

Rigoberto Lara Guzmán: Hello, everyone. Welcome to Databite number 145. My name is Rigoberto Lara Guzmán events producer here at Data & Society alongside my team behind the scenes, Eli Ellie and Amanda, our captioner for today. We'll be spending the rest of the hour together so let's get ourselves grounded. Data & Society is an independent research Institute. We study the social implications of data and automation producing original research to the ground informed evidence-based public debate about emerging technology. Data & Society began in New York city, an island in a network of Hills and rivers in the Atlantic, Northeast known as Lenapehoking, the ancestral land of the Lenne-Lenape people. Today, we are connected online, the vast rate of servers and computer devices. In the United States, much of this infrastructure sits on stolen land acquired under the extractive logic of white settler expansion. We recognize this history and uplift the sovereignty of indigenous people, data and territory and we commit to dismantling all ongoing settler colonial practices and their material implications on our digital world. To learn whose land you are on, feel free to visit the link on the chat, uh, native land that CA to participate in this event.

Um, today's databite features three key thinkers in the field with us today is Brittany Smith, policy director here at Data & Society, Sarah Chander, senior policy at the European digital rights network and Fenwick McKelvey associate professor in information and communication technology policy at Concordia University to learn more about our speakers and read their longer bios, check out the event page on the chat. Now I'll turn it over to our host Jake Metcalf program director for the AI underground initiative. Thank you and see you.

Jake Metcalf: I'm Jake Metcalf. I'm the director of the AI on the ground initiative at data society. Um, and, uh, thank you, uh, to all of our guests for joining us today. Um, and the production team behind the scenes. Uh, the one, one part of the impetus for this conversation today, um, is the, uh, AIG I team has a forthcoming report called, uh, assembling accountability through algorithmic impact assessment coming out later this month. Um, and while this, um, conversation is not directly about the report, um, it does help help us frame the conversation somewhat so that we can see what's at stake in the, uh, different policies that had been proposed for assessing algorithmic systems, uh, and, uh, and a number of countries. Uh, and so, uh, let me share a little bit of our thinking about that before we dive into the conversation on the speakers.

Um, so we noticed that well, algorithmic impact assessment was increasingly bandied about as a potential form of algorithmic governance. Uh, people using the term meant a lot of different things by it. In some cases, an AIA was proposed as a mechanism for exerting democratic control of government software procurement. And other cases in AA was proposed as an internal governance process that town tech companies would conduct anytime they built an automated decision system. Um, and what exists primarily as like a regulatory record or a transparency activity. Um, in other cases, and AIA was envisioned as a type of third-party auditing practice. Uh, what was common across these purposes, uh, is that, uh, impact assessment is a bundling together in a single process of both an obligation to measure what algorithm algorithmic systems actually do. Um, along with an accountability structure that specified as, um, uh, who needed report that information to whom and who is responsible for fixing the problems.

Uh, in other words, impact assessment appears to be an appealing option for algorithmic governance, precisely because it bundles in account of what has happened and who is responsible for it in this panel, we will be discussing these, these particular regulations here on the screen. Um, the us algorithmic accountability act, which proposed in 2019 and, uh, according to news reporting are currently being redrafted to be submitted in the current Congress, uh, the, uh, European union's, uh, uh, algorithm max has a much longer title, but that's, it's short as its nickname, uh, which was, which was, uh, the sort of the first draft of it was put pub made public about a month and a half ago. Um, and the, uh, federal government of Canada's, uh, directive on automated decision-making, uh, and the associated assessment tool, um, uh, it was released in 2019, um, and writing a report to make sense of why we saw so much variability on this topic. We decided to study, uh, impact assessment from other domains, choosing areas that have some structural stuff, molarities to algorithmic systems, okay. That environmental impacts fiscal impacts, privacy impact, digital privacy impact and human rights impact, and look for common features that would help make sense of what AI's might actually accomplish.

Um, and we found was that impact is an evaluative construct. Now, uh, it's, that's a fancy word, but what we really mean is that, uh, an impact is a weird kind of thing to measure. Now, you can't go out with binoculars, you go to telescope or approaching array and find an impact in the world rather impacts only emerge as a kind of thing that you can study. Um, when you have a community that needs to create a common shared object that enables them to coordinate their activities around that object. That doesn't mean that impacts aren't real rather it means that they are only real when a community needs a proxy that allows them to talk to each other. And if you look to mature complex types of impacts that assessment, it becomes clear that many different types of expertise and measurement practices and measurement tools, um, can go into the construction of impacts. Uh, for example, an environmental impact assessment report might draw on archeologists. Fluvial, geomorphologists civil engineering, there's toxicologists, urban planners, and a lot more, uh, all of those experts have distinct measurement practices and distinct tools, uh, yeah. Impact as an object is capacious enough to construct a capacious enough of a construct, uh, to accommodate all of those different types of measurements, but that the patients in this means that there are significant stakes and who gets to involved in choosing what is measured inside of impact.

Okay. Um, and so we use this phrase, um, co-construction, uh, impacts and accountability, relationships are co constructed is our claim. And co-constructed is really just a social science-y way of saying they're built together. Um, so don't let that throw you off. Uh, what we mean here is that the core feature of any impact assessment is who reports what to whom the, what an impact assessment is not really separable [00:09:30] from the who, uh, from who is obligated to explain the what. Um, and so here, we're leaning heavily on the definition of institutional accountability offered by the organizational sociologist, mark bogans. He says that central to any formalized accountability structure is the relationship between an actor and a forum. The actor has to present, uh, an account of a system or an event to an independent forum as empowered to render a judgment or force changes for algorithmic impact assessments.

