**DATA&
SOCIETY**

Databite 134: Origins of Trust and Safety

With Alexander Macgillivray and Nicole Wong, hosted by Robyn Caplan

Recorded: 07/08/2020

Robyn: Hello, everyone. Welcome to Databite 134: Origins of Trust and Safety, featuring Alexander Ma gill way, known as "amac," and Nicole Wong. My name is Robyn Caplan, a researcher here at Data & Society.

For those of you who don't know us yet, Data & Society is an independent research institute studying the social implications of data and automation. We produce original research and can be multidisciplinary thinkers, to challenge the power and purpose of technology in society. You can learn more about us through our website at datasociety.net.

So before we begin, I ask you to join me in acknowledging the Lenape community on whose land Data & Society was founded and what we now refer to as New York City and the exclusions and erasures of many indigenous peoples. Acknowledgments hold us accountable to committing to the process of working to dismantle the ongoing legacies of settler colonialism. They center questions that include, what does it mean to live in a post and neo-colonial world? What did it take for us to get here? And how can we address our part in that history? I also ask you to join me in recognizing the activists protesting in the streets, in city council meetings, in workplaces, and around the kitchen table to support Black liberation and to end racism in all its forms.

And now a little more about our speakers today. Alexander Macgillivray, or "amac", was Deputy General Counsel at Google, General Counsel of Twitter, Deputy Chief Technology Officer under Obama, and co-founded the Trust and Safety Professional Association and the Trust and Safety Foundation Project. He's also on the Data & Society board.

Nicole Wong was VP and Deputy General Counsel at Google, Legal Director of Products at Twitter, and also served as Deputy Chief Technology Officer in the Obama Administration. Learn more about the TSPA in the chat links.

So, this Databite came to be from a Twitter conversation, as all great Databites do. Alex and Nicole had just written a blog post together that they published on Alex's blog, bricoleur.org , everybody should go check it out, on the origin of the product counsel role. It's an interesting look at a lot of the early legal issues that tech platforms didn't even know they had to consider when creating new products for launch. So when they posted this, I had a million more questions that I responded to his tweet about, not just on product counsel but on the history of trust and safety more generally. So Alex suggested we do a Databite, along with Nicole. And I'm really excited to get to speak with them both today about the role of product counsel, but also about how it connects to trust and safety and other values and ethical concerns that platforms have to navigate.

So, we're just going to get directly into this Q&A. I want to start just by kind of centering this conversation and explaining a lot of the different terms that I'm already using. So you're both working for Google in the mid- 2000s in different roles when you created the role of Product Counsel. Can you tell us a bit about what you were first hired to do, how the job shifted, and how the role of Product Counsel came to be?
Nicole: Alex, you got there first. Do you want to start?

Alexander: Sure. So, I had been -- I was lucky to go to law school at Harvard Law School. There I was helping teach a couple classes on internet law. Google's basic pitch to me was, why don't you come in, and you can just do everything that was on your syllabus in the Internet Law class, and you can do that, you know, for real at Google, which at the time was a small but growing search company, really only had search. I was very excited about that prospect. And the title they had for me was IP Counsel at the time, which never really fit.

Nicole: And I joined Google about three months or so after Alex did. I think we must have both been in conversation with Google at the time. I had started in the mid-90's as a First Amendment lawyer. So I mostly represented online publishers, a number of the Northern

North American media publishers. Then when the commercial internet really started to take off in Silicon Valley, I started to represent sort of internet pure play clients like Hotmail and a number of others. That shift was, for me, sort of taking my First Amendment media expertise and shifting it online, which is a lot of where my grounding comes from. I was a partner at a law firm and was pitching Google's then- general counsel like, hey, I could do some stuff with you on how to launch your products. He said, well, why don't you draft up a job description. So what I did was drafted up all the thing its really love to do and none of the things I hated to do. That eventually became the job that I took, which was called, at the time, Senior Compliance Counsel. They had their minds like, well, the product should be compliance. I think Alex and I lobbied on saying nobody invites the person with compliance in their title to the product meeting. So we're going to have to realign the titles, which is how Product Counsel came to be named.

Alexander: Nicole's description, now I finally know why I kept being the employment lawyer, even after she joined, until we had an employment lawyer. Nobody likes to do employment law. So that's great.

Robyn: That's great. So what's the job of Product Counsel? What role do you play in the life cycle of product development for tech firms? And what are some of the issues that you address?

