We're all set John. John, you're on mute.

John, are you able to connect to audio? I am, can you hear me? Yes, now we can. Good afternoon. On behalf of the American Geographical Society council, our members, and the staff,

it's my pleasure to welcome you to this fifth Location Tech Task Force Blue-Ribbon Panel; Data Quality and Building Trust: False Negatives, False Positives, and Policing/Surveillance.

For those of you participating in the conversation on WebEx and those that are watching our livestream on Facebook, it's great to have you joining us this afternoon.

We're proud to partner our EthicalGEO initiative with the Henry Luce Foundation to investigate the societal implications of geospatial technology and location tracking.

Mobile-based location applications have become ubiquitous in our society.

As all of you know, they've changed the way we live our lives in a very short period of time.

There are however problematic and unanticipated effects of using this technology.

To better understand the ethical implications of its use, we have provided this platform to frame the discussion and to address these issues as they are already impacting our lives on a daily basis.

COVID-19 has put a spotlight on the concept of using mobile tracing and surveillance to fight the pandemic.

Around the world,
the utilization of this technology to fight the coronavirus is being employed to various
degrees and already governments and people worldwide are faced with the issue of
compromised privacy and what that means as we go forward. Over the past

our first four Blue-Ribbon Panels met and looked at ethical implications of mobile
location technology and the impact on vulnerable publics from an international
perspective, and from the unique American experience.

In addition, we had a panel of national security leaders who focused on mobile tracing
technology, and its use in national security and democracy. Just a couple of days ago,
we heard from state and local leaders who shared their invaluable experiences with us.

In the case of all the panels, the discussions were eye-opening and extensive.

We also had the opportunity to hear from Ambassador Samantha Power. She added
the human rights aspect of the use of mobile technology.

Today, we appropriately turn our focus towards data quality and building trust, a topic
noted by many of the participants in our previous panels.

Over the next couple of weeks, we will continue to convene Blue-Ribbon Panels and
Leadership Spotlights and collect testimony from other perspectives to service the
basis of information and data that policy makers can use to help and guide us into the
future.
most out of today's panel. For those viewing on desktop computers or laptops, we recommend that you customize your viewing by hovering the mouse in the top right of your screen and selecting the icon in the middle.

During our Q and A session later, to ask a question, hover your mouse under the arrow and click on the question mark icon in the gray bar at the bottom of your screen. For those of you using a tablet or mobile device, select the icon with the three dots, which will then allow you to select the question mark icon to submit your questions to our panel.

Now, it's my pleasure to introduce Dr. Christopher Tucker, the Chairman of AGS and our moderator for today's session. Chris, it's a pleasure to welcome you and your panel.

Thank you, John and I thank you to the team for putting all this together today. There's a lot of moving parts behind the scenes that no one ever gets to see, so thank you for your efforts.

Our society, AGS, has been convening people from government, industry, academe, and the social sector for the past 169 years since 1851 around the vital issues of the day through geographic lens. And in both the vital issues of the day have changed vastly since 1851, in many ways, maybe in some cases not, and the nature of geospatial technology has also changed considerably over the years. Last year we began our EthicalGEO initiative to understand, better understand the ethical implications of geospatial data and technology on our society and with COVID-19, issues of contact tracing and the new roles of location based technologies.
in combating COVID-19, we partnered with the Henry Luce Foundation to help us explore the ethical implications of mobile location tech in an era of COVID-19 and beyond, particularly with an eye towards the next challenge, the next threat that will face our society.

So, in that context, I could not be more pleased to have our speakers that we have with us today and let me introduce each of them before we get started.

Mr. Albert Fox Cahn is the Surveillance Technology Oversight Project, (S.T.O.P.’s) Founder and Executive Director. He's a fellow at Engelberg Center on Innovation Law and Policy at NYU, NYU School of Law, and a columnist for Gotham Gazette. He began S.T.O.P. because he believes that the emerging surveillance technologies pose unprecedented threats to civil rights and the promise of a free society. Dr. Hyeong-Ah Choi is the Professor of Computer Science at George Washington University.

Her research lab centers its activities on graph and optimization algorithms and network problems and on analysis and design of algorithms relevant to the modern,
massive data problems that we face. Her current projects include the development of low complexity algorithms to find underutilized wireless channels and cognitive radio networks,

47

and the analysis of modern, online education, content, and finding optimal learning paths. Ms. Faine Greenwood does consulting work

48

that focuses on ethically integrating new technologies into existing organizations and processes with clients, including the World Bank and the American Red Cross.

49

She is also a humanitarian and technology author. As a researcher at the Signal Program at the Harvard Humanitarian Initiative,

50

she focused on the use of unmanned aerial vehicles and spatial information in humanitarian aid. Faine is also a regular contributor to ‘Slate’ and ‘Foreign Policy’. And last is Dr.

51

Julia Stoyanovich,

52

who is Assistant Professor of Computer Science and Assistant Professor of Data Science at NYU. Her research focuses on responsible data management and analysis practices, operationalizing

53

fairness,

54

diversity,

55

transparency, and data protection.

56

She established the Data Responsibility Consortium

57

and serves on the New York City Automated Decision Systems Task Force. So, with that, I will pass it over to our first speaker. Each of our speakers will be speaking for roughly 7 to 10 minutes,

58
and then we will move on to a discussion between our speakers. And then we'll open it up to the questions from our online audience.

So again, please follow the guidance that John Konarski gave you earlier. At any point in this presentation feel free to upload your questions to WebEx.

So again, thanks for joining us, and we'll begin by passing this over to Albert. Albert?

Well thank you so much for inviting me to take part in this conversation, and the issues we're dealing with here today involving privacy and technology, and the deployment of mobile tracking in response to COVID-19 and beyond, it really could not be more important and impactful and how it will shape, you know, decisions here in the United States and around the world. You know, just as a brief introduction, if we could go to the first slide?

Alright, so I am the founder of the Surveillance Technology Oversight Project, an organization here in New York that fights police and state-level surveillance. It is a mix of litigation legislation, public education, grassroots, community,
organizing and programs designed at trying to tackle the most destructive uses of surveillance and the ways that public surveillance can give new powers to police and to immigration officials, while analyzing the unique harms that these technologies pose to historically over-policed communities. Next slide.

So, at S.T.O.P. we've been quite concerned about the deployment of exposure notification systems that transform individual smartphones into potential diagnostic tools to detect if someone has been exposed to COVID-19. We believe that this use of mass surveillance is both unproven, but potentially misleading in the data it provides, and that it creates a massive new surveillance capacity. We've looked at, primarily at the Bluetooth API that was developed by Google and Apple, trying to analyze the privacy impacts of that data, but also its limitations, the lack of contextual data, the lack of understanding whether a phone is registering another device's proximate simply because the two individuals are in adjoining rooms or the same room. Are they outside or are they inside? These are the contextual data points that mass surveillance will provide us, whether it's through Bluetooth GPS or any other form of mass location tracking.
We also are quite concerned about reliance on end-user devices, such as smartphones, because that can easily bake in-bias by allowing us to depend on the performance of those user's devices when we know that many, many users, because of socio-economic status or other factors will be using very different phones, very different types of wearables, and this will lead to a very inconsistent and potentially discriminatory data set.

Next slide. Specifically, we're concerned about the fact that we see incredibly low smartphone penetration rates among low income elderly individuals, especially those in communities of color, creating a situation where those communities who have suffered the most under COVID-19 would potentially be least served by these exposure notification systems.

We also are quite concerned by the fact that there are a variety of factors that potentially could impact the performance of Bluetooth as a proximity tracking system, everything from housing density and housing stock to the age of the phone, power settings,
whether a phone is being held in the hand, or whether it's being placed in pants pockets or a purse, these factors that potentially would lead to higher rates of false positives and false negatives for the exposure notification system, have the potential to replicate the sorts of bias that we've seen with AI when deployed in the law enforcement sector, as well as a number of health applications. Next slide.

Crucially, we challenge the question, the premise that these technologies are opt-in. Google and Apple in particular have highlighted the fact that a user would have to install an app and click 'I agree' on terms of service prior to using one of the exposure notification systems. But this isn't purely a legal question, this is, sorry, this isn't purely a technical question, it's also a legal question, because what happens if that individual is told that they will lose their job, or if they will not be able to attend school, or they won't be able to access public accommodations or mass transit without clicking that box and saying 'I agree'.

We currently don't have protections against the sort of coercion that would undermine the premise that these technologies are consensual. And when we are laying the foundation for mass coercive surveillance, it also normalizes some forms of mass surveillance that we oppose that stop. Because they are often used as a form of state power against those communities that have the least power to fight back. Next slide.

And we're also quite concerned about the fact that many of the novel tracking tools that we see being deployed in response to COVID-19 have no scientific track record to actually demonstrate that they're effective,
there's no peer reviewed evidence,
there's no independent evaluation.
There's often times with AI driven systems,
no independent auditing for bias, and so with some of the systems we see being proposed for location tracking, the privacy impact and the equity impact is really quite disturbing.
One example is facial recognition. We've documented at least one college that has looked at installing facial recognition kiosks as a way of determining, as a way of tracking students as they go around campus and using that persistent location information as a contact tracing technology.
As, you know, facial recognition is an increasingly controversial technology in part, because of its demonstrated bias against communities of color, particularly darker
skinned,
black women.
And because of the underlying errors in these technologies
and because of the way they're currently being challenged,
we're quite skeptical that many of the vendors that are entering the COVID-19 contact tracing space with location based systems, whether they be biometric or whether they'd be smartphone-based are actually going to be able to live up to the promises that they're making.
But, crucially, these sorts of invasive systems also undermine community trust. They also arode belief in the legitimacy of contact tracing more broadly
and they potentially make it more difficult for public health authorities to use evidence-based measures,
such as traditional manual contact tracing,
especially manual contact tracing by culturally competent individuals who are rooted in the communities where outreach is taking place. Next slide.
perfect. And perhaps the most distressing concern we see in this space is the lack of protection against police misuse of data collected in COVID-19 contact tracing and as part of other public health tracking efforts. Under federal law, a bill called HIPAA, there is simply a requirement that police obtain a subpoena, not even a warrant, prior to obtaining contact tracing data, and that's only in those cases where HIPAA applies. Oftentimes when private vendors are the ones using contact tracing, who are deploying these systems, HIPAA doesn't even apply that minimum threshold, meaning that firms will oftentimes be completely free to sell this information to police or to other entities without any judicial oversight, we believe that we need additional state and federal protections to ensure that public health location data is never weaponized as a policing or immigration enforcement tool. We do not believe that it can possibly serve both functions, that as soon as you begin to weaponize that data against over-policed communities, you'll make it impossible to gain the community trust that is necessary to have an honest conversation about potential plans of exposure and people who may have been,
who may have contracted COVID-19. That is why we've been part of an effort here in New York State to pass contact tracing privacy legislation that would ban police access to this data but more needs to be done at the federal level as well because states on their own cannot block the sort of abuse by federal law enforcement that immigrant rights groups have feared now. And with that, I will bring this portion to a close. Thank you again for letting me participate in this panel.

Thank you Albert. I think that point that you ended on is just so important and I'm looking forward to exploring it in the Q and A, you know, if contact tracing data ends up being used to kind of feed a larger over-policing regime, that people, particularly those most vulnerable, are going to be less and less willing to play a role in actually resolving these contact tracing leads, right? They'll either not want to be contacted, or they're going to try to stay off the grid. So, and that's, kind of, starting from a place of lack of trust.

So, I think, as we've said, found in many of our panels, it's very hard to build trust and, kind of, a digital trust environment, and it's so easy to tear it down, so how do we explore the nuances there? So thank you. Next, I'm going to move over to Professor Choi for her introductory comments, Professor Choi? Hey.
Yeah, I'm going to briefly talk about the current technology being used and the potential direction to go. So, next slide please.

