AspenAIElections_March2024

Fri, Mar 29, 2024 1:05AM • 2:48:16

YouTube Video: https://www.youtube.com/watch?v=GGH2O6gkUD0

NOTE: This was generated by Otter ai and may contain mistakes. Check any quotes to the video.

SUMMARY KEYWORDS
ai, elections, people, information, companies, platforms, disinformation, democracy, tech companies, fake, content, talking, artificial intelligence, taiwan, happen, country, hear, understand, tools, regulation

04:42
Hello, and welcome. I'm Karen yar hamilo. I'm the Dean of the School of International Public Affairs and the addley Stevenson Professor of International Relations. Well, I cannot think of a timelier or more important topic to discuss Then the role of artificial intelligence on democratic elections around the globe, and we are thrilled to be partnering with Aspen digital on this effort. Over the course of the current year, about 2 billion people across 80 countries are expected to go to the polls to cast ballots. Carrying out free and fair elections is difficult even under normal circumstances. But in the current information ecosystem, it is even more challenging. As we know, the health of a democracy is linked to the integrity of information, and those who wish to spread lies and myths and disinformation have never had the tools to do so at the speed and scale artificial intelligence allows. No wonder that democracies are facing stiff headwinds as autocrats from Moscow to Myanmar challenge the geopolitical order and are using AI to tighten the grip on power. And I vividly recall why what Maria ReSSA one of IGPs Carnegie distinguished fellows told last year's graduating class of SEPA students. She described the lack of government oversight on this technologies of these technologies as dooms days clock for democracy. And we've already seen disturbing evidence of the dangers posed by AI to the integrity of elections. In Slovakia for example, there was deep fakes audio of a candidate who appeared to be conspiring to rig the election. In endo in Indonesia, deep fake technology was used to portray Schwartz Schwartz Suharto, the country's long standing dictator who's been dead for 16 years telling people to vote today. And this is only the tip of the iceberg. These misuses of technology are all happening as governments and tech companies are scaling back the resources to police fake content, and fight disinformation. Today's panels will examine their four topics, including risks to 2024 elections worldwide lessons from recent global elections, and role of responding and responsibilities of the tech companies. We'll also talk about implications for this year's US elections in particular, and I know many of you are interested in this. We are fortunate really fortunate to have so many of the world's leading experts and practitioners on this subject with us here today. I am hopeful that today's discussions gives us some reasons to be optimistic for the future of AI and democracy. Although knowing some of the panelists, I'm not sure about the optimism part. Were already seeing some signs of progress. And we should mention them Europe has passed promising new regulations to rein in tech companies, and increased transparency and accountability. And here in the United States lawmakers in 32 states have introduced 52 bills to regulate deep fakes in elections. And we have a lot of work to do here in the United States

and elsewhere. And I hope that this form this conversation will move us forward in thinking about the problems, the challenges and the potential solutions. So before I turn it over to our co host, I just want to thank everyone again for coming and participating in today's panels. I want to thank the staff organize this. Look, this event is the embodiment of IGP. As an organization, we are discussing here important topic that crux that cuts across so many of our global challenges. Here, it's HIPAA, we're focusing on technology and innovation. And this is a big part of what we will be looking at and doing. We're looking at democratic resilience. And here this is exactly in the intersection of these two. And I'm very excited about the projects we have coming up in the pipeline, looking at those with our faculty, with our fellows, and definitely more to come. This is again, we're doing this bringing together the best of the private and public sectors to generate new ideas and solutions that are based on data and evidence. And that is why the the Institute of global politics was created. This is exactly the vision that Secretary Clinton and I had when we started this. How do we take this on? What can we do to help and this convening cannot be more relevant and important today. And with that, it is my great honor to introduce our co host of today's event, Vivian Schiller, one of the highlights in the of this work that we're doing here at IGP, is how we are joining forces with incredible partners on meaningful issues. And when Vivian and I first met, we knew immediately that we wanted to collaborate, we wanted to work together, we were really passionate about the same topics. And I'm so glad that we were able to do this and get to this point that we are doing this event today. Vivian joined the Aspen Institute where she is the Vice President and leads Aspen digital, just four years ago, after a long career in the intersection of news media, and technologies with stents that included president and CEO of NPR, head of New York, times.com, Chief Digital Officer for NBC News, and global head of news at Twitter, among other things, please join me in welcoming Vivian Schiller to the stage.

11:34

Thank you so much, Karen. I remember all too well, our meeting in your office. It wasn't even really that long ago, it was in the fall. And it was just this instant spark. And we're like, yes, we have to do this. And well, here we are already. Anyway, I just I on behalf of the Aspen Institute in Aspen digital, we just are so honored to partner with SEPA on this important conversation. It's gatherings like this today that remind us that really a key measure of any democracy is its capacity to absorb and adapt to change. In the short life of this country, which is actually I was reminded younger than the life of this university. People in good faith have come together to meet moments of tremendous disruption, we are capable of it. We are capable of marshalling our ideas to mitigate the bad, and particularly to harness the good that reforms into reforms that reflect and preserve democratic values. It's important for us to remember we are capable of doing this, particularly as we today draw on the example of global elections that have happened to date, in an effort to learn about the AI driven challenges ahead for this country's elections here in the US. In this way, we are really engaged in a quintessential American tradition, forming a democratic response to emerging technologies. So AI advances, it's really important to say at the get go have tremendous, tremendous potential for good. And we also know that those who would seek to disrupt free and fair elections, because they have tried before, are eager to enlist AI tools in their ongoing effort to undermine trust in democratic institutions, to pollute our information environment, and to distract or discourage voters. This is true ahead of this November's elections, and it will undoubtedly remain true for the foreseeable future. So Language Tools, for example, can be co opted to mislead minority communities. audio tools can spoof the voice of a candidate or an election official, to de mobilize or disincentivize voters. Fake images or videos can be used to deceive the public at

particularly critical moments. And to me what's even more alarming is even if voters are not fooled, we now have to prepare to navigate a world where it's easier to completely dismiss evidence based reality. We could be living in a world where everyday people will simply stop believing anything they see anything they hear. And to me honestly, that's more terrifying than anything else I can think of. It is part also of the autocrats playbook. So this is what we're up against. In the words of our next speaker, it will take a village, in this case to build social resilience that resists the pull to become a suspicious and stubborn people. We will have to insist that the truth is knowable and worth knowing. And we need to learn to trust again, institutions, information and each other. The Demak democracy practitioner and experts here today are helping to define what democratic resilience will mean in the face of each of these challenges. And by the way, I should mention in addition to our speakers, there's so many folks I know in the room here who are working hard at these efforts, even if they're not on stage, and we thank you for being here. I am confident that we have what is it takes the all of us the we this country to meet this moment and to secure a democratic future in the age of AI, as surely we must. I am now pleased to introduce our first panel. Our first panel is called setting the stage risks to the 2024 Global elections. And with us, I am so honored to introduce Secretary Hillary Rodham Clinton. She is Professor of International and Public Affairs, Columbia University's of for Columbia University, the 67th, Secretary of State and former senator from New York IGP, faculty advisor and board chair, a very Rova, the Vice President for values and transparency with the European Commission. Hello, Madam Vice President, Maria ReSSA, my friend of many decades Nobel Peace, Prize winning journalist, co founder, CEO and President of Rappler IGP. Carnegie Distinguished Fellow and moderating is the spectacular journalist Gillian tett columnist and part of the editorial board for the Financial Times, I welcome you to the stage.

16:37
Well, thank you very much indeed, Vivian, and the dean for that wonderful introduction, which sets out the issues, I should say that I'm personally absolutely thrilled and honored to be moderating this panel. Because first of all, I am a journalist who cares deeply about the truth, at a time where it's under threat, an American citizen who's deeply worried about the election that's coming up. And also when I'm not a journalist, I'm attached to Cambridge University, King's College, as an anthropologist, trained in digital ethnography, and passionately believed that the only way to handle AI responsibly, is to add a second AI, which is anthropology, intelligence, to understand the social impact. And this is what this afternoon is going to be all about. And I'd like to start perhaps, with you, Secretary Clinton, and ask you I think we first talked many years ago about the terrible threat of misinformation. I happened to have a background in Soviet studies and had seen some of the absolutely nutty misinformation that was going around in the Russian media many years ago, well before 2016, that were betraying you, as some kind of sheet devil out of Indiana Jones, stalking the world, which was, you know, incredibly corrosive. And since then it's got worse and worse and worse. How alarmed? Are you about the upcoming elections? And is there anything that we can actually do to stop this tsunami of misinformation?

18:12
Well, Julian, thanks so much for being here with us and for moderating it. And I know you cover these issues and have a lot of, you know, experience in trying to make sense of them. And I think that anybody who's not worried is not paying attention. There's more than enough reason to be worried about what we've already seen. But certainly, I think, as we're here today, doing this panel and having these other experts and practitioners speak to us, there are literally people planning, how to enter rupt

interfere with distort elections, not just in the United States, but around the world. And so if I could just focus on the United States for a minute. What Julian's referring to I think, was really motivated, by my time as Secretary of State doing the job that I was asked to do by President Obama, representing our values, our interests, our security around the world. And in the fall of 2011, after Putin announced that he would be coming back as president. They had elections for the Duma, and they were so blatantly fraudulent, I mean, videos of people throwing away ballots and people stuffing ballot boxes. This was not made up. This was, you know, very clear, undeniable distortion and interference. So as Secretary of State, I said that, you know, the Russian people deserve better they deserved free and fair elections where their votes would be cast. First and counted appropriately. And having literally nothing to do with me. When the news came out about what had happened in those elections, 10s of 1000s of Russians, particularly in Moscow, St. Petersburg, a few other places, went out into the streets to protest to protest for their right to actually choose their leaders. And it totally freaked out Putin. And he actually blamed me publicly for the reaction in Russia. And that was the beginning of his efforts to undermine and take me down in, you know, very real time, starting before the 2016 election, but certainly picking up a lot of steam and impact during that election. And it was such a an unprecedented, and really quite surprising phenomenon, I don't think any of us understood it, I did not understand it, I can tell you, my campaign did not understand it. There, you know, the so called Dark Web was filled with these kinds of memes and stories and videos of all sorts, you know, portraying me in all kinds of, you know, less than flattering ways. And we knew that, you know, something was going on, but we didn't understand the full extent of the very clever way in which it was insinuated into social media, if it had stayed on the dark web, you know, whatever, you know, maybe a couple 100,000 people would pay attention, but this jumped into how we communicate. And the only thing I can say about it is like say two things about it. One, it worked. You know, there are people today who think I have done all these terrible things, because they saw it on the internet. And they saw it on the internet in their Facebook feed, or some, you know, Twitter this or snapshot Snapchat, that they were, you know, following the breadcrumbs. And what they did to me was primitive. Yeah. And what we're talking about now, is the leap in technology that we're dealing with, you know, they had all kinds of videos of people looking like me, but weren't me. And they had to keep whoever that woman was with her back to the camera enough so that they couldn't actually, you know, be found out. Now, they can just go ahead, they can take me. And in fact, they're experimenting. I've had people, you know, who are students and experts in this tell me that, you know, they're once again, because they've got such a library of stuff about me, they're using it to practice on and see how more sophisticated they can get. So I am worried because, you know, having defamatory videos about you is no fun, I can tell you that. But having them in a way that you really can't make the distinction that Vivian was talking about. You have no idea whether it's true or not. That is of a totally different level of threat. So I think, you know, we're setting the stage in this panel, and we've got, you know, two people who really understand this deeply with our other panelists.

23:26
Well, thank you very much indeed. And I'm glad you went back to that extraordinary story since post 2011. Because I think most people don't know about it. I like most people when I first saw those images very early on, because I do speak Russian. Last, I thought it was so ridiculous. And boy, was I wrong to laugh. It's extraordinary how this has spread and how much more virulent it is today. And I'd like to turn to you, commissioner and ask you. When you look at the problem, you've just had an election in Slovakia, where as we just heard earlier, there was aI used to manipulate the vote seemingly

successfully. Europe has a lot of elections coming down the track this year. The European Commission has taken a much more aggressive stance than Washington in trying to stand up to big tech and control maybe not aggressive enough, but control the angle so your face you can tell us if you agree or not. But at least you've tried to challenge big tech in some aspects of their responsibilities. How concerned are you that this year's election in Europe will be undermined by AI?