The actor would be in the system developer and the forum would most likely be a regulatory agency or perhaps a party responsible for procurement. Other types of

impact assessment can have different accountability relationship structures for environmental impact assessment. The actor is the developer proposing the project, and the forum is a governmental body that has to consider public input and approve or deny the proposal for human rights impact assessment. The actor is the company which wants an assessment of the consequences of their business practices. And the forum is usually just public opinion. Um, so thus, when we say that accountability relationships and impacts are co-constructed, we're indicating that what will be measured as an impact is not at all settled, rather it is always evolving in response to the accountability relationships that have been empowered, uh, to determine what needs to be reported when a report is adequate.

And when a solution is satisfactory and some domains, the content of impact assessment reports have changed significantly over time by examining impact assessment or streams from other domains. We developed a schema that identifies the core features of impact assessment, which we call the 10 constituent components. And the previous slide, we already discussed two of the core components, actor inform and impact. Uh, now's not the right time to go into detail about all the others. You'll find extensive discussion of them in the forthcoming report, including likely failure modes for algorithmic impact assessment. But what I would like to point out is that when it comes to policy-making, there are some distinct categories here. Policymakers are well-equipped to make decisions about the formal structures of impact assessment regimes, the who, and the, when such as actors and forum, the catalyzing event that triggers the need for an assessment or the source of legitimacy, such as legislation or regulatory decision, but the questions of what and how such as the methods of assessment and the process of public consultation are inherently a matter for evolving community consensus. If policymakers either overdetermined these components or fail to leave room for the public interest to be expressed, the industry will largely be able to determine those conditions in a one-sided manner.

And so our call to action is that even if regulators can't rigidly prescribe an exhaustive list of impacts and mandate a set of tools to assess them and developers can't be left to regulate the impacts of their own products themselves, they nevertheless remain stakeholders in an AI process alongside others. As we've seen regulators, I needed to establish a forum to set clear guidelines. And when AIS are called for and to convene a process for deciding what, when in that assessment is adequate, when companies are needed to provide detailed descriptions of their systems, specifications, or access to their systems so that they can actually be assessed. But outside auditors who have developed critical technical skills for, to evaluate systems, performance characteristics are needed to, as our critical scholars who have developed techniques for investigating the social and environmental implications of algorithmic systems, and most crucially communities and advocacy organizations are needed for their expertise on actually living with algorithmic systems and encountering their impacts, the methods for assembling these disparate forms of expertise into an assessment process will need to be worked out through practice and content contestation, but policymakers can create the framework through which these perspectives can come together.

Stakeholders can wrangle with each other, and the interests of those who operate agnostic systems can be balanced with the public interest, which brings us to today's discussion about the different forms of algorithmic impact assessment proposed in the U

S, EU, and Canada. These proposals have a fairly wide range of accountability structures, but they're all facing a similar challenge. How can reporting about the impacts of algorithmic systems provide a, some openings for protecting the public interests? What will it take to ensure that developers do not have a totally free hand for determining how their systems will be measured? And with that, we will return it to our panel discussion.

Um, so once again, thank you panelists for joining us for this discussion. Uh, my first question is for, uh, Sarah, um, as I mentioned earlier, the EU commission recently published the first draft of a major regulatory framework for AI called by the shorthand, uh, the artificial intelligence act. Uh, it includes an obligation for developers to, uh, assess their products for risks of harm, for risk of harming core human rights. Um, can you share with us some insights into how the EU and villages, the process of moving from early rulemaking to the creation of actual measurement and accounting practices, um, and whose voices will be included in that process?

Sarah Chander: Yeah. Thank you so much for the invitation and for the question. Um, so following up from what you just said about all the different types of impact assessments, I really have a feeling that the EU's legislation will add even more complexity to this question about what, um, what they, what are constitutes or impact assessments and all the different forms that they can take. And specifically, because I think that that the EU's proposal is really quite limited in many regards, especially echoing a lot of what you've said about the determination of industry and how far they get to set the terms of these prior assessments of algorithmic systems. Um, but before I go into that, I would just say that the, for those that aren't really following the EU's policy debate, essentially what we've seen over the last few years is that the European Union, which is the regional block, um, comprised of a number of governments has basically set, uh, its own regulatory proposal, uh, to regulate AI.

And the context in which that's happened is that we have had sort of, so two draw objectives, um, that the EU has proposed to try and meet with this piece of legislation on the one hand being promoting the uptake of AI and regulating the market for it, or called the EU single market, but also on the other side, protecting people's rights. And just to have that context. But again, I think many of the things that I will say will allude very much to that tension between those two have dual objectives. Um, but then a little bit of clarification. So the proposed rules in the use, um, artificial intelligence act, which is still at the stage of proposal, it was released in April, essentially. There is really not a, such a requirement on developers to assess for risk of human rights, essentially, they rather than using the term impact assessment, they, um, use, uh, the term risk, uh, what they, it's not necessarily limited to human rights, but, um, under the proposal providers or developers of AI systems, I'm using the term providers, cause that's the 10th and all users, um, they have to create, uh, a risk management system by which they have to assess the potential risks, um, that, um, the AI system that they're developing might pose.

