worry about them. Should Gab cross over into the realm of illegal behaviours then they would be subject to law enforcement in the related jurisdiction. Federation is a good model, he



believes, as it makes it easy to communicate across instances, and importantly users can leave an instance yet retain their follower accounts – without giving up their “digital social capital” as they would if they exited Twitter, for instance. Federated models also have the capacity to allow for development of alternative content moderation practices, including franchising some of it out to third parties. However, federated models do have some network effects. For instance, the top five instances have 80% of the followers, but that is far less concentrated than Facebook.

What is, or should be, technologies role in norm creation or norm enforcement?

Masnack asked the panel to comment on technology’s role in norm creation or norm enforcement. In the current environment of centralized platforms, norm creation is outsourced which presents challenges and issues. It is not only about freedom of expression, but also freedom of association online. The example of Gab and how Mastodon instances decided to deal with them is instructive. One of the hopes is that the decentralized systems can create communities which will mimic real world interactions, and that if they can design the technology properly, to create the right incentives, the norms will follow. Masnick asked if that can work, and what is the basis of norm building in these kinds of communities? Is it technological or what are the other incentives?

Velez said that is an issue she confronts in her work and something they did in the BlueSky group.

Getting people to adopt new norms in decentralized spaces is very difficult, and hence the tools must be very user friendly in order to get people to adopt them and make the behavioral changes. According to Velez, we need to make tools for what we already do and that is big part of what’s been missing in this very incentivized space where everything is advertiser driven. There are a lot of tools for things that help advertisers and that track people’s behaviors - i.e. tools made to encourage people to spend money, but we are lacking tools that incentivize for things like staking credibility and other positive human actions.

Panelists noted the artificial quality of the online world with so much anonymity, where users do not know if they are interact-

ing with real people or bots, and this artificiality allows for negative behaviours to flourish. Feerst observed that this is a hybrid human tech problem, where some of it is novel. One aspect of social media according to Feerst is that “you can selectively repeal laws of the natural world, or enhance them. You can have ephemeral speech or you can make it non-ephemeral, you can make it reach more people or fewer people.” The artificial quality of the ecosystem has the ability to alter how communications normally take place in the natural world. He asked how can you take the wisdom of the natural world, and build tools in a way where humans are able to both perceive things and act on norms in ways that give them agency, but also the awareness of the scale and resulting impact of their online behaviour. Even in decentralized ecosystems, if you can prevent harms on the individual basis – for instance in the Gab example or if people select specific filters not to see certain types of content - harms may be happening in other parts of the system that need to be monitored. As Feerst explains it, there are “these larger second and third order effects around what norms and normalizations you’re allowing and encouraging that you should not fail to think about even when you’re tracking specific harms to people in a careful way.” To him it is important to think about both positive and negative enforcement when designing online networks.

Keller stressed that content moderation is expensive, and hence there is an economic problem that needs to be solved for these models to work. New incentives must be created for providers, if ad revenues are gone, and we need tools can we give to providers to reduce the costs of potentially redundant moderation in different nodes. In instances where moderation tools are shared, Keller warned that we need to avoid what Evelyn Dueck has coined “content cartels,” where there are limited arbiters determining who is and is not a terrorist, for instance, without the necessary oversight and transparency.

What is the impact of current regulations on decentralized networks?

Mike Masnick asked what is actually happening at the intersection of the regulatory space and the policy space which will enable any of this to happen or prevent any of this from happening. He is concerned that a lot of the regulatory proposals and

policy proposals that are out there today are so focused on how to limit Facebook, that they are not considering how they might also kill some of these ideas before they ever really have a chance to take off. He asked if any of the panelists had feedback on how to structure the regulatory environment in a way that actually enables decentralization?

Keller pointed out that there are many competing pressures in the regulatory space, such as with privacy, competition, interoperability, and content. The solutions all require trade-offs between competition and privacy values, with an overlay of the content concern which makes it challenging but all those variables need to be part of the conversation or there will not be a resolution. Masnick agreed that the issues are interconnected and that moving the needle in one area will impact other areas significantly. Even more troubling is that in the policy making space these ideas are developing mostly in silos. Velez added that another big issue is that regulations being established in the west are now affecting people who do not have a voice at the table, such as the populations in the Middle East, China, Hong Kong and Taiwan.