Nicole: So, I think the fundamental part of the job is that you get the product out the door in the most robust and responsible way you can. So it's where the legal part comes in, but it matches to what is the product or the company mission. At least in the time that we started doing this, it was a combination of the IP trademark and copyright, not so much the patents. We left those to the patent attorneys in the department. Consumer protection, content regulation, privacy. Eventually, it started to include competition as one of the components of doing it. But, you're looking at that product, hopefully, from its very start of design and working with the teams to build it in a way that is both compliant and meets the needs and the specs of the product and continue to help maintain it as it goes.

Alexander: It might be worth adding this is against the backdrop of the opposite of that, because that seems really reasonable, but the opposite of that, which was mostly going on, which was basically that the products were responsible for trying to figure out which legal issues they had, and figuring out which lawyers were the specialists in those product issues. So you might be a product manager and say, oh, I think I have maybe a privacy issue. Then you would go find the privacy lawyer and work through that issue with the lawyer. Or you might think, oh, I might have a copyright issue, I should go to a copyright lawyer. So this was a more product focused and user focused way.

Nicole: I know some of the people in the audience maybe be Product Counsel now or thinking about becoming one. Part of the job is you kind of need to know all the things or at least be able to spot them across that range. And if you're really doing it right, you not only understand sort of where your position with the law, but also where that product is positioned in your company is written up and where your company and the product sit within the overall regulatory landscape. So it's multidimensional chess in terms of what does it take to get the best possible product out the door, which is super challenging, as both Alex and I experienced a lot of sleepless nights on, but also one of the best parts of the job is being that close to it.

Robyn: And basically, you're trying to figure out all of the legal issues for legal issues you haven't even necessarily considered yet as part of these products, which is interesting. So what's the difference between Product Counsel and Trust and Safety? How do the two roles relate, and how are they different?

Alexander: Yeah, the way I think about Trust and Safety is very much that these are the folks who are establishing what is okay and not okay on particular online environments. That's both behaviors, and content, and actions. There is a fair amount of overlap. Trust and safety tends to be non-lawyers, although there are a lot of great lawyers doing Trust and Safety, and Product Counsel tend to be lawyers, particularly if they're doing the attorneying of a particular product. But there is a lot of advising back and forth. There are people within the trust and safety teams who really understand in a visceral way because they've been doing it for years and years and years how products are likely to be abused. So

you would often go to them and say, oh, I'm working on this new product, how do you think it's going to be abused? How do you think it will impact particular communities? That type of thing. Similarly, a Trust and Safety person might say, oh, I'm trying to figure out a way to distinguish between this particular type of content and that particular type of content. I know that the law has tried to distinguish between those two types of content. Is there something that a Product Counsel might know that would be helpful?

Nicole: Yeah, I totally agree. They're so closely tied, and one of the reasons I think that a lot of legal departments brought the Trust and Safety teams into the umbrella of legal, is because the Trust and Safety folks are one of the most essential tools you have for risk management if you're the lawyer. I think it's on two fronts, right. One is that I used to tell my newspaper clients, like your best chance of avoiding a lawsuit is the first time you pick up the phone and hear the grievance of the person on the other side. That's your best shot. And that's what Trust and Safety is. It's hearing the person on the other side, the user on the other side tell you what's wrong with your product. That's actually super important risk mitigation part of any company. Then separate from that, and you see it come out in the transparency reports that a lot of the companies do now, is you now have a dashboard of how is your product doing. So, in terms of thinking about, is my product flawed, is some part of it failing, which part of it is failing? The Trust and Safety teams' reports on what are they hearing from the product, they're like first responders in diagnosing what needs to get fixed. So sitting in the legal department, I was on the Trust and Safety team. One of my key go-to people in terms of trying to diagnose what's going on with a product.

Robyn: Just a follow-up from that, is Product Counsel the ones changing the policy in the end?

Alexander: It really depends, and different companies do this differently. My belief is that the people who are on the ground actually seeing this user interaction firsthand, the people who are closer to the facts, that those are the best people to also empower to change the policies. Sometimes it'll be Product Counsel, sometimes it'll be Trust and Safety, at least in my experience. Often, certainly the teams I work with, we wanted to make sure the Trust and Safety people were empowered not only to actually do the policy but to -- just like the

Toyota assembly chain, to be able to pull the cord and say, hey, this policy doesn't make sense when it is applied to this particular set of facts and people. Because at the end of the day, there are real people at the end of the line. So we should do something different. There had been a number of times when that ability to do something that was not our current policy and to reshape the policy because of the facts on the ground were extremely important.

Robyn: So, beyond content concerns and Trust and Safety, what's other work being done by Product Counsel that's often not thought about? I want to kind of expand our view of what we're thinking about the role being here.