So, a brief overview of contact trace technology is there are basically two categories. One is actively reporting geographic locations either manually, or the actively tracing by the cellular tower can actively trace cellphones to locations. Or the other one is, the most famous being developed by Apple and Google, which is, is tracing exact location, rather than the two devices within the certain proximity is then the and they're using that information for their applications.
in the example, the geo-location based technology, they've asked some, this is currently mostly used, such as, including the MIT has developed and the North and South Dakota has developed their own app and another application in South Korea, which I had a recent visit to South Korea and I can, I like to share my experience of how they're doing in the next slide. And the second technology, which is, I'm going to get into. So, in the contact tracing in South Korea now, I just actually, I came back from Korea last week and I realized why they can maintain such low cases and basically, their contact tracing and monitoring system is very strict. In a sense, I believe with the SARS outbreak of 2002-2003, right? After the Congress has passed the law, allowing the government in the case of a pandemic / epidemic,
government, allow government to completely monitor and trace the individual's data. So, based on this, they can quickly act on the contact tracing issues. So, for example, if a person is tested positive, then they immediately isolate the person. And then they do the interview doing the contact tracing and then verifying using the CCTV by doing a whole entire major city is covered by completely covered devices. So they're able to check and review the CCTV data, and to make sure what the person has been around in the last few days, and even detect whether a person has used a credit card with a certain location at a certain time, and then try to review other people who were there at the time, and at the location, and that they're contacted. So that's the level of contact tracing they're talking about and they want the foreigners arriving to the airport, then you have to sign a certain agreement in the downloaded app, where you are going to be assigned a certain unique ID, and then you have to be quarantined in fourteen days, either in a place that's government designated,
or if you have a place to stay alone,

or the traveler, traveling together family members,

if they approve it then it's in the location for fourteen days,

very strict monitoring. And then you have to go to the nearby location to get the test

and then basically the test result is available within

twelve hours.

So, if you test positive, then you will be brought into a certain location, complete
isolation. And if you're negative, you still have to stay fourteen days in quarantine in the
location. Next slide, please.

And then, during the fourteen days, you have to report your temperature each morning
and evening through the app. And if you're late, a couple of hours late, you're going to get

a certain alarm, so you have to report it in time. And then they call you, they visit you at
random times, random days and then to make sure that you are in the location.

So everything is monitored through your cell phone. So, what if you leave the phone,

what if you take the phone outside?

Then that part is the easy part, they can quickly detect the IP address of the

Wi-Fi

you are supposed to be

in,
or the cell tower,
you are supposed to be connected if you stay in the location.
So that's how they check. So, what if you leave the phone inside, you just go out? Then what they do is basically, the phone, they monitor your phone activities.
So, if your phone is not active for an hour or a couple of hours, they immediately send the alert alarm using, through the app, and they ask you to push the buttons.
If you don't, it means you are not with your phone, so, if you don't comply with all the requirements,
there's a possibility you could be deported.
So this is the kind of level of contact tracing and monitoring,
which is one, very extreme and probably is not realistic in most Western societies.
So this is my experience from South Korea, how they are doing. Next slide please.
So, when we are talking about using the technology for contact tracing, it's really heavily depending on the smartphone technologies.
So, in the location, petitioning is concerned either using the GPS or using the direct connection to the cell phone. So,
is not the most accurate,

00:27:10.404 --> 00:27:21.115

but modern smartphone use is called Assisted GPS where the GPS data as well as the

base's station information,

00:27:21.115 --> 00:27:33.144

because even if we are not using the phone, always, our phone is kept connecting to

the base's station so that the base station knows where your phone is located so that

you can receive an incoming call.

00:27:33.595 --> 00:27:40.974

So, using that data, they try to get their location better, because, as we know the

location is not very accurate.

00:27:41.095 --> 00:27:55.704

It could be within my ensemble reading, the coverage area of the tower. So that's not

the most accurate. So, next slide please.

00:27:58.015 --> 00:28:12.444

So, the other option is using the Bluetooth, which at Google and Apple, is being
developed. So basically you put what Google and Apple platform is,

00:28:15.295 --> 00:28:16.194

the person,

00:28:16.255 --> 00:28:23.545

each one will be assigned a certain ID, so every 24 hours,

00:28:23.575 --> 00:28:25.494

then within the ID,

00:28:25.765 --> 00:28:26.125

then,

00:28:26.125 --> 00:28:30.894

within my understanding is that within 10 to 20 minutes,

00:28:31.194 --> 00:28:45.565

there will be, another code will be generated from your ID, it's basically search and key
encrypted 120bit key is generated using the key without revealing

00:28:46.914 --> 00:28:48.234

your identity,

00:28:48.505 --> 00:28:57.265
then you send out your own beacon and there's somebody nearby with the Bluetooth communication range.

If they recognize each other, it's called a contact.

So, in that case, those two contacted that information, the pair will be sent to the cloud then, based on how you implement it, using that data, for example, one of them is positive, then that information can be allotted to the other person.

So then how the app is being developed, depending on how you want to use it, but this is basically the platform.

So now using this Bluetooth proximity, sensing technology of course, there are many issues. First of all, it could be, we are talking about a contact really, is within the six feet range. We are staying 5 minutes or 4 to 10 minutes at the CDC guidelines. Can we get that level of accuracy? Probably not at this point. And if there are some classes between, then there could be a false contact, may be generated.
So, the more, the improvement needs to be done, but this is the basic technology. And potentially under the future 5G environment using the millimeter wave signals the location can be more accurately localized. So, that's the one promising technology we are moving toward. And next slide please.

So, from the perspective of public safety concerns, as we already know, public safety alerting system, so there is an emergency alert or amber alert is already there, but, as we know these are the simple alert text messages. So, text messages being sent to each cell phone.

However, after the 9/11, congress has approved and now are developing a nationwide broadband public safety network. It's being developed, being deployed. Next slide please. So, right after the 9/11, congress approved a nationwide broadband network and in 2012 FirstNet has been established and AT&T is in charge of deploying nationwide broadband network, who's operation is under the certain organization.
and under the Department of Commerce, the NIST is in charge of managing all the research development for the Public Safety Communications Research Division, is in charge of developing the research. and G.W. is part of that research effort. So, after the FirstNet launching in 2018, about two years later, now, most of the nation, country is being the broadband network is being deployed, but of course it was 4G / LTE network and now it's evolving more and more toward the 5G, adding more capabilities. So that initially, this was solely for the first responders to use as we all remember, right, after 9/11, all the towers were gone, so the communication was a disaster, in order to prevent something else, the extra enhanced network tried to develop, but now, looking into the pandemic situation, this public safety network is one really good source to be two to be utilized in order to deal with such pandemic cases. Next slide please. So, the main issue, when we consider the public safety, or has the services concerned, especially communication, the contact tracing is really based on the technology-based
issues and then there are two. So instead of, imagine, instead of the simple text messages, there can be more real-time video streaming may be very useful, for example, imagine an application, just like a traffic navigator system, we are going from one location to another location. There we look at the map and then it will show you if you follow that route, there is a lot of the traffic you want to avoid. So what if we have that kind of app showing, you are going to a restaurant, I want to go to the restaurant, but if you check around the area, a certain area. What about there are more positive cases, more contagious issues there to be avoided and so on. So, this is all if. A lot, but everybody, some data is available, people are putting their own information there. So, that part is, assuming then, what can technology do? What technology can do is using that data, then she says some app real-time that disseminating all those information, then, based on this,
every individual user can make a decision about their daily life.

So, if we are envisioning, that kind of a direction, then, now there are two issues, one is the coverage.

So we can not just serve only a certain area, we have to provide a service to everybody in the nation.

which means, even if a certain area where the, for example in the rural area where the coverage may not be covered well enough or at the edge of the cell then your signal may be very weak,

but still,

you need to get some service even if you are in the location. The second part is,

that kind of real-time data, that requires a certain peak rate so you need to talk about a certain peak rate per second,
So certain coverage and the actual flight time, which means the delay concerning two issues, the current 4G / LTE technology may not provide the options.

So instead, next slide please.

So as you can see in this figure, one cell tower base station, sending the signal, it's called the unit case, one at a time.

So sending one signal to the left-most cell phone, which is too far away. So, the signal is very weak. In the middle is a political signal, but on the right side there is a building, so the signal is bad.

So next slide please. So basically, such a unique case has what we want.

So, we want to cover everybody within a certain amount of the bit rate. So, next slide.

So, there is a technology under the 4G / LTE

which most of the cellular operators are currently available through most cellular operators, and the, where it's called a two-base station in this example.

In other words,

instead of a base station,

completely synchronized manner, sending out the same message to surrounding area,

which means everybody around there will receive the message, is called through the cast.

So that on the left-hand-side, the phone, two signals coming from two different cell towers, is actually the same.
So, even if one is weak, the other is better, so they combine the signal, it can enhance the receiving quality. Same as on the right-hand-side when one is blocked through the building. Since it, the other signal can help so that that phone can also have a better reception at that point. That's the way to concept all the multicast using one is, the cell tower sends the signal through the casting, everybody receives it, but from the end point, it is narrowly received from one cell tower, but is going to receive everybody participating in that broadcasting scheme. So that a receiving signal will be improved. So, next slide please. So, clearly you see the big advantage with two, one is a resource which means wireless spectrum will be saved, because it will either be shared by everybody around. Second of all, even if the one, the user is having the bad signal when it is connected to only one cell tower. But since you didn't receive the signal from multiple base stations, the receiving signal will be significantly improved, therefore the coverage will be improved.
Next slide please. So this is that application scenario.

I just mentioned that using the pick is, all users will submit their infectious status, daily basis,
all users will allow their device to upload their live location in real-time,
or allow a complete tracking.

And, of course, we're assuming that enough of a population
is sufficient, users are there, then using all those informations then if the wireless technology allows such a massive amount of
broadcasting allowed, then
someday we may be able to get that kind of application similar to tracking navigators. So that's the kind of application we are looking at. And next slide please.

So,
we have done some
initial, basically, this is the part of the research we have done for the Public Safety Networks Project but
as this COVID pandemic situation arrives,
we start to look into different angles. So we have done some initial experiments
and showing that with some baseline scenarios,

one urban scenario, rural scenario, in each of the scenarios

the multicast environment significantly outperformed the unicast alone.

So, can you skip the next few slides? I'm not going into those details, but we can move on.

Move on one more, yeah, just stay there. In the rural area,

for example,

in the X axis,

if the required repeat rate has increased

the unicast alone,

which is the blue line,

the coverage will be significantly dropped while the multicast can still achieve good performance.

So this is our initial experiment featured to show. And of course, there are many more to look into, and next slide please. So,

in terms of the future direction of the technology,

we believe a public safety network is a tightly incorporated in health systems,
especially there are many nice feature software numbers from the public safety network from the 4G network to the 5G network,

and tracing, contact tracing is concerned,

so many issues being considered,

but our small focused area is really,

how can we achieve to improve the coverage as well as the flying time, which means to minimize the delay. So we believe multicast, the so-called eMBMS technology is a very promising technology, so this is probably continue to,

we are planning to continue to work on this area for a while. The next slide? Yeah.

So, we have in the last few years, working with NIST and continued to work on several topics.

But this is one of the topics we're focusing on, is the multicast research.

So, I think it's going to be interesting to see how that can be really helpful for the pandemic situation. And thank you. Great. Thank you Professor Choi.