Session II: How to Get There from Here? Regulatory Requirements & Necessary Standards



Kate Klonic, Associate Professor of [Law at St John's University](#), moderated the second panel of the day. To preface the discussion, Klonic highlighted that there has been a lot of power shifting in the legal and cultural landscape of the internet environment around the world. She asked the panelists to share their perspectives regarding these historical transformations—and their influence in the current regulatory frameworks that exist in the United States and Europe—from their points of view and in light of their disciplinary backgrounds—some as tech experts, activists for freedom of expression in Europe, or public policy specialists.

Andrew McLaughlin, Co-founder of [Higher Ground Labs](#) and Board Chair of [Access Now](#), started his intervention by saying that 25 years ago there were two main axes of conflict in the world of internet policy: free speech vs. censorship and access and privacy vs. surveillance. McLaughlin noted that these axes still prevail today, but that he would add two new ones: concentration vs. competition and the borderless of the internet vs. sovereignty and territorial jurisdiction in policymaking. McLaughlin also highlighted a change of paradigm regarding the regulation of the internet. While 25 years ago, it was not possible to regulate in a speech protective way, and the idea of regulating the internet meant the possibility of killing it, with all its economic and socio-political benefits, nowadays the discourse is: “internet sucks: It ruins everything, we can’t make it any worse so let’s start regulating it as best we can”. Further, the tech companies that were disrupting the status quo back then are now

the incumbents. He observed that the “big goliaths to be brought down are themselves the internet companies that we cheered and championed as upstarts a generation ago.” Now it is individuals and collectively public interest which is challenging the concentration of power. For the Co-founder of Higher Grounds Labs, the European Union has established itself as the world’s lead regulator, aware that it is hard to limit their rules to just Europe and hence their rules “will be global by default”. McLaughlin explains that Europe has a strategic motivation to assert European policy preferences over the internet in order to avoid being a powerless actor between China and the U.S. Europe is heavily dependent on China for hardware and the U.S. for software. Meanwhile, the United States has transformed from a coherent player, regarding internet regulation, to a “chaos monkey” with scattered regulation. When McLaughlin looks around he feels “everything is getting worse, there is more censorship, less diversity, more walled gardens, less interoperability, more surveillance, less openness, and the rise of a new regulator across the ocean.”

Cory Doctorow, Special Advisor at [Electronic Frontier Foundation](#), disputed the current narrative or myth that the internet has gotten us “where we are”, and is responsible for so many of society’s ills. He described the internet as the nervous system of the 21st century to underscore the internet’s prevalence in many central tasks of life, and not just as a place where “whatever I don’t like” happens. In that sense, for Doctorow, any regulatory framework must take into account all equities.

Doctorow pointed out that many of the problems under discussion are not inherent to the technology, but rather to the market concentration we did not see coming. Antitrust enforcement is now a growing concern as these companies have cemented their dominance in such a way that they have interrupted the cycle of competition. The corporate oligopolies themselves have independence now equal to states. Doctorow highlighted that market concentration in the tech industry has created “giants with economies of scale” that can be weaponized by states. This is resulting in a dangerous situ-

While 25 years ago, it was not possible to regulate in a speech protective way, and the idea of regulating the internet meant the possibility of killing it, with all its economic and socio-political benefits, nowadays the discourse is: “internet sucks: It ruins everything, we can’t make it any worse so let’s start regulating it as best we can”.

ation where the tech companies are not only doing business with authoritarians, but setting up offices in those states with employees in them. This puts the employees under threat by state security services when those states “enact policies that say things like you may need to spy on your users, or have to censor content that is harmful to the government. This creates a circumstance of great vulnerability and acts as a spur to accelerate the ‘splinternet.’”

Barbora Bukovská, Senior Director for Law and Policy at [ARTICLE 19](#), discussed ARTICLE 19’s advocacy work related to the Digital Services Act (DSA) and the Digital Markets Act (DMA). Bukovska explained that while the DSA addresses content, it does not specify what content needs to be removed, rather focuses more on transparency and the procedural rights and the role of Civil Society in enforcing these regulations. Alternately, the Digital Markets Act provides tools to deal with very large on-line platforms. It defines gatekeepers in terms of their turnover and the users, and then establishes certain obligations, including interoperability, a very important component especially for messaging platforms.

ARTICLE 19 put forth a proposal for “unbundling” to address not only market power, but also content moderation and hosting services. Currently, these dominant platforms offer the hosting services and content moderation services as a bundle which has economic value, but it also locks users into using these platforms. The result is that there are high costs for users to switch platforms, such as losing data and connections. ARTICLE 19 proposed that large platform unbundle these two services and allow competitors to provide content moderation services on their platforms. Bukovska argued that would be better for market competition and for the users who would have a more control over the content moderation. Regulation, however, would be required for this model to work. Decentralization would be needed on a contractual level as well as on a technical level, through interoperability. A level playing field would need to be created for the alternative players to provide the content moderation services in a human rights compliant way. Although the unbundling proposal did not make it in to the DSA, ARTICLE 19 is not abandoning it as there is interest in the model in the global south.