Nicole: In the Trust and Safety role, which I think is broader. Right now, the current media focus is so much on content moderation. I think that's where a lot of the attention is, and appropriately so. But there actually tends to be a whole spectrum of things, and it will depend on the company and its product. If you think about what is Trust and Safety, then the key question to deciding how do you organize and hire for that thing is, well, whose trust are you trying to earn, and whose safety are you trying to protect? So, content moderation might be one, if you're a platform with a lot of content on it, user content on it. But there's also spam and fraud. There will be law enforcement requests or other requests for personal information. Some companies include security of the product, if that's a core part of their offering. So I think it will vary depending on the nature of the product that you're trying to deliver.

Alexander: Yeah, another way I think about it is often user support and Trust and Safety are somewhat similar. In user support, you really are trying to be as efficient as possible, and you might sacrifice some of your efficacy by just being more efficient. In Trust and Safety, I think it's the flip. Often decisions you're making are so important to the users and to the brand, and as Nicole said, to the risk of the company, that you want to make sure you're doing it right as much as possible, and you also want to be efficient, but you, in general, don't want to sacrifice efficacy for efficiency.

Robyn: So, within your time at Google, when did you both have to start considering content concerns as part of your role? So, what was the product, and what were some of the earliest concerns that you had to address?

Alexander: Maybe I'll go first on this one, Nicole. I think mine predate yours. For me, it was in -- I think it was 2002 before I joined Google I was an outside counsel. Google had just removed from web search a site called xenu.net, which was a Scientology criticism site that was also publishing a bunch of the Scientology secret texts. The Scientologists were well known and extremely early adopters of DMCA take-down notices. So these are copyright notices where you send an email or a letter or a fax sometimes, if we're unlucky, to the platform saying that someone on the platform has violated your copyrights. In this case, the Scientologists were complaining that Xnew had published specific Scientology texts, and Google, in response, had taken down the whole Xnew.net site. And Slashdot which used to be a big community place, had just gone crazy because this was a mistake. This was something Google didn't have to do. It was against generally what Google was doing at the time, which is they weren't taking down a lot of search results. And no one could figure out why they had done it. They were trying to decide what to do about all that. So that was when -- I had worked with the brilliant Wendy Seltzer at Harvard. She came up with this thing called "chilling effects," which is now the Lumen Database of takedowns.

It was the first time that we suggested to Google that a good way of handling these things would be to be very transparent with the users as to why a particular thing was being removed from search results and to put in the search result page, hey, there's a result missing, and here's the complaint that got us to take down that result. That really helped the feedback loop be better because, of course, these platforms are going to make tons of mistakes, and those mistakes are going to have significant, real-world consequences. In this case, making it much harder for people to find this entire site of Scientology criticism. But making it faster so that you could unearth the mistake and correct it. Google quickly put back up most of the Scientology pages, except the ones that were claimed to have violated copyright, was another helpful thing that came with the added transparency as well. So that was really my first time with Google and doing that type of thing.

Robyn: So, this is like an early data void almost.

Alexander: It was scary because Google was just starting to be the thing that most geeks used to search the internet. So, it felt like, especially to the geeks, that Xnew disappeared, and that there was a company that had the power to make something disappear. Of course, that's not what was happening, but that's what it felt like. I think throughout the platform's history, they're sometimes slow to realize the amount of power they have. Not just actual power but apparent power within communities. So being more transparent was a helpful way to at least start the explaining of why particular things were happening. It also is classic in that so much of this early stuff was copyright related. And that was because a lot of the copyright lawyers got into the game of sending notice to take down really early. There was a clear law on both what had to be in one of those notices and what the companies should do if they wanted to maintain the safe harbor that the DMCA, which is section 512 of the Copyright Act provided.

Nicole: Yeah, I'll build off that. Candidly, the start with Google is so far away from me now. I'm not sure I would remember the first content issue I dealt with for them. But I think, just to frame this full conversation, it's worth thinking through content moderation is not new. Content moderation goes back to message boards, system admins, and Steve Jackson games, taking stuff down because it wasn't appropriate on the site or for whatever reason users didn't want to see something on the site. I remember in one of my early clients was Craig Newmark. So, Craig of Craigslist was his own sys admin. He personally responded to every customer email, including ones about content problems. So these are sort of the origins of Trust and Safety, really handled by people who were sys admins in a lot of ways, or later on customer support people. So, I'm so pleased to see it develop into the profession that it's about to become or starting to become now of people who are trained to think about content issues and the variety of ways that bad actors can be on your platform.

