Regulating the online sphere

Jacob: Did you know…

Flora: That for the past two decades, various laws around the world, most notably Section 230 of the US' Communications Decency Act, and the EU's e-Commerce Directive have been in place to mostly shield digital intermediaries from liability for illegal content on that platform? Such laws also mean that companies have not been legally obliged to monitor their platforms for infringing content.

Jacob: However, did you know that in response to mounting concern over the years about the misuse of the internet and the exploitation of online platforms for criminal purposes, this has begun to change?

Flora: And that terrorist content is one of the sectoral arenas where much in this battle is currently being waged?

Jacob: This is Tech Against Terrorism. I'm Jacob Berntsson.

Flora: And I'm Flora Deverell. In recent years, governments around the world have moved to create laws and regulatory frameworks that seek to combat illegal or even harmful speech online, and to hold digital platforms accountable for what is on their platform. In 2017, Germany notably introduced the Netzwerkdurchsetzungsgesetz or NetzDG law which stated that platforms with more than two million users have to remove illegal content within a certain timeframe. Once notified, will be heavily fined.

Jacob: Since then, various governments in Europe and around the world, including the UK and Australia, have launched their own proposals geared towards regulating the online sphere. Of course, regulatory approaches vary, with some focusing on setting performance targets for content removal, others holding platforms accountable for having certain procedures and processes in place, and yet others requiring companies to restrict certain forms of speech, even if the speech is not illegal.

Flora: Terrorist content is often a key focus for this regulatory attention. In 2018, the EU published a draft proposal for an EU-wide regulation on preventing the dissemination of terrorist content online. The regulation remains in the negotiation stage between various EU bodies at this point, however, we know that amongst other tenets it will likely introduce a responsibility for proactive measures, as well as a one hour deadline for companies to remove terrorist content following removal orders from member states.

Jacob: Such laws, including NetsDG and the EU's proposed regulation on online terrorist content have earned both praise and criticism. Many welcomed the laws as well needed responses to the perceived impunity with which tech platforms have, according to some, inadequately dealt with illegal content online, and the associated danger for our societies and democracies. Critics, particularly those amongst civil society and human rights groups, have raised concerns that they may lead to draconian censorship of the online sphere, with worry and consequences for fundamental rights, such as freedom of expression. Regardless, it is clear that what happens in the online counter terrorism legislation space, we'll see effects that go beyond just online terrorist content, and it's likely to influence online speech more widely.

Flora: Such state led online regulation spells the end for the pure self regulation approach that tech companies have been pushing for years. Whilst all of this has been going on, tech companies have continued to innovate their approaches to self regulation, endeavoring to prove that they are taking concerns about ethics and accountability seriously. A prominent example is the introduction of Facebook's oversight board, announced in 2018 and set to become operational this year. This board is designed to be an independent body that will oversee questions of what content should or should not be on Facebook and Instagram. Another example of self regulation with regards to counter terrorism is the Global Internet Forum to Counter Terrorism, a coalition founded by Facebook, Twitter, Google, and Microsoft in 2017. The GIFCT has, together with us, Tech Against Terrorism, worked towards improving industry wide responses to counter terrorism through a mix of technical and policy oriented capacity building.

Jacob: In many ways, regulation of the online sphere is coming to a head. Different initiatives continue to be deployed by different sectors, including that of state led regulation and industry led initiatives, while especially mandated law enforcement bodies and internet referral units, or IRUs, submit referrals for content take down to tech companies under their own terms of service. It remains somewhat unclear at this stage how this all fits together.

Flora: Nevertheless, to help us think through the components and impact of these approaches, we are joined by two very special guests, Evelyn Douek and Daphne Keller. Evelyn is a lecturer on law and S.J.D. Candidate at Harvard Law School, and Affiliate at the Berkman Klein Center For Internet & Society, studying international and transnational regulation of online speech. This is a topic that

Evelyn, and Daphne too, frequently writes on for various publications. Daphne is director of platform regulation at Stanford Cyber Policy Center. In the past, Daphne has been the director of Intermediary Liability at Stanford's Center for Internet and Society, as well as Associate General Counsel for Google. She has worked on groundbreaking intermediary liability litigation and legislation around the world. Evelyn, Daphne it is an honor to have you here. Many thanks for joining us.

Daphne: Thanks so much for having us.

Evelyn: Yeah, thank you.

Flora: So first of all, as a means of introduction to this debate, it's noteworthy that neither of you are counter terrorism researchers per se, however both of you, as well as other academics researching online speech, have written extensively about policies and regulations with regard to online terrorist content. I was wondering if you could elaborate a little bit about why you think that terrorist content is such a key focus of a lot of these regulatory debates, or what impact could what happens in this space have on online speech and the internet as a whole. Daphne, would you like to start?