Bukovska said that despite the progress made with the DSA

and DMA, there is still a long way to go. Enforcement will take time to operationalize and it is a complex system to implement. If it does not work, then the legislation will be just an empty shell. She observed that the GDPR was another revolutionary piece of legislation which was adopted but enforcement is seen as a failure because it was not properly resourced.

Zoe Darmé, Senior Manager of Search, Government Affairs, and Public Policy at Google, mentioned that tech companies have undertaken significant efforts to diversify the people that work on content moderation and privacy issues. Nowadays these topics are not only the purview of lawyers or public policy professionals, “but also of engineers and product managers”. This is incredibly important because the policy intent behind most of the regulations is about driving trust. Public policy concerns are now being integrated more closely into product development done by engineers. For her, that’s a luxury big companies like Google can afford. Subsequently, Darmé expressed concerns regarding the impact of the costs that new bills around the world to regulate platforms can have on new market players, resulting in barriers (algorithm transparency, local presence, turnaround times, out-of-court dispute mechanisms). While Darmé expressed optimism in the preparedness of big tech companies to build products that comply with regulations and that can be trusted, she also highlighted that these regulations are motivated “by certain large platforms” but are not taking into consideration “what comes next”, such as new startups or decentralized networks.

Regulations are motivated “by certain large platforms” but are not taking into consideration “what comes next”, such as new startups or decentralized networks.

Is regulation the solution?

The panel discussed challenges related to the implementation of the proposed regulations and the technical challenges. McLaughlin was not optimistic that the US would “consolidate a set of policy objectives and then sit down with the European Union and figure out how to come up with a harmonized regulatory system that vindicates everybody’s shared goals as best [it] can.” However, he did support US skepticism about some of the pillars of the EU approach. For example, in his opinion, “unbundling has a spotty track record of achieving competitive markets. That is be-

cause in the unbundling world you require if not the consent, you require the cooperation of the incumbent in order to make unbundling work. What we saw across the EU telecom sector was that the incumbents were extraordinarily good at complying in some ways with unbundling requirements but making it meaningless as a practical matter. I think the U.S has some deep skepticism about versions of interoperability that require the cooperation of the incumbents.” McLaughlin sees “adversarial interoperability,” the kind of interoperability that doesn’t require the consent or the cooperation of the incumbents, to be the better approach.

Klonick observed there is a huge pushback in various parts of the world to take back the power over the internet as they feel a form of imperialism through design and infrastructure has been imposed on them silently by tech companies. That pushback is one of the reasons why there are more than 40 different regulations being proposed around the world, which will contribute to the splintering of the internet and pay into the sovereignty concerns of states.

McLaughlin warned that civil society must be involved in the regulation process so there is some public scrutiny to avoid capture of the process by vested interests. If you want to get “there,” I would not start from “here,” he explained that “if you want to make good internet regulation, I would not start by asking a highly concentrated sector and its concentrated critics what regulation should look like because you will get a big second media marketing code.”

Is better technology the solution?

Kate Klonick asked the group, will more tech (like Web3) or increasing reliance on end-to-end encryption give us more solutions or more problems, or both? Bukovska replied that with Web3 there is very little understanding of human rights or freedom of expression, and protecting both of those is not only about negative obligations not to infringe rights. States also have a positive obligation to create an enabling environment for human rights and freedom of expression to flourish. We need states as they play an important role in implementing and enforcing the laws to create a better economy and the necessary structures. She stressed that technology will not solve everything and that Web3 does not operate in a void. There are social and political influences which need to be taken into account for decentralization to work. McLaughlin said that we cannot regu-

late our way out of these problems, nor can we blockchain our way out. We do need more and better tools and a middle path.