I feel like we got two great lessons there. In your recent experience in Korea,

I can only think that it must have made Albert, kind of, cringe,
issues with kind of comprehensive surveillance infrastructure here in the United States.  
but yeah, that comprehensive infrastructure that's been built in Korea has been so effective in countering COVID-19 that it really gives us something to think about, right?  
There was a real benefit to having that infrastructure in play, at least in the Korean context.  
So, it seems to me, we're going to have some great discussion in the Q and A around, kind of, the important cultural frame that these technologies are deployed within, and how we think about that.
And then your second half of your presentation on a public safety alert system  
that you've been so involved in, you know, I find it interesting, and I mentioned this before, in other panels, as we were organizing these panels,  
we found that many of our friends in Asia, in particular  
I'm reminded reminded of the conversation with colleagues in Taiwan when they thought about COVID-19,  
they immediately thought of it in the terms of a natural disaster,  
and it was their responsibility to deploy mobile location technologies in a public safety framework in response to a natural disaster.  
Which is a very different way of thinking about it than we thought about it in the United States, it's health surveillance, and how that leads into law enforcement.
So, I think that's going to be also a very useful fodder for our discussion today. So, thank you so much for introducing those very important conversations.

So, next we have Ms. Faine Greenwood and if you're ready Faine, we'll pass the mic to you. Sounds great. Yes, I'm ready. Great.

So, my name's Faine Greenwood, I work on drone technology as a specialist. I'm really interested in how to use elegy both efficiently and effectively for things like disaster response, transportation, monitoring and crisis response.

I'm also really interested in the ethical and privacy dimensions of journey technologies. That's something I've been working on for quite a long time, and really figuring out how we can balance the many benefits of drone tag versus some of the negative downsides.

And I think drones are also, kind of, a fascinating lens in which we can look at many issues related to privacy and location data because they are such, kind of, a hot topic. But a lot of what comes up here can be generalized to other technologies.

So the next slide. So I want to talk briefly about how drones have featured in the COVID-19 crises.

And as you probably noticed in March and April, when the full size of the crisis was really becoming known around the world, there was a tonne of drone-hype out there, things like delivery, monitoring social distance, checking temperatures from the air, all kinds of ideas were thrown out there and it's now July and basically, none of that's translated into useful COVID-fighting strategies and really, the hype has kind of died down, I think people are realizing what challenges with meaningfully implementing this technology for COVID-19 fighting and data collection. Next slide.

So, I think one of the most egregious cases of this relates to using drones to find fevers, which is something that was talked about a lot again, a few months ago and it's kind of died down a bit.
I think it's still happening, but it's something that has, kind of, fallen out of the news cycle.

So what the word is in India, Italy, China, other countries have been using drones equipment with thermal cameras to look for people with fevers, to more or less of an extent.

Certainly, I found news stories, indicating that these were being used operationally and the problem with this is that it doesn't work. There's a number of technical and ethical problems with this, and, you know, Albert,

what you were saying earlier, it has that issue of there is no scientific validation. There's no proof of any, there's no evidence whatsoever that any of this is effective at all, which is a huge problem. Next slide.

So, some of these issues include distance; standard thermal cameras that are used to screen fevers, they're only used maybe 1.2 or so meters from the subject and they focus on the face.

because it turns out that the area very close to one is the best place to capture human body temperature with the camera. Now

of course drones operate way farther away from subjects than 1.2 meters and actually clear the manufacturer of these most popular thermal cameras used on drones told me that they do not recommend using thermal cameras in this way.

that is not something they recommend

or validate.

You know,

the distance thing is bad enough, drones are too far away from people to read temperatures in the meaningful way. Drones move when they're in sky even when a
drone is hovering in place, it's still kind of wobbling about a bit and a lot of these things are being marketed as a way to quickly scan crowds,
which indicates you're going to keep the drone moving and that makes it even less likely you would have any hope of being able to accurately assess temperature, and was also the calibration needed to calibrate thermal cameras in many cases that's really hard to do with the moving drone.

And of course, there is again no scientific evidence that anybody is doing this and feel it's in any kind of validated way. And, I mean, there are issues with thermal cameras for fever screening that are stationary and are close to the subject.

So you can really imagine the technical issues with this compound if you're looking at using a drone to find fevers.

Next slide. And some other companies claim they can use drones to detect human heart rate, breathing rates, and even coughing and people on the ground using computer vision systems.

These systems don't rely upon thermal sensors. They rely upon analyzing aerial video to detect biometric signs.

But, unfortunately, all of the same problems apply and these methods are either less tested in the real world than infrared screening is. I mean, one of the main companies that was promoting the technology again, there was, I could find zero evidence of validation in the real world.

There's a ton of problems with accurately identifying these biometrics and people from a meaningful distance. So they had done some testing but the testing is only from a maximum distance of a few feet away.
Nothing like you would see with an actual drone in a real world setting and actually, I think the town of Westport, Connecticut considered rolling this technology out back in May but they retracted it almost immediately.

There was some public outrage and the disturbance over this but also because I think there are just so many problems with the system itself and little proof they'll actually work, which makes it, in my opinion, extremely irresponsible to roll them out in the real world in the context might have serious, serious implications for people's freedom, their health, and also, of course, these systems can be used in incredibly unequal ways that will reinforce racial discrimination, gender discrimination and class, there's all kinds of things happening here that can be exacerbated. Next slide.

And again, facial recognition is yet another widely feared, widely valued potential drone-use case that I could find no evidence of actual uses in the real world. Last theoretical talk about facial recognition drones,

in the context of, like, police drones, and I should also mention here that my work very much focuses on small civilian intended drones. I'm very focused on drones that are marketed to consumers that may be used in some cases by police forces.

But these are police forces, almost exclusively use that exact, use drones made by consumer manufacturers.

So they're almost exclusively using the same drones that civilians use for things like, surveying, transportation, infrastructure, for photography, for search and rescue, humanitarian aid. They're all the same small consumer drones.

That's what, kind of, what we're talking about here, when we're discussing police drones, of course,

quite separate from military drones or even some federal government-used drones,
which are very different capacities, are usually much larger and flies for vastly longer,

have much more specific technology aligned with them.

So,

again,

I also can't really find real one-use cases,

of facial recognition

military drones either,

but in the context of consumer drones that are used by police are by other local law

enforcement,

the

technology is not there.

It's not anywhere close to there. Again,

there's very little that, I can find some technical papers and some evidence that

researchers are working on the stuff, but there's a ton of challenges related to, just like

with the thermal drones, distance from the subject,
the movement of the drone, the sensors you're using, and as we know, accuracy, of facial recognition tech with people who are very close to you, who are on the ground, is lousy.

There's been many, many documented cases of facial recognition tech, being nowhere near as accurate as advertisers claim it is.

And, of course, again, as Albert mentioned, these, the missing of identifications become much more dangerous when it comes to women, people of color, again, female, yeah, so, female people of color are much more likely to be identified by facial recognition technology and basically, every issue of a missing identification that will be applicable to the stationary camera would apply with a drone, vastly more,

but no one's even doing this in the real world so far as I can tell. So, that's an interesting, so that's something we need to keep ahead of and keep an eye on, but it's not really happening.

But if it does happen again, with no sense of the accuracy of what the accuracy of identification, we risk, we really risk false identification. And actually there was a case, I think, this actually happened last year or a few months ago, but it came to light. There was a black man in Detroit that was indeed falsely identified by a stationary police surveillance camera as being implicated in a robbery. He was obviously not responsible, but he was wrongly matched with another person.

So we really actually had our first real world example of these mistaken identity cases using police surveillance, so it's all very dangerous stuff. Next slide.

And here's another area where our drones are being used by authorities of law enforcement primarily, using drones for public communications.
speakers that can issue instructions, kind of, the Chula Vista police in California back in April or May were actually using these talking rooms to communicate with homeless people in settlements in remote canyons. They thought it was a good way to run. Obviously I find this very problematic, but they thought it was a good way to avoid putting officers at risk by using the drones to communicate. I definitely think there's a lot of problematic areas here, I think there's a certain blind spot amongst law enforcement where certain people are much more comfortable with drones than they are. I think the main, real problem here is the distance these drones are being used for social distance monitoring or monitoring protests, which I'll get into in a minute. They create an ambiance of surveillance which makes it seem like authorities are watching more than likely actually are. The interesting thing about small drones is that they're actually nearly invisible, much more so than police surveillance cameras, much more so than higher flying aircraft. A lot of police forces are using manned aircraft like helicopters and even small manned aircraft to track surveillance but we're used to those, drones are, kind of, novel and they make us really think about the fact we're being watched. Small drones are actually not great tools for persistent surveillance which means tracking over a long period of time. The most battery life you're going to get out of these small drones, you're looking at maybe 35 minutes, so that's not persistent surveillance as we understand it. And what you can tell from a drone that's been in the air for a long time, that it's confined to a very small space, so it's not really functioning, it's essentially an easy-to-move stationary camera. So, they're not great tools for persistent surveillance but since people don't know that, people assume persistent surveillance is happening even when it's not which is a big problem. Next slide.

So again, and getting into the protests, which happened after the murder of George Floyd, in late May, in the United States, police have been really stepping up area surveillance of the protest movement. And that includes things like police drones, which are many, many more, and more police station, police entities in the United States are acquiring. Lots of manned aircraft, CBP, Center for Border Protection and Border Patrol in the United States actually used the federal government Predator military drone in one case to monitor protests in the Minneapolis area, and again, tons of manned aircraft.
I think one, if you were, kind of, following the protest earlier in the summer, you would have noticed that there were helicopters, often, multiple helicopters operated by law enforcement at pretty much every protest. Here in Boston, 

I was watching the maps, the flight monitoring maps and was really amazed to see they would scramble a helicopter basically, 

even for a really small protest, 

which I find really disquieting because again, 

it's, 

really normalizing aerial surveillance and yeah, 

this is n't just about drones, it's all kinds of surveillance. 

For example, 

in Baltimore, police are actually, they're testing some persistent surveillance using manned aircraft over the city of Baltimore, flying these small aircraft with super zoom cameras over the city. 

Basically, every day for extended periods kind of a test, 

testing for surveillance. It is really disquieting. Police justify this as the need for situational awareness. So it really is a form of pre-crime monitoring 

in the sense that it's using drones to ensure that you've got jobs, you're able to see what's going on, in a way that others can't. And it's also disquieting. I'm speaking of somebody who's a drone pilot myself.
I have a journalism background in addition to the work I do. They’re also dominating the airspace because journalists and activists use drones very widely around the world these days, a really popular tool for, kind of, journalism, again, for reporting.

And if there's a police drone in the air, there's a police helicopter there, and they're absolutely not putting up my own drone, right? Because there's a major risk that I'm going to be accused of interfering with their drone or their helicopter.

So, which means in adsense, that the existence of this persistent, kind of, standardized police air surveillance means that they own the skies. That means that you really can't benefit from using a drone to gather a different aerial narrative, it doesn't work because it is unsafe for you, put your own aircraft up there, which are, like, kind of neutralize a lot of the power of small consumer drones to give your average person, the access to the aerial view. And that's, kind of, that's pretty worrisome to me.

because again, drone video has been used to really expose police abuses at protests, a lot of activists standing on Native American reservation. The Dakotas were using drones to document some of the stuff, and that's becoming really increasingly hard.

And, of course, the authorities are also putting up no fly zones around protest areas.

to keep drones out, so that actually happened in Portland recently. So these are the big, this is a sense of all these constraints on revenues with the public is concerning and there's also very little regulation around police drone use.

So, 18 States at the moment, have a warrant, require you to get a warrant to use a drone for police purposes, for law enforcement actions. The other states do not require a warrant.

I mean, I think there are slightly more than 18 States with some kind of rules around police drone use, and the rules differ markedly from state to state. There's really no centralization at all. There's very little standardization on what happens to the data.

the drone collects. So, I just wrote a story last month where police in Minnesota used a drone to monitor the search for evidence of rule breaking at a nude beach.
So, the beach was not officially a nude beach, people would use it in that way. And the police use the drone to capture imagery of those who were not wearing tops or shirts.