But I agree with Amac. I think the early content demands we saw, the disputes were user versus user very often, like claims of defamation or claims of fraud or scam or things like that. There were a lot of what we call pump-and-dump message boards in the early days of the internet, trying to do sort of financial schemes online. But eventually, because the

DMCA was so effective on getting content down quickly, you ended up seeing a rise of copyright complaints and DMCA demands. I think the other areas where things grew a lot were like having to deal with child sex abuse materials, was also one where the law was super clear, where nobody wanted to have that on their site anyway. So, you saw the development of teams both to identify and quickly remove, and the tooling that was necessary to do that generate very quickly in the Trust and Safety areas.

Robyn: So, kind of building off of that, Google rapidly expanded into a lot of new areas all around the world. I've spoken within you about this in the past, Nicole. You spoke about kind of the process of going into new areas and needing to learn new legal systems. So, I was wondering if you could both speak a little bit about what that was like and what it was like to suddenly kind of be the mediator between different cultural conflicts, if you can.

Nicole: Yeah, when we've talked before, Robyn, I describe this convergence of things that happened in like the mid 2000s, which is I think the largest of it was the ability to do real images and video and deliver it quickly across the globe, really came into its own. Part of that was about YouTube and other services. But the ability to share content visually changed the game in a lot of ways. It was also a time when a new generation of countries joined the internet and became real in terms of a market. So, you think about the first generation starting in the mid 90's of countries that had real internet markets, it would be the U.S., Canada, the EU, Australia, and Japan. So maybe ten to twelve countries, all of which had fairly similar rule of law norms, fairly similar principles around free expression and privacy. But by the mid-2000s, you see internet penetration happening in a whole host of new countries that include: Saudi Arabia, China, Russia, and Vietnam, Turkey. All of a sudden, you're in a bunch of different legal but also cultural norms about what's okay for people to say to each other and behave with each other. That, to me, was like this turning point of how do we really be a global platform that is genuinely available to people who don't agree with each other on some very fundamental things?

Alexander: And lawyers are often -- like, we're designed to be really bad at that, I think it's worth pointing out. We are put through an education that is, at most, country specific, and often state specific. The exam at the end of our time is about a particular state's law in

the United States. So many of us don't have -- first of all, it would be absolutely impossible to be an expert on every local law in every jurisdiction all over the world, which is what you kind of need to be. Second of all, that's just not what the going in assumptions of the law schools are or necessarily of lawyers. So getting up to speed quickly on all of these things, and not just the legal norms, but cultural norms, was something often that a person who was actually living in that country or part of a group that the area of law was important to or the cultural norm was important to would be much better at telling you what was okay and not okay, than a lawyer based in wherever we were based in, Palo Alto, California.

Robyn: That's super interesting. So at the same time, Alex, lawyers are kind of cultured into a specific American view point, but you've said that content decisions beyond what was just legal and not legal were not dependent on First Amendment norms, per se. Can you speak to some of the other value tensions you and your teams grappled with when you tried to determine the right policy for each product?

Alexander: Yeah, I always think of the starting place of this as being more about what the product was trying to do and the value that users were hopefully going to get out of it, that we would deliver to people. That really, to me -- and the example I use, usually, is Google web search is a little bit like a library. As a user, I really want Google web search to have pretty much everything in it, and I'm going to be mad when it doesn't have my weird little piece of archive. So I really like ultimate frisbee. If it doesn't have a particular funny rule for a game that I like to play in the rain that I had seen before, that is going to make me angry. At the same time, I don't care about college football, but Google search having a whole bunch of college football is not going to make me angry. And there are pieces that more of the country doesn't want. For example, some alt-right screed that its existence in Google web search, at least for me, does not upset me.

At the same time, Facebook is a little bit more like a dinner party, or at least it was back in the day. I'm less concerned about it having every little thing that I personally am interested in, and I'm more concerned with it not having the things that I don't care about. So if I go to a dinner party and someone is just droning on about American college football, that's not a good experience for me. It's okay if nobody ever talks about ultimate, right? That's fine. I

understand that I'm weird. But the American football thing, fine. And if someone is talking alt-right baloney, I'm leaving. So the types of things you might want to get out of a particular product, to me, were more what drove how we thought about the types of policies in terms of content regulation that we would be interested in, against of course the backdrop of what is legal and illegal and the First Amendment giving us a lot of space to play. Same with freedom of expression in many other countries. But that's just the backdrop to a narrower question of these policies aren't policies for their sake. They are designed to further a particular set of uses of the product or the platform. At the end of the day, hopefully deliver a certain type of good and minimize a certain type of harms to people. I think when I was getting into that disagreement about whether the First Amendment was so controlling, that's what I was really thinking through. It's not like we were getting into particular First Amendment doctrine on thinking up the content regulation. Rather, that was the backdrop rule, and we were trying to design something that would work.