Now, uh, this is very broad and I'll talk a little bit about like the, the problems with that in a minute, but first, like the first step before you even get to this provider self assessment aspect of it is that inherent in the ease, um, framework is that the European commission itself, which is the executive arm that like bureaucratic arm of the European

union has basically decided that it will regulate only high-risk systems. So we see this in other contexts this term, but it's a very centralized determination of what is a high risk system. So determined by this civil service by this executive arm. Um, well, obligations on developers for AI systems are only for what the commission, the European commission determines determines high risk systems. Essentially it spells out a couple of use cases by which what constitutes high risk systems, um, AI systems, uh, intended to be used for real-time or post remote biometric biometric identification of people is one of them.

So facial recognition, uh, is one example for that using public spaces, but also, uh, risk assessment systems, um, measuring the risks on individuals and also use of AI systems that determine access to education. For example, these are some examples of these cases that the European commission has done determined high risk. And we've seen how that a lot of these case studies have played out in already like deployments that have been, but across the rest of the world. So only the type of rules I'm going to explain now apply to those type of systems. So it's already predetermined what rules apply to which, and therefore risks, which type of risk assessments have to happen only with these type of systems. Will the rules that I'm about to explain I'm applying then? Yeah, there is something called a conformity assessment, which I think is probably closest to this idea of impact assessment.

I can't explain their role in detail because it gets a little bit technical and there are a lot of them, but essentially under this conformity success assessment providers of AI systems have to do a bunch of things. Firstly, they have to put in place a risk assessment system. And the EU says that that should be some sort of continuous, it's a reflective process. Um, it goes on throughout the entire life cycle of a high-risk AI, the system, um, and it requires, might require systematic updating, um, the provider or the developer must identify and analyze the risks, but it doesn't necessarily speak to human rights generally, but any risks, um, if the system is used in coding to its intended purpose and under conditions of reasonable foreseeable misuse, um, other requirements that the providers have to put in place are things like, um, having, taking certain, they take evidence practices, a lot of language about having relevance, so relevant design choices, formulating relevant assumptions, et cetera, that the developer must examine possible bias.

They must show Dr. Technical documentation and record keeping. Um, they must ensure a level of transparency to the institution that's deploying the system. The system must be designed to allow for human oversight and the systems must be robust and accurate. And the levels of accuracy, um, should be declared alongside the company instructions of use and that the provider must give to the deploying institution, or whoever's putting the system into practice. Now that that's a law and that's already quite technical. But in essence to say, I think what we have to think about, Hey, is this is putting forward a very specific model of risk assessment, um, to your point, Jake, about industry. I think this is probably one of the best examples of how you can really concentrate this question of impact and measuring impact and risk with industry because under the EAs process, for the most part or high risk systems, there are a few exceptions, but all high risk systems are self assessed for conformity by the providers themselves.

So not only do we not necessarily say that much about human rights, but actually all of the potential risks that, um, we could encounter with these systems are basically to be foreseen and to be mitigated determined and defined by the provider themselves by the person or the company making profit from the system red flag already. Secondly, like there's questions about the metrics that the EU puts forward. So like I was saying before, things about relevance, things about accuracy, even things about bias, these are metrics and maybe on the face of it seem objective, but actually I think they are very subjective and practices. They can mean many different things in different contexts and they require a value judgment by the provider themselves, by the person that developing the systems. They also, I think, and I think this is something that we've seen with other impact assessment systems, all of them perpetuate an assumption that the harms that we might foresee, um, in this process of impact assessment or prediction can be mitigated by the provider themselves and that they can be predicted by the provider themselves.

Um, back to the point about, um, I think the term was like you had in your report, a specification dilemma. What about their systems that when used as intended, they harm certain people in certain communities, this whole process doesn't speak to that at all. Actually all systems can be fixed and all systems can be fixed through technical technical means under the E's process. There's also this whole issue about developers. Um, focusing on developers is one thing. And I think maybe pointing to this question of who has agency, this might be valid to some degree, but also we probably we've seen with many of the biggest algorithmic homes cases that we've seen across the world. Many of the harms relate to the context of use the context and with these systems are deployed and that those types of harms and whether or not we can predict those types of harms will be very different depending on which police force uses them, depending on which educational institution is deploying them and the context, the harms and the potential implications allow you to be very different in, depending on the context abuse yet all the responsibility is with the providers, those developing the system in obstruction, basically.

Um, the last thing I'll say related to that, to the use processes that we see that, that this focus on risk very broadly is not necessarily exceptionally compatible with a broad conception of home. So by centralizing the determination with the provider, we also see that we're obscuring your broader questions. Yeah. Like who gets a say in what the potential harms of the system are that often neurology knowledge is experiential it's lived and yet there's process with pretty much all of the emphasis placed on the D the people developing the systems and profiting from these systems. I mean, in essence, make this a concentrated and top-down process, the people whose like whose voices and who I'm like this question about like, whose voices are heard in that question of determining harm, and even does this allow us to prevent, um, um, I think that when you then become obscured by those harms, that sort of narrowly determined by providers and private companies, essentially, we also miss questions about what about harms that are not just about individuals, but also experienced by collectives and also society as a whole.