Which is one of those, you know, the argument is that they're in a public place so they're not entitled to that privacy. That's not really how people operate. People don't expect to be monitored by a drone when they're at the beach. They don't really, aren't really expecting it

even in the context of protests, so there's a lot of new expectations around privacy, a lot of 4th amendment issues. There's a lot of murky and unsettled issues,

but I think the main takeaway is we really need much more regulation around what law enforcement can do with drones, but also with helicopter surveillance data, with surveillance data from manned aircraft,

we really need much more legislation around that. Next slide. And again, please, and in terms of the public trust theme in this conversation, this is another huge issue right?

So, police, often, justify using drones to de-escalate conflicts with the public and they don't really recognize in many cases that the public perceived drones as contextual and vary based on the concept of privacy and contacts.

So it really matters how we use the drone and how people see you use the drone. I mean, there's some public opinion research that says the people here, I'm not sure, likely to approve of drone-use for search and rescue. They're not likely to approve a hobbyist neighbor.

They're much more likely to approve the search and rescue than at police surveillance and protest. So, they really need to come to this understanding that it's less about the drone as an object in many cases. It's more about what are you doing with it?

And is there community consensus around this use case?

And is there community knowledge of what happens to that information that, so those are all really important and losing that trust means that you're going to lose access to
your drone program in law enforcement because in communities people get really angry about drones.

People get really, really unsettled by drones, the ways that they don't with helicopters, they don't really about surveillance cameras. So they're very visible and maintaining trust is essential.

if we're going to put together a drone program. Next slide.

And again, to the other theme of this talk, about drone data that is incredibly dangerous, whether you're talking about thermal drone data to determine if someone's got COVID or are producing drone data to determine if someone's committed a crime and also to false positive positives and false negatives are intensely problematic. In the case of COVID, if you're using a drone to detect fevers and you don't realize that that drone is incredibly accurate, you might falsely identify people as sick, or you might falsely clear someone as healthy.

And as we know that there's a lot of discrimination baked into these things, it's entirely likely, that some of what be more likely to pull someone aside because they're suspected of having COVID, if they are from a minority group or another demographic.

Of course, there's also bad or confusing data, again it's open to interpretation and that ambiguity leads to a lot of space for unethical use cases.

That's what we really need to be aware of. Next slide.
Overall, again, the hype monster. We’re really in the middle of a hype cycle with drones and again, I’ve been very relieved to see that a lot of hype around drones for COVI, which has mostly died down, and I was pretty concerned for a while there.

I think not to say that there aren’t still problems, but really to keep an eye on this. And I think my main message here is to be very skeptical of any of these claims about using drones for COVID-19 or using drones for law enforcement purposes. We have very little, real world evidence.

We have very few test cases. It’s older, experimental and, of course, that also means be suspicious of data that is reported to be collected by a drone if it’s used to ID people for COVID-19 or to ID people who have violated laws.

Be very suspicious, and we really need to focus more on the things we know drones are really good at, grounds for creating maps, for collecting aerial data, those are the things we should focus on at the moment, in the context of what’s going on in the world.

So, again, they’re not quite detecting COVID-19 from the air. And, while they are pretty good at collecting aerial images or protest movements, they’re also not going to be good at recognizing the cases from the air.

And now we have to also use this moment, this kind of newfound attention on aerial surveillance, on tech surveillance, to rekindle discussions about drones that really need to be taken up again about how we regulate the stuff.

How do we define privacy alarms from aerial drones that are really poorly defined right now and we really need way more research that really attempts to define and articulate these harms so we can do a better job of regulating them while still maintaining access to this great technology.

Drones are fantastically useful, and it would be really terrible if we allowed misuse of the technology and unethical uses of it to take the analogy away from everybody else. That would really be awful so that’s why it’s so important.
You've got some great turns of phrases in there, you know, you're kind of warning that we've been creating an ambiance of surveillance, in a way, normalizing aerial surveillance and other forms of tech surveillance where it makes people, kind of, assume that things are being observed, things are being watched that may actually not be being watched, right? So that's both giving false comfort and maybe even false discomfort to folks just because of this overwhelming ambiance that's being set up. But, yeah, within that, whether people are comfortable with it or uncomfortable with it, there's a lot of misidentification. There's a lot of bad and confusing data, precipitating ambiguity into these situations, whether it's for policing, surveillance, service provision. Yeah, so I think you pulled a lot of threads that I'm looking forward to bringing out in discussion here, so thank you for that.

So, last, before we go into open discussion, we're going to hear from Professor Julia Stoyanovich, and if you're ready, I'll pass the mic to you. Yes, I think I'm ready. Great. Hi, everybody. No, the first slide please. Although you want to move things. It's a pleasure to be here and in the next few minutes, I will offer my thoughts on what I'm calling Data Equity Systems.
This is a slightly more general concept then what we have been discussing and focusing on here today.

531
01:08:47.125 --> 01:08:56.364
But I will do my best to ground the main points and the main thoughts under this umbrella of Data Equity Systems in Geography. Next
532
01:08:56.364 --> 01:09:04.704
please. So, I was asked to show a map, to keep you all engaged, and so I obliged. Here's a map.
533
01:09:05.694 --> 01:09:13.614
It's a map of New York University's campus. So I teach at New York University,
534
01:09:14.034 --> 01:09:21.265
and what we have been speaking about and thinking about institutionally is how to re-open safely.
535
And this is going to be the thread that runs through my presentation here to make things concrete. So, once again, the campus of NYU,
536
01:09:30.029 --> 01:09:38.965
we, like many institutions, are planning to re-open in the fall and we're grappling with how to re-open safely and responsibly. Next.
537
01:09:41.604 --> 01:09:56.574
Zooming into the map, we see that NYU has presence throughout Manhattan, mostly in Lower Manhattan, along with significant presence in Brooklyn. And what is shown here are just the educational buildings where we have classrooms and labs.
538
01:09:57.744 --> 01:10:12.564
But we must also consider when we think about the re-opening, where the members of the NYU community actually live, and where they would be commuting from when coming to school or to work, and we will be coming back to this last point in a bit,
539
01:10:12.564 --> 01:10:17.005
so please, just hold that thought and remember that I did show you a map. So next.
540
01:10:18.805 --> 01:10:20.454
So as we plan for reopening,
541
01:10:21.534 --> 01:10:25.375
and as we consider how to implement specifically,
data collection, data sharing, and data analysis mechanisms in support of these re-opening efforts, we need to consider three aspects, broadly speaking, and in no particular order really. So, the first, I'm going to say is ensuring the safety of everyone on campus. The second is being cognizant of the rights and preferences of individuals and these include students of course, but also educators and administrators and support staff. Especially support staff don't get mentioned a lot, but they absolutely should be mentioned quite a bit when we speak about these topics and these rights some preferences pertain to their own data and to the engagement and participation modalities that these individuals want to have and can have on campus. And now, the third point is that as we re open, we should be creating systems that are efficient and effective at meeting both of these goals, the goals of safety, and respecting the rights and preferences of people being impacted. And in all of these goals, we need to reason under uncertainty. This is both collectively and institutionally, and also individually, yes, participants. And this uncertainty is what makes everything particularly difficult. And it's difficult, these are both operational, we don't know what the future holds, and also ethical, we
don't know what the risks are and what the potential benefits are of the choices that we make.

So, let's go to the next slide. And here, we're going to think more closely about data equity and in particular how it pertains to these benefits and the risks that I mentioned.

So more generally, I like to think about data collection, sharing and analysis, specifically, in terms of these benefits and risks, and in terms of who it is that reaps the benefits and who shoulders the risks, who are these stakeholders really? We must be very careful to ensure that the risks are not born disproportionately by some populations, said populations and that the benefits therefore are not enjoyed disproportionately by some other staff populations.

And there is, of course, a rich history of imbalance of power in society that does, has been translating into disparities in this context.

And many of these are also in the public health domain, the same domain in which we are operating today, due to COVID-19.

There are some stark historical examples that I want to go into, kind of, the most immediate one that comes to mind is this study, please look it up. If you're not familiar with it,

it's an absolute disaster in terms of these failures of balance.

one set of issues was brought up by a student during a recent public conversation that we had at NYU about re-opening and that is that individuals with disabilities experience online learning very differently.

So,
Then do other students, but they also have different types of risks, depending on their disabilities that are due to impersonal engagement.

We also need to think about how individuals who are experiencing housing insecurity are hearing with respect these benefits and the risks. Are they actually able to productively engage, for example, in the remote instruction?

So, kind of, to generalize this, when do we think about data equity we should be accounting for the needs, particularly of marginalized communities, of communities that we don’t think about historically and that typically experience many of the risks.

So this thinking links closely with the notion that again, that I believe should be central to how we’re thinking about all of these systems and that is this notion of data equity.

Even in the best of circumstances, underlying structural inequities and access to healthcare, employment, and housing show up in the data record, and they are propagated through decision systems, alternated or otherwise, to become reinforced by policy.

And under the current regime of haste, we not only reinforce existing structural inequities, but we also risk introducing and normalizing nuance.

And the previous speakers all gave examples of this, ample examples. Right, so these, for example, might arise if contact tracing apps and certificates of immunity become the norm.

And importantly, as is underscored by the pandemic, there are significant risks associated with not collecting and not sharing data.
If we don't know how the disease is evolving and if we don't know how it's impacting different demographic groups, for example, then we won't be able to control outbreaks and we won't be able to mitigate inequities, and we just want to be able to re-open in any meaningful way. So difficult as these topics are, we must engage with them. In a nuanced and productive way that will allow us to act rather than just, kind of, sit back and be paralyzed by this fear, and therefore bringing additional risks to society by inaction. Next slide please. So in my ongoing work, I focus on building data equity systems. These are systems in which data collection and analysis and use are guided by the risk benefit trade-offs to a range of stakeholders. These systems embed ethics and legal compliance by design. In the next few minutes, I’d like for us to focus on a particular aspect of this, and that is data sharing and collection as it pertains to our re-opening efforts. We have different modalities of data collection, sharing, and use that are appropriate for different scenarios, and I will discuss these with you in a minute. But before we dive in, let us think some more about this data equity concept. Next slide.

So, what exactly is data equity? We consider four classes of issues.

One of these, we call representation equity, and this is increasing the visibility of underrepresented groups that have historically been suppressed in the data record. For example,

confirmed COVID-19 cases require testing, and there can be racial disparities in both the availability of testing and the desire of individuals to be tested, which itself hinges on an abundance of trust or of this trust of members of a particular group in the healthcare system leading to systematic biases in collected data.
Albert spoke eloquently to this set of concerns earlier today. So, trust of course, is a central issue here and then in particular, it impacts how well certain groups are represented in the data.

The second dimension we call linkage equity, that is facilitating linkage across data sets to ensure access to features that help expose and quantify inequity. For example, if features, attributes such as race and income are not recorded along with the data, then it becomes hard or impossible to discover systematic biases that may exist, let alone correct for them.

But if we don’t know what performance is of our metrics, of our solutions, with respect to particular groups, then we want people to mitigate this and we want to know that we aren’t being biased.