24:37
Thank you very much. And thank you for inviting me because this is really an honor to be in such panel. Well, how worried I am I don't have any right to be worried I have to act because I am the European regulator. And I don't know whether we were aggressive but what we did was maybe not aggressive but necessary, because we already have fun Good data good analysis showing that most of the elections in the EU member states affected by at least Russian propaganda and Russian hidden manipulation. These days yesterday, the Czech and polish Secret Services disclosed the data and the facts about the the Russian propaganda and disinformation affecting several elections through domestic parties. Yeah, they need, Putin doesn't, cannot do it directly from his mouse to the ears of me.

25:45
He cannot do it, right. That's not Putin,

25:48
you'll never know. He cannot do it directly from him, his mouth to the ears of European people. But he needs simply the allies in our member states. We had to take measures. And we are taking measures before European elections, which you ask about, for two reasons. We do not want Mr. Putin to start winning elections in the EU, because the purpose of his propaganda is clear. It's a message stop the support of Ukraine. And he knows that we are all democracies. So he has to do it through our people. So the purpose is absolutely clear. So we take measures agreement with the platform's to remove deep fakes in the campaign. So the aim should be very much limited, or at least to label aid. We have measures involving the civil society for enhanced fact checking, we have a course on the independent and public service media to take care of the facts. Because what we speak about is the protection of evidence based truth. And Madam Clinton, it's interesting how we politicians try to avoid the word truth, because we will be immediately accused of having our Subjective Truth doctrine. So when I speak about the truth, I speak about evidence based truths. We speak about effects. And we we really believe that our set of measures might should have a real impact on the campaign's that they will be fair that they will be transparent that they will be free of hidden manipulation by AI. And maybe last comment why I am dealing with that I am Commissioner for values. Can you imagine the shock at the beginning when I got from Musa Fonda Lyon, this portfolio of values? What does she mean by that? And so it's protection of rule of law, democracy and fundamental rights. And I would add also the protection of the evidence based truth, because the destiny of the society which stops valuing the truth, is to live in live. And this is what we don't want in Europe.

28:18
Absolutely. Well, that's particularly pertinent and potent for anyone emerging from Eastern Europe, Eastern European countries. I'm just curious before I turned to Maria, have you been the victim of Miss Information yourself? Oh,

28:32
my God, I can write any attacks. Well, I could ask for on a feed. I see the attacks on both these women, by the way, you know, our attackers just combined. So yes, you've been under? Yes,

28:42
I am. I am under attack many years. That's why I also canceled the Facebook account. And at the end of the year, I might get out of politics. Totally. So believe me, I will cancel everything. I know on Twitter and Instagram. Yes, I am under permanent attack. I was a one out of two women mainly mostly attacked in Czechia. And the other one was, was Angela Merkel. US for a I have no complaint because there was just one case of me having Lara Croft body. I liked it. So, but maybe we should not joke. make jokes about that. Because we see a lot of really harmful things against the girls and women. Yeah, and that's why we also adopted recently the directive on violence against women, which can which contains the chapter on digital violence. It's, I think, the first ever in the democratic state legislation that we are defining that and that is So in AI act, we define these kinds of practices as something which has to be punished. Because we also have to see crime and punishment in practice. Absolutely.

30:12
Well as someone who's worked in the Muslim world a lot, and one of the things that horrifies me is how female activists are being silenced by the use of AI to create pornographic images that are so shameful, that it makes it extremely hard for female activists to continue in that culture. It's absolutely horrific. It really is. Maria, have you been attacked? Oh, my

30:35
gosh. Well, first of all, it's nothing compared to what both these women have had. And you know, I think for our American audience, just Vera drover is not only handling her values, but she also has the portfolio of Margaret divestiture, which means she is the most powerful woman regulating tech right now. Right, that's, and that's part of the reason I switched seriously, the last time I was on this stage with Hillary, I was attacked by her attackers. And every time I'm on stage with Vera, we also get excited

31:09
that we love each other we do.

31:12
We're all gonna get attacked, the minute we get off stage.

31:15
And I think the hard part is you don't know what it's like until you are attacked. And that's part of the reason I would like some of the men from Silicon Valley to actually trade places with us for a day or so. So have I been attacked? Yes, it is a prelude bottom up, you say a lie a million times, it becomes a fact for me, it was 90 hate messages per hour. And then a year later, the same thing that was seeded online, became cases that were filed by my government against me, very slowly, you know, the 21 investigations became 11, criminal charges became only two left after seven years. Right. So we fought

it. But I think the real impact of this and in you've talked about it, but Russia is really the pioneer and and the EU selections, the major democracies around the world are having elections this year, where the EU, where America goes, it's really scary for the rest of us in the Global South, because you're not even acknowledged you're being manipulated. If you're a woman, gender disinformation, is using free speech I II, information warfare to pound you to silence If you are in a position of power, if you're a journalist, if you're a human rights activist, if you're a student who stands up for, you know, this whole thing of woke, like we kind of jumped into it. But there are information operations that seed a lot of this. So this is, this is the world of non the world of lies, let's not even call it it's a world of lies. It's a world of personal personalization. Which is, and I see so many faces here from because because you're gonna hear from David gravich. I see Katie Harbath, who is also in the Philippines. So please ask the questions. But more than anything, you can be attacked. And it's not just about being attacked. It's the fact that we have lost agency. We live in different realities, right? Personalization, when you're talking about buying sneakers is, you know, okay, fine, you're, you're gonna get recommended sneakers, because you look for sneakers, that was a long time ago. Now, personalization means that I will give you your reality, I will give you your reality. But even though we're in the same shared space, we have 100 plus realities. That's called an insane asylum. That is the world we live in today.

33:51
Absolutely. Well, it's no accident that there's four women sitting on this panel right now, because it really is a strong gender issue. And thank you, Maria, for pointing out that, notwithstanding the inward looking nature of a lot of American and European politics today, it's not just a Western issue. In many ways. It's actually harder to tackle in the emerging world right now. Which is, you know, very alarming, but we're going to hear a lot later on about what can be done to counter this. Would you like to share any thoughts? Maria, do you have thoughts about what you'd like to see to fight back?

34:28
I mean, for Americans get rid of section 230. Because the biggest problem we have is that there is impunity. Right? Stop the impunity. Tech companies will say they will self regulate self regulation comes from news organizations when we were in charge of gatekeeping the public sphere, but we were not only just self regulating there were legal boundaries. If we lie, you file a suit. Right now there's absolute impunity in America hasn't passed. Anything I do that the EU won the race of the turtles and filing legislation that will help. It's too slow for the fast lightning pace of tech. Right? The people who pay the price are us, us this young generation. I was just with Vivek Murthy. And you know, the Surgeon General of the United States didn't file his report until May last year. Hillary was probably ground zero for all of the experimentation. What kind of different world would we live in? If he had become president? I mean, she won't say that. But I will write like, I

35:39
would think that. Secretary Clinton, would you agree that the first step is to abolish section 230?

35:46
It certainly is, among the first steps, you know, I think it's very difficult to be as upset with the tech companies as we are. And I think rightly so, since they were granted this impunity. And they were granted the impunity for a very good reason back in the late 90s, which is we didn't know what was

going to happen, we had no idea where they a platform, kind of like a utility, which sent content through it. And therefore, you know, they didn't have the kind of liability and you would go underneath to see where the content came from. Were they content creators? Did they have a duty, either to warn or prevent? I mean, nobody knew anything, because nobody had a real sense of what was happening. Well, now we do. Yeah, and shame on us that we are still sitting around talking about it. Section 230 has to go we need a different system, under which tech companies and we're mostly talking obviously, about the social media platforms operate. And I for one, think they will continue to make an enormous amount of money if they change their algorithms to prevent the kind of harm that is caused by sending people to the lowest common denominator every time they log on, you've got to stop this reward for this kind of negative, very lent content, which affects us across the board. But I will say it is particularly focused on women, the empowerment of misogyny online, has really caused so much fear and led to some violence against women who are willing to take a stand, no matter who they are. Are they in entertainment? Are they academics? Are they in politics or journalism, wherever they are, and the kind of ganging up effect that comes from online, it could only be, you know, a very small handful of people in St. Petersburg, or Moldova, or wherever they are right now who are lighting the fire, but because of the algorithms, everybody gets burned. And we have got to figure out how to remove the impunity, impunity, come up with the right form of liability, and do what we can to try to change the algorithms. And the final thing I would say is, we also need to pass some laws that understand that this is the new assault on free speech. You know, in our country, people yell free speech. They have no idea what they're talking about half the time. Yes. And they yell at to stop an argument to stop a debate to prevent legislation from passing. We need a much clearer idea of what it is we are asking governments to do businesses to do, in the name of do no harm. And free speech has always had limitations, always been subjected to legislative action and judicial oversight, and we need to get back into that arena.

39:02
Right, Commissioner, I can see you frantically scribbling notes. If you were you are officially the leading regulatory turtle. What would you do?

39:13
Well, I remember last year, when I was in Davos, I said similar things as you madam Clinton about maybe the United States will also have to move towards less impunity or no impunity online. You cannot imagine you can but I received from Republicans, I was afraid that I will be somehow wanted here as somebody who is committing horrible crime, but maybe for for the EU. It is easier to legislate the digital space because look at the situation Are we, while the United States have to make a big, big jump? We were a count of ready for for that. Because Count with me illegal content, hate speech, child pornography, terrorism, violent extremism, racism, xenophobia, anti semitism, we have all these things in our criminal laws. For decades. This is nothing new. So when when we started to think about how to legislate the digital space, we in fact set what is illegal offline has to be handled as illegal online. So we didn't create any kind of new crime. It was just pushing the existing law to the digital space. So that's why for us this, this era of adaptation was maybe easier than in US where you really have to do bigger, bigger jump. And if you if you let me say two more things, impunity is wrong. Crime and Punishment missing in the digital sphere is is another crime, I have to say. And we have to also adapt as the society I would like to still be alive when I will see strong rejection from the society that this is not acceptable. We don't like it. If in that system, we are confronted with hate speech and dirty content, we will simply

move somewhere else. So also for the digital digital companies, it will be a strong signal that they should not let their business damaged because they need users. Yes. So this societal reaction is still missing. I think that it will, it will take some some more years. Last comment on violence against women. We see women disappearing from public space. And here we speak about politicians and journalists mainly. And we had a conference. Maria was in Brussels slot last month. And one shocking thing came out that run here I speak about the politicians when the political parties want to win elections, they are attracting women to come because they are well, products, good products to sell. Yes, sorry. We speak about women. Yeah, in campaigns. But then when the women take the temptation and become politicians, the same political parties are not honest enough and courageous enough to defend them. So I see cases of women who are horribly attacked with horrible words like slow so we can president. Yeah. Nobody's defending her. Yeah. So should we remain alone? With that? I think that there should be also some health care reaction from the political parties, and from the newsrooms as well. Well,

43:12
thank you. Well, sadly, very sadly, we are at a time you set the scene fantastically. It take away three key points. One is that if women were running the world, I think there will be quite a different tone and sense of urgency to this debate. Secondly, these issues of misinformation are not entirely new. I mean, they go back a decade, but they have dramatically accelerated in recent years and AI is threatening to make it worse, and we have no time to lose because of the impending elections. And thirdly, we cannot duck the question of what is happening with the tech companies and their responsibility. If we want to move forward to some kind of if not solution than containment. We'll be hearing from tech companies later on today. We'll be hearing from another other a number of other voices about this vital debate. But in the meantime, can you all please show your thank yous to them so great.

45:13
Please welcome back to the stage Secretary Hillary Rodham Clinton. Hi, how are you?