Um, and also what about those harms that are not just, um, it can be carriage, can be not just those that can be characterized by infringements of rights well, by impact. So for example, um, predictive policing systems are increasingly foreseen to be used by the

European union taking very much, a lot of inspiration from the U S how far to harms like the systemic ties over policing of racialized communities for him to impact assessments, particularly if they're done by the people and the companies developing them, the systems themselves, or even broader than that. Um, we are really learning very much from the work of Cedar Garces and I guess by lane and also many others, um, like how do, how does the increase result to algorithm systems? Um, how has that fundamentally transforming public institutions and can that be seen as a harm? So we're increasingly seeing the economic impacts of the promotion of AI decreasing the capacity of public institutions, uh, increasing the reliance of privately developed services.

And especially when we're talking about things like educational provision and private and public and policing and things like this, we're essentially having huge consequences from democracy. I think I'll leave it with that because essentially what we're doing is we're coming back to this underlying ethos of the ease regulation being to facilitate the single market, facilitate the market for AI in Europe, essentially the freefall flow of AI as a product or a service, um, and therefore promoting the uptake of AI without really any framework to speak to the harms, the broad, very broad harms that might come with that.

JM: I think that's a great transition actually for Brittany's question, um, which is about, uh, how regulation of algorithmic systems might be divided up, um, essentially by regulated markets. Um, so, uh, Brittany, um, in both the EU and the U S we see the responsibility for handling impact assessment reporting, uh, is based on a regulated industry. So for example, a finance F a financial service company that use algorithmic tools, um, in the U S I would have those tools [00:28:00] regulated by the securities and exchange commission, um, not by a general purpose algorithmic assessment office. Um, is that likely to be a successful model? What about like the specialized expertise you need to, to understand algorithmic systems, um, what might be missed if these laws are constrained to already regulated industries, um, that are controlled by their specialized regulators?

Brittany Smith: Sure. So I will attempt to answer, but to do that, I'll go on a little bit of a journey first. Um, so with the algorithmic accountability act that you have already referred to, uh, which was introduced in 2019 and may be reintroduced again soon, uh, if it's passed, it would direct the federal trade commission to require large firms to conduct assessments for new and existing high-risk automated decision systems. Uh, the bill focuses on algorithms that make decisions or facilitate human decision making from sensitive information. And that would include analyzing or predicting sensitive aspects of people's lives like their work performance, their economic situation, health or personal interests in companies would have to assess whether the high-risk automated system in question poses a significant risk, both to the privacy and security of consumers. Uh, and if it results in inaccurate, unfair, biased, or discriminatory decisions impacting consumers, neither the house, nor the Senate bills advanced in 2019, but reintroduction this year might play out differently.

Um, as the problem hasn't gone away in any form. And I think most importantly, as Rouhani has noted perfectly that black communities are already living in the future of a tech Sophia when it comes to policing, or when it comes to all kinds of algorithms that are making life and death decisions for people. And on top of the steady drumbeat of

tech driven harm and abuse, the Biden administration has relevant executive orders and policy ambitions around science and technology. So the environment is absolutely still conducive to tech accountability work, uh, which let's not gloss over this either is being led in part by black legislators. So, but one specific way that this act has been imperfect as other people have pointed out, is that it could slip into creating hierarchies or risks that are disconnected from the lived experience of people on the receiving end of the harm.

And to resolve this, I think we need tighter and more specific definitions, and we need more clarity on impact assessment methods, especially to ensure that they're done in the public interest and prevent them from being co-opted by companies who are more than happy to grade their own homework. Uh, and so I think organizing regulation around industry-specific applications or creating categories of risk based on the use case or type of data involved, I think it's interesting at the very least, because it helps us as researchers and advocates to move out of the conversation space, where people are saying that AI is simply too complex regularly. It's just literally not, um, dev Rajiv MIT tech review piece from last December, um, on how our data and code systemic racism makes this point perfectly. She wrote that the machine learning community continues to accept a certain level of dysfunction as long as only certain groups are affected.

And that's top of mind for me, when I'm thinking about the trade-offs involved in regulatory approach, when people say, okay, this approach, isn't perfect and therefore none are going to succeed. What I hear is that some communities aren't worth the trouble of protecting in relation to economic growth and our breakthroughs, the legislation isn't perfect, but we also don't want to see legislative efforts to devolve into box checking exercises, um, when these issues are incredibly dynamic. Um, but I think this represents a concrete step forward that we haven't really seen yet, and like crucially, our governments have a responsibility to try. Um, so now I'll get to the question. Uh, I'm not sure yet what I think about something like a general purpose algorithmic assessment office. I think on one hand, as in this case, the FTC is remit is already very broad. And I think the debate is often about whether they have the tools they need to enforce powers they already have, or that they are given in new legislation.

It may not be that specialized ultimately to understand who built a thing, what they're using it for and who they're using it. On one of the other hand, I get that this is hard to work out when there is so much specialized knowledge per industry. And so many industries are being encouraged to adopt AI, which is a whole different panel like AI for the sake of AI, or again, for the sake of pearls or competition with China, uh, is worth assessing and detailed because it creates a permission structure for harm that we obviously shouldn't accept. And I think that's all the more reason why we need to enforce asking basic questions about what did you build and why did you do it? And who did you consult? And [00:33:00] I think the answers will be very telling regardless of who those answers are sent to. Um, uh, although I do think they should be public, but the answers will be very telling. Um, and so will the non events that are harder to measure, but that might, uh, absolutely happen where having to disclose this information precludes harmful project from happening at all. So I'll stop there.