Daphne: Sure. Well, I mean, I think the reason why the threat of terrorism is such a driver in this space is fairly self-evident. I mean that it is a tremendously dangerous and frightening thing for society is that they want to respond to with laws that can cut down on the risk of real world violence. And so, because we have seen so many horrible episodes of real world violence coming from extremists over the past years, it's not surprising that that has been a huge pressure point in the law regulating content on platforms.

Flora: Absolutely. And Evelyn?

Evelyn: Yeah, just to echo all of that, obviously it's been an immense area of public and regulatory focus, in particular because of a very high profile issues. And I guess one of the reasons why I've written about it specifically in my work, and thank you for noting upfront that I'm not a counter terrorism researcher, I think that's important for the rest of this conversation, is that it is an area where it's one of the early spaces where platforms really were pressured to step up and take more responsibility for the content that appeared on their websites, and then sort of that then becomes a model, or an example when we start talking about what they should do in other areas.

Daphne: And I would add it's been interesting to see, at least in the part of the debate that I've seen around the terrorist content regulation, just how little the work of real security researchers is being cited. I think there is an easy assumption that we make that if horrific content weren't online, or if recruitment content weren't online, that would correlate to a decline in real world violence, but from what I was able to find digging into the work by real security researchers, which again, that's not me, there's actually a whole lot of debate about this, and a lot

of questions about whether, by taking things off of mainstream platforms, we drive people more into echo chambers, or make it harder for law enforcement to see what's going on and go after offline crime and violence. And so there's a really important evidentiary piece here that I think can get glossed over too easily in the rush to get rid of violent offensive and often, frankly, illegal content online, the connection to real world violence remains really complicated.

Jacob: Yeah, I think that's a really good point, and that's also something that we stress at Tech Against Terrorism often, that there is a debate over this sort of causality between viewing content and committing a violent act of terror. And also very happy that you mentioned the notion of content removal and what sort of unintended consequences that might have in terms of displacing the threats across to other parts of the tech ecosystem. But moving on, so obviously a number of national laws aiming to curb the spread of, amongst other things, online terrorist content are already in place, for example in Germany and Australia. However, one law or regulation that will certainly have particular impact is the proposed regulation on online terrorist content currently prepared by the EU. Daphne, if I could just ask you to sort of explain exactly what this law is about and what potential impact do you see it having?

Daphne: Sure. So this law, so far, exists in three drafts from the Commission, Council, and Parliament, and the EU. And it is in trialogue negotiations right now to reconcile the three. So there's some difference between the two, and I'm kind of going to blur them together, but it's worth noting, in particular, the Parliament draft really, in my mind, is much better than the others and has a lot of important protections in it. Some key questions in the law are, first of all, what the definition of prohibited terrorist content is going to be, obviously that's very fraught, and also how European law and platforms should handle conflict in those definitions from one EU member state to another, from Spain versus Hungary, for example.

Daphne: Then the parts that I think are more squarely on point for the work that Evelyn and I do have to do with empowering national authorities to tell platforms to take things down. And these national authorities might, in the sort of worst case scenario, or least civil liberties protective scenario, they might just be local police. They are not probably going to be courts. They're going to be something faster and with less process than that. And they can do two things. One, they can decide that content is illegal and order a platform to take it down, which might have to happen as quickly as in one hour, even for very small platforms and platforms that don't have an EU presence. They can also refer content for removal without deciding that it violates the law on the basis, instead, that it violates a platform terms of service. And this is very similar to what internet referral units do now. I know we'll talk about that in a minute.

Daphne: So in that case, they're sort of a government actor, police or whoever it turns out to be, applying not the law, but instead standards created by platforms to tell them to take down content that might actually be legal. We've got some troubling examples, for example, of law enforcement telling the internet archive

to take down a bunch of things that turned out actually to be parodies, or actually completely innocuous material in some cases. And then finally, once authority has issued an order to take something down, it can follow up by telling a platform that it has to build a proactive monitoring effort, so it has to apply a filter to keep contents down in the future. And that is the piece that has been the focus of a particularly large amount of opposition from civil society in Europe. It's been critiqued by three different human rights rapper tours from the United Nations. There has been a lot of concerns about mandatory filtering leading to disparate impact for different groups based on their nationality or religion. So there's been a ton of focus around the filtering piece.