Nicole: I think that's totally right. But it does inform -- because so many of these very dominant platforms were built here, right, by founders who grew up in an environment where the First Amendment and this notion of a marketplace of ideas and a bunch of the things that come with that is built into their product and their vision for their product. So, a lot of the work that I do these days is helping companies think through what are your content rules, but fundamentally, what are your principles? You start with what is the mission of your company and who do you serve, right. So, to me, where part of some of the difficulties that U.S. based companies have had, they've had these enormously ambitious missions. Yahoo! Was the portal where all the world could come and find what they needed. Google was to organize the world's information and make it useful. Facebook's is to build community and bring the world closer together.

Robyn: Building a global community, I believe, right now.

Nicole: I'm not sure what it is today, but this global scope of bringing everything into one place. I think that actually puts them in this difficult position of trying to mediate a world which actually has many differences that's not going to get along. So, if you're trying to create a platform where all of that's going to work together, that's just an enormously

difficult task, if not impossible task. This is sometimes where I get into conversations with clients. Like, you don't have to be that. You can define your product in a way that serves your users but doesn't have to run into some of these more difficult questions on content. So, I think LinkedIn is enormously successful, being the professional network it is. Some things are not appropriate at LinkedIn, and everyone is okay with it being taken off because that is not what they are there for. I think Pinterest is in that same bucket. Disney gets to be Disney. I think we need to expect that and encourage that. The platform can be the narrow thing, not everything.

Robyn: So, in response to that, in my work I'm seeing platforms trying to grapple with a lot more of the difficult, gray-area concerns, concerns that might not be legal issues but are definitely ethics issues and definitely issues where the public interest has raised some important -- where the public interest is at stake. They're kind of leaning on some American values. They're using terms like "legalism", "precedent", "free speech", "federalism", to explain what they do. Can you give a little bit of insight into why legal teams would fall back on that as they're trying to navigate these, kind of, messier areas of content concerns? So, looking at basically the Facebook Oversight Board and this Supreme Court type function, where they're going to be creating rules. That rule system is going to serve as a system of precedent. So how can creating those structures help decide these issues going forward?

Nicole: Just to take the Facebook Oversight Board as an example, to me what the value of it may be as it comes together is, one, transparency, the way it understands complaints and processes complaints. Maybe a court system or a quasi- court system feels like the right way to air those concerns as well as lay out the argument in a clear way. That certainly has a bunch of lawyer bias in it, by the way. And my feeling, and this gets a little more technical, these are platforms trying to mediate a lot of cultural values and norms. I think that's a really good thing to have happen. We don't do it often enough, and we don't -- we aren't talking about it in an honest way, about how hard it is. So, it creates -- those kinds of structures create a forum for that. What gets us closer to is some sort of norm system we can agree on. And we need a place to have that conversation.

Robyn: Anything to add, Alex?

Alexander: No, I thought that was great.

Robyn: So lastly, before I move on to my final question, where is the impetus coming for tech companies to address broader ethics concerns? Is it coming from pressure from advertisers, media, users, advocacy groups? Or does it come internally within the company?

Alexander: It's funny. I would put a group you didn't mention at the top of my list, which is it's coming from employees. In particular, engineers. I think at a lot of the companies that Nicole and I have worked at, engineers are the most powerful folks at the company. These changes often do come from them. Now, granted they come from them because you can look at all the diversity and inclusion surveys of these companies, and Silicon Valley remains overwhelmingly white, overwhelmingly male. There are plenty of work the companies need to do in terms of diversity and inclusion. So it's not all coming from them because they just don't have the experience, the lived experience to catch a lot of the important ethical concerns. So a lot of it is coming from users and advocacy groups, and stakeholder groups, and just trial and error. Then it's percolating back up.

At a lot of the companies, there's no such thing as a Department of Ethics. Sometimes, like in Google's case, that was very aggressively attempted to be something that everybody was involved in and then maybe legal was a backstop. I think other companies -- the one I really -- the model I really like here is Dropbox. When Ramsey Homsany was there as the General Counsel, he had an ethics title that was separate and apart from his legal title. I think the good thing about lawyers and ethics is that we have struggled with some of these questions in the past. The bad thing is that we often fall back to what is legal being what is ethical, when actually they are two very, very, very different things in a lot of cases. So, I liked the fact that Dropbox actually had two different titles as a way of saying if something happens that is perfectly legal but grossly unethical, that is still Ramsey's job to make sure it doesn't happen. And having some combination of the Google any person can pull the lever and say, hey, I think what we're doing here is wrong, but there was a buck stops here person that Ramsey was at Dropbox, I think some combination of that is probably a better way forward.