JM: Thanks, Brittany. Um, uh, for Fen, uh, as, uh, we have, uh, as we've already discussed, establishing accountability, relationships is a critical aspect of any impact

assessment process, um, who reports to who is just as important as what gets reported. Canada has had some challenges in implementing its directive on automated decision-making, uh, particularly around the matter of actually getting federal government agencies to conduct, uh, the impact assessment process required by [00:34:00] that rule. Um, can you explain how the reporting structure is supposed to work and how it has broken down? And to what extent do you think Canadians, um, can actually make use of this reporting and transparency requirements in this rule to exert some power over the algorithmic systems that the government uses.

Fenwick McKelvey: Thank you. And I just want to begin by acknowledging that I am right now in Montreal at Concordia university, and it's also located in unseeded indigenous land. So the Ganagehage nation is recognized as the custodians of the lands and water on which I gathered today. Uh, Jojohoge or Montreal is historically known as a gathering place for many first nations today, it's home to a diverse population of indigenous and other peoples when we respect that continued connections with the past present and future in our ongoing relationship with indigenous and other peoples within the Montreal community. And I, and I say this, and I begin this because I think this is fundamental. I think a really important part of what to emphasize about the possibilities of thinking about algorithmic impact assessments and how we push forward, our understanding and our literacy about AI and particularly point to the work being done at the indigenous futures group here at Concordia university and particularly their indigenous protocol and artificial intelligence position paper, which I think it's really critical for, for describing, at least for me, how questions of relations and thinking about these ideas of relations to technology can really, I think also be a useful and algorithmic impact assessments as a building on that indigenous epistemologies.

And I think that's one clear way forward for me and something that I'm really thankful for that my own work in the internet demons has kind of found a connection to some of this work with. So I think that's, to me, I think at some of the forefront of the field and a part on maybe you've returned to, I think, you know, specifically, I really liked Brittany's comment about talking about the, uh, permission structure that enables harm and the use of persuasive harm from AI. Because I think Ken is an interest in case of a, of a country that really trying to see if itself as a middle power has doubled down on AI and AI as part of an industrial strategy. So I think when I'm talking about the content context of AIA is I think it's important that this is something that comes as part of tremendous investment on the part of our government in artificial intelligence.

And in particular ways that, you know, is kind of dubious it'll particularly in the sale of element AI to service now in the United States, you know, which basically meant huge investments in public resources. And then, uh, the fact that that basically didn't live up to expectations and it really point to the work of Anna Brenda show at McGill university. Who's got a great report about just documenting the amount of investment the Canadian government has done, uh, in artificial intelligence. And I say that context is important because then it helps us understand what's going on with this algorithm again, impact assessment, automated decision making, because the calls and the push for better AI regulation come from the fact that we see our federal government making these investments in AI. And we don't see the same levels of investments in governance structure or regulation for AI. And that's, I think particularly important because the same concerns that we might have about where AI is being deployed and echoing the work of

Rudolf Benjamin about the idea of the new Jim code, where these technologies are being experiments are developed on.

I mean, we have concerns about being used in policing at the border, and it also went to some of the earlier work in Canada on top by Tom kudos, Cardoso talking about just the risk assessment being used in the Canadian prison system and how that is discriminatory against our indigenous peoples. And so really there's a context here for thinking about better algorithmic, accountability and governance. Well, before billion dollars investments in AI, the overarching strategy of the Canadian government largely has been to model best practices. And so it was an early adopter and using algorithmic impact assessments to change procurement for AI at the federal government level. And it's important to recognize that like to EU, you know, the Canadian federal government has limited agency over the provinces. And so this was I think, a very con a contained at the federal level at the government level and trying to inspire companies to adopt this AIA toolkit, which, uh, Jake alluded to and to see that being picked up.

And what it means in practice is that it's kind of twofold is that departments looking to adopt an AI system or algorithmic systems should undertake is AIA tool, which is public online. And that vendors should also be checking to make sure that the departments they're selling to have undertaken that AIA. And the expectation is that these AI's are public and that they're being put online. And so what has come up is that the AIS are premiere primarily a risk assessment and very much very standardized form, which is interestingly and open source. And so there's a lot of credit to be given to the work being done about open government, no government initiatives that have led us here, but also I think kind of a disappointment about the modesty of this proposal and the fact that we haven't seen an uptake. So thus far there's concerns that the treasury board hasn't been able to require other departments to take up and use this tool.

So we're not sure at the level of which other departments and this came out. So some of the reporting and the globe and mail that the department of national fence was implementing AI to help ease a TA diversity hiring issues, which raises, I think a deeper question of why, and when AI is being used to do address problems of diversity, uh, is that the cheaper solution, um, and that, you know, there wasn't an uptake in the government to use us and that AI tool wasn't being done. So in the first part, what we see is that there's a limitation of who's actually adopting this tool. The second concern, I think the one for me is that there hasn't been a corporate buy-in from anything I've seen here. So this tool is being out there, it's put forward as best practice. We don't necessarily see corporations.