***Greetings from UNESCO & The Way Forward
in Multilateral Regulatory Policy Regulatory
Requirements and Necessary Standards - Guilherme
Canela Godoi, Chief of the section of Freedom of
Expression and Safety of Journalists, UNESCO***

Guilherme Canela Godoi, Chief of the section of [Freedom of Expression](#) and [Safety of Journalists](#) at UNESCO highlighted the important role of social media platforms in connecting people while also warning about their potential as vectors for the dissemination of harmful content to democracy, such as hate speech, disinformation, and misinformation. In that sense, Canela asserted that companies have the responsibility of being transparent in the implementation of their policies, technological developments, and review procedures. Canela also said that any regulation of online platforms must be aligned with international standards on human rights and be protective of both people's integrity against serious harm and information as a public good. At the end of his intervention, Canela announced UNESCO's global conference "[Regulating Digital Platforms for Information as a Public Good](#)", which will be held in February 21-23, 2023. The goal of this conference is to develop a model regulatory framework for digital platforms, through consultations with key stakeholders, that protects freedom of expression as a human right.

Any regulation of online platforms must be aligned with international standards on human rights and be protective of both people's integrity against serious harm and information as a public good.

Session III: Business Viability: Decentralizing Power and Opening Up Competition

Farzaneh Badiei, Head of Outreach and Engagement at the [Digital Trust and Safety Partnership](#), moderated the discussion. Badiei stated that the aim of the conversation was not to convince the audience to believe in blockchain or Web3, but rather to discuss how businesses in this new digital ecosystem can help to protect freedom of expression. She started the conversation by asking them about the role of alternative technologies in the current socio-economic landscape.

Michael Lwin, Managing Director & Co-Founder at [Koe Koe Tech](#), claimed there's a lot of warranted skepticism about Web3. Renowned intellectuals like Paul Krugman, for example, have criticized new technologies related to crypto and blockchain. Lwin explained that although there are many scams in the world of Web3, new technologies in this sphere can foster a better environment for freedom of expression and human rights in the Global South. He pointed to the use of crypto wallets in the context of Myanmar's authoritarian regime, where inflation runs amok, there's no bilateral support, and the military junta oversees all banking. Crypto wallets make it difficult for government officials to control or surveil money transactions—thus allowing the funding of democratic resistance movements in the Asian country and strengthening privacy and freedom of association—and enable people in Myanmar to transact with “inflation-resistant USDC stable coins”.

Crypto wallets make it difficult for government officials to control or surveil money transactions—thus allowing the funding of democratic resistance movements in the Asian country and strengthening privacy and freedom of association—

Alison McCauley, Chief Advocacy Officer at [Unfinished Labs](#), reiterated the notion that when it comes to new technologies such as Web3, narratives get very polarized: on one side there are doom-day skeptics and on the other there are the people on “hopium” that believe technology can solve all problems. McCauley advocated for a middle position that reconciles a healthy dose of skepticism with aspirations for technological innovation.

McCauley highlighted the use of protocols to enable a “universal social graph no longer owned by a specific application”—that is a network or user’s database that is universally accessible by design and feeds from other applications or networks that benefit from it— to free social networks from private control. Decentralized blockchain technology is instrumental, as McCauley noted, to this goal. McCauley added that a universal social graph could incentivize meaningful connections around the world —by allowing global discoverability— and encourage a truly global marketplace.

DAY 2 – TUESDAY, OCTOBER 4, 2022

Opening Remarks & Presentation of the Second Day

Catalina Botero, Consulting Director of [Columbia Global Freedom of Expression](#) and Co-Chair of the [Oversight Board of Meta](#)

Day 2 of the conference began with the opening remarks provided by **Catalina Botero**, Consulting Director of Columbia Global Freedom of Expression and Co-Chair of the Oversight Board of Meta. Botero explained that the sessions on the second day would focus on understanding “the different regional approaches to platform regulation, in particular to content moderation”. She noted that, currently, there’s little to no consensus about basic and essential questions about regulatory strategies on content moderation. For the Co-Chair of the Oversight Board, there seems to be only some degree of agreement “that platforms should have special duties when it comes to transparency”.

To frame the discussions to come during the rest of the day, Botero posed four essential questions that are still the subject of debate. The first question asks what responsibilities and rights (or privileges) platforms should have. For Botero, different judicial decisions issued in countries like Germany, Argentina, and Colombia provide an appropriate starting point for the development of better regulatory solutions because they understand the novel nature of internet platforms. The second question asked by Botero concerned what law should regulate content moderation. In the context of rising authoritarian governments worldwide, she considered International Human Rights law provides an appropriate legal framework to protect the internet’s architecture, the rights of users, and freedom of expression.

In the context of rising authoritarian governments worldwide, she considered International Human Rights law provides an appropriate legal framework to protect the internet’s architecture, the rights of users, and freedom of expression.