The third dimension is access equity and that is about providing for equitable and participatory access to data and data products across domains and levels of expertise. For example, an institution’s or a state’s COVID-19 case data will be shared broadly, and I will speak about what some of those modalities might be for sharing in a minute, but it should be shared, rather than held stately by the state’s own government since combination with other sources can lead to better verification and greater insights. The final dimension I will speak about is outcome equity and that is monitoring and mitigating unintended consequences for any groups affected by the system after deployment, directly or indirectly.
01:19:03.024 --> 01:19:05.604
contact facing apps may facilitate stigma,
599
01:19:05.635 --> 01:19:11.664
harassment or retribution in the case of positive diagnosis with disproportionate effects
600
01:19:11.875 --> 01:19:19.494
on marginalized communities. So this is another dimension that we should keep in
mind here. Next slide please.
601
01:19:20.395 --> 01:19:34.255
So, as our first complete example, suppose that our goal, and again, here I'm speaking
on behalf of an institution such as NYU, is to help schedule when people can use
elevators or to access a classroom or a lab.
602
01:19:34.944 --> 01:19:47.574
And to do this, I'm showing you two buildings on NYU's Tandon School of Engineering's
campus, the Brooklyn campus where I teach, and as you can see by the height of these
buildings, you do need an elevator to get to a classroom or lab.
603
01:19:49.015 --> 01:19:52.975
So, back to our goal, we need to figure out who can take an elevator when.
604
01:19:53.340 --> 01:20:07.104
Right and so all we need to do in this case is to collect counts, to count the number of
people who are using an elevator without any personally identifiable information or PII.
Once
605
01:20:07.104 --> 01:20:15.324
we know how many people are in the elevator, we can tell a particular participant
whether they should now go and try to take the elevator to get the class, or if they
should wait.
606
01:20:18.234 --> 01:20:32.604
Participating in data collection here brings an immediate benefit to the data subject,
they can now use an elevator or get a desk at a lab. Further, because this collected data
is arguably not very sensitive,
607
01:20:33.024 --> 01:20:42.354
we can afford to very carefully share these aggregate accounts to help the institutions
schedule things more globally at the university level.
608
01:20:42.954 --> 01:20:52.225
Not just at this particular moment, whether I can take an elevator but what's a good
strategy for elevator scheduling as a function of demand. Next please.
609
A promising technology for data sharing, for sharing data carefully, is a set of techniques known as differential privacy, or DP for short. DP techniques aim to strike a balance between the level of privacy that we afford to individuals and the amount of utility.

How correct the accounts actually are in our example of elevators and lab desks.

To protect an individual, we will be enabling through the process of data sharing, plausible deniability and to derive utility, we will learn something about the population, but we will have to add noise, to be able to keep this plausible deniability. So, next slide please. Let me illustrate the spirit of DP, differential privacy techniques with an example here. So let's play a game of truth or dare.

I'm going to ask you a potentially incriminating question and this is a yes / no question. So, here, I'm asking you, whether you went out drinking over the weekend and let's pretend that this is an incriminating question somehow, that you don't actually want to be found out, and the answer is yes. I want to get the sense of the proportion of this group who answers yes, while at the same time, not compromising each of the individual’s responses by giving individuals plausible deniability, as I will explain in a minute, we incentivize participation and data collection because you feel like you're helping us gain insight, while at the same time, not putting yourself individually at risk.

So the game goes like this; answer the question in your head, did you go out drinking over the weekend, and then flip a coin. You can flip a physical coin, or you can type into Google 'flip a coin', and it will give you this nice app and it will flip a coin for you. If the coin comes up tails, then respond truthfully, this is denoted by C1 on my slide, it's the first coin flip.
If it comes up heads, then flip the coin again. The second coin flip is denoted C2. If the second coin comes up heads, then respond YES, no matter what the truth is and otherwise respond NO.

So, this mechanism, these two coin flips, is a so-called randomized response mechanism that has been invented decades ago to collect some incriminating information about health status, actually in the armed forces.

I can observe the answers that you gave me, I can observe from this population, however many of you told me YES, and however many told me NO.

And I know because of the way that the process is designed that the fraction of YES answers is the fraction of the group who did go out drinking, \( \frac{1}{2} \), that's \( \frac{1}{2} \) in my slide on the bottom left + \( \frac{1}{4} \).

we, I, can observe the answers that you gave me, I can observe from this population, however many of you told me YES, and however many told me NO.

From this equation on the left, on the bottom left, we can estimate the fraction of the group who in fact, went out drinking over the weekend as \( 2 \times A - \frac{1}{2} \).

I'm not expecting you to follow precisely all the details but the reason that I gave you this is just to give you a quick, kind of, a feel for what these randomized response mechanisms look like.

So this particular mechanism is actually differentially private, and the reason that I gave it here is because, first of all, I personally always like to learn something that I didn't know before when I listen to a talk.
And secondly, it kind of helps to take the magic out of this differential privacy concept that we hear about quite a bit today and I hope that you enjoyed playing the game, that's also important.

So, for more information about this and other technical, responsible data science topics, please take a look at the course that I developed and I'm teaching and where all materials are publicly available, that URL that's listed here;

dataresponsibility.github.io/courses. Next slide please. So now that we understand how differential privacy works, let's zoom out to why we want to use a mechanism like this in the first place.

One modality is to collect noise data from individuals. And this is represented on this slide by drawing the boundary between trusted and trusted, right, and the response to the curator line, and we can do this with mechanisms, exactly, such as the one that I just provided to you.

This is known as local DP, local differential privacy, meaning that we will know approximately, we as the curator, will know approximately how many people are in an elevator at any given time, but it's going to be close to the actual number. Next
Another modality is to collect data from individuals precisely, but to add noise before we share that aggregated data. So this is something, for example, that the US census is doing. They’re collecting your responses precisely,

but before releasing aggregate statistics, they add noise at that curator-to-analyst boundary.

So this is another way to use techniques, like, differential privacy.

Next slide please. We spoke about collecting cones, that is the number of people in a classroom or in an elevator at a given time. As another concrete example,

suppose that our goal is to collect sufficient information about people who are in the same room, like an elevator, or a room, or a lab, or a classroom at the same time to be able to contain a future outbreak.

If that is our goal, then we need to record the bit more than just cones. We actually need to record who specifically is entering and leaving the room and at what time.

And this information is sensitive, therefore, it’s very important to limit the sharing of this information, to only share it on a need-to-know basis with specific accountability structures in place.

For example, an accountability structure would be to allow a third party entity to audit who accesses the information, when and for what purpose?

We should also be sure when we design these systems to destroy the detailed records and keep only the aggregated accounts as soon as we no longer need the detailed records.

Another thing to consider is that we may need to only keep records if folks who are in the room are there for longer than some specific period of time, like if they are in a room for less than five minutes and everybody is wearing a mask, we don't need to keep that information,
right? So, there's nothing absolute here. The message that I want to, once again, reinforce, is that of data collection, data sharing and data analysis to be in the service of the benefits, but with an eye on the risks.

So, if there's no benefit, then we should not be collecting it, right? We should really be discarding it. Next slide please.

As a final example, we may ask ourselves, so whether it's worthwhile to implement full-on contact tracing on campus and here again, we should again ask ourselves, who would, what would we gain from collecting such fine-grained information of the sorts that the previous speaker was explaining how cellphones do handshakes, right? So really, really fine-grained information. I'd argue that it's not much benefit really that we would derive, but at the same time, it's too much risk. So too much risk without much benefit. Please click.

So next slide, and this is because people spend a lot of time off-campus, and click again.
And they take the subway to get there. So we wouldn't be able to get a clear overall picture of their potential exposure to COVID, and yet we would institute something very close to surveillance.

And so, you know, given these considerations, given that we only have a very local view of people's location behavior that does not cover their entire 24/7 life, I would not be in favor of such a solution. Next slide please.

Another crucial ingredient is allowing people to opt-out of data collection and sharing while at the same time, enabling them to productively participate in university life, even if they decide to opt-out.

So what does opting-in actually mean here?

It means that we should support meaningful, remote participation in courses, in our case and meaningful social interactions between the students without requiring them to be on-campus in person.

And in that case, they would also opt-out of data collection, sharing and analysis. And this essentially amounts to broadening this concept of the campus community.

And this is another point that our students speak very eloquently about when we discussed re-opening with them, that they really need ways, meaningful, engaged with us, and with each other, even when they are remote. Next slide please.

So, in summary, my recommendation is to work very hard, but also very quickly because we don't have that much time to identify data collection, sharing, and analysis solutions that bring tangible benefits, and where we can have a handle on the risks. When we develop these approaches,
which would be in close contact and consultation with the people whose data we are about to collect, to understand their opinions on the benefits and on the risks. And more generally, to create an environment of trust in which these data collection, sharing, and analysis efforts are not seen as a sinister social experiment, but rather as a piece of a productive approach to re-opening, that is to everybody's benefit. Thank you. Thanks so much Julia. I mean, I love that you end it with, you know, we want to avoid having it appear as though it's a sinister social experiment. I think that encapsulates so much, just in that statement there. Well, you brought so much to the conversation here. I really appreciate the premise on differential privacy. You're right, as we were organizing these panels, it came up continuously, US Census, some of the major platform companies, everyone who is trying to collect massive amounts of data to solve a particular problem, but yet wants to, you know, abide by the right ethical guidelines and principles, that seems to be a tangible, you know, kind of strategy for protecting privacy but, and I'm sure it'll come up in the Q and A. So, I think this a good time to, we prepared for the audience, we prepared a number of questions that all the panelists have actually seen. So, I figured we should go through some of them, particularly in light of Julia finishing on the note of trust. Trust is hard to build and it's just so easy to destroy, particularly in, kind of, the impersonal, kind of, digital relationships, and using digital platforms.
So, building trust with citizens and consumers around the use of their most intimate data, you know, where they are every minute of the day, can be an enormous challenge.

And technological glitches, or procedure errors, or perceived misuses of the data can create irreparable harm.

How should governments and companies think about building trust with those citizens and customers whose mobile location tech they hold? I mean, that's a big question.

Hopefully, it's an open-ended question that lets you, kind of, approach it through your own perspective. I don't know who may want to pick up on that first. But Albert has his mic hot, so maybe I'll just pass it over to him.

Yeah, well there is one part of that that really stuck out to me, which was the premise that governments and companies would be the ones holding this data because I think that one of the key privacy safeguards in all of these different approaches is minimizing the amount of data that's held in the first place, as Julia was talking about, you know, really data minimization and especially on the government side, is one of the most valuable safeguards we have, and, with any of the sort of tech-assisted
contact

01:34:18.055 --> 01:34:28.345
tracing and exposure notification technologies that we're talking about, whether they're Bluetooth or WiFi or cellphone based or any other technology,

01:34:29.034 --> 01:34:41.395
absolutely none of them will work unless you have buy-in from the overwhelming majority of the public and to get that buy-in you would need clear guidelines on how data is preserved and how it's deleted.

01:34:41.725 --> 01:34:53.244
And I think tools like differential privacy are incredibly powerful in providing the mathematical protection for that data.

01:34:53.484 --> 01:35:02.364
But it's also a very difficult democratic issue of how do you get popular buy-in for something that is mathematically so,

01:35:03.204 --> 01:35:10.914
you know, sophisticated and, you know, even, I know a lot of people with advanced degrees who don't fully understand

01:35:11.994 --> 01:35:14.064
how differential privacy operates.

01:35:14.755 --> 01:35:28.194
And so I think it's a strategy, but the most powerful one we have is just simply say, that this data won't be held, except where it is clearly linked to a beneficial use case,

01:35:28.194 --> 01:35:34.074
and only for the amount of time needed for that specific use.

01:35:35.069 --> 01:35:46.284
Thanks, Albert. Do you want to jump up on that? Yeah. Okay. Let's go with Faine, and then I'll go back to Julia. Yeah, I completely agree with that.

01:35:47.574 --> 01:35:55.435
And in my mind, really what we need is legislative protections, because it's very much a social collective trust thing,
and in the US context, we really lack a lot of meaningful data protection laws or ways to defend ourselves.

I, you know, as I watched the social media companies, the big companies hoover up data over the past decade, I just have very little trust that anybody is going to self regulate who's holding this stuff.

And I'm especially distrustful about, obviously, private companies and entities and the main thing that will make me feel better speaking for myself here, but I think others share the sentiment, would be stronger laws and regulations that would oppose meaningful penalties and meaningful, meaningful impacts on people who use that unethically. And these just don't really exist unfortunately, that's the thing right now.

Great Faine. Julia, yes? Yes? Okay. So to follow up on both of these points by Albert and by Faine, which, of course I agree very strongly.

One other piece that we did not bring up is the competence of the operators of these data collection and data analysis methods, right? So very many of these methods are going to be predicting things.

And prediction is bound to sometimes feel we cannot foresee the future, right?