Robyn: That's very interesting. So Alex, you just co-founded a professional association for Trust and Safety. Nicole, you're an adviser. Can you tell us a bit about what you're trying to accomplish with this organization and who's allowed to join? Is it just policy, or is it the thousands of people who are often enforcing content rules as well? And what issues are you hoping to address with this?

Alexander: So, there's two things that we founded. I should say this is myself but really Adelin Cai, Clara Tsao, Eric Goldman, and Denelle Dixon. It's a group of people. The Trust & Safety Professional Association is designed to be focused on our members, which will be the Trust and Safety professionals writ large. So that's everyone from the person working as a contractor for another company that works for another company that works for Facebook doing content removals, to the people who are setting the policy back at Facebook headquarters, just to pick on Facebook for a moment. But it's also the Trust and Safety teams at places like Etsy and Airbnb, places that are often not thought of in the same context but are doing this Trust and Safety job. To me, the biggest thing there is that's a job that has been going on for a lot of years. There are people who are really, really, really good at it and have developed a lot of experience and wisdom. It would be great if there was a place for them to share their experiences and practice, and a place that would help support them through the many unique challenges that they have, everything from how do you look at hundreds and thousands of pieces of child abuse material and deal with that on a personal level, to how do you deal with career advancement in this profession? So that's the Trust and Safety Professionals Association. We just got off the ground. We have tons of work to do, including finding an executive director, which we have a job rec on the site.

The second part of it is the Trust and Safety Foundation. This is a project of the Internet Education Foundation. There, our focus is broader than just the people who are doing the work but more broadly into Trust and Safety, how do we help the bigger population understand it better, how do we get more research about it out into the world, and that's another sort of exciting part. It's a little less developed, but we'll be developing that more over the coming months.

Nicole: Yeah, I have to say I am so excited about the work that Alex and Adelin and Clara have been doing. We saw a similar thing happen with privacy professionals. For those who may not remember, there was a time where a privacy lawyer wasn't a thing. That's not a job. It wasn't really until it was created as a job title it was given its own professional association that you started to see both people being able to find their network, a community that could discuss best practices, and it gave them standing within their companies to be able to be the trained person who could speak wisely to an issue that the company was going to face, not just to do the job but actually be the internal expert. I feel like Trust and Safety is in that place now and very much needs it. Again, because I think it came out of such disparate places from customer support to legal analyst to whatever. It will be boosted by this notion of shared best practices and the standing of a profession that is trained and has expertise in a particularly difficult area that we actually really do need strong guidance in. The one thing I would add is it's not just about the substance of the work. For those who have done content moderation, it is hard work and it is emotionally hard work. I think the standards around what does the profile of a content moderator look like, how do we take care of them to make sure they can have longevity in the profession, all of those are the types of things that an organization like TSPA can weigh in on as a collective. I think that's enormously important.

Robyn: That's fantastic. Okay. So I have more questions, but we have some in the Q&A. I'm going to move on and ask you questions that have been voted up by our community. The first one is from Tarleton Gillespie. In what ways do a company's competing priorities push back against things that Trust and Safety teams see as important? How do those conflicts typically get resolved? When does Trust and Safety tend to win that argument and when do they not? How has that changed over the time that you've worked in this field? You weren't going to get an easy one from Tarleton, guys. All of the questions we should have addressed.

Nicole: Going back to what I was just saying, I think that Trust and Safety has frequently not had the seat at the table that they should in the development of products for companies that are trying to push hard on getting something out into the world. I think their increasing rise as an important constituency that speaks for the user has been really

important. So my experience is it depends a little bit company to company who wins, but if you're doing it right, the path that is best for your users is the one that should win. I wouldn't lay that necessarily at Trust and Safety or engineering or sales, necessarily, but that's one of the roles of Trust and Safety, being able to speak to that user advocacy.

Alexander: I guess the only thing I would add there is I think from the outside, often people put Trust and Safety and revenue on somehow opposite sides of the table. I have never seen that happen. Not have I never seen revenue beat Trust and Safety, but I've never seen the battle, because when Trust and Safety is trying to figure out a particular problem, their revenue implications, honestly, have never come up in the discussions that I've had. Now, I haven't been at every company, blah, blah, blah, but I think the bigger issue is often just the types of things that a company is working on and the time of engineers to launch stuff is constrained. The question is how do we rank the various things that we need to do for the users and for the community at large? That's where, as Nicole suggests, I think Trust and Safety is much more getting a seat at the table these days. I think the more enlightened companies are actually giving Trust and Safety engineering resources to make fundamental product changes rather than just policies. Which is really great.