Botero then proceeded to ask who should regulate platforms



and provided examples of different regulatory models: self-regulation, independent regulation, judicial regulation, etc. The final question Botero made, questioned whether platform regulation should be focused on design decisions and the systems of implementation of their policies or concern exclusively “the *ex-post* process in which decisions are made regarding whether specific content is removed or kept”.

Botero mentioned she believed a strong regulatory wave was approaching and expressed her hope about the influence of the Oversight Board’s doctrine on new regulatory bodies.

Session IV, Part 1: Regulating Online Speech: Between the First Amendment and the Digital Services Act

Pamela San Martin, Board Member of the [Oversight Board of Meta](#) moderated the first panel of the second day. She prefaced the conversation by mentioning that the Digital Services Act (DSA) had just been approved by the European Council and expressed her feelings of uncertainty regarding the reach and implementation of this legal framework. To further contextualize the discussion San Martin also wondered about to what extent or degree the DSA would coexist with existing regulations in Europe and influence or model legal frameworks in other parts of the world. San Martin also problematized the DSA's aim of protecting the digital space from illegal content, considering that such category varies depending on the applicable norms: International human rights standards, local laws, terms of service, platform's community standards, among others.

Agustina Del Campo, director of the [Centro de Estudios en Libertad de Expresión y Acceso a la Información \(CELE\)](#), began her intervention by questioning whether there's a consensus regarding the issues or problems "we are seeking solutions for when trying to regulate social media". Del Campo mentioned that until very recently regulatory proposals and legislation regarding content moderation were "content oriented", but recently there's been a shift in which the focus is now on transparency and the nature of the processes and implementation systems used to regulate content. The director of CELE considered that the need for transparency was an issue in which there was a general social agreement when talking about content moderation regulation. Nonetheless, she mentioned that there's no consensus about what we need that transparency for or what kind of transparency we are looking for. Del Campo pointed out three common issues she found in her research work on comparative legislation about content moderation: 1. Most of the regulation is not content neutral; 2. There's a *harms* approach underlying regulation that ignores legal categories of speech and is focused on the harm; and, 3. There are no oversight processes or clear guidelines regarding the implementation of these new legal

There's no consensus about what we need that transparency for or what kind of transparency we are looking for.

frameworks. Del Campo also highlighted the regional Latin American standards of protection to freedom of expression—where prior censorship is forbidden and the discussion regarding content moderation has been advanced by domestic courts—and extended an invitation for regulators and stakeholders to understand its nuances when thinking about content moderation regulations.

Matthias C. Kettemann, professor of Innovation at the [University of Innsbruck](#) and Research Program Leader at [Leibniz-Institute for Media Research | Hans-Bredow-Institut](#), started his intervention by underscoring that he didn't believe there is a unique "European approach" to content moderation. Regulations regarding content moderation in Europe, for professor Kettemann, are a co-creation that draws from multiple sources, especially from global principles. He noted that over the last few years tremendous changes have occurred in the online environment. For example, platforms are suddenly more engaged in content governance about public elections, misinformation regarding health-related matters and armed conflicts in the world. Kettemann underscored what he perceived to be a value-oriented approach in European legislation about content moderation. As such, transparency is key for civil society to understand whether social platforms abide by the Rule of Law and act in a non-discriminatory manner. For Kettemann it was important to keep in mind that transparency should fulfill an essential mission: "namely to ensure that we know what platforms are doing". Kettemann also highlighted that self-regulation will continue to play a significant role in the context of the DSA which serves the purpose of granting autonomy to social platforms and fostering innovation.

Self-regulation will continue to play a significant role in the context of the DSA which serves the purpose of granting autonomy to social platforms and fostering innovation.

David Kaye, Clinical Professor of Law at the [University of California](#), Irvine, and former [UN Special Rapporteur on the promotion and protection of the right to freedom of](#)

[opinion and expression](#), highlighted two areas of innovation present in the DSA. Kaye mentioned that the transparency obligations —included in the DSA and that social platforms must follow— are very interesting. For example, Kaye noted that transparency towards users is especially relevant since it encourages companies to provide statements of the reasons for specific content actions. David Kaye also talked about risk assessment as another point of innovation of the DSA. Risk assessment, as described by Kaye, will require annual independent audits on platforms and incentivize an environment in which platforms will have to implement mitigation policies in response to systemic risks. Kaye showed concern as to whether the DSA could push “a divergence between European human rights and emerging global norms for freedom of expression” and wondered about the impact and influence the DSA —a regulatory initiative adopted by global leading democracies— could have on regulatory frameworks in other parts of the world. Kaye also reflected upon how the DSA could affect regulations to come in America and other parts of the world.