And I'll talk about that a bit later, we don't see corporations actually doing these AI algorithmic impact assessments. So there's also this question of whether the soft touch approach is actually effective and particularly the modesty of it. The fact that we haven't seen a buy-in for something I think, which is so, uh, simple to do, which is, you know, really a box checking exercise and a critique of it, you know, raises real doubts about, you know, the effectiveness of algorithmic impact assessments. And I think, you know, the final point, and this is something that my, uh, work, my student, um, Nick Gertler is working on in their, their master's project is actually looking at the structure of that and how really the EIA is something very equivalent to what Sarah was talking, but as a risk

assessment. And how do we, you know, think about a nebulous term, like impact is also very nebulous in comparison to risk.

And, and whether we can think of this tool and this kind of questionnaire is actually a proper tool to put forward, um, matters of kind of assessing risk or kind of documenting that situation. And I think that really speaks to the, the, I think the limitation of the Canadian public at present is that there was a hope that we'd be seeing these risk assessments. This should be sparking journalism and commentary and public discussion about what the impacts were being used. And I think a wider concern about where, you know, if you're using AI, why apply it in certain high risk applications like hiring border security justice, and that I think would have really, you know, cultivated a conversation a bit more. And I think the disappointment, the lack that you don't see these tools being done is that it really, I think is speaking to a deepening concern about public literacy on AIA as wider speculation of overstating, the potential and possible applications of AI, which really put it in a very precarious policy-making context. Cause it's easier to talk about the, the, the deep part, the, the, the abstract terms that, that rather than the kind of real and particular harms that is algorithms and AI might be doing presently. And so, I mean, I think that that's kind of the, the sad context of Canada, but I think one that really tampers how, um, we, you know, we can understand the impact of algorithmic impact assessments here.

JM: Great, thank you. Uh, is very illuminating. Um, and I think that's, that's a great lead in, um, for asking Brittany, um, about her experience on managing impact assessments in a corporate environment. Um, so drawing on your experience of managing human rights, impact assessment, um, previously, uh, what do you think are the major challenges in conducting them, especially with regard to integrating the experiences of people most impacted by algorithmic systems. Um, and, and what do you think, like, what do think about the matter of knowing when an impact assessment is successful or even adequate?

BS: Right. Um, I, and I thought the questions about human rights impact assessments in the Q and a too. And so I will see how much I can do here. Um, but I think there are tons of challenges, uh, and, um, the work is super rewarding, but in my experience, it was also extremely difficult for out for reasons that we've already outlined so far. I think to start even making the case for this kind of work, uh, is not easy. When you start trying to operationalize the rhetoric of ethics, you start getting a lot of pushback. Um, and that's all I'll say about that point. Um, and I think actually in my experience that pushback manifests in the lack of resources where companies are allocating budget and headcount is a strong signal of what they care about. And I was a team of one, like less than one extra because running human rights like that, even my primary job, um, this is not something you can just tack on to the policy team or the team it's a different job entirely.

Um, another challenge is about how accountability is structured in a corporate environment. And I mean, this both in terms of designing and executing the criteria we use to do an impact assessment and who is responsible for the outcomes, um, of that assessment on the criteria question, this is really difficult because as we've already said, this is entirely vague and allowed to be determined by individual actors, even, uh, in the business and human rights space, where there is a clear set of guidelines and

expectations from the UN, it's still confusing how to determine whether you and GPS apply or don't apply to your specific situation. And many best practices are written with hardware or applications in mind. But for me, the question was, uh, how does this work in a technical research environment where we're talking about fundamental, theoretical research and a possible harm that might occur when the research is combined with data sets or compute that don't even exist yet?

What are we supposed to do? And I'm not precluding, uh, not researching when those are non-existent, but, uh, this is also a part of the problem that the research train is happening. And you're expected to come in very late into the process and try to find a way to assess the impact. Um, the UN's B tech project is working on this exact same thing, as I understand it, like they're trying to develop authoritative, um, guidance and resources for implementing you and GPS and thinking specifically about their application within tech. Um, I'm almost done on the responsibility point if maybe obvious, but we haven't incentive alignment problem here as well. And that's because when you're operating under the same org level mission and working on the same project, you might still have misaligned incentives. I'm here to identify and mitigate risks, including potentially canceling your project.

And you're here to get it done and publish it as quickly incompletely as possible. Um, and what was not happening at the time that I was doing this work is the consideration of human rights impact of the work rather than, or fringe or additional process to go through before the work is published. I also had some curious experiences trying to get people to think of, um, human rights, impact assessment processes as generative, the impact assessment and human rights due diligence processes can help us uncover ways. I work my actually advanced human rights. Um, I worked a lot with our team on interpreting the right to share in scientific advancement and its benefits and ways that human rights impact could guide our research planning, help us prioritize positive impact, determine which areas to research and so on. And I think the vagueness and the lack of clarity at the intersection of these skills, doesn't have to be a doom and gloom conversation, but there's room for creativity and exploration too.

JM: Thanks, Brittany. Um, I have one last question from me to panelists before we move into Q and a, um, and, uh, I'm going to ask it, uh, to Sarah. Uh, and even though it's an expansive question, if you can keep it brief, then we can move on to Q and a. Um, so, uh, how can civil society organizations like Edgery, um, and community advocates put pressure on the legislative and rulemaking process, uh, behind algorithmic accountability laws and assure that they actually have a voice in the table and how assessment gets done? What are some like concrete asks, um, and some, some calls to action that we can be engaged in, uh, in the near future to ensure that, um, the voices of people who are most effected by algorithmic systems, um, are integral to the assessment process?