Robyn: Yeah, I think we're starting to see a little bit of a limit of it's the business model critiques. Some of the content concerns from my research, it's very hard to tie directly just to the business model. So a question from Jeanna Matthews. Someone may have the right -- she wants to get your reaction. Someone may have the right to free speech with their one mouth and body, but do they have the right to 20,000 bots to amplify their message? Do they have the right to use sock puppets to caricature those on the other side of an issue in a way that makes them appear ridiculous or extreme?

Nicole: Yeah, I think the way I've been hearing this conversation come up is you may have a right to speech but not reach. I think that's actually a really important area of conversation, both for platforms and for the users that are on it, which is how do we think about the use of bots on a system and what can we do technically to prevent it. I think that's actually a pretty hard thing to do. Not for all of it but for some of the more targeted campaigns. I also think platforms need to think really hard about how they make their

engines of surfacing content, what it surfaces and how it surfaces, and how much user agency ought to be involved. So I think that gets to a technical question as well.

Robyn: We have a really interesting question from Stephen Fox. How do we build accessibility features into product development rather than having them fall to a specific team, whether that's specifically Trust and Safety or not, to try and fix the gaps after launch? Should Trust and Safety be the guardians for inclusionary design?

Alexander: That is a great question. I don't think I'm going to have a good answer on the last bit, whether Trust and Safety is the right place for inclusionary design. They do have a lot of the skills for it, but not always the understanding and the lived experience. I think it's hard to replace that. I think honestly, if we did a better job in Silicon Valley of having people with accessibility issues in our teams, we would do a better job of building accessible products. And there's no real way to get around that as the best thing we could do. The second-best thing we could do is talk more to people who have those issues because they're going to be the best people who are able to tell us what to do in terms of making our products more accessible. For a lot of the stuff, they're not unsolved problems. We're not even doing the basics. So I don't know the answer to how to better prioritize getting the very basic stuff that everybody should understand, should be done in products in terms of accessibility done.

Nicole: Yeah, Alex. I don't know if you remember. It's kinds of foggy for me. I think at one point for Product Counsel, you have a checklist. Have you looked at the copyright issues? Have you looked at the privacy issues? I thought accessibility used to be one of the things we would look at. If you're Product Counsel and involved in the design stage of a product, then in theory you're raising at that point the accessibility question. I do think in my experience that frequently gets lost because there's not a continuing person on the team for development to make sure it actually crosses the finish line. But if you were designing the process well, you're hopefully catching it earlier and not after the thing goes out the door.

Alexander: Yeah, and I do think, speaking personally here, that regulation in this space might be a good idea.

Robyn: So, Beverly asks, what kinds of backgrounds are helpful, in your view, going forward to build robust Trust and Safety teams in organizations? Is this a place where philosophy grads can go?

Alexander: Was that part of the question, or was that your add-on?

Robyn: That was my add-on. I may or may not have a degree in philosophy.

Alexander: I mean, the biggest thing -- the biggest driver I've seen in terms of people being successful here is empathy, attention to detail, and a fairly -- an ability to deal with very difficult things and very difficult conversations, because this job is just very hard. There's no sugar coating even some of the easier ones, which are people yelling at each other online. So I think those three things would be what I would look for. Within the team, having people with a diverse set of backgrounds would be extremely useful. In terms of some of my personal regrets over time have been not pushing harder on that particular lever.

Nicole: Yeah, I totally agree with that. It sounded like the question was asking what kind of educational or experiential background you need. I think it's a little less of that and more about the traits you have, which I agree with Alex. Attention to detail, empathy is huge, and a resiliency. You get yelled at, and sometimes completely unjustified, but yelled at, and have to not only be able to take that and transform it into something productive, but sometimes escalate it within either your managerial ladder or into the legal department or engineering department and give them some bad news about, like, this is really bad and we should fix it for the user. I think that kind of leadership is also one of the things you end up looking for, for managers in those functions.

Alexander: It's also something we're trying to explore at the Foundation and at the Association, the Trust and Safety Professional Association. Are there particular courses or things that would be useful for people to start to learn to be able to get into the profession? And we developed a bunch of case studies of some of the things that have happened

already. This is not a blank slate. So that we can learn from our mistakes and not make them again and again.

Robyn: So, question from Sareeta. The story you told about sys admins to Trust and Safety is about making informal, tacit knowledge tacit. As the point about diversity suggests, that tacit knowledge is white and male. Even those doing the labor of content moderation, e.g. applying the rules or not. How do you value the labor and expertise of these latter people and Black and Brown people more generally going forward? Second question, the term "user" is very broad. How do you determine which user to listen to when they're saying conflicting things and have differential access to power? States versus community groups.