45:31
Well, if you're not depressed we'll get you there. I could not be happier to have these extraordinary panelists follow up on the setting of the stage because now we want to get a little bit deeper and understand the implications for the upcoming US elections. And we have four amazing panelists. Jocelyn Benson is the Secretary of State of Michigan, and she's been in the eye of the storm since well, before the 2020 election, by far, but you know, since then, certainly one of the real leaders to try to understand what was happening. Michael Chertoff, the former Secretary of Homeland Security, co founder and Executive Chairman of Chertoff Group, and, you know, Michael really has just a depth of experience about dealing with originally it was online radicalization and extremism. And now of course, based on his knowledge of that set of threats, he understands, you know, we've got a, you know, we've got to face what's going to happen in the elections. Dara Linden mom is the Commissioner of the Federal Election Commission of the United States. And as such, you know, she is part of the group that is trying to, you know, make sense of where money is being spent and what's being done with it and the impact that it is having. And Anna makan Zhu is the Vice President of Global Affairs at open AI. And we really are thrilled that she's here with us, because clearly, open AI along with the other companies, you know, is forging new ground. And a lot of it is very exciting. And frankly, and a lot of it's very concerning.

So part of what we want to do is help sort that out, particularly as it does possibly affect elections. So Michael, let me start with you. Because, as I said, you you really were on the front lines, when you were at the Department of Homeland Security, and leading efforts to understand and prevent the use of the internet at that point to provide outlets for extremism and the radicalization of people. You know, and now I think there's legitimate concern about hostile foreign state actors, not just Russia, there are others who are getting into the game, why not it? It looks like it works. So you know, join the crowd. But we're now worried that they will use artificial intelligence to interfere in our elections this year. Can you explain for not just our audience here, but the people who are watching the live stream? You know, both the downsides, as to how AI can be used by our adversaries, but also what can we do to protect ourselves?

48:32
All right, thank you. And thank you, again, for reading this secretary. So let me say I mean, I think in this day and age, we have to regard the internet and information as a domain of conflict. Actually, if you go back historically, even 100 years, it's always been true, and that our adversaries have attempted to use propaganda and false information to manipulate us. But the tools they have were relatively primitive. What Artificial intelligence has done is equip people to have tools that can be much more effective with respect to the information domain. We've talked a little bit about deep fakes and the ability to have simulated video and audio that looks real. And unlike Photoshop, or some of the things some of us remember, from years ago, this has gotten to the point that it's very, very difficult, if not impossible, for an ordinary human being, to tell the difference. But I would actually argue that artificial intelligence has capabilities and risks that go beyond that. What artificial intelligence allows an information warrior to do is to have very targeted misinformation, and at the same time, and it's not a contradiction, to do that at scale, meaning you do it to hundreds of 1000s maybe even millions of people. What do I mean by that? I In the old days, again, 1020 years ago, if you sent out a message that was incendiary, you affected and positively maybe induced our belief by some people, but a lot of other people would look at it and go, Oh, this is terrible, and it would repel them. So that was an inhibiting factor in terms of how extreme your public statements were. But now, you can send a statement tailored to each individual viewer or listener that appeals only to them and nobody else is going to see it. Moreover, you may send it under the identity of someone who is known and trusted by the recipient, even though that is also false. So you have the ability to really send a curated message that will not influence others in a negative way. And the reason I say it's it's scale, you can do it millions of times, because that's what artificial intelligence does. So I think that has created a much more effective weapon for information warfare. Now, in the context of the election, in particular, what do we think worried about? Well, I think obviously, one experience we had, we saw this in 2016, with Russians assisting the Trump campaign is there can be an effort to skew the votes to a particular candidate or against the candidate. And we've seen that now we saw that with McCrone in France, in 2017. We've seen it in the EU, and other parts of the world. But I would actually argue that this year, we're facing something that in my view is even more dangerous. And that is it will be an effort to discredit the entire system of elections and democracy. You know, we had a defeated candidate who I won't mention their name, who was talking about a rigged election. Now imagine that for the people who are an audience for that, they will start to see videos or audios that look like your persuasively examples of rigged elections. Now, it's like pouring gasoline on a fire. And we could have another January 6. And I understand that the reason our adversaries like this, is because more than anything else, they want to undermine our unity of effort and our democracy. And

in a world in which we can't trust anything, and we can't believe in truth, we can't have a democracy. And that's, I think, going to lead to a third consequence, which will be very dangerous. We're talking about how do you distinguish and teach people to distinguish deep fakes from real things. And the idea being, we don't want to have the misled by the deep fakes. But I worry about the reverse. In a world in which people have been told about deep fakes, do they say everything's a deep fake, therefore, even real evidence of bad behavior has to be dismissed. And then that really gives a license to autocrats and corrupt government leaders to do whatever they want. So how do we help counteract that? Well, I mean, there's some technological tools, for example, there is an hour an effort to do watermarking video and audio, where genuine video or audio when it's created has an encrypted mark, such that anybody who looks at it can validate that it is real, and it's not fake. More than that, we've got to teach people about critical thinking and evaluation. So they can cross check that when you get a story that appears to stand alone look to see what are the other stories is anybody else picking it up? And we need to actually establish trusted voices, that are deliberately very careful and very scientific about the way they validate, and test things. And finally, I think we've got to teach even in the schools, and this is gonna start with kids, critical thinking and values, what it is that we care about, and why truth matters, why honor matters, why ethics matters. And then to have them bring that into the way they read and look at things that occur online. This is not going to be an easy task. But I do think we need to engage everybody in this process, not just people who are professionals, and make it part of the mandate for civil society over the next year or two. Thank

54:28
you so much, Michael, that was incredibly helpful laying the, you know, the groundwork for what we need to be thinking about. So, Dara, what what is the Federal Election Commission doing to try to set up some of those guardrails on AI fueled disinformation ahead of the 2024 elections?

54:49
Well, thank you for having me. First of all, it's an honor to be a part of this really important discussion. So to your question, the short answer is that the FEC is fairly limited in what it can do in the space. But there is hope on the horizon. And there are different ways that things are developing. So just to lay the baseline despite the name, the Federal Election Commission really only regulates the campaign finance laws and federal elections. So the money in money out and transparency there. But last year, we received a petition for rulemaking asking us to essentially clarify that our Fraudulent Misrepresentation regulations include artificial intelligence and really deep fakes. And we are in the petition process right now to determine if we should amend our regulations if we can amend our regulations. And is there a role for the FEC and these campaign finance regulations in this space, our language is pretty clear and very narrow. So even if we can regulate here, it's really only a candidate on candidate bad action. So if one candidate does something to another candidate, that is all that we could possibly cover, because of our statutes, and that's unless Congress expands that, but all is not lost. And there are some pretty great things that have come out of this. And one is what happened during our petition process. We received 1000s of comments from the public and from many other institutional actors, including a lot of the smaller tech companies and organizations that don't often have a seat at the table. But here, it was really an open forum for them to bring their ideas to light. These comments were insightful. They were creative. And I it is my hope that Congress and states and others looking at this will read all of these comments as they try to come up with possible creative solutions here. And in

addition, Congress could expand our limited jurisdiction. If you asked me three, four years ago, if there was any chance Congress would regulate in the campaign space and really come to a bipartisan agreement, I would have laughed. But it's pretty incredible to watch the widespread fear over what can happen here. We had an oversight hearing recently, where members on both sides of the aisle were expressing real concern. And while I don't think anything's going to happen ahead of November, I see changes coming. And there's a bipartisan discussion. Senator Klobuchar is leading this Senator Warner, and they're thinking about ways that they're they can do something, these are really on the really only the deep fakes space. It's not the the misinformation disinformation that's underneath it all. But this discussion of AI and how AI is so at the forefront of everything that we're discussing in this country, I think it has brought more to light, this misinformation, disinformation, and the ways that the information gets disseminated, that it is bringing that discussion out. So things could change. I'm hopeful.

58:10
Well, I really appreciate your talking about that, Dara, because a lot of people say well, who oversees elections, who tries to make sure that our elections, you know, don't go off the rails, and we don't have a lot of these problems. And as you just heard, it's not the Federal Election Commission, their mandate is narrow, and they try to, you know, make sure people who are contributing to elections have the right to do so and candidates are spending appropriately. So much of the work about regulating elections is done at the states in our country. And we're so fortunate to have Jocelyn here. Because as I said, in introducing her, she really has been at the forefront of trying to figure out how to protect our elections to make sure they have integrity. And Michigan has recently tried to regulate artificial intelligence. And I want you to tell us about that elect, that legislation and any other actions that you are taking on behalf of your state and that you know, other states are taking, but maybe just start Jocelyn with just a quick introduction of what you've been facing. You are elected in what 2018 2018. And, you know, if you remember pictures of armed men storming the Capitol, because they didn't like what the governor was doing about COVID and Michigan was at the real center of all of the, you know, crazy theories that were put forth in 2020 about the election. Give us just a quick overview, and then tell us about what your regulation intends to do and what else needs to be done. Thank

59:52
you. And thank you, Secretary Clinton for inviting me to be part of this really important conversation to me. We cannot protect The security of our elections if we don't take seriously the threat that artificial intelligence poses to our ability as election officials to simply ensure every vote is counted, and every voice is heard, and that citizens have confidence in their democracy and in their voice and in their votes. And that's our goal in Michigan, and in several other states all around the country, we are coming off of being in the spotlight and 2020, rising to that occasion, but also seeing very clearly and living very clearly, when people with guns showed up outside my home and dark night in December, and I'm inside with my four year old son trying to keep us safe. And that's real. And they showed up there just like they showed up at the Capitol on January 6, because they've been lied to and fed misinformation. And now we're facing an election cycle where those lies will be turbocharged through AI. And we have to empower citizens to stand with us and not being fooled and pushing back on that misinformation and those lies. And therein lies both our opportunity but also the real challenge. How do we, in a moment where the adversaries to democracy are focused on sowing seeds of doubt, creating confusion and

chaos and fear in everything they do, and now have this new emerging technology that is day by day essentially getting stronger and more, you could perhaps say effective at being poised to accomplish is accomplishing those goals of creating chaos, confusion and fear in our democracy in our voters minds. How do we respond to that at the state level and throughout our country as citizens by by giving each other certainty and confidence that our democracy will stand just as it prevailed in 2020, and every time before and since, but also that we can be equipped every single one of us to have clarity as to how to respond when we get this misinformation. So in Michigan first, first, we set up the guardrails and several other states have done this too. And we do hope the federal government joins us in banning the deceptive use intentionally deceptive use of artificial intelligence to confuse people about candidates or positions or how to vote or where to vote or anything regarding elections. So we've drawn a line in the sand, it's a crime to intentionally disseminate through the use of AI deceptive information about our elections. Secondly, we've required the disclaimers and disclosure of any type of information generated by artificial intelligence that's focused on elections. So for example, one of the things we're worried about is and we know because of AI, it can be targeted to a citizen on their phone, getting a text saying, Here's the address of your polling place on Election Day, don't go there, because there's been a shooting. And stay tuned for more information that's going to invoke fear, again, goal is fear, right? It's gonna invoke fear and a citizen, with the disclaimer and disclosure in place requirements, they have to be disclosed, this has been generated by artificial intelligence, it's still not sufficient, but it is a key piece of enabling us to push back. But the other side of that is we need to equip that citizen, when they receive that text to be fully aware as a critical consumer of information as to what to do, where to go, how to validate it, where are the trusted voices. So in addition to passing these laws, we are setting up voter confidence councils building out these trusted voices, so that faith leaders, business leaders, labor leaders, community leaders, sports leaders, education leaders, can be poised, even mayors and local election officials to be aware and push back with trusted information. And so it's layers upon layers of both legal protections and partnerships to equip our citizens with the tools they need to be critical consumers of information. And then in everything we do between now and in every election, but certainly leading up to November, helping to communicate to every room we're in that it's on all of us to protect each other from the threat of AI in regards to our elections and in many other spaces as well. And while we as officials will be working to do that, we're also trying to communicate to citizens, this is a moment that's going to define our country for years to come. And we all have a responsibility in this moment, to making sure we're not fooled. Our neighbors aren't fooled. Our colleagues and friends aren't fooled and equipping all of us with the tools we need to push back and speak the truth, value that honor and integrity and help define our country moving forward based and rooted in those values.