So, and then the question becomes, who takes responsibility and who is accountable for these mispredictions that we make, who is accountable for the mistakes? And a computer or a data set cannot be accountable, they lack agency, right?

It has to be a person.

So, whenever we have these environments in which data is collected, and then the predicted task is processing the data when there are errors,
the human operator has to be competent enough to understand what the process was.

They have to be competent enough to endorse it, right, to take responsibility for these actions and Faine, at some point, said that we should not be using drones with facial recognition

because the technology simply is not there. Even if the technology were there, but we did not have sufficiently well-educated operators of this technology, and these are some of the cases that we see in policing, in the criminal justice, right?

The conditions are simply not met for the responsible use of this technology. Technology alone is insufficient.

And if I could just quickly add on that, a hundred percent agreed. It took, you know, after the industrial revolution, half a century to develop the liability framework, so, we can actually hold companies accountable through tort law for the industrial decisions that were in place, were making, and I'm fearful that we'll see a half century of legal ambiguity on automated decision systems before we finally get the measures in place to hold developers, operators, and others accountable, but with regard to what Faine said on the legal side, yeah, that that's something we've also been pushing strongly, I think this also maps in nicely to the example that Julia gave of the US Census, right?
add
in noise to that data set is because we have a federal law that for more than fifty years that said that is punishable by,
I believe,
five years in prison to release identifying census data for any improper purpose.
Police can't get this data, not with a warrant, not with any other process.
And if we really want any sort of mass data collection or data analytics to be a part of the response to COVID-19, we need to have equally robust legal protections against the misuse of that contact tracing data for any other purpose, other than the public health response.
You're muted Chris.
Professor Choi? Did you want to chime in on this question at all? Oh, no, it's a third topic.
I don't know,
It is, kind of, a situation. Just a regular citizen, including myself, is it really a fair, desperate situation. So, there are people willing to sacrifice their practice basically, personal information, if it helps for the public health management is concerned, people are willing to provide it because they are very sensitive patients, there are certain cases. So, the majority of the citizens are willing to provide their own private information. I believe. So we are really talking about a certain small segment of the issues that needs to be really seriously considered the data way without really at leadership level. So we disagree. We have a disagreement on the panel and that is really great, I think, because otherwise it's dull. So I think that maybe in Korea there is a small section of the population that disagrees. In the US, it's very sizable, and it's not a random sample of the population. It's not just some lunatics who have conspiracy theories. People have reason to disbelieve the government's data collection efforts, right? And people, in fact, mistrust the government's data collection efforts, particularly from marginalized communities, for good reason, right? So I don't think that this is an issue that we can dismiss as a statistical fluke.
It's a fundamental issue in our society here. No, I'm not saying, but I just happened to see the Google timeline reported today. They basically, without actively, knowing to me, they really have a complete trace of whereabouts last month. Full detail of every location on the time. So this is the kind of data only to ve, and for me to I have to, I mean, at least for that part of that. But I understand, I agree. This is a really serious issue, because especially the government who's going to use what's gonna happen, including the other countries, including Korea. People worrying about, at this point, because of independence, everybody wants to solve the issues, they're all willing to sacrifice certain things, but everybody is worrying about what's going to happen and what government is going to do with the data. So, this is the real major issue, trust is a real major issue. Yeah, go ahead, Albert. So I was just going to push back slightly to say that the polling I've seen on this shows
the majority of Americans actually opposed or strongly opposed the use of this sort of tracking. According to a Washington Post poll from April, 20% of users definitely wouldn't use it, 30% said that they probably wouldn't use it.

We've seen low adoption rates in the majority of countries that have had this technology.

And I think that, I do think that there's,

I personally tend to push back on the framing of utility versus privacy because I don't think we've actually established that there is a counter-failing public health utility for this data collection,

but that we've only established that there's a clear privacy harm. Yeah, before I cross over to Faine, I just want to kind of point out, you know, sometimes you hear a social cohesion argument, right?

You know, in Korea, in Taiwan?

Yeah, there's, kind of, a different cultural frame, but a different, kind of, level of social cohesion than we have necessarily in the United States, or some other nations.

And I think it's also, you know, things that you hear is, you know, in Korea and Taiwan, for instance, they have a comprehensive response to COVID-19 where contact tracing technology is upiece.

And that's born of previous pandemic with SARS for instance, where they have an experience, it didn't necessarily go as well as they wanted,

and there's been a lot of public deliberation and open rethinking of what they've done and how they want to do it in the future.

And in the United States, in a way, we're kind of walking into this cold, you know, it's going to be interesting what the discussion is in the next pandemic in the United States, based on lessons we've learned as a society. Maybe we'll learn no lessons.
Maybe we'll come at it with the same cultural assumptions, but sometimes learned experiences do lead to public considerations. I don't know, Faine?

Yeah, no, I mean, I agree with you completely Albert, that these discussions around contact tracing, especially in March and April, were very much around, oh, well the predating assumption was that the contractors will work and be effective and the question is, really,

are we willing to accept the trade off to get to utility social contract tracing. I think people are realizing more and more that that's actually a false frame.

We really don't have any much in the way of solid proof that the contract tracing does a public health utility, as you said, Albert, and therefore the argument really shouldn't be, oh, do we want to give up our privacy for this benefit?

It's we're giving up our privacy for an unknown benefit that, for a number of reasons in the US context, especially may not work very well. Not to fall into US assumption but that I do think, in this case, we are quite different in our lack of centralized approach to COVID-19 and horrible cultural society factors,

I mean, it's really hard, in my opinion, to overstate the impact of what's been happening since May too with police funding protests because so many people are now, because we're having this massive, and desperately needed might I add,

societal debate about police and how much we trust the police and law enforcement. I think a lot of people were also feeling a lot less trusting in US law enforcement than they did even a few months ago.

And not just because of COVID-19? Exactly. Because of police brutality. So I think there's,
again,

I'm someone who's followed data privacy issues for many years

now,

but even

I definitely notice there's a shift in how I think,

and which is,

of course,

made me feel even less inclined to be interested

in any way,

shape or form doing any of my data law enforcement,

because I don't trust them and I think, again,

that the societal trust is really low right now,

and that's going to make everything a lot more difficult.

Right. Hyeong-ah, your mic is on.
No. Okay. Well, this made a great conversation. I'd love to move on to the next question we had and it's really central to the data quality issue. You know, we often garbage-in garbage-out.

It's a well accepted kind of aphorism within the technology community. In the context of COVID-19, false positives, false negatives in testing our illustrative example, given that such false test results will then be tied to mobile location data and the technologies that are tracking you within contacts tracing,, if we implemented them in a contact tracing regime, how should we think about the ethical issues around data quality?

You know, when it's not just the issues of how good is the location data, but, like, you're going to be bucketed, based on a test, and the test may be 30% error prone. How did we think about those things? I mean, I know when I recently had a COVID- scare or, you know, a potential contact, you know, we went and took multiple tests, just because we know that this test is 30% wrong and there's going to be multiple of us. You know, how do you, even personally de-risk, what you know are imperfect data, when you're going to have to suffer the consequences of living by that judgment, right? So, I don't know, I kind of see Julia's eyes tinkling there. So go ahead Julia. You need to call on others first sometimes also otherwise I will lose this, right?
So, I really like this question, and also the way that you framed it, because you’re framing this as a set of ethical questions.

but they also are just firmly technical questions, and this question shows how the ethical and the technical cannot be de-coupled these days, right?

When we talk about data, when we talk about accuracy of prediction, when we talk about taking multiple uncorrelated signals to try and reveal the truth, if there is such a thing, right? Whether you taking multiple and correlated tests, these are all deeply technical decisions that we make, and that we kind of hope the jointly will help us operationally some of our ethical thinking, and some of the values that we embed into the systems so yes, we absolutely are gonna make mistakes. The more uncertainty there is to the data, the less likely it will be useful. Absolutely, we have to have people making decisions, policy decisions, or even individual decisions, whom to quarantine, how to distribute resources, who understands statistics, at least at the rudimentary level, right?

Or they need to have people that they can talk to and whom they can then explain what their questions are, and then these decisions will translate because here, what we have built are these very complex human engineering, kind of, environments, right? Where machines, and by machines I mean data, and predictive analytics and all that stuff, and people have to inter-operate so that people can make decisions that make sense.

And again, they can take responsibility for and therefore, people have to understand machines. So this is really just down to, you know, you need to understand what an error bar is,
you need to understand what uncertainty is and there's no magic. Your algorithm, whatever it is, it should not be telling you YES/NO, it should always be giving you confidence in the results and you should know how to read these intervals and mistakes do compound in the worst possible ways, so if there is a 30% error in any one of your components that you can throw out the results.

So, those comments, I remember when I first ended up in the tech industry, one of the older, wiser technology veterans, he just said, Chris technology is not magic, right? And I was like, he felt the need to say that, because it's just so easy to come through it with the assumption that magic is going to happen and things will happen automagically, and they don't. So, Albert, I saw your mic go.

Yeah, and I agree with Julia I think this is so crucial because, as I touched on briefly in my presentation, with a lot of these contact tracing solutions, we're looking at using hardware that's already owned by the public, right? Building on their existing smartphones, which is baking in the inequality in smartphone ownership patterns.

So, for example, those who make under $30,000 a year, have smartphones, maybe 71% of the time,
and then over 95% of the time, when people make over $75,000, and you see these same inequities, find out over age brackets, based on educational attainment and based on race.

And so,

what you end up is that you'll be creating this data dashboard, that decision makers rely upon that's going to be constantly skewed by the data points, coming from only a selected subset of the community.

And we see a real risk that you'll reinforce health inequality because, for example, if you have wealthier, white communities running an exposure notification system, more frequently, and you see more potential points of contact there and you see more potential spread there based off of the system, you could easily imagine a scenario where decision makers are bent over investing in health resources for that community, not because of need, just because of how that need is mapped by the system.
And so that's why I'm very skeptical of using an exposure notification system, even if I believe that the underlying technology was a suitable proxy for exposure, I would be very wary of running it on end-user devices when we see this much variation and who has access to these devices. Great.

Faine? Yeah, I just wanted to build upon what you were saying, Julia, about technological competency and I got what you said earlier as well. I completely agree. That's a huge issue.

And in the context of the fever-seeking drones, for example, one of my biggest questions was, how are you going to train law enforcement officers who accurately interpret this thermal data, and we have a number of examples of the facial recognition expand number of news stories where police are supposed to interpret facial recognition as a likelihood of match, but they don't interpret it. They're interpreting that it's a guaranteed match. The ambiguity is really not coming through to the end-use of this technology or not. They're not understanding the ambiguous nature of these results and that is going to lead to inequitable outcomes.
I mean, that's obviously a huge challenge, because it really depends on who you're asking to implement. If you're asking, again, a police officer, who was not trained in the stuff to implement.

If you're asking a social worker, a public health worker, you're not going to especially trained people, which means that they're rolling out tools that really cannot be meaningfully used by people without special training or not provided that training. You're simply giving people false sense of confidence and it's going to create mayhem.

So, that's something we really, that training piece of that knowledge, that technology is useless without a strategy for how you're actually going to train people to interpret that data.

I work on implementing drone programs and that's something we see all the time. Organizations buy a drone, they don't devote any time or energy to learning how to interpret the drone data. They don't interpret.

They don't really figure out a strategy for storing the drone data or using the drone data for decision support, and the drone ends up sitting on a desk collecting dust, best case scenario. Worst case scenario is they keep on using it, they use a battle to interpret the data. Which is also what we see.

So these are the same, you know, it's the same pattern we're seeing with this COVID-tracking technology that it's definitely super worrisome.

And also, in terms of inequities, is another point I'd like to add, which is not super tech-related, but is really meaningful. I think, actually, it is tech-related. Google, so, entering the early stage of the pandemic, Google was working to set up proven testing centers, and they were requiring, and I wrote a story about this earlier in the year.