Nicole: Yeah, those are great questions. In the first one, I think we need to diversify this work force. I think that's one of the things I would hope TSPA would be focused on. I think finding people from a wide variety of backgrounds and experience, including Black and Brown people, is essential. In my experience, and I think it's because maybe of the time I was most deeply in this, a lot of my focus was on hiring a team that had language capability. You're serving 170-80 countries, and you need dozens of languages. So, in my brain I was trying to figure out the global diversity of my work force, so that I would understand the content issue in India versus Turkey, versus Australia. But I think what has clearly become the case in the last several years is we actually need to diversify the teams within the countries or their regions.

Now I've forgotten the second one, but it was important. Oh, who's the user? So Alex may have a different take on this. I think one of the most difficult but interesting spaces that I've seen in content moderation in the last few years is trying to account for the people who are not the users on your platform, trying to account for harms that happen off your platform because of something that starts on your platform. So, I think who's your user depends a lot on your company mission and what your product is for and all that. I think one of the really interesting things to wrestle with for Trust and Safety teams these days is how do I deal with harms that happen outside of my service in particular, not even for users of my service.

Alexander: Amen to what Nicole said. Particularly the distinction between users and non-users hopefully being something pushed these days as these platforms become more and more impactful in the world. I would say one of the things that happened as we launched, that we did as we launched, the Trust and Safety Professional Association, is even in that launch, we didn't do a good enough job of having representation from African-American and Latino and Latina communities. We're working on that, thanks to some very good folks who pointed it out to us. But that's an important thing to happen and is important in the profession as well. The only thing I would push back a little bit on is at least my understanding of this profession is that women are actually fairly well represented within the profession and in positions of power within the profession. There may still be systemic problems with how we treat the input from those women in power in the profession, but the representation is certainly there. So I'm not sure saying that most of the sensibilities are white male, at least from within the profession, quite gets to the root of the issue as well.

Robyn: So, there's a ton of really great questions in the chat, but we are at time now. I don't even know if we have -- if you guys have some quick closing remarks about your hope for the field going forward, I'd love to hear them now, before we close up.

Alexander: Nicole, why don't I go first because you're more eloquent than I am. Then you can close the whole thing off. I think one of the things that I am most excited about, Nicole spoke about the marketplace of ideas and sort of the fundamental framing of the best way to encourage wonderful discourse, including political, being this First Amendment or freedom of expression type of framing, which the marketplace of ideas has been historically a very big part of. The questioner was asking a question, on the bots especially, that I think is one of the avenues we've seen for the limits of that type of understanding of the best ideas will win out and get the most air time and become the dominant ideas. I think that's often not the case in our current society. So some of the things that I'm most excited about is as we have many more people from many different backgrounds, and I don't just mean racial and ethnic here, I also mean just academic backgrounds, coming at these issues, we are seeing, I think, and will see a fundamental rethinking of the marketplace of ideas, of how freedom of expression ought to be thought of in this new digital space. That's what excites

me most, sort of rewriting some of the basic DNA about how we think through ways of having great conversations online.

Nicole: That's so good. So, I could nerd out on this all day, so thank you Robyn and Data & Society and Amac for putting this together. It was such a great discussion. This is such a hard time to be doing this work, especially with what we're calling the sort of tech lash. It feels to me like all the companies are on their back foot, in a really defensive crouch trying to manage through it, at a time when I actually think it's most important for us to be having an honest conversation about how hard the job is, about the real social questions that we have to answer. The truth is, there's a bunch here that's terrible and that tech will never fix because tech didn't create. Misogyny and racism and anti-Semitism and state warfare, those are things that are sort of outside of what tech can fix. So what I would love to see is a credible voice bringing that conversation to the table but also with real ideas about what we can fix. Like, we don't have to have an engagement al algorithm that accelerates all of this terrible stuff. We don't have to create filter bubbles that exacerbate polarization. And so, to me, the professionalization of a Trust and Safety group and all the lawyers and all the other folks at the companies who are working towards these products have an obligation to really engage in that conversation, but do it honestly. And not pretend they have the tech solution, that it is just one more AI algorithm away. Let's have the honest conversation about what we can and can't fix.

Robyn: I think that's a fantastic way to close this out, thank you so much Nicole. And I want to thank everybody for joining us, and thank you again to Alexander and Nicole for sharing your expertise. Please visit our website and join the Data & Society list for updates on future programming. Thank you again to everybody.