1:04:28
Well, I am a huge fan of what you and your attorney general and your governor have been doing. And I think it would be great if you could get some help to model this and I'm hoping maybe some tech company or some foundation will talk to you afterwards because we need to show this can work. I saw Michael nodding his head. I mean, we've got to get you know if if this is a fight against disinformation, we have to try to put up guardrails. So you all To have to flood the zone with the right information to counter the negativity that is out there. So I hope you can implement that. And we can then all learn from it. Because it's going to be not a problem that goes away after this selection. So Anna, you've been sitting here, you've been sitting through the first panel. Now, you've heard our other panelists, and you are, you know, truly at the epicenter of this because, you know, at chat GPT, you all are moving

faster than anybody can even imagine. Sometimes, I think probably yourselves about what it is you're creating and the impact that it will have. And this is obviously the Ground Zero year, this is the year of the biggest elections around the world since the rise of AI technologies, like chat, G GPT. And so, can I ask you, do you agree what you've heard from the panelists about the dangers, but then tell us what you're doing at Chet G at open AI? To try to help safeguard elections? You know, give us your assessment? Are we overstating it? Are we understating it? And what can be done? And how can you help us do it?

1:06:19
So I think what's been really interesting to me listening to your first panel, and to my co panelists here is that so many of the ideas and the concerns we are already integrating into technology. So if I could just say that the one piece of good news is that, unlike previous elections, none of us are coming into on in terms of the tech companies, election officials, even the public and the press, we're not coming into this unprepared. You know, this is especially true for me, because I was actually working at the White House on the Russia portfolio in 2016. So this has been top of mind for me from day one in the job. But open AI, you know, a relatively young company, this is something that's been top of mind for us for years, in fact, GPT two, which was several years ago, and you know, quite, you know, embarrassing compared to what exists now at the time, it was state of the art, it could produce paragraphs that were texts, like a human could write. And even then, we thought, Oh, well, like the possibility for this to be used to interfere with democracy and electoral processes, very significant. And so we made a decision then not to open source it. And it was quite controversial in the research community, but it was because we have this in mind. So, you know, that is that is, you know, ahead of 2016, we were not having panels like this. And so I think in general, we are much more prepared as a society and we are working together and open AI is working with the National Association of secretaries of state, and with social media companies, because one key thing to remember is that this there is a real distinction. So we are not dealing with the same kinds of issues at AI companies we are responsible for generating or what we do is generate AI content rather than distributed, but we need to be working across that chain. And, but in terms of you know, of course, as many have mentioned here, and as I hear almost with in every interaction with policymakers, deep fakes, are a very serious concern. And so for us, we have Dali, which is an image generator, and let's just ban you know, we do not allow to generate images of real people, period, but in particular politicians. And now we are implementing something called ctpa, which is a digital signature. And the great thing about ctpa, is that it's not just AI companies, you know, this is the New York Times and Nikon and, you know, the BBC, so it's going to be an ecosystem, where there's actually a tool across the ecosystem that's going to work to help journalists and social media companies identify if a piece of content is generated by AI. Obviously, this is not, you know, a complete solution. But this is not, you know, this was not the case a year ago. So already, we are much more advanced in our ability, as you know, the entire ecosystem to deal with these issues. But we also have, you know, threat investigators, we recently just took down a bunch of state actors who were using our tools. So we and so these two pieces of cooperation across all of the players and all of the state of the arts interventions that we're building, I think mean that for you know, right now, the kind of thing that you described, Secretary Chertoff is not possible with opening AI tools where you cannot connect them to a chatbot to fuse information and targeted at voters. But you know, we're constantly evaluating what other kinds of threats does this technology create that are novel? And I would just kind of wrap up with there is, of course, I do have optimism, otherwise I wouldn't be working with open AI. One of the things that

these tools have the potential to do is create access to education for new segments of society. And so, you know, there's a potential these tools to actually help create a citizenry that is more educated and more aware, which I think is a really key aspect to a healthy democracy. And it can be used, you know, for Secretaries of State that are incredibly busy in back offices, and it is a bit of a race between the positive applications of these technologies and negative ones. So I am, you know, this is why it's so fantastic that, for example, the the EO by President Biden really works to strike that balance.

1:10:32
Well, we have only a few minutes left, but I just want to ask each of the panelists, you know, what steps can governments, obviously national, state, local, in our, in our country, the private sector companies, particularly the tech companies, the AI companies, the platforms, nonprofits, any anyone that you think of as to what they could and should do to try to, number one, ensure the integrity of this upcoming election, but then for the longer term? What are the changes we need, and if you start with them, I think

1:11:06
it goes back to what I already mentioned, which is that really close collaboration with a AI companies, social media companies, election officials, civil society, really working together to address this problem and sharing best practices and sharing knowledge. Because if this is a whole of society problem, and no single actor is going to be able to be fully effective and solving it.

1:11:33
I think the education component and pushing to find trusted sources is key. The technology is going to change, there's going to be new technology in the future, no matter what the government does, or what the tech companies do, we need to strengthen the trust that people can build and interested, you know, forms of of news, and I think we're seeing some of some of that starting to change.

1:11:58
I would say in addition to those suggestions, information sharing, when there is an indication that something is coming, that's part of a wave of disinformation, to share information among all the stakeholders, including federal and state and the public is very, very important. Now, I want to say I know that there's some litigation now where some states have tried to make it illegal for the government to share information about disinformation with the platform's because they argue that that censorship, I personally think that's nonsense. I think what you're doing is giving information that's helpful and not doing anything that's harmful. Yeah,

1:12:37
I agree. I think and particularly for, for philanthropy, and Foundation's to really invest in entities and partnerships, that are focusing on this education and sharing of information and building more collaborative partnerships and teamwork around this, all of that, I think has to be be the foundation for every entity making their first priority, protecting citizens from the ways in which AI can be negatively used to harm their own voice and their votes in a democracy, and to recognize that, that our adversaries to democracy have figured out how to divide us and D mobilize us and deter us from believing in our voice through the use of AI. And so our response needs to be similarly collaborative national in scope, and focused on empowering citizens and partners all across every arena and sector

tech and beyond to be a part of the pushback and the protection of our citizenry, from this threat to our democracy.

1:13:41
Well, I can't think the four of you enough and maybe out of this panel will come that kind of cooperation, let's try it out. Let's see, let's get you know, open AI, Facebook, others together with people like Jocelyn and Michael, who have you know, a lot of depth and you know, what Dara knows from, you know, the where the money flows are that she sees, and let's see if there can't be some cooperative effort between. Now in this election, if we don't try, we know what's going to happen. We know what's going to happen. And I think that, Michael, you made a great point, we need more transparency and openness. You know, and that should be declassifying information as quickly as possible. So it gets out in the public. And frankly, governments need to get it out and not ask for permission because it can influence the conversation going forward. But I think this idea of collaboration, it's always better in a democracy to have collaboration bring people together. So let's see what we can do to follow up. Thank you all very much.

1:15:00
There will now be a short break in the program. We will resume at 2:50pm. Thank you

1:29:33
Hi, everybody. Boy, those were two fabulous panels. I hope you agree. And of course, there were many mentions of those darn tech companies. So now we have the tech companies up here. And boy, are we going to grill them sorry, you

1:29:50
guys are terrible.

1:29:53
No, no, we will actually what I want to start well, let me first of all, let me introduce our let me introduce our panelists. We have at the end, David Agron vich, who is the Director of Global Threat disruption at Mehta. Yasmin Green, who is the CEO of jigsaw. Jigsaw is a unit of Google that addresses threats to open society. And Clint watts, who leads Microsoft's threat analysis center, which is part of customer safety and trust. So where I where I want to begin, actually, with each of you, because all three of you in slightly different ways, are focused on looking at the threats and the risks that you're seeing across your platforms or across society, I think maybe even more. So in your case, Yasmin, so I want to begin by you sharing with us what you're seeing. And again, particularly with relationship to the use of AI, when it comes to information, deception, or other forms of AI deception. And Yasmin, I'm going to begin with you jigsaws a little bit of a different, a different animal here, because you really are looking at societal changes, and what kind of interventions you can make not even necessarily, via your platforms to, to effect changes. So give us a little bit of a sense of what you're saying,

1:31:14
Okay, what I wanted. Hello, everyone, I wanted to actually pick I think the panelists before did such a good job of surveying the landscape, including the threat. So I wanted to get a bit specific and build on

what was said. So there was, we talked about trust in the last panel. And one of our observations about the trust landscape is not that we are in a post trust era, because as humans, we have trust to a sixth, we have to make decisions, we have to evaluate things. But it's not that trust is evaporated, it's that it's migrated. So trusted is much less institutional, and much more social. And I think that's really important as we think about the risks posed by generative AI. So we did an ethnographic study with Gen Z, to figure out how young people going about and what trust his or her mistakes they have online, and how they go about evaluating information. So I want to just do a survey of this room. And I think we have a good generational mix here. But just by by show of hands, how many people have read the comments underneath news articles? You say that's about maybe half two thirds. I gotta say, I don't read the comments. I thought like our collective coping mechanism for the Internet is that we don't read the comments. Okay, so Well, I'll tell you, who reads the comments. Gen. Z, Maria. And the interesting thing is not that they need them as much as as, as when and why when do they read the comments? They often go headline. Well, one understudy here, if anything, which I appreciate it, and I love you. No one else I'd rather speak the words from my mouth and Maria ReSSA. But headline, comments, and then the article. Why would they be doing it in that order? Because and this is, according to them. And this research that we did, they want to know if the article is fake news. So you see the inversion there, like I would have the article as the journalist being, you know, the authoritative curators of information. And they are interviewing experts who are authorities, and Gen Z. And I think increasingly, we are going to the social spaces to look for signal, and we kind of throw out the term information literacy, and we wait up to instead of like information sensibility, they're looking for social signals about how to situate the kind of the information, the claims and the relevance to them. So it's like, you know, we had famously had the term alternative facts. This is alternative fact checking, you know, and we shouldn't be really concerned. And it's relevant to generative AI, because one of the things that we maybe emphasize less than we should, because it's a threat that's coming around the corner. In addition to synthetic content, we have synthetic accounts, we have accounts that are going to be we talked about this earlier, but that you know, these human presenting chatbots. And what we're seeing with one of the products that we offer, Jigsaw is is the most popular free tool for moderating comments spaces. So we have billions of comments every day that go through our 1000 partners, and we hear about synthetic accounts that are there and posting. They're not sending you crypto, they're not spreading disinformation, they are active. What are they doing? They are building a history of human like behavior. Because in the future, yes, it's going to be really important for us to evaluate wherever we can to do detection to evaluate whether something's a deep fake. When there's a deep fake Where do you think people are gonna go to check that They're gonna go to other people in the social spaces, the signal. So we need to invest in humans and also invest in, in ensuring that the human presenting chatbots and not do not have an equal share of influence that

1:35:13

fascinating. So, so synthetic identities, not just synthetic content. Yeah. Fascinating. Clint, so you I've known you for a number of years, you've been now with Microsoft for what, two, almost two years, two years, but you have been sort of doing this kind of deep digital forensics for quite a long time. So you've seen that sort of trajectory of history, how we've seen seemed seems things seem things evolved since you know, 2016, prior to 2016 Until now, so give us a little sense of what what you what change you have seen, particularly since generative AI is sort of taken off in the in the wake of the launch of chat GPT and what risks you're seeing today? Yeah,