Evelyn: Drawing on Daphne's previous writing in this space, and the work of many others, Danny O. and Seth Kreimer and things, the problem with that kind of scenario is obviously, if you're a platform, the incentive structure that it creates is when you get this kind of notice or pressure from a government, your incentive is to take it down. And so the real concern there is that the individual speech interests that are at stake aren't going to get a proper hearing.

Jacob: Yeah, that's really good. And I mean one concern that we have had is, for example, the one hour removal deadline, given that a lot of the smaller companies that we work with might struggle with that, given that they might be run by one person, so what happens when they're asleep? Literally, that could be sort of a very practical challenge for this specific piece of legislation. Daphne, you mentioned internet referral units earlier. What are they, and what impact do they have in this space?

Daphne: These are pre existing entities, and honestly you might know more about them than I do, so please correct me if I'm missing anything here, but they exist within, for example, the London Metropolitan Police, and within Europol. They're basically employees of law enforcement whose job is to look for online content and then request platforms to take it down, again, not based on it violating the law, but based on it violating the platform's terms of service. And there's some really interesting legal questions, what in the US would be constitutional questions, about that mechanism.

Daphne: For example, in the UK, there was a human rights audit of an analogous system for child sexual abuse material, that Internet Watch Foundation effort, where the human rights auditor said, "Look, if police want content taken down, they have to go through a court under UK law. But by the police referring something in a way that's not based on law and that kind of passes it through a private actor based on some other standard it's avoiding a legal process that would otherwise apply."

Daphne: So there's been a lot of concern from civil society about these internet referral units for that reason. And also just more broadly, because you can think of it as a big systemic shift in authority to regulate speech from saying it's democracies and governments and elected officials and courts that decide what's illegal, which is sort of how things have been for a very long time under the rule of law

to saying, "Well, no, that stuff doesn't matter so much. The rules that matter are the ones created by private platforms and we're going to deploy government officials to enforce those rules."

Evelyn: Yeah. And just something that I want to emphasize about sort of the pieces that we're talking about here is we can talk about these sort of regulations or initiatives in isolation, but it's important to emphasize how they are sort of all interrelated. Jacob, you mentioned the Australian regulation earlier, and then these preexisting examples of the IRUs and the child sexual abuse material databases. They all sort of are informing the conversations that we're having now about what to do next and often are held up as examples of what to do next. So that kind of history, it really is more of a trend rather than just sort of discussing the current EU regulation in isolation.

Flora: And just building on what both of you said, I'm particularly interested in something I saw you say, Daphne, which I thought described it quite well as this being a Faustian bargain. We asked platforms to take down that, which is so-called awful but lawful, instead of developing speech laws. And that means that we at the same time are relinquishing our constraints on platform' power in the same way that we might have rights-based constraint of a governance power. And I was wondering if you could, I mean, you've touched on it earlier, but if you could explain what you meant by the rule of terms of service, which I think you were talking about in relation to the proposed EU regulation. So what do you mean by the rule of TOS and why is this in your view a reason for concern?

Daphne: Sure. Well, I mean, the starting point is that platforms, and I'm thinking of big popular platforms like YouTube or Facebook here, but platforms use their terms of service to prohibit a lot of speech that is protected by law, certainly in the US with its broad first amendment protections, but really many places in the world. And that's because that's what their users want for the most part. Most people don't want to go onto a platform and be confronted by barely legal images of violence or threats or bullying or racism. There's a real commercial reason and consumer demand reason that platforms do this, but that also turns the terms of service into this incredibly powerful mechanism for deciding what speech can be online and makes it easy for governments to say, "Well, let's just leverage that," which we've seen in the case of IOUs.

Daphne: And I don't know how we get out of that bargain really, because I don't think very many people would seriously propose that every platform should have to carry every legal piece of speech that would turn them into cesspools and drive users away. And so I think it's unavoidable that there's private power being used to constrain lawful speech. And we just need to really think harder about how that intersects with government power, how that intersects with competition and consumer choice, what role the organizations that Evelyn has written about as content cartels play in this, what value transparency might have in correcting for this, because it's just, it's very complicated and there's not a single clear way forward.

Jacob: We're sort of leading up to this, the question of self-regulation and the sort of standards that companies set for themselves in this space. So I think for many people, it's very clear what government-led regulation is in this space, but what exactly does self-regulation mean in regards to regulating online terrorist content?