But they were requiring people to present ID to get tested. So, and also have a Google account, so they were really mediating access to testing through the lens of these requirements that are going to leave people out.
I mean, if you were an undocumented worker and you don't have a laptop, you don't have a Google account. You don't have a smartphone are you going to be able to get access to testing? So yeah, these are all these, like, tech blind spots that are coming up that are really being exposed by COVID.

Great job. Hyeong-ah, do you want to chime in on this question or should I move to the next one?

So, while I'm listening to everybody, it's all really important issues, but this is where we are.

So, what can we do as it is now so, for example, the testing Richard inaccuracy 30% is unacceptable for sure. While we agree the United States is dealing with 30% of inaccurate test results, while other countries can do much better isn't there something really, somebody else has to take care of it?

And then we try to combine the technology on contact tracing, has to have a certain confidence of their level of test accuracies.

So this is a really sad situation we will have to deal with at that level. That's how I feel.

No,

I think that's interesting,

you know,

sometimes I feel like these conversations end up being,
you know,
896
01:57:53.755 --> 01:57:59.579
compounded distrust because of compounded technical,
897
01:57:59.574 --> 01:58:00.505
bureaucratic,
898
01:58:00.505 --> 01:58:01.225
legal,
899
01:58:01.225 --> 01:58:03.024
administrative failures.
900
01:58:03.625 --> 01:58:06.685
You know, sometimes if you replay the tape and you wonder,
901
01:58:06.930 --> 01:58:21.265
what if we had in the United States, a meaningful proactive, robust public health
response, what if we had widely available testing, you know, at Day 1, and, you know,
what
902
01:58:21.265 --> 01:58:33.564
if, you know, contact tracing technologies already existed, had been, kind of,
socially-vetted, right?
903
01:58:34.225 --> 01:58:46.765
Based on a prior previous experience, there were strong public laws around it? It just
makes me think that, like, in the United States, we've managed to enter this pandemic
having nearly everything wrong,
904
01:58:46.975 --> 01:59:00.715
and therefore, it's difficult to have a meaningful discussion about how certain
technologies might be used positively. So I take your point,
905
01:59:00.715 --> 01:59:13.555
and I kind of feel that frustration myself sometimes. And it's unclear, I mean, it's clear
that progress can be made and how to make it is unclear, I guess. But I will roll to the
next question and
906
I'll probably come back to you Hyeong-ah just because of the topic of your
presentation. Even if we did everything right,
geographic precision and accuracy of mobile location data can vary widely whether
you’re traveling indoors,
01:59:31.079 --> 01:59:40.074
dealing with urban canyons, underground scenarios, even rural areas where the
coverage is dicey are problematic.
01:59:40.350 --> 01:59:49.255
And of course, everybody’s using different pieces of technology that require the
infrastructure to have different strengths.
02:00:03.324 --> 02:00:12.085
This can create massive blind spots, inaccurate geo-location and incorrect assessment
of space, temporal co-location, and I think you went into that Hyeong-Ah in your
presentation. In that context,
02:00:13.944 --> 02:00:25.194
how should we think about the authoritative nature of this data when drawing
conclusions with massive ethical ramifications?
02:00:25.854 --> 02:00:40.194
So, a part of leaders’ big effort is going on to really improve the accuracy of the location.
02:00:40.975 --> 02:00:47.515
So, especially, all the different layers, including the very low wireless communication,
using the millimeter wave in the 5G setting.
02:00:47.604 --> 02:01:00.204
the location estimation. So again, this is all tied up together from the top level and as
the management is concerned, is all part of it. So,
02:01:01.404 --> 02:01:11.994
again, somebody in the leadership has to have a big vision to put the major effort
investment to improve the data technology direction
02:01:12.354 --> 02:01:19.885
so that they can be really meaningfully utilized in these management issues.
02:01:20.609 --> 02:01:29.994
So, again, it’s a lot of frustration. I feel it, everybody feels it, and this is where we are.
02:01:31.404 --> 02:01:37.435
Great. Faine? It all comes back to the societal trust question, right? I think.
Like, in the Korean context, I think most people seem to agree this stuff is good and useful and I think it makes a lot of sense in terms of accuracy.

But, you know, when you said that, my mind, as an American, immediately went to, oh great, now the police will find it even easier to track down.

And it's just unfortunate that we have to think this way, and, you know, in the US context, which is, of course, they're not familiar with. But yeah, that's again where my mind goes immediately that there'd be so many we have to cycle trust and better leadership.

We would feel a lot, I feel more comfortable, but that's not where a lot of us are at. And I think again, a lot of this kind of feels like, you know, you can out-tech yourself about failures of leadership.

You can't out-tech yourself on structural social failure. We just can't create, you know, it's that one, tick, right? So, it's solving the COVID-19 crisis. The clever technological solutions won't work if leadership is completely failing to lead and that's unfortunately, I think where we're in the United States right now. Yeah, Julia, please. Great. So, let me try to put this around a little bit

and I'm going to put on my engineering head firmly, which I never really took off, but maybe it was invisible a little. So, rather than saying, and I think we all agree that there are lots of dangers here, right? We get databases faulty where operating systems and people don't really know what they're doing.

What would it actually take? What are some positive examples of where we have been able to use data and technology productively in the public health realm?

And absolutely, technology does have a place in helping us, find problems and helping us construct solutions. Absolutely, technology is not going to be the only thing, right, I mean, I'm the first one to say that.
we need to also have informed operators and informed consumers,
930
02:03:23.244 --> 02:03:26.545
et cetera but two things I'll say, one. is that
931
02:03:26.545 --> 02:03:27.685
in the medical domain,
932
02:03:28.284 --> 02:03:30.864
we have been more successful than elsewhere,
933
02:03:30.954 --> 02:03:31.614
for example,
934
02:03:31.645 --> 02:03:35.875
in the use of computational tools to diagnose diseases.
935
02:03:36.625 --> 02:03:41.904
And the reason for this is that first of all, doctors have a code of ethics.
936
02:03:42.864 --> 02:03:56.994
Doctors are very well educated. They understand that their implications of a mistake
are catastrophic for the patients, and for their own careers, right? So, there is this
understanding of responsibility already. Some doctors, many doctors know statistics.
937
02:03:58.284 --> 02:04:11.244
And so this is, kind of, an example over an environment where the stakes are high, but,
kind of, the human operator really is very well educated and conscientious.
938
02:04:11.274 --> 02:04:24.175
So, and also the patients, they have on the immediate incentive, the benefit to them to
share their data. Especially if it's the case of a rare disease, let's say. It's tremendous,
right? So, we have an environment from which we can learn.
939
02:04:24.685 --> 02:04:35.604
We can look at this domain and we can say what went right there, what went wrong,
what were the difficult pieces? It wasn't the technology that was particularly difficult. I
mean, it was difficult to build, right? But that's not the hardest part.
940
02:04:36.024 --> 02:04:41.664
It's really figuring out how to navigate this patient-doctor-machine interface. So that's
one point.
941
02:04:42.654 --> 02:04:54.534
The second point is that absolutely, there are, we have to figure out how to use data
and predictive analytics to help us mitigate and control an outbreak such as COVID-19.
This is not the first time we're seeing this, we saw Ebola, we saw SARS, right? We saw the Spanish flu, of course, that was different technology, not modern technology, but these things unfortunately will keep coming back, so we have to figure this out.

And to start, I think we need to identify specific cases where data collection brings, very viably, brings the benefit where we understand the risks, and we can outline the scope of a technological solution.

And in my three examples that I give, the first two were of this kind. Counting people in an elevator or figuring out who was at the lab and may have gotten infected.

because somebody else who was also in the lab at the same time was tested positive for COVID. These are all cases that we know how to handle without a lot of exposure. If we take the right combination of technology and data and educating and,

you know,

building trust and listening to people about their privacy concerns,

this can be done.

So we need to start with examples where again, we have a solution for a problem that emits an engineering solution, where we can ask a question,

we can design an experiment and force a viable hypothesis that we can then test and see whether our proposed solution works or it doesn't work.

And then, based on that engineering scientific argument, we can actually institute these solutions to great benefits in society. I don't think that benefits risk privacy, utility, or forced dichotomies. I think they are correct. It's just in some domains.)
But if we can write this down, we're golden, right? We have the scientific method on our side. So, let's just think as engineers here. Let's stop pretending like these tools are magic or like, they're evil, right?

Think about how we don't produce snake oil. How do we produce concrete solutions that we can defend? And this is the first step.

Start at the bottom and build. Before I pass to Albert, I think that's part of the interesting thing in Korea and Taiwan,

most of the people I've talked to are engineers, they've been applying that scientific method within their social context, right? Within their institutional context, their legal context, but they've been working hard to answer those kinds of questions,

and I feel like, at least in American society, we just kind of weighted into it willy-nilly.

Anyways, Albert, you had your mic on. Yeah, I think there are a couple of important threads here. Julia's plan,

I definitely agree about, sort of, delineating the different use cases, in the way

I tend to think that we have purely medical applications of surveillance and technology and AI that reads a very different set of concerns from dual-use technologies

that can also be co-opted by government for policing, for immigration, for things of that nature. And so I think when you're talking about deploying

data analytics, machine learning in the, you know, in trying to develop therapeutics and trying to analyze the, you know, the actual way that a novel stream is operating, things along those lines.

It raises a very different set of concerns and heuristics than when you're talking about spatio-temporal

co-location, because that impacts such a broad array of privacy interests, but other civil rights interests, associational, free speech, right to travel, all of those things.
And I would say that even if we were in a situation where we had a technology that was proven to help in identifying potential exposure that one key differentiator between us and South Korea is that we lack the public health prerequisites to use those sorts of tools. In New York, people are waiting more than two weeks to gain access to COVID-19 testing. People do not have reliable ways to isolate when they are presumed infected. These are the basic public health measures that are indispensable to an effective response that we certainly have seen done very well in South Korea. And I think we do ourselves a disservice if we look at the value add of mass surveillance as a part of a public health response without looking at all of the other forms of social investment and response that were needed to make that tool effective. You're muted again. Can you hear me now? Faine? Your mic's on, I didn't know if you wanted to jump back on that or? Well, yeah, I just wondered, again and I agree with what everybody said regarding the need for, and actually another point there is that the healthcare data again, there's HIPAA in place. There's some sort of legal constraint upon the use of the data that it's not, might not necessarily apply in the case of these contact tracer apps, so that's certainly a level of uncertainty that we need to be resolved. But, yeah. But again, it is really important from a social context that's operated and from a legal context that's operated. Unfortunately, they lack a lot of the basic elements we would need to actually process this data in a meaningful way.
So in terms of the mass-laws hierarchy, unfortunately, we are lacking the basics right now. Yeah. Nobody's going to help us to leap-frog unfortunately.

So before I go to the audience questions, I've got one more that I'd like to get to, you know, in the world of AI, in which, I think, several of you have picked up on, there's been a lot of wrestling with the issues of bias in the deep-learning training sets. Training sets derive from large volumes of mobile location data, no doubt, share the same patterns of bias..

I think you guys had picked up on that, even if it just starts with basic ownership of handsets, right? So, when using various, let me call them geo-AI methods,

how can authorities and communities achieve some transparency into the sense-making part of the process, the sense-making that's being derived from this mobile location data? What other issues should we be concerned about? I don't know, Julia has that tinkle in her eye again though, but I don't know if she wants it, go ahead. Yeah, I can let somebody else speak first but, whatever. Yeah, so what does interoperability mean is a way I might interpret. Pardon the pun, your question. This depends again for who you are explaining and for what purpose. I do think that we're, if we are to be successful, and I'm going to consciously put this positive angle on everything that I say now because I want us to end on a positive note that, you know, things are possible, if people actually work. So what would it take? If things are to be successful then we need to understand how to give members of affected communities insight into how decisions are being made about data collection and data analysis that affects them and how to allow them to challenge the particular processes, the particular practices. And there is this confusion when we talk especially about deep-learning, I think that because deep-learning is just so complex and even the designers of theses algorithms can't really understand how they work exactly, means that there is absolutely no hope for us of explaining to anybody what’s going on, and I think that’s false. Our genetic code is also very complex. The literal program is complex,

but that doesn't mean that we cannot attempt to diagnose and treat disease, right? I mean, that that would be ludicrous if we just give up, because we don't understand the syntax of the program.