1:35:59
so it's, it's interesting in terms of timing, it was 10 years and two months ago that we encountered our first Russian account and impersonated an American, that would later go after the election. I'm sorry, we were trying, you know, and we're working from her house. And we use a tool called Microsoft Excel, which is incredible. If you've ever checked it out. Now we use Microsoft. So that's a major change in 10 years. And and what's interesting is in 2014 1516, it was testing it on the Russian population first, Ukraine, Syria, Libya, it was battlefields and then it was taking it on the road to all the European elections and the US election. And so watching what has transpired in that, there's often a little bit of a misunderstanding about how much things have changed in 10 years in terms of social media. Speaking of Gen Z, Gen Z, would you read more than 200 words, I bet you would watch 200 videos. So that's one of the biggest changes in 10 years with the technology. And that's not just about Gen Z, that's about my generation. Everybody, older video is king today. And if you're trying to influence by writing a hot 9000, word, blog, you're, you're running uphill, like with with a lot of weight on your back. So you know, our monitoring list in 2016 were Twitter or Facebook accounts, linking to Blogspot. In 2020, it was Twitter or Facebook, a few other platforms, but mostly linking to YouTube. And today, if you go to it, it's going to be all video, any monitoring list any threat actor so my team tracks Russia, Iran, China, worldwide. We've got 30 on the team. We do 15 languages amongst the analyst. And we're mostly based here in New York. And nine months ago, we did a dedicated focus on what are the what is the AI usage by these threat actors. And so we have some results of our research so far. And what I would say in 2024, there will be fakes, some will be deep, most will be shallow, and the simplest manipulations will travel the furthest on the internet. So in just the last few months, the most effective technique that's been used by Russian actors has been posting a picture and putting a real news organization's logo on that picture, I'm sure David, he'll be able to tell you more about this distributed across the internet that gets millions of shares or views. There have been several deep fake videos in and around Russia, Ukraine and some elections. And they haven't gone very far. And they're not great. Yet. This wall change. Remember, this is March. So things are moving very quickly. So what I would note is just looking at a few things, there are five distinct sort of things to look at. One is the setting, is it public versus private in public settings, and I would love David's take on this. When you see a deep fake video go out crowds are pretty good. Collectively, it's a nun. We've seen that video before. I've seen that background. He didn't say this. She didn't say this. We've seen Putin Zelinsky deep fakes and the crowd will throw the real video out and it kind of dies. The place to worry is private settings. When people are isolated, they tend to believe things they wouldn't normally believe. One, anybody remember COVID When we were all at our house, it was very easy to distribute all sorts of information, it was hard to know what was true or false or what to believe. And people had totally different perceptions of the pandemic. The second part is in terms of the AI the medium matters tremendously. Video is the hardest to make. Text is the easiest. Text is hard to get people to pay attention to video people like to watch. Audio is the one we should be worried about. Ai audio is easier to create because your dataset is smaller. And you can make that on a lot more people. It takes a much smaller data set and you can put it out and there's no contextual clues for the audience to really evaluate. So when you watch a deepfake video you go somebody I know how that person walks, I've seen how they talk, that's not quite how it is audio, you'll give it a discount, you'll say, Yeah, on the phone, maybe they do sound like that, or that's kind of garbled, but maybe that is where to look at. We've seen that in the Slovak elections, we've seen that with the robo calls around President Biden, Indonesia, we've seen these

sorts of examples, there was a deep fake video that used Tom Cruise's AI voice, he's probably the most faked person, both video and audio around the world, that's tougher to do. And that kind of comes to the other thing to look for is there's a intense focus on fully synthetic AI content, the most effective stuff is real, a little bit of fake and then real. Blending it in to change it just a little bit. That's hard to fact check. It's tough to like chase after. So when you're looking at it private settings and audio with a mix of real and fake. That is, that's a powerful tool that can be used. A couple other things sort of to think about is the context and the timing. Many of you probably saw information was totally incorrect about the Baltimore bridge tragedy this week, right? People immediately, you know, rush to things and when you're feared, or there's something you've never seen before, you tend to believe things that you wouldn't normally believe. So imagine it's a super contentious event, or there's some sort of an accident or a tragedy, AI employed in that way, can be much more powerful tool. To do that. You have to have staffing, you have to have people, you have to know the technology, you have to have compute, and you have to have capacity. That's not a guy in his basement on the other side of the river, folks, that is a well organized organization with technology that has the infrastructure to do that and is ready to run on something instantly, ie the Russian disinformation system, which doesn't hire just about 10 or 20 people, we're talking about 1000s of people that are working this nonstop and around the clock. And as we know, in all of the governments around the world, there are just 1000s of people working to counter disinformation day in and day out, right, we stay up till two in the morning watching. We're just not set up the same way. And so that gives them a strategic advantage. 10 years ago, we were tracking two activity sets of Russia that ultimately went for 2016. Today, my team tracks 70 activity sets tied to Russia. So that just tells you in terms of the scale worldwide, and the way things are going, that's something to look for. The last thing to think about is knowledge of the target and Secretary Chertoff brought up a great point is, if people know the target, well, they're better at it deciding whether something is fake or not, if you've seen it over and over again, but if you don't know the target Well, or the context, well, they are not as good at it. So there's always the presidential candidate, presidential candidate will be a deep fake, and it will change the world and make her heads explode. Probably not. But if it's a person who's working at election spot, somewhere out in a state, and a deep fake is made, or maybe they're not even a real person. It's these contextual situations that we have to be prepared for in terms of response. So our team is setting up I we work with Google and and meta, and I would just tell you is my experience being on the outside of tech. And now being in 10 years ago, when I notified tech companies about the Russians going after the election, they told me I was an idiot, and that no one would believe that. Now I work at a tech company, and we do exchanges all the time. So I would just like to point out, I feel like we've got great relationships, Yasmine, and David, we've worked together for years, you know, on different projects. So I think that's something else, it's quite a bit different today.

1:43:38

That's great. Thanks, Clinton. And, David, I want you to sort of pick up where we're Clintus leaving off, obviously, any additional context that you can provide, in addition to what he said about what you're seeing out there. But then I also want you to address something pick up on something that Clinton mentioned, which is, it's one thing when it's, you know, a big splashy, deep fake, that's, you know, all over public forums. So those can easily be up debunked. And I agree with you, the big, spectacular deep fake of one of the major presidential candidates is unlikely to have huge impact. But the stuff that

you can't see, because it's on messaging platforms. That's what we worried about. So talk about what you're seeing there. Absolutely.

1:44:15
And I think building a bit on what Clint Clint had mentioned around what we're seeing from threat actors around the world. So our teams have taken down now, a little over 250 different influence operations around the world, including those from Russia, China, Iran, but also a number of domestic campaigns from countries all over the world. Maybe the key three things that we're seeing, in addition to the trends that Clint mentioned one, these are increasingly cross platform cross intranet operations. The days of a network of fake accounts on Facebook and network of fake accounts on Twitter, somewhat, you know, closed ecosystems are gone right now. I think the largest number we've ever seen is 300 different platforms implicated in a single operation from Russia, including local forum websites, things like next door but like for your neighborhood. I as well as more than 5000 just web domains used by a single Russian operation called doppelganger that we reported on last quarter. So what that means is the responsibility for countering these operations is also significantly more diffuse right platform companies don't just have responsibility to protect people on their platforms like the work that our teams do, but also to share information. I think Secretary Chertoff mentioned this in the last panel, not just sharing information amongst the different platforms that are affected, but with civil society groups and with government organizations that can take meaningful action in their own domains. The second big trend to think that we've generally been seeing is that these operations are increasingly domestic and increasingly commercialized. It's their commercial actors who sell capabilities to do coordinate what we call coordinated inauthentic behavior. Disinformation, for hire something unknown Maria's organization has written a lot about the Philippines, in the commercialization of these tools, democratizes access to sophisticated capabilities that used to be basically nation state capabilities, and it conceals the people that pay for them, it makes it a lot harder to to hold the threat actor accountable by making it harder for teams like ours or teams in government to figure out who's behind it. And then the third piece is that we're increasing so to the use of AI is that much like Clint mentioned, I think we've generally seen AI, I would say cheap fakes or shallow fakes are not even AI enabled. But just things like Photoshops are repurposed content from other events, mainly being used by those sophisticated threat actors, Russia, China, Iran. But where we do see AI enabled things like deep fakes or text generation being used by scammers and spammers. Now, that's not to downplay the threat. Scammers and spammers are arguably some of the most innovative people in the online threatened, they move the fastest, they're the least responsive to external pressure, because they just want to make money. And they often are in jurisdictions that aren't going to do anything about them. What I what we should all be alert to is the tactics and techniques that the scammers and spammers use being adopted by more sophisticated actors over time. So if you want to look to see what's coming, that's where I would be looking to see where things are coming. Now, what can be done about it, what's working, what isn't working, especially some of the examples you use some of these AI enabled capabilities being used in smaller, more private settings. This is where things like some of the watermarking and again, by watermarking. Here, I mean more what Ana makhanda was talking about. So technical steganographic watermarking, that can't be easily removed, to identify whether content is authentic or was created by an AI system can be perpetuated by social media platforms, right. So if a company that produces AI content, which met is one of those is willing to be part of that coalition, make sure anything that our models produce is discoverable as AI generated, then when it shows up on Twitter, or shows up on our own platforms, or

shows up on snapshot, it should carry through those standards. And so there was compact at Munich amongst many of the tech companies, Microsoft was part of that as well, Google is part of that, the more we can raise the bar across the industry, to require companies to be building in these capabilities early. Before we get to the point where the bad things have already happened, the more we can actually build meaningful defenses. I one thing from the last panel that really stuck with me was so when Ana was at the White House dealing with Russia policy, I was in the US government on the security side, also dealing with Russia policy. And we were chasing after the problem at that point, right, it had left the station, we have an opportunity now to start building these safeguards in as this technology is taking off. So I'm happy we're having this conversation now. And I thank everyone who pulled this together, because it's an incredibly timely time for us to be building this. Thanks.

1:48:32
I want to just stick with you just for a second David and talk a little bit about go a little bit deeper on messaging platforms. So of course, Mehta owns one of the most used significant largest, private encrypted messaging platforms in the world, which is WhatsApp. So much of what we know is that is traveling that could be these kinds of synthetic messages no matter what form factor they are, video or text or images, or audio travel through WhatsApp, can you how how do you think about ensuring that those platforms do not become vectors for for this kind of harmful synthetic content around elections? And what are you doing about that? And also about the open parts of WhatsApp as well?

1:49:22
Absolutely. There's some really exciting, I think, integration between some of the technical standards that we've talked about things like steganographic watermarking, that can be programmatically carried through on platforms, and ensuring that robust and reliable encryption remains in place for people all over the world so that their communications can't be spied on by governments or particularly in authoritarian regimes. So one of there's, I think, two different tool sets here. One is ensuring that as platforms whether it's WhatsApp or signal or anyone else who's building these point to point communication tools, that we're building tools for the people who use the platform to it identify and report problematic content things scamming spams, but also things like disinformation. And also that we're building in technologies as the industry uptakes more of the safeguards around AI systems, that can be programmatically propagated in our own software without really needing to break fundamental encryption, right. So you can imagine the future where we can get all of these companies that produce AI images, or AI generated text, to sign up to watermarking standards. And if that content ends up being sent through one of our platforms, that the watermark can be carried through without having to have someone in the middle saying, Oh, that right there, that's AI generated. And so I think that's actually one of the several reasons why some of these technology standards are so important, and can hopefully be enshrined not just in industry agreements, but also in many of the regulatory conversations that are happening. Because there is there is a world in which we can, and I think it's really important to retain fundamental encryption standards, while still making sure that we are doing our due diligence and our responsibility to protect the broader information environment. Well,

1:51:02
certainly though, there are things that medic can do to keep these kinds of messages from going viral even while protecting encryption. Yasmin talk a little bit, I'm gonna ask you both to talk a little bit about

what Google is doing to eliminate the risks and stop the spread of AI generated election. misleading information. Yeah,

1:51:24
quickly, just this idea of, at the you know, at the origin origination of the content, trying to kind of stamp it in a way that is enduring so that it can be identified as, as as synthetic is really important. And that's ongoing work. One of the things that I think is interesting actually is that is kind of actually just refusing to provide the Jenai service, when the stakes are as high as they are when there are election queries. So now it's a kind of like, it's a new it's, you know, in a lot of people kind of understand intuitively that there's a tension for technology companies and wanting to make the the experience for the user safe, but not creating so much friction that they don't want to use the product. So it's interesting, for example, now, if you go to Google's generative AI product, which is Gemini and you search for something election related, it will give you a non answer, which is actually pretty crappy feeling, you know, but they send you to search. Instead, they say, go to search, and there's research by person research that shows that people want an authoritative source. I think this is interesting thinking about this tension between authority and authenticity. You know, those are the mental models that we have from, from the last decade of search, and social media. It's like, if it's coming from an institution that I trust, or even Google search, you know, I'm, I'm, there's a lot of trust there. So the stakes are really high, you better get it right. Or if it's social media, if it's coming from my friend, there or my social network, then I trust them. Of course, Jeremy via AI is neither of those. It's not authoritative. It's not summarizing what the internet says, and giving you this destination of something that's authoritative. And it also sounds like a human, but it's not a human that you know. So I think we're in an we don't have mental models to deal with, with generative AI output. And at the moment, you know, I think it's an interesting demonstration of a commitment to trying to put election integrity first, is actually giving users a pretty bad experience of the Genesis. So you're

1:53:26
defaulting to sending people to search, which gives it more reliable while you're still sorting this out, we are quickly running out of time. So Clint, just tell us what Microsoft is doing.