SC: Yeah, definitely. I'll do my best to answer quite complicated question quickly. Um, so yeah, then I think the first things first, it might not seem so concrete, but I think it's essential to do something meaningful is actually recognize that there is a real risk that if impact assessments, however, they, in whatever form they come, if they're very centralized, if they're exclusive very exclusive and don't involve socialized communities, other marginalized communities. And also if they have no real teeth, um, as to like

fundamentally change how systems operate or actually even stopping their use to Brittany's point, then essentially we have to recognize that they might have an enabling function. And also there is an ongoing process of industry co-opting these tools basically, um, further the uptake of their products and services basically. And we need to really take seriously this potential performative function of impact assessments.

If we are actually then to go on to make something useful of this. Um, I say this because we have to think about impact assessments also in terms of the burden of proof question. So most, very institutional processes that focus on proof of harm, whether it's potential or actual, they definitely end up concretizing this burden on communities to prove, um, and document harm, whether it's potential or actual. Whereas I think that there's like a, of scholars that have already worked on this and we have Benjamin, um, do you then the algorithmic justice league [inaudible], they've all talked about in some regard this idea about the fact that electric systems like any other opaque systems imbued in systems of power and structures of power, they're more likely than not to continue existing patterns of exclusion and oppression. And so we, I think a civil society really need to have think about the claims that those deploying AI systems are making and producing our systems are making, and also have a question about efficiency.

So most systems are optimized for efficiency, right? However, what does efficiency mean, particularly again, coming back to this question of the specification? Yes. Dilemma efficiency for some means injustice for others. And so, whereas like a D a system might be measuring impact in terms of how well it's optimal for efficiency. In other cases, that's talking, we're talking about optimizing injustice, basically. So all this to say, I think that civil society really needs to, first of all, have a question about how far can impact assessments be a mechanism for real redress and accountability. Um, thinking in the EU context, we need to, like, we need to have discussions basically about whether we are very, we have a lot of faith in them, um, impact assessments, which are focused on providers before the market stage, as opposed to developers, uh deployers before the use stage. And if we can think about that or pivot in the EU context, there might be ways to actually start harmful practices in use.

If we actually then don't put too much a burden on civil society to have these really like evidence concrete, um, things that they're putting, putting forward to show harm that actually hasn't even happened yet. And even in context of structural discrimination, even when such hands have happened, yet, we have very difficulties to evidence it because structural discrimination is not an easy thing to evidence. So I think there's a lot of questions and we need to push role, um, reversing the burden of proof also my last point. And I'll be very quick. It's like, I think like very frankly, we need to not put our eggs in one basket and think about other strategies beyond impact assessments. So again, coming back to the amazing work or [inaudible] algorithmic justice league, carceral tech, net tech resistance network, all in the US have really thought about this question of abolition and applying this concept of abolition to tech, not just talking about, okay, we need to, we need to stop all tech in certain angles, but also okay.

What our arguments for alternatives for such systems, what are the criteria and for the democratic engagement, can we put on governments to actually both were a process of justifying to us as communities, and as people affected, show us what this will bend, how this will benefit for us before use, rather than putting something into use, and then

asking us to prove the harm after the fact once the harm has already been done. So switching up all these presumptions, I think, is something that we really need to think about. Um, a lot of Andrew's work has been focused on this question of prohibitions too. And we actually like forced the European committee there you're being commissioned to put in a prohibition, um, provision in its legislation. So there is also, I think, some hope that can be done by advocacy and pushing for as strong as legislative processes as possible.

Um, but also I think in that process, and this is for all, this is a recommendation from me from all of academics and civil society that are working in a digital space or that I've had expertise on tech is that we really need to do better work to situate ourselves. Like what we're talking about here is not necessarily technical it's about structural harms and the people that know best about those structural harms are the people that work on structural discrimination, the people that work on climate justice and re situating ourselves, not as experts on the topic, because it involves tech, but engaging people that have that standing experience of working about how to dismantle [00:53:00] structural inequalities and oppression. That's I think where resources need to be shifted to, to, and attention needs to be shifted to as well.

JM: Great. Thank you. Um, for Fen I'd like to, um, give you the first, uh, audience question, um, and specifically like in reference to the indigenous protocol, that excellent indigenous proposal, uh, protocol article that you shared, uh, earlier. Um, how can we, uh, so this question is from Catarina, um, uh, do any of these proposals envision addressing, uh, cumulative harm, especially to already disadvantaged communities, um, and how, like this seems to be one of the shortcomings of focusing on high-risk processing? Um, it also seems to be a methodologically, very challenging. Um, so just briefly as briefly as you can, um, what would it look like to actually do assessments to that considered cumulative harm and not just constrained risk measurement?

FM: Yeah, I, you know, I don't necessarily have the answer for that. And I think that the, one of the things to me is that I, I wanted to emphasize is that algorithmic impact assessments remain in opening. And I think that that's all that's, what's in play right now. That's the politics at stake is that it very easily could be closed and it very easily could be turned into something where, um, it's instrumentalized in such a way that it, that it loses that opportunity. And I noticed also in the chat, there was lots of discussion about, say the re the idea of a rights-based assessment. And I think, you know, when I think of what the limits of my knowledge and where I'm questioning seeking, it's like the neatness of a right space approach to talking about specific human rights courts. But I also recognize that that's a very, um, abstract European tool and one that also has a CA capacity of abstracting it out.