Session IV, Part 2: Regulating Online Speech: Between the First Amendment and the Digital Services Act

Jamal Greene, [Dwight Professor of Law at Columbia Law School](#) and Co-Chair of the [Oversight Board of Meta](#), mentioned that this session would be a continuation of the conversation about the DSA held in the previous panel, regarding the varying responses to online harm and the different modes of regulation —from self-regulation to government regulation. To start the discussion, Greene posed questions to the panelists about the convenience of finding an optimal regulatory model vs. allowing different models to coexist and complement each other, and invited the speakers to reflect on the impact of regulatory models for content moderation within the constitutional framework of the United States in which freedom of expression grants protection mainly from the State.

Founder and Executive Director of [Justitia](#) and the [Future of Free Speech Project](#), **Jacob Mchangama**, argued against the preconceived notion that social platforms don’t implement measures to combat hate speech. The research he presented, for example, showed that out of 2,400 comments that Facebook flagged



as hateful, only 0.46% “would actually violate the Danish Criminal Code”. Mchangama also stated that the internal policies of social platforms, such as Facebook and YouTube, —against hate speech— were more expansive than legal instruments found in the Danish legislation or the ICCPR since the list of protected characteristics is larger. In light of this, the Director of Justitia expressed concern regarding social platform regulations such as the DSA, since he considered that state intervention could be counter-productive to the promotion of freedom of expression, especially in Europe where Article 10 of the ECHR and Article 11 of the EU Charter provide less protection to freedom of expression than other international human rights instruments, such as the ICCPR. Mchangama also noted that his concern and skepticism is shared by users at large, whom —according to a recent survey— prefer content regulation to be carried out by platforms rather than governments.

Suzanne Nossel, Chief Executive Officer at [PEN America](#), asserted that there are three forms of social platforms’ regulations: traditional or top-down, self-regulation, and co-regulation —as seen on the DSA. For Nossel, none is optimal. Traditional regulation fails to understand the distinct and unique nature of social media platforms and seems to equate them with

State intervention could be counter-productive to the promotion of freedom of expression, especially in Europe where Article 10 of the ECHR and Article 11 of the EU Charter provide less protection to freedom of expression than other international human rights instruments, such as the ICCPR.

phone companies. Self-regulation—which Nossel believes is often synonymous with non-regulation—has paradoxically, proven to work: how platforms voluntarily dealt with misinformation and disinformation during the Covid-19 pandemic, for example, was deemed as effective by Nossel. Nevertheless, as noted by Nossel, self-regulation is problematic since it exists only as long as companies see “a business interest in doing so”. Nossel mentioned the Oversight Board as an interesting experiment in self-regulation in which both the potential and limitations of this form of content moderation can be seen: On one hand the Board has been able to address many issues concerning Meta’s lack of transparency, while in other aspects, users and experts still crave a regulatory body that can put its thumb on the scale in other matters outside the Board’s current mandate.

When talking about co-regulation, Nossel referred mainly to the Digital Services Act. She described the DSA as an ambitious and interesting experiment in which transparency obligations and the participation of civil society are promising, but whose results can only be analyzed and understood in the long run.

Martin Fertmann, Researcher at the [Leibniz-Institute for Media Research | Hans-Bredow-Institut](#), focused his intervention on what he considered novel aspects of the DSA: Statements of reasons, internal complaint handling systems, and out-of-court dispute settlements. As described by Fertmann, the statements of reasons provision of the DSA demand that every single content moderation decision—like demonetization and deletion or demotion of content—must be accompanied by an explanation to the user provided by the platform. Fertmann considered that the DSA also imposes obligations on social platforms regarding the need for effective internal complaint-handling systems to the extent that violations of due process requirements can trigger damages in favor of affected users. Fertmann highlighted that under the DSA class action suits can be lodged to make these kind of claims, which then begs the question to what extent this mechanism can be used to address systematic failures.

Lastly, Fertmann mentioned that under the DSA, social platforms that meet certain thresholds (more than 50 employees and 10 million euros turnover or more) “have to engage in good faith with an out-of-court dispute settlement mechanism”, whose decisions are

not binding. For Fertmann this is a weird compromise that raises questions about how any social platform could engage in good faith with an out-of-court dispute settlement mechanism while continuing to ignore its decisions. Similarly, Fertmann criticized the fact that it is likely that the decisions issued by these dispute settlement mechanisms, in their case-by-case analysis, will not serve to reach or further systemic changes in platforms' policies.