1:53:35
We're gonna get him this time, certainly get the Okay, our work just just conceptually, the Russian concept of reflexive control, if you're familiar with it, is you conduct an attack on your adversary, and then they attack themselves in response. That's somewhat what has happened over the last 10 years. They're winning through the force of politics rather than the politics of force. They're more than three nation states that will probably do some sort of election influence and interference. My team is designed to focus on the Russia, Russia, Iran, China, you know, absolutely. You'll see that in our November report, we have another report coming out this election focus on this one, I think the key point is that you have to raise the costs on the adversary at some point, rather than raising the cost on yourself to function as a democracy. And so there are lots of things we can do in policy and tech, and we've won those at Microsoft, and we do data exchanges amongst ourselves. But ultimately, we've got to say, there's a hack here, there's a leak here, and it's coming and we're anticipating we're going to be out in front of it the next time. It's inoculating the public, it's inoculating the public, it's also raising the cost for actors to do that sometimes that is methods and platforms, you know, communicating, so we include

controls, but a lot of it is awareness, communicating to the European governments communicating the US government. This is what we're seeing because we can see it better oftentimes from the private sector than the public sector. See you are sharing that info Imagine we do. Yeah, if it's if it's something that's impactful, our nation state notification system.

1:55:05
Okay, well, we're out of time. So thank you so much we could have gone on much longer, appreciate it.

1:55:33
Please welcome to the stage Columbia SEPA Professor Anya Schiffrin.

1:55:45
Thanks, everybody. It's so good to be here. I'm Anya Schiffrin. And I direct the technology and media specialization here at sepa, where we are all abuzz. I see a lot of our students are in the room. We've been talking all year about AI and the elections and disinformation. And it's really been fantastic to have the secretary here, as well as our dean and Maria ReSSA. And so many of us who are involved. We've really been Yes. So this builds very nicely on some of the other events that we've had. I was lucky enough to get invited to Vivian Schiller's Aspen event in Miami in January, Tom Asher was there as well, where we laid out, it was incredible wake up call what the threats were. So we heard a lot of what we've been hearing about today, that audio deep fake would be a real threat in the US that there would be certain pain points during the election, such as the counting would be really dangerous and or risky. And then also Alondra Nelson and Julia Angwin, worked with IGP and brought together election officials from around the country. And they came here in February to game out scenarios with Maria ReSSA. And they were the kind of people like Dara Lindenbaum, they were so meeting them was so emotional, and so inspiring, and to hear the stories of the death threats and everything else. So I'm really glad that for this panel, we're turning a little bit to the international situation. And as we were all preparing for our classes, in December, in January, we were reading about how the 85 billion people were going to have elections this year and dozens of countries and you know, we've got to really watch out the US is not just the only place where this is happening. And so I was just thrilled that IGP decided to bring in some international voices to this discussion, because I think we have a lot to learn. So we've got Ethan to from Ai labs and Taiwan. And of course, Taiwan has the reputation for being really good at all of this, right. It's like, you know, you got it. You had Audrey Tang during the COVID pandemic, you've got all that public diplomacy. And we all know you're used to China. So we're looking forward to your expertise of your political Millay just gotten elected in Argentina. So we're gonna have to hear about how disinformation played into that are not. This is obviously a country that's extremely polarized. So no surprises there. And this is the first time that I'm meeting Dominica Ha Joon, I think you're going to be bringing in the perspectives from Central and Eastern Europe. So I think maybe I think your election was first Taiwan. So maybe we'll start with you. How did it all play out? Where did you have the same problems that we keep hearing about from the other panelists?

1:58:23
Yes. So I can introduce my institute a little bit. So we have en la PLRA first open air Research Institute in Asia 20, open Azalia we are a little bit younger than open AI. So what we do is we do the transparency responsible trustworthy AI evaluation, including the information manipulation. So for

example, during the pandemic, we use the artificial intelligence to know is there a troll can internet is manipulated information against Taiwan. And during Taiwan presidential election, we can reserve billions of activity that fund online social media and law including our Facebook Twitter, PTT,

1:59:11
mainly a threat from China. For internal to

1:59:16
in we have PTT last internal Taiwan playfulness fun by me in 1995. And Facebook. Of course, Twitter is also one of the major prefers in Taiwan. And Taiwanese people also look into WeChat Twitter. Tick tock, that's a big topic recently. So in Taiwan, we also we also observed the information manipulation, this social media or cross poll, during the election, there are a lot of choice activity, which means that is the Facebook just mentioned this a collab is in your Sunday behavior. So if we use artificial intelligence, we can identify those People HLA are not real human. They do appear together these appear together, these communities have false information together. And they like to like reference the PTO nice show video is a trending topic this year. In the past, we can see a lot of information minutiae in tax policy, we see a lot of Shell video and Shell video we'll have the show video and the YouTube also have the she'll be done. But usually the show video on the YouTube was already fun to talk about very

2:00:36
the intro, just what we were hearing about about cross platform and a lot of video and audio as well. Could you tell us the source? Were you also able to track down the source of this?

2:00:45
Yes, so using the artificial intelligence stuff so so we know this defect so like, there are a lot of video they have the same narratives, ballet use the land use different backgrounds, different boys, ballet brain essential and fraught into, like YouTube and talk platform that try to influence how people feel in Taiwan. So so. So do we use the artificial intelligence. So we use the speech recognition language understanding, then by identifying the troll can then we know, last stroke and they are now real human humans, then we can cluster the story layer trying to spread. So during the livestream, we can carry understand, for example, we know when was the DC measure for C measure the party's candidate, for example, the party was was coming. When Taiwan present, really United States, that was a very first huge activity of the toy activity happening. And then another peak is when Joe Biden said, when Taiwan is under stress, they were in those Taiwan, then there we see a speck or inflammation when you pray is the same, that United States is helping Taiwan to develop bio VoIP, they try to destroy the narratives. Also, yeah, so

2:02:13
this is basically China's the source your I feel like your yes,

2:02:17
go according to our understanding they are a lot of they will the troll con and the social media level, try to emphasize the military Strait and China. And the way you look into the state of Fenian media. So we can compare, we use artificial intelligence, we can group the same narrative together, then we can see

the troll cannon, the Facebook, Twitter, for example, lay what a call the narratives, and the channels they meet.

2:02:48
Fred Oh, great, well, very interesting that you're able to do that kind of detection work. And I know you've put out some really interesting reports that everybody can read who's interested. Dominic, I wanted to find out from you is this. Does this sound like what you're seeing in your part of the world sort of video narratives kind of being spread out from state actors? Audio? What what? What's the sort of state of play where you are? Yeah.

2:03:14
So also, just through an intro, tell

2:03:16
us about your organization. Yeah, absolutely. I have some more detailed

2:03:20
news myself. So I come from globe, sec. It's a think tank, which was founded in Central Europe. We were founded in Bratislava, Slovakia. But we covers basically all countries in Central Eastern Europe. And we now have offices in Kyiv, and Brussels and in DC. And I am leading the Center for Democracy and resilience. And we were founded in 2015. So right, shortly after the annexation of Crimea, and the invasion of Ukraine, because we started seeing the floods of information manipulation, and disinformation across Central and Eastern Europe, primarily coming from from the Kremlin. And at that time, of course, this was mostly limited to a few pages, that that spread pro Kremlin propaganda, and it was very visibly pro Kremlin or pro Russian. But then, of course, as it has been mentioned, during the first panel, I think the tactics, of course, have evolved tremendously. First of all, the, especially the Kremlin incident in the context of Central and Eastern Europe has been able to build the networks and the proxies. Right. So right now, the Vice President mentioned it of the European Commission, it's not that much about the Kremlin interfering directly, but it's, it's through the domestic actors, political actors, websites, social media pages, etc, etc. And I come from Slovakia, and we had elections in September. And it was a peculiar case because we could actually see both direct and indirect interventions from the crowd. When direct, a lot of the countries in the EU have this political campaigning silence, which is which means that from one day to two days prior to the elections, you cannot do any campaigning basically. And during this period, there was a press release by the Russian press agency, saying that the US was going to interfere in the elections by doing everything they can to for a pro Democratic Progressive Party to win. And just so that you're aware how it was, is the two parties, one rather nationalistic populist with some very strong pro Kremlin figures was running first and just and progressive, liberal pro Democracy Party was running second. So this was released. And among very similar time, a deep fake audio and inserted into a synthetic made audio was also released on telegram by an account which was probably a wife of a Slovak political representative, who is currently being prosecuted for spreading Russian word propaganda. So attribution in this case is quite difficult. But this deep fake has spread through telegram on Facebook, and has had 1000s of shares on Facebook, despite the fact that it was quite an Wowsie audio that if you listen to it carefully, you would actually see that it wasn't that it that it wasn't true. The problem is that despite the fact that it's quite a small case,

and quite a small country draws several important lessons. First, is that we really need some red lines, clear lines to be made when it comes to generative AI prior to the elections. Because I do agree that labeling is important. And watermarking labeling is a way to go definitely. But what if there is a content spread 24 hours before the elections, and it's made by Kremlin based or Beijing based company, which doesn't require such watermarking? Because this is going to be a consensus among the Western based companies. Is there an ability to stop this? If this is 24 hours prior to the elections? Are we going to ban it are we going to take it down? I think that there are also some red lines that have to be defined, and I think the Michigan law can be could be a way forward. Second, is that these measures also have to be clearly defined for social media platforms, because what has happened in Slovakia specifically is that there were around 70 pieces of AI Generated Content identified, and around half of them stayed aligned 15 were taken down and searching relabeled or something like that. So there is quite an inconsistency in treating these these these cases. And of course, we need to treat some of the cases on on specific on a specific basis, whether they, whether they can whether they are talking about election manipulation, or the narratives of rigged elections, which actually most of these big fakes we're talking about. So this is a common tactic. That has been said,

2:08:27
I hope we'll have time to talk about regulation. But I know heavier was saying, you know, we've obviously some of our alumni, a lot of people have been very involved in tracking Russian disinformation in South America. And I've heard you were saying have year that actually that's not really the problem in Argentina. So that's quite interesting. What is the problem? I know Malay won with a lot of youth vote, I guess, hold they very high unemployment. And inflation has been upsetting people quite a bit. Exactly. What what's economics and what's information?