Evelyn: Yeah, so as, as Daphne was just saying, often platforms' terms of service really are where a lot of this occurs and a lot of the regulation and most important decisions get made about what stays up and comes down. So over the years, platforms have been developing ever more intricate and detailed, well, the major platforms, at least intricate and detailed rule books around what they will take down and leave up. And then also expanding slowly on the transparency measures around how they enforce those rules. But really it is still very, very opaque. And there's sort of not a lot of visibility into exactly what's occurring or the accuracy of the figures that they release and things like that. So that really is the big area for improvement in self-regulation at the moment.

Flora: So as we're recording this on the 13th of May, Evelyn, last week, there was big news for those of us monitoring this debate. And that was that Facebook announced its 21st members for its new oversight board. Evelyn and Daphne as well, you've written extensively on this board based on the charters that Facebook has published on to this date. Evelyn, could you tell us a little bit more to start off with about this oversight board, what it is, what impact it will have and how significant it is that a company like Facebook is pursuing this model in this overarching regulatory debate?

Evelyn: Yeah, absolutely. You're right. It was an exciting moment as someone that's been tracking it for the last 18 months or two years. When I was sort of starting to focus on this, I heard a story about physicists that were doing dissertations on the Large Hadron Collider and then their PhDs were put on hold when the thing broke down for two years. And over the past two years, I've kind of felt like that with the oversight board. So it was exciting to see it finally sort of having faces and names and things and look like a real institution.

Evelyn: So the idea, I mean, characterizing what this thing is, is tricky. It sort of looks like a court and the idea will be that it will hear appeals from Facebook's content moderation decisions and can overrule them, but it can also issue policy recommendations. It won't have binding authority except in the individual case before it. And so, the question obviously is going to be how much more broadly will Facebook implement that across the platform? But really, the focus of this thing is exactly what Daphne was talking about earlier is that terms of service are becoming so important in this space and in many cases, much more important than government rules and laws. And so we really have this problem where these extremely important speech regulations are being promulgated by these private unaccountable companies and probably governed by business imperatives rather than any sort of democratic or legitimate or rights-based incentives or values.

Evelyn: And so the hope is that this body, by checking Facebook's decisions, will be able to inject some of those values back into the way that content moderation occurs on Facebook or at the very least provide more transparency.

Flora: Yeah. And, but it's interesting because the remit once thought to be broader, will only focus on content removed, I think I'm correct in saying. And obviously the oversight board will only be able to get to so many cases per year and it will be individual pieces of content, even though it will have parallels of course, across the whole platform potentially. But what do you think about this focus in terms of the remit?

Evelyn: Yeah, so that really is the biggest disappointment to me over the past 18 months of the board's development. It's jurisdiction or its sort of mandate has been steadily narrowed over time. So we were promised this supreme court of Facebook and then sort of over time, it's gotten smaller and smaller and now it's maybe looking more like a local district court or something along those lines or the department of motor vehicles. So I am really hopeful that if we keep the pressure up that Facebook will eventually expand the scope. They tell me that it's coming. They've publicly said that they do want to expand the scope to especially leave up decisions because often those are some of the most controversial.

Evelyn: So that's yet to be seen, but I really do think that it's important if this experiment is going to work, that it has input and oversight over really the most important content moderation decisions that Facebook makes and not just some matters off to the periphery of the core business model. I should just say, I should just clarify all of those matters could come before the board from startup as long as Facebook decides to refer it to the board, but that's obviously not a really strong model of oversight if it is up to Facebook when leave-up decisions come before the board.

Flora: Do you have anything that you would like to highlight on this topic, Daphne?

Daphne: Well, I would say first that the super duper experts on the board are Evelyn. We're so lucky to have her here to talk about that and US academic named Kate Clinic and they just know it inside and out and in a way that the rest of us can't compare to. But I've thought about it a lot sort of systemically the role that it plays. I wrote in the Atlantic about how this is not really a supreme court. It's not a real constraint. It's not the kind of democratic or rights-based limit on power that we think of when you hear phrases like supreme court and that's part of the fallacy and bargain. At the same time, I'm excited about it. This is a bold experiment and it's an attempt to solve this problem where nobody really wants Facebook making these hard calls about our speech, but nobody really wants governments doing it either. even if we could reconcile what all the different governments want.

Daphne: And so it's an attempt at a different way forward. And I do think that Facebook has incentive to make the board pretty independent, to make it not too powerful, not give it too many things it can decide, but make it pretty independent for the things it does decide. Because for Facebook, this is a wonderful opportunity to finally pass the buck and take the hard decisions where everybody's going to be mad no matter what they do. Pass them to a third party and then say, "Don't blame us. The board decides." So it will be really interesting to see where this goes, even though its impact is not nearly as broad as I think a lot of people expect.