Session V: International Human Rights Law as the Basic Framework of Meta's Oversight Board Decisions

Co-hosted in partnership with the Oversight Board of Meta

Joel Simon, a Fellow at the [Tow Center for Digital Journalism at Columbia](#) Journalism School, moderated this conversation. To frame the discussion, Simon referred to the lack of consensus regarding the purposes of content moderation regulations and the challenges they pose to freedom of expression considering the rise of authoritarian leaders in countries such as the United States and Brazil. He also asked questions to specific panelists regarding the impact and effect of the decisions issued by Meta's Oversight Board: Whether they could be regarded as some form of jurisprudence and their influence on Facebook's policies on content moderation.

Joan Barata, [Intermediary Liability Fellow at the Stanford Cyber Policy Center](#), presented his recently published paper titled [The Decisions of the Oversight Board from the Perspective of International Human Rights Law](#). Barata and the other members of the panel commented on the paper's findings. He described the role of the Oversight Board as a private regulatory body with the mandate to review content moderation decisions on Meta and make policy recommendations to the company. Barata also highlighted the Board's role in the disclosure of the content moderation criteria used by Facebook and Instagram, when it compels these companies to provide detailed explanations about what triggered certain content moderation decisions. Barata further stressed that the Board has the obligation to assess the impact of removing content based on standards established by international human rights

law which protects freedom of expression. Barata said that although the Oversight Board profusely cites universal human rights standards when issuing decisions, it would be pertinent in terms of legitimacy if the Board could also refer to regional standards, especially those related to the region affected by the specific case. To conclude, Barata pointed out that the Board should be able to make an assessment, in the near future, of the protocols that guide content moderation and also considered “there’s some work to do when it comes to assessing freedom of expression” against rights such as privacy and dignity.

The Director of the Program on [Platform Regulation at the Stanford Cyber Policy Center](#), **Daphne Keller**, underscored the positive role of the Oversight Board and the credibility and expertise of its members. Nonetheless, she considered that this private body shouldn’t assume the role of being the sole or main arbiter regarding content moderation under the DSA in Europe, in part due to her commitment to decentralization and the need for diverse arbiters. Keller also described the tensions that would arise if an organization such as the Oversight Board—with a global scope and whose decisions are based on global human rights standards—were to become an arbiter under European human rights law. Namely, she expressed concern over the normative differences between regional and international instruments that protect human rights, and the dangers of a “monoculture” regarding freedom of expression and content moderation.

Although the Oversight Board profusely cites universal human rights standards when issuing decisions, it would be pertinent in terms of legitimacy if the Board could also refer to regional standards, especially those related to the region affected by the specific case.

Founder and Director of the [Dangerous Speech Project](#), **Susan Benesch**, drew attention to the limited capacity of the Board—which has issued 29 decisions in the last two years—to moderate content on a large platform, such as Meta, which takes down more than one million pieces of content daily in the process of enforcing its community policies. Although the Board, Benesch said, makes a thorough process of selecting its cases to establish guiding principles in their decisions, this is simply not enough to answer many questions still lingering regarding “equity, inclusion, and discrim-

ination”. To cite an example, Benesch asked if it was possible to know if there’s discriminatory treatment regarding content posted by black people, women, and Muslims in comparison to hegemonic groups in society. Benesch affirmed that Facebook is the world’s largest speech moderator, even surpassing the regulatory power of states like China, and as such demanding answers to these questions was essential for the enjoyment of freedom of expression. She also questioned if, under international human rights standards as applied by the Oversight Board, Facebook could issue regulatory content moderation decisions that limit freedom of expression based on the protection of national security or public morality, and whether that would be convenient.

Kate Klonick, Associate Professor of Law at [St John’s University](#), described how the Oversight Board’s decisions impact Meta’s content moderation process. According to Klonick, there’s a team at Meta in charge of implementing the decisions and policy recommendations issued by the Board and reporting back to the Board on the course they will take. For its part, the Board has a group in charge of “evaluating the effectiveness and implementation” of its recommendations to Meta. This back-and-forth process, Klonick noted, has fostered transparency and accountability within the company. Klonick also mentioned that the idea that the Oversight Board can only provide a glimpse of all the content moderation at Meta is unfair when taking into account the number of interactions that happen in daily life that never go to court or are decided by the highest court of a given jurisdiction.