2:08:59
Yeah, well, that was one of the main things that we observed. In my previous work, I worked at an organization called Axios. Now, I used to be the global director of policy there, and we were able to see how these issues evolve around the world, right, and called my attention in Argentina, that there were quite quite specific characteristics that I haven't seen anywhere else. Right. And, for example, one of them is that we found when I did, I also was working on on an independent research on the Argentinian elections. We haven't found any specific, like, clear evidence, or at least an initial evidence of foreign intervention. Most of what happened with the online social movement that brought me late to power was quite organic. We didn't see much of fake accounts or troll centers, we didn't see much of foreign intervention, which was very interesting, you know, in the sense of that, you As you mentioned before, Argentina is a country that has been divided for a while now. You know, it started with kitchen tourism a couple of couple of presidential periods ago, you know, and this kind of like internal political division was really, really strong from the beginning. So when the one of the candidates appears and makes a proposal that goes against the status quo, there's another like half of the country's ready to engage. And there is a need to engage, there's something that we detected, which is very, very clearly a resonance of certain kinds of messages and words, and ways of communicating online that resonated with the people. So I would say that the main aspect of the campaign formulae is the sentiment of anti politics, people are not only disenfranchised, or disillusioned with politics, they are offended with politics, and with politicians, you know, it's a personal thing, it's a real, really a reaction movement. And these people who don't talk like politicians, don't look like politicians, you have seen his, you know, his

looks, or the way he conducts himself, and so on, they are really, really attractive for the kind of people and especially younger people who, as was mentioned before, by by Mrs. Green, you know, they have a different way of understanding information. And this idea about moving from institutions, to people as the source of, of authority is something that is really resonating with, with the, with the population, and it's easy to understand, for example, in Argentina, our military dictatorship lasted until the 80s, which in internet times is the Middle Ages, but in historic times this yesterday, right? So this idea of not trusting institutions, the government being a potential, you know, source of oppression, and violence. And also at the same time, this idea of like, institutions that are really young, that haven't had the time to really, you know, get a good basis in our society together with corruption and other kinds of things. It's, it's a terrible mix that is ready, it's like the fertile ground for any of this, you know, populist leaders to just appear and, and have a lot of following. So I will say that that was one of the key points. Of course, the lack of regulation for for for platforms is a problem. countries like Argentina are second class citizen, second class users for some platform. And

2:12:23
that's what I also wanted to talk about is how much agency do you have? You know, it's, it's great to hear all the companies talk about the new standards, but let's face it, if they voluntarily started doing content provenance, they would set the standard for everybody. So you know, sitting around saying, I'm waiting for regulation, you can also model with best practices. And then I'm thinking precisely, I mean, we know from all the reporting that's been done, including by many of the people in the room, that that your countries, you know, have less moderation, they have nobody to call their you know, minority languages aren't properly, you know, moderated are looked after. So I'm wondering your perspective on and I obviously, Dominik will talk to us about the Digital Services Act, but we are going to sit back and wait for the big companies to change their policies to start doing content authenticity. What can you do on your own? And I feel like, you know, Brazil is really been leading the way for Latin America in terms of regulation, but other countries haven't been. And I'm curious to know whether, you know, regulation comes up in Taiwan as well, because it's not something I know that much about, do you want to go first and then we'll hear from Ethan and then we'll talk about DSA or

2:13:37
so. So when we talk about regulating Asia, Taiwan, ever had a failed case, try to regulate the platform by being fed, they say people label information manufacturers say even go against freedom of speech. So it's like the US. Yes, so. So, Taiwan, we just recently published information manipulation about the talk, maybe you can go to our website, invert me.cc Look into law, a lot narrative is pretty similar to what we happen in Taiwan before

2:14:11
Could you tell us like what, what regulation had been considered and what were the forces that defeated it?

2:14:17
So, I will say, if we go to the definition of faith checking and the content moderation, that there is you will have a lot of challenge because people will say, not against the freedom of speech, and also fake check what is fake? And that will be a lot of debate. Yeah. So So in Taiwan, instead of where so now we are

instead we are talking about fake checking, all content moderation. We are talking about how we can disclose the information manipulation,

2:14:50
right, of course, so I'm being given the sign that we only have five minutes, okay. But this is of course, what's happened to the US, right. We were told you can't regulate but we can at least do media literate See, and fact checking. Then it turned out that even teaching media literacy was controversial. And all the researchers were doing the tracking are all about getting subpoenaed. So that just like in the 1930s, when Columbia was also pioneering where the space is getting pushed, so that's really interesting. I'm definitely gonna go to your website. As soon as this is over, have your any conversation about regulation, then we're going to finish optimistically, DSA but very

2:15:23
quickly, I think that another untapped resource, Brasil is a great example. Another anti untapped resource is the Inter American system for human rights. The freedom of expression standards contained there are widely accepted across Latin America and across America, in general, it's a very good mix between a more regulatory strong, strongly oriented stance from the European side and more allowing, let's say, First Amendment standards in the US. So I think that there's an interesting middle way to work there, there is jurisprudence from the Inter American Court on Human Rights, for example, on indirect means of affecting freedom of expression, one of them, for example, could be the unduly interference and you know that some of the external actors or sometimes social media companies themselves do on the discourse of people. So there's a lot to grow there. And of course, there's a lot to do in terms of electoral regulation, modernization of the bureaucracy of the electoral Commission's giving more power, more agency to them. Bolsonaro, for example, was stopped very, very quickly by the electoral authorities.

2:16:23
Interestingly, he's been banned from participating for really much longer than

2:16:27
at least his whole electoral regulation. It's not too many. Okay, get the last

2:16:31
word, because I know we're running out of time. Oh, PSA, how do we feel? Is it helping? So

2:16:37
to DSA targets, illegal online speech, and I think this is a very powerful legislation in a way that doesn't target disinformation. Because there you're on a very systemic risk. So you're coming up with a plan. But when it comes to illegal speech, I think that it is making progress because they're actually requirements for the platforms to issue regular reporting, which is helping us and it's per country basis. So this is important because in languages like Dutch, Slovak, Czech Hungarian, you actually have to see what has been done. Because we didn't have this information before. So in this sense, it's really good. And I

2:17:17
think it will start to kick in because I know that the different countries are still staffing up.

2:17:22
Yeah, it has, it has started already. No, no, but

2:17:26
when will we all notice that? Oh, well. Okay,

2:17:29
so there are already reports out, so you can already check those out. So so if you do a bit of research, you will notice it. But in terms of other platforms, for example, you can already like, report illegal content. What I'm worried about is the platforms that are not cooperative. So if there's so many, so much exchange of information between Facebook, Microsoft, Google, that's amazing. But then what about Telegram for example? Yeah, right, which is the source of extremism and also pro Russian propaganda and all the mind content very

2:18:01
much so and there's been so much in so much interesting stuff anyway, we could go on all day, but I certainly don't want to get in the way of the next panel, which is gonna be really interesting. So thank you very, very much. Yeah, and Ethan in Dominica and hopefully

2:18:21
got out that way.

2:19:26
Please welcome back to the stage secretary, Hillary Rodham Clinton and joining us virtually IGP Carnegie distinguished fellow Eric Schmidt.

2:19:45
First, we are so delighted to have Eric Schmidt with us. Especially because he is as you just heard, one of our Carnegie distinguished fellows at the Institute of global politics, and he has been and meeting with students and talking to faculty about a lot of these AI issues that we have surfaced during our panels today. And of course, he wrote a very important book with the late Dr. Henry Kissinger about artificial intelligence. So we're ending our, our afternoon with Eric and trying to see if we can pull together some of the strains of thinking and challenges and ideas that we've heard. So Eric, thank you for joining us, you look like you're in a very comfortable but snowy place. And I wanted to start by asking you, what are you most worried about with respect to AI in the 2024 election cycle?

2:20:50
Well, first, Madame Secretary, thank you for inviting me to participate in all the Columbia activities. I'm at a tech conference in AI conference in snowy Montana, which is why I'm not there. If you look at misinformation, we now understand extremely well, that virality emotion and particularly powerful videos, drive voting behavior, human behavior, moods, everything. And the current social media

companies are weaponizing that because they respond, not to the content, but rather to the emotion, because they know that things that are viral, are outrageous, right, you know, click crazy claims get much more spread. It's just a human thing. So my concern goes something like this. The tools to build really, really terrible. Misinformation are available today, globally. Most voters will encounter them through social media. So the question is, what are the social media companies doing to make sure that what they are promoting, if you will, is legitimate under some set of assumption?

2:22:03
You know, I think that you did an article in the MIT Technology Review, fairly recently, maybe at the end of last year. And you put forth a six point plan for fighting election, misinformation and disinformation. I want to mention both because they are distinct. What were your recommendations in that article to share with our audience in the room and online? Eric? And what are the most urgent actions that tech companies particularly as you say, the social media platforms, could and should take before the 2024 elections?

2:22:47
Well, first, I don't need to tell you about misinformation, because you have been a victim of that, and in a really evil way by the Russians. When I look at the social media platforms, here's here's the blunt fact, if you have a large audience, people who want to manipulate your audience will find it and they'll start doing their thing. And they'll do it for political reasons, for economic reasons, or they're simply needless, there are people who just want to take down powerful figures because they don't like authority. And they'll spend a lot of time doing it. So you have to have some principles. One is you have to know who's on the platform. And in the same sense that if you have an Uber driver, you don't necessarily know that Uber drivers name and details, but you can be quite sure that Uber has checked them out because of all the various problems they had in the past. So you trust Uber will deliver you a driver that is at least a legitimate driver. Right? So that's sort of the way to think about it, the platform needs to know even if it doesn't tell you who they are, that they are real human beings. Another thing you have to know is where did it come from, and we can technologically put watermarks on the technical term is called steganography, where you use an encryption technique and you mark where the content came from. So you know roughly how it entered your system. You also need to know how the algorithms work. We also think it's very important that you work on age gating. So you don't have people below 16. And those are relatively sensible ways of taking the worst parts of it out. I think one of the things that's happened since I wrote that article is if you look at the success of Reddit and their IPO, what they did, they were reluctant like everybody else in my industry, they were reluctant to do anything. They brought in a new CEO who shut down entire sub Reddits of hate speech. And it improved the overall discourse. So the lesson I've learned is if you have a large audience, you have to be an active manager of people who are trying to distort what you as the leader are trying to do

2:24:51
that Reddit examples of very good one because you know, I don't have anything like the experience you do, but just as an observer And it seems to me that there's been a reluctance on the part of some of the platforms to actually know it's kind of like they want deniability. I don't I don't want to look too close, because I don't really want to know. And then I can tell people I didn't know, and maybe I won't be held accountable. But I actually, I think there's a huge market for having more trust in the platforms,

because they are taking off, you know, certain forms of content that are dangerous in however you define that. And your your recommendations in your article, focus mostly on the role of content, distributors. So maybe go a little bit further, Eric and explaining to us like, what should we think about? And maybe more importantly, what should we expect from Ai content creators and from social media platforms that are either utilizing AI themselves or being the platforms for the use of generative AI? How do we think about protecting our elections? And doesn't matter whether it's a social media platform, a big AI company or even open source developers? Is there some way to distinguish that? Well,

2:26:21
it's sort of a mess, as the previous panel discussed, and the reason it's a mess, is there are many, many different ways in which information gets out. So if you go through the responsibility, the legitimate players, the authoring tools, and so forth, all have a responsibility to mark where the content came from. And to mark that it's synthetically generated, that seems kind of obvious. In other words, we started with this, and then we made it into that. And there are all sorts of corner cases, like I touched up the photo, well, you should record that it was. So you know that it's an altered photo, it doesn't mean an evil way. But that's an example. The real problem here has to do with a confusion over free speech. So I'll say my personal view, which is I'm in favor of free speech, including hate speech that is done by humans. And then we can say to that human, you are a hateful person, and we can criticize them, and they can listen to us, and then we hopefully correct them. That is my personal view. What I am not in favor, is is of free speech for computers. And the confusion here is you get some idiot, right? Who is just literally crazy, who's spewing all this stuff out, who we can largely ignore, but the algorithm then boosts them. So there is absolutely liability on the social media platforms responsibility for what they're doing. And unfortunately, although I agree with what you said, the trust and safety groups in some companies are being made smaller and are being eliminated. I believe at the end of the day, these systems are going to get regulated and pretty hard. And the reason is that you have a misalignment of interests. If I'm the CEO of a social media company, I want to maximize revenue, I make more revenue with engagement, I get more engagement with outrage. So one of the ways to think about is why are we so outraged online? Well, it's partly because the media algorithms are boosting outrageous stuff. Most people it is believed or more in the center, and yet we focus on and this is true of both sides, everybody's guilty. So I think that what'll happen with AI just answer your questions precisely, is the AI will get even better at making things more persuasive, which is good in general for understanding and so forth. But he's not good for the standpoint of election truthfulness.