Evelyn: As someone that has been in the hot seat and worked on these issues from the inside, I wonder whether you would have really appreciated having something like the oversight board there to sort of offload some of the more difficult questions to.

Daphne: Sure. I mean, as long as it's a question that's entirely discretionary to the platform. It gets much more complicated for questions where there's some countries saying, Russia saying, "Hey, our law requires you to take down this pro LGBTQ content." And in the bylaws for the board, my understanding is Facebook says, "Well, if there's a legal obligation, then the board doesn't get to decide." So there's this whole category of really difficult human rights questions, where there actually is a government claiming there's a reason something has to come down. And if Facebook interprets those as legal questions, rather than terms of service questions, then they don't go to the board.

Daphne: There's also, Nicole Wong raised this on the Lawfare podcast, which Evelyn and I were both on last week. The really interesting cases will be ones where Facebook's business imperative conflicts with what the board wants them to do. Maybe Facebook wants to go into Malaysia as a market. I'm not sure if they're already in Malaysia, and to appease the government or to be popular with users there, they'd like to take something down. So there's money on the line in setting policy a certain way, but the board tells them not to. Those will be the really interesting decisions where there's more of a potential actual conflict between Facebook and the board.

Evelyn: Yeah, I'm actually really glad that you emphasize that first point as well about the fact that the board explicitly will not be considering content that Facebook is taking down under local law, because that's obviously going to be relevant to the audience of this podcast. Because for many terrorist content questions, that's going to be a legal question. The oversight board won't have oversight over that, and it certainly also won't have oversight over Facebook's determination that it is a legal question. That also itself is not necessarily an easy call. We saw this in the Soleimani... when there was the assassination, there was this dispute around Instagram's decision to take down posts expressing support for the assassination, which it originally said it had done under US sanction laws. Then they sort of reneged and backtracked on that. Those kinds of questions are not things that the oversight board will be talking

about. So there's still a lot of accountability and transparency work that needs to be done in this space. It's just not what the oversight board is doing.

Daphne: That makes me think of another interesting intersection with law, which is, as a platform, you can be in a situation where there's some government saying, "Hey, I have jurisdiction over you and my law says you have to take this down," and you want to maintain the position like, "No, you don't, you don't have jurisdiction over me. Because I, you know, I don't have people or assets on the ground in your country," or whatever, or "Your law does not require this. I disagree with your interpretation."

Daphne: But if you don't want to get in a big public fight about those things, the easiest thing to do is say, "But my terms of service require me to take this down anyway. So I'm not acknowledging your power, government, but I am going to take it down." That's, for example, what was the upshot of the famous Yahoo France case back in the 2000s, was that Yahoo denied that France had power to make them take down Nazi paraphernalia, but then decided to do it voluntarily anyway under their terms of service, and the case went away. So the board in a way takes away that out that platforms have, because if this question goes to the board and the board's like, "Nope, your terms of service don't require you to take this down," then suddenly that legal conflict that had been avoided becomes very sharp and pointy again.

Evelyn: I just want to add that that scenario that Daphne described, where a platform gets pressure or receives an order from a government, and it doesn't want to concede power, but then takes it down under its terms of service. Then the board steps in and says, "No, you need to change your terms of service," or "This isn't in violation of your terms of service." That's actually a positive case for the oversight board, because then that gets kicked back to the government and we have the proper democratic discussion about whether this is illegal content and whether it is something that should be taken down. In that case, the actor that gets held accountable is the government, which is often where this kind of accountability should fall. So that is a positive case for having something like the oversight board being injected into this system.

Daphne: Yeah. I totally agree with you, Evelyn. You and I were in a Twitter back-and-forth with some people, some law professors, about whether it would be appropriate for legal clinics to use law student labor to bring cases before the board. I think a lot of people's response was, "No, we're not going to defer to the Facebook rules and structure our legal clinics around it." But I think in a case like this one, like what you just described where it's really teeing up a question about state power, that might be an interesting and legitimate use of clinic resources.

Evelyn: Yeah. For the record, I think it is far too early to be discussing clinics for the oversight board. I just want to see it hear a case first. Then we can think about the next steps.

Jacob: Going back to self-regulation in the counter-terrorism space or the online counter-terrorism space more broadly, obviously at Tech Against Terrorism, we work closely with bodies like the Global Internet Forum to Counter Terrorism or the GIFCT, which is perhaps the most significant industry body for self-regulation of terrorist contents. We also support the wider tech industry in improving their capacity to respond as well. That includes things like terms of service, which we've discussed at length already.