Monroe Price, retired Adjunct Full Professor at the Annenberg School for Communication and the Joseph and Sadie Danciger Professor of Law and Director of the Howard M. Squadron Program in Law, Media and Society at the Cardozo School of Law, asked if there were alternative frameworks for the protection of



Rachel Wolbers
@rachelwolbers

It's not very often we see five @OversightBoard Members together in one room! Thanks @ColumbiaGFoE for putting together such a great conference! #OnlinePublicSphere2022

@SuzanneNossel @jamalgreene @SnMartinPamela @cboteromario @EvelynAswad



11:44 AM · Oct 4, 2022

freedom of expression, other than international human rights law, that the Board could adopt. He also questioned whether Meta's search for legitimacy—a political economy—could have determined or shaped the establishment or operation of the Oversight Board. To him, Facebook's adoption of international human rights law is a decision in which the status of Facebook in society and its credibility played a big role since, according to Price, Facebook sought to confer the Oversight Board legitimacy where States had given none.

Monroe Price was joined in his intervention by **Joshua Price**, a professor at the [Toronto Metropolitan University](#), who argued that the Oversight Board's support by "a broad consultative network" legitimizes it. For Joshua Price, decisions issued by courts or agencies gain legitimacy as long as they are able to dialogue with significant constituencies and stakeholders. Structured efforts undertaken by the Oversight Board, such as including comments from the public in its decision-making process, are evidence of dialogue and the construction of an epistemic community.

Session VI: El derecho internacional de los derechos humanos como marco de las decisiones del Consejo de Supervisión de Meta (in Spanish)

Co-hosted in partnership with the Oversight Board of Meta

Joan Barata, [Intermediary Liability Fellow at the Stanford Cyber Policy Center](#), moderated the last panel of the conference. To frame the discussion, Barata underscored the importance of discussing the work of the Oversight Board in Spanish, considering the importance of the Board's decisions for Latin America. Following this, he described the nature of the Board as an independent body in charge of assessing Meta's content moderation decisions in light of International Human Rights Law. Barata highlighted that the Board's decisions go beyond content moderation (to remove or keep content) as they can also suggest policy recommendations or set out criteria for the interpretation of Meta's Community Standards.

Susan Benesch, Founder and Director of the [Dangerous Speech Project](#), stressed the important role of oversight bodies in the private sector to protect users and the general population. Benesch reiterated several of the ideas she mentioned during the previous session in English. She expressed her intention to convince the audience of the importance of understanding if, or to what extent, Meta could be discriminating against marginalized groups through its content moderation decisions. Benesch suggested the creation of a new regulatory body, in charge of overseeing this type of issue, composed of engineers and computer scientists.

Benesch questioned again whether Facebook should be entitled to limit freedom of expression, under Article 19 of the ICCPR, to protect national security, public morality, or the reputation of others. For her, this debate needs to be further developed when discussing private companies.

We live in a context of ever-increasing demands for social platforms to undertake more tasks and responsibilities. The task of localizing international public law instruments to the corporative world, by understanding its nuances and differences, is a very important contribution, from the Oversight Board to the regulatory landscape.

Director of [Centro de Estudios en Libertad de Expresión y Acceso a la Información \(CELE\)](#), **Agustina del Campo**, focused her intervention on two main points: The emergence of the Oversight Board and the challenges to come for the online ecosystem in relation to content moderation. Regarding the first aspect, del Campo mentioned that we live in a context of ever-increasing demands for social platforms to undertake more tasks and responsibilities. In that sense, the Oversight Board was created in an effort by Meta to legitimize its operations and to receive feedback on the way the company was intervening in the public debate. Del Campo believes that the task of localizing international public law instruments to the corporative world, by understanding its nuances and differences, is a very important contribution, from the Oversight Board to the regulatory landscape. Del Campo then pointed out as a challenge the way the Oversight Board fits within what should be a wider regulatory system since it is necessary that other stakeholders are able to “look” at things the Board can’t while it’s focusing on solving specific cases.

Carlos Cortés Castillo, Colombian journalist and lawyer and Co-Founder of [Linterna Verde](#), underscored the importance of the existence of the Oversight Board as a relevant experiment that has allowed researchers to gather concrete evidence about content moderation problems. That said, he expressed concern about how the Oversight Board can, with its design limitations, keep up enforcing content moderation in light of the ever shifting and growing array of new products offered by social platforms. Cortés described, for example, that most content on Instagram, recently, consists of *stories*. Even if the Oversight Board, Cortés argued, included within its scope this type of ephemeral post, the point of restoring or removing content after a specific juncture has already elapsed would be innocuous.



**THE
FUTURE
OF
FREE
SPEECH**



Global Freedom of Expression
COLUMBIA UNIVERSITY



JUSTITIA