2:28:43
Yeah, that that is exactly what we've heard this afternoon is that, you know, the sort of authoritativeness and the authenticity issues are going to get more and more difficult to discern, and then it'll be a more effective message. And, you know, I was struck by one of your recommendations, which is kind of like, it's, it's a recommendation that could only be made at this point in human history, and that is to use more real human beings to help. And it's almost kind of absurd that we're sitting around talking about, well, maybe we can ask human beings to help human beings figure out what is or isn't, you know, truthful, but how do we incentivize tech companies to actually use human beings and how do we avoid the exploitation of human beings because there's been, you know, some pretty, you know, troubling disclosures about, you know, the sort of sweatshops of human beings in you know, certain countries in the Global South, who are being you know, driven to make these decisions and it can be quite, you

know, quite overwhelming. So, when you've got companies, as you just said, gutting trust and safety? How do we get people back to, you know, some kind of system that will make the kind of judgments that you're talking about? Well,

2:30:12
speaking as a former CEO of a large public company, companies tend to operate based on fear of being sued. And section 230 is a pretty broad exemption. And for those in the audience, section 230, is that is sort of the governing body on how content is used. And it's probably time to limit some of the broad protections that section 230 gave, there are plenty of examples where someone was shot and killed over some content, where the algorithm enabled this, this terrible thing to occur, there is some liability. Now we can we can try to debate what that is. But if you look at it, as a human being, somebody was harmed, and there was a chain of liability, including an evil person, but the system made it worse. So that's an example of a change. But I think the truth if I can just be totally blunt, is ultimately information and the information space we live in. You can't ignore it, I used to give this speech where I would say you know how we solve this problem, turn your phone off, get off the internet, eat dinner with your family, and have a normal life. Unfortunately, my industry and I'm happy to have been part of that, that made it impossible for you to escape all of this as a normal human being. Right? You're exposed to all of this terrible, and filth, and so forth and so on. That's going to ultimately either get fixed by the industry collaboratively or by regulation. A good example here is let's think about tick tock because tick tock is very controversial right now, it is alleged that a certain kinds of content is being spread more than others, we can debate that tick tock isn't really social media tick tock is really television. And when you and I were younger, there was this huge practice over how to regulate television. And there was a something called an equal time rule. And ultimately, it was a sort of rough balance, where we said, fundamentally, it's okay, if you present one side, as long as you Prince present the other side in a roughly equal way. That's how societies resolve these information problems, it's going to get worse unless we do something like that.

2:32:24
I agree with you 100%. In both your analysis and your recommendations, and in the very first panel, we talked about, the need to revisit, and if not completely eliminate, certainly, dramatically revise section 230. It's outlived its usefulness. I mean, there was an idea behind it back in, you know, the late 90s, when this industry was so much in its infancy. But we've learned a lot since then. And we've learned a lot about how we need to have some countability, some measuring of liability for the sake of the larger society, but also to give the direction to the companies. I mean, these are very smart companies, you know, that you spent many years at Google, these are very smart companies, they're going to figure out how to make money. But let's have them figure out how to make a whole lot of money without doing quite so much harm. And that partly starts with dealing with section 230. You know, when we were talking earlier, about, you know, what AI is aiming at, you know, the the panelists were all, you know, very forthcoming and saying, Look, we know, there are problems, we're trying to deal with these problems. You know, we know from even just the public press that a number of AI companies have invented, tools that they've not disclosed to the public because they themselves assessed that those tools would make what is a difficult situation a lot worse. Is there a role you think, Eric, for another was the Munich you know, the statement negotiated at the Munich Security Conference, which was a start? But is there more that could be done with a public facing statement, some kind of agreement by the AI

companies and the social media platforms, you know, to really focus on preventing harm going into the election? Is that something that's even, you know, feasible?

2:34:25
It should be. The reason I'm skeptical is that there's not agreement among among the political leaders. Of course, you're a world's expert on that. And the companies on what definition what defines harm. I have wandered around Congress for a few years on these ideas. And I'm waiting for the point where the Republicans and the Democrats are in agreement on from their local and individual perspectives, that there's harm on both sides. We don't seem to be quite at that point. This may be because of the nature of how President Trump Works, which is always sort of baffling to me. But there's something in the water that's causing a non rational conversation. It's just not possible. So I'm skeptical that that's possible. I obviously support your idea. I think the other thing I would say, and I don't mean to scare people, is that this problem is going to get much worse over the next few years, maybe, or maybe not by November, but certainly in the next cycle, because of the ability to write programs. So I'll give you an example. I was recently doing a demo, the demo consists of you pick a stereotype stereotypical voter, let's say it's a Hispanic woman with two kids, she's, you know, she has the following interests, you create a whole interest group around her, and she she doesn't exist, it's fake. And then you ask the computer to write a program in Python, to generate 500 variants of her different sexes, different races, so forth, and so on different ages and backgrounds who commingle the same voices. So the ability to have AI, broadly speaking, generate entire communities of pressure groups that are in fact, virtual, it's very hard for the systems to detect that these people are fake. There are clues and so forth. But to me this question about the ability to have computers generate entire networks of people who don't exist to act for a common cause, which may or may not one that you and I agree on, but he's probably influenced by either national security for North Korea, or China, or Russia, or are influenced by some business objective from the tobacco companies, or you name it. I worry a lot about that. And I don't think we're ready. These the, it's possible just to hammer on his point for the evil person who inevitably is sitting in the basement of their home and their mother gives them food at the top of the stairs to do this on their computer in a day. That's how powerful these tools are.

2:37:06
Okay. Well, let's try to, you know, bring it back a little bit to where we are here at the University in this, you know, great setting of so many people who have a lot to contribute, and working in partnership with Aspen digital, which similarly has a lot of convening and outreach, potential. What can universities do? What can we do in research, particularly on AI? How do we create a kind of, you know, broad network of partners like we're doing here between IGP and Aspen digital, and, and we began to try to do what's possible to educate ourselves educate our students in combating myths, and disinformation, with respect to elections.

2:38:04
So the first thing we need to do is show people how easy it is. And so I would encourage every university program to have to have students actually try to figure out how to do it, obviously, don't actually do it. But it's relatively easy. And it's really quite an eye opener was an eye opener for me. And I've done this, you know, for as long as I've been alive. The second thing I would do is there are there's an infrastructure, that would be very helpful. The best design that I'm familiar with is blockchain based.

And it's essentially a name, it's a name and origin for every piece of document independent of where it showed up. So if everyone knew that this or that this piece of information showed up here, you could then have provenance and understand how did it get there? Who pushed it? Who amplified it that would help our security services, our national security, people don't understand? Is this a Russian influence campaign? Or is this something else? So there's technical things and then there's also educational things? I think this is only going to get fixed. If there is a bipartisan broad consensus that taking the edges, the crazy edges, the crazy people, and you know, I'm talking about and basically taking them out. I'll give you an example. There was an analysis in the last in COVID, that the number one spreader of misinformation about COVID. Online, was a doctor in Florida, who was like 13% of all of it. And he was very clever. He had a whole influence campaign of lies trying to convince you to buy his supplements versus getting a vaccine. That's just not okay, in my view. And the question for me is, why was that allowed by that particular social media company to exist even after it was pointed out? So you have a moral framework, you have a legal framework, you have a technical framework, but ultimately it has to be seen as it's not okay. To allow this evil doctor for profit to allow people to get to, essentially to mislead them on vaccinations.

2:40:06
Well, just to follow up on on that, I mean, I don't at all disagree about what has to happen if we're going to end up with some kind of legislation or regulatory framework from the government. But is there if they were? So if they were willing, is there anything that the companies themselves? Could do, as I say, if they were willing to that would lay out some of the guardrails that need to be considered before we get to the consensus around legislation? Of

2:40:36
course, but, of course, the answer is yes. But the way that the way this actually works in a company is you don't get to talk to the engineers, you get to talk to the lawyers. And the lawyers, as you, as you very well know, are very conservative, and they won't make commitments. So it's going to require some kind of an agreement among the leadership of the companies of what's inbounds and what's out of bounds. Right. And, and getting to that is, is a process of convening and conversations is also informed by examples. So I would assert, for example, that every time someone is physically harmed from something, we need to figure out how we can prevent that, that seems like a reasonable principle if you're in the digital world as I am, right. So working back from those principles is probably the way to get started. It's not going to happen. Unless there's agreement, either it's forced on them by the government, or there's agreement by the CEOs, the best way to achieve that, in my view, is to make a credible and detailed proposal of where the guardrails are. Right? And, and what it means what I have learned in working on this, is you have to have content moderation, when you have a large community, these these groups will show up, they will find you because their only goal is to find an audience to spread their evil, whatever, whatever the evil is, and I'm not taking sides here.

2:42:01
Well, I think the guardrails proposal is a really good one. And obviously, you know, we hear at IGP, Aspen digital, the companies who are here, others, the researchers who are here, I mean, you know, maybe people should take a run at that. I mean, I, you know, I'm not naive, I know how difficult it is. But

I think this is a problem, we all recognize, it's not going to get better if we just keep wringing our hands and fiddling on the margins, we have to try something different. So let

2:42:34
me let me just be obnoxious. You know, I've sat through all of these Trust and Safety discussions for a long time. And these are very, very thoughtful analysis. They're not producing solutions in their analysis that are implementable by the companies in a coherent way. So here's my proposal, identify the people understand the provenance of the data, publish your algorithms be held, as a legal matter that your algorithms are what you said they are, right? In other words, what you said you do, you actually have to do reform section 230, make sure you don't have kids on so forth, etc, you know, make your proposals, but make them in a way that are implementable by the team. So for example, if there's a particular kind of piece of information that you think should be banned, right and out, write a specification of it well enough, that under your proposal, the computer company can stop that, right. That's where it all fails, because the engineers are busy doing what they understand. They're not talking to lawyers, too much the lawyers job is basically prevent anything from happening, because they're afraid of liability. Right. And you don't have leadership from the Congress for the reasons that you know, and that's why we're stuck.

2:43:48
Well, that's both a summary and a challenge, Eric, and I particularly appreciate that, and especially the work you've been doing to try to, you know, sort this out and give some guidance. So you get the last word from beautiful, snowy Montana, the last word to kind of, you know, offer that challenge, you know, ask us to respond, to follow up on what you've outlined as at least one path forward and try to do it in a, you know, a collaborative way with the companies and and other concerned parties.

2:44:25
I do this as the snowstorm is hitting in behind me. Look, I think that the most important thing that we have to understand is this is our generations problem. If this is a huge this is under human control, there's this sort of belief that none of this stuff can get fixed. But you know, from your pioneering work over some decades hear that with enough pressure, you really can bend the needle, you just have to get people to understand that these problems are not unsolvable. This is not quantum physics, which is impossible to understand. It's a relatively straightforward problem of what's appropriate, and what's not the AI algorithms can be tuned to whatever society wants. So my strong message to everyone at Columbia and of course, all the partners, is instead of complaining, which I like to do a great deal, why don't we collectively write down the solution? Organize partner institutions, try to figure out what the how to get the people in power to say, okay, I get it. Right, that this is reasonably bipartisan, it makes society better. There's this old rule about Gresham's Law, which is that that bad speech drives out good speech, which is why the Internet is a cesspool. I used to say that and I would say, and since I don't like to live in a cesspool, just turn it off. So the problem you have, and this is especially true for young people, the damage that's being done online to a women and so forth, it's just horrific. Why would we allow this in modern society, we can fix it, you just have to have an attitude. I've tried to fund some open source technology in this area that are better tools to detect bad stuff. It's gonna take, it's gonna take some concerted effort. And I really appreciate your secretary, your your attention on this. Somebody's got to push.

2:46:21
Well, you and I, let's, let's keep going, Eric, and I'm so grateful to you. And I hope you have a great time in the snowstorm and whatever else comes next. But let's show our appreciation. Derek Schmidt for being with us. Thank you so much. Thank you all. Well, I think we have a call to action, we just have to get ourselves in the frame of mind that we're willing to do that. And even writing something down will help to focus our, you know, mind about what makes sense and what doesn't make sense. So we're not going to let you off the hook, we want to come back to you. We want to have something come out of this. We can talk about this meat about this till the cows come home. But in the meantime, as Eric said, and I agree, it'll just get worse and worse. And we have to figure out how we can assert ourselves and maintain the good and try to deal with you know that which is harmful. So please join us in this effort. And as I say, we will come back to you and seek your guidance and your support. Thank you all very much