Jacob: So we definitely think that this is a good approach. However, this of course does not mean that industry-wide, self-regulatory approaches are without their challenges, or that self-regulation should be the only answer going forward. Evelyn, this brings me to a piece that you've written about some of these challenges, called The Rise of the Content Cartels. If I could just ask you, what do you mean by the term 'content cartel' and why are they, in your view, potentially problematic?

Evelyn: I am really looking forward to hearing your thoughts about this because I personally think it's one of the most difficult issues in content moderation, and something that we don't necessarily talk about enough, which is when do we want cooperation and collaboration across the industry, and when do we want the idea of multiple marketplaces of ideas competing to have more diversity online? What I mean by 'content cartels' is arrangements between platforms and sometimes governments to work together to remove content or actors from their services, and without adequate oversight. That's why I use the phrase 'content cartels', because I do think that there are good cases to be made for collaboration in certain contexts, but it's really the lack of adequate oversight that, to me, makes these arrangements pernicious in the current forms that they're occurring.

Jacob: Yeah. I guess one thing I would say on this issue is that from our experience of working in this space, I guess you could, to some extent, make the argument that a lot of governments has 'allowed'... quote-unquote 'allowed'... industry bodies, like the GIFCT, to take the lead on this issue. One of the examples I would raise here is the issue of definitions and standard-setting of the term 'terrorist content'. So whilst the international community, to some extent, has failed to provide clarity on this term and on the 'terrorism' concept more widely, it is understandable that the companies working across a myriad of jurisdictions see it as preferable to build consensus across the tech industry to help tackle these threats across their platforms. I guess in that sense, you could argue that models such as the GIFCT or other industry approaches, whilst acknowledging that there is room for improvement, of course, is the best current alternative. I'm curious, Evelyn, to hear your thoughts on that specifically, or if there are other realistic models that you might see as preferable.

Evelyn: No, I definitely... I have some sympathy with that argument. I think in certain areas, if you want a standard for what content is available online, then you want a standard and you want that to apply uniformly. I think also, one of the biggest

cases to be made for these kinds of institutions is the help that it can give to smaller platforms in particular. A lot of this kind of content moderation can be extremely tech or resource-intensive. If you then set standards that you expect platforms to comply with, smaller platforms aren't going to have a chance of developing the kind of technology that the big platforms have, to enact those standards across the platforms. So it really is a matter of forcing bigger platforms to help the smaller platforms, it's pro-competitive in a way.

Evelyn: So I definitely... I have a lot of sympathy with a lot of those arguments. Really, my concerns are around the lack of transparency and accountability. With the way that these are happening at the moment, we don't really have as much visibility as we need over what exactly the definitions are that are being applied. We also don't have a lot of transparency around who is making these decisions. So if you have these collaborative arrangements and it looks like individual platforms all making their decisions for themselves, but in actual fact what it is, is it's a few big platforms, and therefore a few individuals within these big platforms, who are taking decisions for the rest of the internet or whoever else is in these arrangements, I consider that quite problematic.

Evelyn: I wrote this paper and I tried to carve out a middle path, and I got slammed from all sides on this. I do want to say that there are people who think that we need much more collaboration, otherwise the internet is doomed and this is never going to work. Then there are people who think that collaboration is always going to be a threat to civil liberties, and there's just no way that this could ever be legitimated. So I don't think it means I'm necessarily doing everything, anything, right. So many people disagree with me, but I do think it means that it's a conversation we need to be having much more.

Jacob: Yeah, I definitely agree with that. I think it is a really valuable contribution and even for us at Tech Against Terrorism, our work to really consider what we're doing and really have these issues in mind as we go forward, trying to encourage better self-regulation in the industry on this issue. So I think it was a great contribution to the debates. I think it's really good that we can discuss it at length here. Daphne, is there anything you would like to add?

Daphne: Yeah. First, I'll add that everyone should read Evelyn's paper because it's fabulous. It's really, really good. It tees up a bunch of issues that just have not been discussed enough. Then beyond that I would say, I think it's useful to think about, in terms of content cartels or centralized databases for filtering, or coordination across different platforms, it's important to think differently about different kinds of content. So if there's content that is 100% illegal every single time, regardless of context, such as child sexual abuse material, which is never going to be legal in news reporting or whatever the exceptions are in other cases, that's the strongest case for coordination. I think almost everyone recognizes that as the strongest case.

Daphne: Then if there's content that is illegal sometimes, but legal other times, depending on the context in which it appears, like a clip from an ISIS video that

is material support of terrorism under US law in one context, but is used for news reporting or research in another, that gets incrementally harder and that's, I know, something, Jacob and Flora, that you deal with every day. Then another step down from there is content that really isn't illegal, but that lots and lots of platforms prohibit, and it's a real waste if they're all replicating the resources to assess it.