If people just have these intractable rights, you don't actually need to do the consultation or the work to understand those relations. And so in many ways, I look to say the indigenous protocols and thinking about relations and that, that work of relations and something I'm learning. Uh, I also would emphasize that, you know, some of the work in feminist standpoint theory, and, you know, I look to Louisa Moore's work here and talking about partial accounts and trying to say of like the algorithmic impact assessments as a specific tool, and to get into, uh, you know, Catarina is I think specific point I, you know, this gets into, you know, what do we mean if the impact assessment is

cheap and quick versus expensive and fulsome. And I think that, you know, often what I'm most aware of now is the fallacy of participation and the idea that if participants aren't compensated and rewarded or allowed in such a way that the, the consultation process itself trying to address, you know, what I, I wouldn't, I don't know, the term is welcomed a lot of harm, but address is just the kind of disadvantages that participate and create barriers in the policy process presently, unless there's actually building that in then many of these measures are kind of failed to begin with because there's not any way to have people actually meaningfully participating.

I think that that's the greatest risk in many ways with all this consultation is unless you have think fair compensation, making sure there's time and accountability, you know, childcare, even these basic things. Those are things that really [00:56:30] are the barriers to participation, unless they're kind of front and center of the design process for these consultations, then it's not actually addressing these cumulative harms. That kind of rate is talking about. I hope that answered your question.

JM: Thank you. Um, there were a number of questions from folks about, um, whether human rights is, um, uh, like an operable framework for thinking about, uh, impacts. And so, uh, I'd like to pose this to Brittany first, but, and then also chime in also, um, our human, our human rights frameworks, uh, an optimal way to sort of ground and impact assessment. Um, even one, it's not like specifically the human model on human rights impact assessment. Um, what, what are some of the pluses and minuses around adopting that?

BS: I think, um, I think, uh, a not ideal answer is that I don't see any reason why not. I think not enough work has been done on any of these issues so far to say, like we have to rule that out entirely for some big reason. I think that using human rights as the basis for an impact assessment, whether it's an HRA that's guided by UN GPS, um, is it is a totally viable and useful and productive way forward. My only concern in my experience is that people who people have a very, uh, little knowledge about what human rights are. I spent a lot of time trying to convince people that just because their research wasn't torturing anyone, that they still needed to pay attention in the meeting I was running, or that there was still an impact that we needed to consider that human rights are much broader and expansive than let's not torture people. Let's not surveil people, et cetera. So, um, I think like basic education on what human rights are and how they apply to the technology that we're building is a big part of the discussion and is not currently happening in any meaningful way.

JM: I think maybe we have time for one last question, um, from, uh, Andrew selves, uh, he asks, um, if industry actors are the ones making technology, it's hard to imagine a regulatory structure where they don't have a degree of discretion in what to assess and report. Um, oh, sorry, question moved there. Uh, um, which will allow them to control or undermine the regulation as they need. Um, do any of the panelists have thoughts on how to design regulation to reduce the degree or importance of that discretion by private companies to keep the backend of their systems private?

SC: I can say something really quickly, which is that what we're seeing from the regulation is that part of it is about the metrics. And part of it is about also the emphasis. You play some on developers, technology companies in the, in the legislative

framework itself, right. And I think there is a question around moving towards specific metrics, if you are going to make certain requirements and have that level of responsibility to providers. However, then I think you also have to acknowledge that, for example, where the question of bias in the EU's legislation, we have like a sort of obligation for providers to measure bias. But I think that these sorts of regulatory tools not always best place to put, um, specific specifications of how, what the obligation on a providers should be in terms of bias. There are many forms of bias. There are many sort of homes that can be addressed in that context.

And there's so much complexity there that we have to question whether regulatory ways of the main ways of making certain providers do certain things. I would really say my one way is to not, if the question could be like to not put the sole responsibility on providers, self assessing, I do agree that there's an implied in the question that there is a level of discretion, maybe that there are certain things that only providers can really contest, but I think it also upskills the broader socio-technical question of harm and impact, which doesn't necessarily always emanate from how the systems are designed only, but also the context in which they are deployed in. And for this, you really need to think about other mechanisms which also take into account obligations on what the person, what the system, what the institution deploying those systems, um, should do. Should they engage them, marginalized communities? Should there be an obligation not to curse certain system and use of certain conditions are met things like this, and at least in the UC system, this whole area of sort of criteria and legislative options is really missing.

JM: Um, thanks a lot of Sarah. So, uh, with that, um, I'd like to, um, wrap up by thanking the panelists, um, Fenn, Sarah, Brittany, and then also, uh, Rigo and the rest of the production team, uh, for making this really interesting, uh, panel happen. I know, uh, assessment and measurement are not necessarily always the sexiest topics. Um, uh, but I, I hope that we, uh, made clear, uh, the stakes involved in, uh, having these questions about, uh, who is involved in assessment. Um, and that matters just as much as the what, um, and that can push legislators working on these topics, um, to attend to that more closely, um, keep your eyes peeled, uh, for at, to the end of the month for the report. Uh, suddenly accountability coming from Devon society's AIGI team. And, um, we'll see you all around with that. I'll hand it off to Rigo.

RLG: Thanks everyone. And thank you to the panelists. Uh, you can stay tuned with our work @ datasociety.net. And with that enjoy the rest of your week. Everyone have a good day.