Daphne: If they all need to hire an Arabic speaker and a Chechen speaker and a French speaker and a Japanese speaker to assess the same piece of content, it would be very nice for smaller and resource-poor platforms to not have to do that and to have a way that they can share information. But the information they share can't be this binary bit that says "leave it up" or "take it down". It has to be something much more granular that tease it up for review under the diverse speech policies of the different platforms. That's a very hard thing to build.

Evelyn: Yeah. just to add to that, I think desegregating the kinds of content is really important here, because that's part of the story that I tell in this paper, and that I was getting to earlier when I was mentioning the trends that we see in regulation. Because this is a story of a trend, right? It started in the context of child sexual abuse material. As Daphne said, that's a reasonably easy case to make for the importance of collaboration. Then that, the 'success'... 'success' in quotation marks... of that model got invoked in the creation of the GIFCT and the idea of that model being used. Now we see platforms talking about the success of the child sexual abuse material and the GIFCT model as being used to justify the creation of a similar kind of thing in other areas.

Evelyn: You hear it in conversations around coordinated inauthentic behavior, which I take to mean some sort of disinformation campaign, and things like deep fakes and things like that. So it really is a trend and we're heading towards this future. The alarm that I wanted to raise is we're seeing this expand. Danielle Citron has memorably coined the term 'censorship creep' for the use of a tool in one context being moved to another context. I am expanding on that idea to talk about content cartel creep, where we're just seeing this model go into more and more spaces without necessarily fixing up the problems with it, in the institutions that we already have.

Daphne: I will make another plug for a point Evelyn touched on briefly before, which is the importance of transparency. Right now, literally no one in the world, including the GIFCT members, knows what all the pieces of content are that are represented by the hashes in the database.

Daphne: And that means no one is able to double-check and see if there are mistakes made in things being added to the database in the first place, something going in that is lawful and important speech, but that's getting taken down in an automated fashion, which many people think happened with the Syrian archive, for example, which had this huge collection of videos documenting human rights abuses in Syria, taken down off of YouTube.

Daphne: There's nobody who can check and see if that's happening. There's nobody who can check and see if there is bias or disparate impact based on people's language or religion or regional background, et cetera, in what winds up in the database. And that's a really big problem. We just can't assess whether this is a good tool that should be a basis for future models or a terribly broken tool, or as I'm sure is really the case, somewhere in between with lots of complicated pros and cons that we should understand before we can build on it.

Daphne: But until researchers can see the actual content reflected in the database and not just the statistics that platforms report, there's no way for anybody to reach an informed opinion about that.

Jacob: On that note, the GIFCT is currently preparing for a new independent structure. That will include an independent advisory committee, as well as a range of working groups across different topics. And spoiler alert, we will lead one one of the working groups, one on technical approaches. That should be exciting.

Flora: Moving on, the UN special rapporteur on freedom of expression, David Kay, suggested that the international human rights framework should act as the basis for online content moderation. Indeed, he actually made this recommendation for Facebook's oversight board in a letter to Mark Zuckerberg. What is your view on the suggested approach or what other principles do you think should guide online content moderation, including of terrorist content?

Daphne: I argue with David about this a fair amount. I think international human rights law is a great framework, and what makes it great is not the substance of the law, but rather the fact that it is something a bunch of countries have already agreed on. It's the only thing we have that that has that transnational legitimacy. So, it's incredibly important.

Daphne: And I think when it comes to content moderation, it's really useful as a way to think about what process companies should abide by. Do they owe users appeals, how clear and transparent should they be in explaining their rules in their terms of service or community guidelines, things like that. But beyond that, I think there's a lot of expectation out there that international human rights law will be useful to set the substance of the rules. If there's a take-down involving nude image, how do you balance child protection against free expression? Or how do you balance a partially public figure's privacy rights or rights to dignity against the speech rights of somebody who's criticizing them?

Daphne: There are all these really hard, substantive questions, and the answer in international human rights law is often that there are an array of potentially correct answers. The answer that France would reach about the right to be forgotten is okay, and the answer that the US would reach about the right to be forgotten is also okay. There's this latitude for different interpretations. And I think in the case of Facebook, for the most part, the take down decisions that they're making already fall somewhere within that latitude. So, adding

international human rights doesn't dictate clearer answers than the ones we're getting now.

Evelyn: Yeah, I just want to agree with everything that Daphne just said. I have such enormous respect for David Kay. And I think he is one of the most important thinkers and writers in this space. And I'm a huge fan of international law. I have a sympathetic critique of this proposal. I'm not rejecting it because it's international law and that's not where I'm coming from, but I do really think that we need to have a conversation around the fact that it's not going to answer specific questions. And I can imagine Daphne sitting there with these questions as an in-house lawyer and looking to international human rights law and not being able to get an answer from it because the substance of the law often is not going to be specific enough.

Evelyn: And this was something that one of the international human rights law experts that's just been appointed to the Facebook oversight board was talking about last week, saying international human rights law, they're universal norms, but they have to take into account context. And so, really, they only get you so far because then you need to answer the question still within the specific context. And the other thing is that the online world is just a bit of a different beast and it's throwing up all these new questions that we just haven't really encountered before.

Evelyn: And it's so fast moving and it's going to be dependent not only on the country context, but on the platform context of the particular moment, and also on the implementation capacity of a particular platform and the way that the decision could be carried out in practice because I think that that's a really important factor that needs to go into policy development. And I just don't think that international human rights law jurisprudence at the moment really speaks to those issues.

Flora: Are there any regulatory approaches or areas of focus or tenets that you think will become increasingly prominent or that we should watch this space for, so to speak? Perhaps transparency or definitional thresholds or algorithm amplification, content moderation targets, that that type of thing. Is there anything that you would like to particularly highlight coming out of this?

Daphne: Well, sure. All of those things. Every arc of new internet regulation or legal dispute so far has had the front runners of copyright, pornography, and sometimes terrorism. And so, we've seen a lot of legal action, especially in the EU around copyright and terrorism in particular. So, that's not very surprising. Then, we have a batch of things that are products of the various strange times we live in now, which are questions about political disinformation, health disinformation, and those are incredibly difficult because often the content that's an issue there is not actually illegal. It's harmful, but lawful.

Daphne: And so, figuring out what, if anything, lawmakers can or should try to compel platforms to do there is a huge question and something, for example, that the

UK wrestled with in the UK online harms white paper. Then, as you mentioned, there's increasing interest in regulating algorithms and regulating amplification. I think the nuts and bolts of that are much more complicated than people appreciate, both because there are very competing values we might be pursuing through this regulation, for pursuing the value of truthful news sources first, that might be intention with pursuing the value of having a diversity of sources and perspectives.

Daphne: And that in turn might be intention with prioritizing ranking based on competition concerns or prioritizing them based on copyright concerns. So, even what the goals are are hard to define. And then, the mechanics of it, the number of engineers that a regulator would have to hire to understand, keep abreast of, and enforce rules about the algorithms for all of these platforms is mind-boggling. But that's definitely a big area of focus. People also talk about regulating artificial intelligence, and very often, they just mean regulating algorithms. So, I would put that in the same bucket.

Daphne: And then, the last point, which I really hope is a focus of more regulation going forward is transparency. In the terrorist content regulation, the parliament draft had the best transparency provisions that I've seen I think in any law, including requirements for the government to issue transparency reports about what they're doing, how many of the take-down requests correlated to subsequent investigations and prosecutions? How many of the take-down orders correlated to real world enforcement or turned out to have problems later on? I hope that we will see real progress in transparency requirements, both for platforms and for governments and other actors in this space.

Evelyn: Yeah, I just want to echo Daphne's all of the above comment because I really think we're sitting at this pivotal moment where, in the next five to 10 years, we're going to see a lot of experimentation from different countries and different platforms. And I think we're going to see some pretty radical transformations in this space. There was conversation a month or so ago about whether the tech lash was over because of the pandemic and how reliant we were on platforms and how they were doing a good job. And I did not buy that for a second. I think the tech lash is very much still here and going to be happening. And so, I think that it's just going to be a hugely important decade for online speech regulation. And I just can't finish anywhere else except to echo Daphne's call for transparency. And I really hope that that's a pivotal part of every proposal that comes out.

Flora: Evelyn, Daphne, thank you so much for joining us today. That's it for this podcast episode. Thank you again to Evelyn Douek and Daphne Keller for leading us through this very complex but important subject. In the description we have collected some of the articles mentioned throughout the episode that Evelyn and Daphne have written on this topic. We recommend that you check them out.

Jacob: To listen to our other podcasts, please visit techagainstterrorism.fm. To read more about our work in supporting the global tech sector in tackling terrorist use of the internet, visit our website at techagainstterrorism.org and follow us on Twitter @techvsterrorism. Thanks for listening and see you next time.