So, there's really a distinction between, kind of, syntactic transparency that just says I'm going to dump out all the data, even if I don't care about privacy, I'm going to dump out
all the code. Then you go and figure out what this means to you. Instead what we should be doing,

and I agree here with Faine, that we need a lot of regulatory, kind of muscles to be able to

effect this. We need to be setting up meaningful, auditing, oversight, and public disclosure mechanisms for data collection and data use and this has to do with things like fairness, right?

I mean, fairness is a very low term, but kind of broadly speaking, I would say that we care about how outcomes are distributed. Are we giving more money to hospitals that serve affluent communities or less, right?

And also about how the process affects different people.

So, there's this notion of procedural fairness as well as, kind of, fairness in outcomes, and we need to be accounting for both in the way that we are explaining. Fairness concerns affect groups usually,

but then there are also individual fairness concerns, right? And these have to do with, really, me as an individual being affected, understanding how decisions are being made and why they're made based on this particular data set, maybe it has errors in the way that it affects me.

So, both individuals,

have to have ways to understand what's going on with their data and have to have ways to speak up, but also groups, and representatives of groups have to have ways to understand what's going on. So, intervisibility is a very difficult topic. It's a very exciting topic.

I'm happy to talk more about the kind of, progress that goes on today, and the interface of technology and psychology and law and policy making in this realm.
But maybe the main message that I would want all of us to, kind of, leave with and I know we're not ending it, but it's something that I really want to reinforce, that we each have a responsibility to try and understand what's going on.

We have to think carefully about, you know, maybe it went over my head the first time I heard about differential privacy, but if I am to live in this society, if I am to complain about government works or doesn't, I also have to take it upon myself to figure out what I would do instead, right?

So please educate yourselves on these technological solutions. Speak to technologists, speak to your cousin who studies computer science.

We can all explain things to you and we will do so gladly but it's up to you, a member of the public to understand what's going on, so that you can productively speak up, rather than just, kind of say, oh, this is all broken and no surveillance and of course, I agree, surveillance is bad but what's an alternative? Right. It's on all of us to come up with an alternative.

I'll go with Albert and then over to Faine. Sure, and I think this is a question that Julia and I wrestled with a lot when we were on the Automated Decision Systems Task Force for New York City,

and I think this is, from my perspective, this isn't a computer science issue,

it's a civics issue. In a democratic society we have to have the tools to make sure that the governed understand how they are governed and can hold those making choices accountable.
it means all of the tools of democratic oversight that we see in other facets of
government.

So that means, yes, it means independent auditing. It means oversight hearings. It
means civics education, where we can make sure that the public is better educated on
how these systems work.

It means a deep investment in trying to make sure that every member of our
community understands the way that our choices about our communities are made.

It also means an ongoing national investment in the sort of AI
explainability and auditing technologies that can provide some insights into how black
box algorithms are operating.

But I always push back and say that,
you know, the technology is an indispensable part of understanding how AI works.
These, sort of, technical solutions are important, but it really, from my perspective, it's
much much broader than that.

Yeah, no, I think this is all great. Great insights. So what my work at Humanitarian
Technology feels, we've been thinking a lot about data ethics in the context of disaster
response aid.
a lot of biometric technologies,

a lot of tracking technologies are being tested out right now so which is obviously pretty disquieting thing

ethically, and during disasters, including COVID-19, a lot of technological experimentation takes place we're calling that, kind of, humanitarian experimentation of ethically problematic contexts where emergency, the need are used to justify rolling out untested tech.

And I mean that because the humanitarian contacts we're looking at this from human rights. So basically, what are my rights for someone affected by disaster regarding the data that concerns me personally. So in 2017 we released something at the Harvard Humanitarian Initiative called The Single Code, which is a rights based approach to humanitarian information and it elevated 5 rights people should have over their data, including the right to information and importantly, perhaps is the right to redress rectification, which is that we believe that people should have the right for what you were saying,
Julia to contest

false data about them, they should have, this is also designated in GDPR of course in Europe, people shut out the right to know what's being held about them. They should have the right to demand that data be deleted or a situation be rectified in some way.

And again, that does come down to legislative solutions, but I do like this framing of thinking about data as a rights issue. And when you violate my privacy with unethical data collection, whether you do it on purpose or accidentally, you're violating my human rights. You're violating the human rights of people who are having the stuff visited upon them who have more power than you do in society.

So, yeah, so that's the main thing I'd like to offer is that thinking of our obligations regarding data collection, ethical data use as really, very much human rights issues that pertain to people's safety, their lives and really important stuff.

Hyeong-ah, did you want to chime in on this question at all? Yeah. So, I had a chance to discuss with the people from Google and some other major technology companies. Basically what they say, we can do anything in terms of technology-wise. So, basically, if you know what exactly what we want, we can do it. Technology can provide the solution, but a much bigger picture here is really the policy level, ethical level, society level.
which is made by human beings.

So, in terms of these and where management is concerned. Yeah, the need,

we have to really look at that really much broader perspective. Technology isn't just a part of the tool and technology cannot be the harder it is just to help us as a tool.

So, this is how we should be.

Right, great. So, a few questions have come out, I'll shift the things that have come in over the internet through a few other channels. I see

we've actually answered them from WebEx, but not everybody on Facebook is able to see it. So, you know, some of these things are nitty-gritty, but I think they all, it's useful, useful questions and they'd love some answers.

Why are women of color at a higher risk of being misidentified, I guess with facial recognition specifically

than men? Julia, I saw you did a little text response, but, like Albert may have, I think everybody could have something to add in there.

But, you know, I think to a lot of people that are not steeped in the technology, they just look at that and go wait what, like, you know, how is that such a weird specific thing that we need to grapple with? Julia, maybe do you want to jump on that first?

True, but I already responded, right? It's a very technical answer. There's nothing in particular about women of color that makes their faces more or less difficult for these machines to recognize.

It's just we don't have enough data and the reason that people of color men and women are recognized again with the lesser accuracy than say,
whites, is because these machines are built and tested essentially by white men.

And so the level of light that is required for you to do certain facial features on the white face is different than for a person of color. So if we have more high quality data, it will mitigate this issue.

And also, if we start having a culture of testing accuracy, not just across the board, but in some populations also, and this is particularly important for small subpopulations, there are anecdotal examples of, for example, systems that are, Facebook has this system where they check whether your name is real or made up before you open an account, and it's a highly accurate system. But it makes all of its mistakes in the Native American category, which you cannot really at all detect if you're just looking across the board, right, because it's just such a small fraction of the population. And so that's one answer.
And another is, now that you've gotten this technology to work, now I can fall back onto Albert, and I say, now, do you really want to use it, right? Where is the context of use? You want it to work well, right? Yeah. And then if so, like, what is it going to do?

02:23:51.924 --> 02:24:06.415
But the calculator itself is not broken. You just give it enough examples, it's going to learn. Maybe Albert and then Faine? Yeah, I think that, you know, we often take the position, this is a bit rejecting and controversial that all AI is biased.

02:24:06.415 --> 02:24:12.864
It's a question of just how biased it gets, right?

02:24:13.375 --> 02:24:23.965
You know, because there are countless objective decisions that go into any automated decision system into which training data to get and how do you deploy the algorithm and how the system is designed, who's on the team? All of these different factors lead to the propagation and sometimes amplification of existing human biases and so, yes, training data can be one major factor and

02:24:23.965 --> 02:24:30.534
for example,

02:24:30.534 --> 02:24:39.114
IBM responded to the performance differential and visual recognition for women of color by investing money and getting more training data of darker skin,

02:24:39.325 --> 02:25:09.295
and boosting their performance, but then groups, like mine, will definitely come in and say great, now you've made it even more terrifying in a lot of ways, to use this technology as a law enforcement tool, as an immigration enforcement tool.

02:25:09.715 --> 02:25:12.715
But I do think that, you know,
all of these systems, you know, as magical as a lot of people think of them, they can be very fragile.

They can have a lot of, there can be systems that are incredibly accurate under most conditions, but, you know, for example, facial recognition, you have a system that will be accurate, the vast majority of time,

but if you alternate from one pixel in a photo, it'll suddenly not see the resemblance to the comparison photo.

So, there are all of these breakpoints in this technology and certainly with the human decisions that go into what counts as a suitable technology,

what counts as a marketable version of AI, that is where you get a lot of the human feedback that amplifies that bias. Yeah. And before I pass over to you Faine, there's another question here that I think would be for you,

but I think it gets to the larger point of, like, these individual pieces of technology may not be bad, but the context in which they deploy are problematic,

there's a question here about the thermal cameras you talked about, saying, often they do work at,

like,

but then beyond some specific distance,
whether it's facial recognition or, you know, a thermal standoff, thermal sensor, all these things have some utility and some value, but then when they're, you know, out of the right context, either the data quality itself, or the quality of the underlying training data may not be useful for the context in which the center was applied in the first place, so, I don't know, just to add that on for you Faine? Yeah. And I saw that question and I think, yeah, I went through some of the camera documentation, some of the manuals for these cameras, and that, yeah, the recommendation,
I mean, 
02:27:06.774 --> 02:27:09.594
even the 1.2 meter standoff at the airport, 
1105
02:27:09.625 --> 02:27:23.604
I would be somewhat suspicious that even the stationary camera being used is longer 
distances in an airport, and again coming to our talks of our discussion of 
implementation and training, who installed that camera? 
1107
02:27:24.000 --> 02:27:36.745
Those interpreting the data, are those people actually doing it right? Is the camera, you 
know, did the airport install the camera the proper distance? Those are all very valid 
questions and it’s entirely possible that airport cameras are not that accurate either. 
1108
02:27:37.194 --> 02:27:39.174
I imagine a drone camera would be even less so. 
1109
02:27:40.764 --> 02:27:55.014
Yeah, I think I actually know them from the earlier discussion about women of color, 
black women, and by spatial recognition, I guess this also comes out of the need for 
diversity in technology in the utilitarian sense. On the back, ethically, 
1110
02:27:55.014 --> 02:27:57.655
it's the right thing to do. It's also the right thing to do because 
1111
02:27:58.315 --> 02:28:10.614
if you have a very homogenous group of people, designing technology, working on 
technology, they’re just going to have ethical gaps and miss things. Risks from 
algorithms are going to apply to them 
1112
02:28:10.614 --> 02:28:20.245
personally. You could really insulate yourself against these problems by making sure 
you're bringing in first, the number of voices and if it's possible. 
1113
02:28:20.245 --> 02:28:30.174
If there are things that I am guaranteed to miss due to where I am in position in the 
world, that someone else will pick up on and that's an incredibly important for building 
and designing technology in society is that diversity. 
1114
02:28:31.825 --> 02:28:43.854
Hyeong-ah, do you have any thoughts on this question? Okay, so I'll move on. So we 
have another question from Facebook. Are there any concerns that, we talked about 
differential privacy in the context of the Census? 
1115
02:28:43.854 --> 02:28:58.165
So, are there any concerns that the Census’ introduction of differential privacy will obstruct data accuracy for public advocate, and academics, while the government and big tech companies will somehow retain the granular data?

I think there's kind of two questions in there. But, Albert your mic's hot already.

You went there first.

Yeah,

I definitely want to highlight the scholarship that's been out there on the possibility that differential privacy and the Census will mask the experiences of certain minority communities,

retaliate and HDSI at Harvard did a conference on this last year.

There have been a lot of folks looking at it because when you're using differential privacy models,

you can end up in situations where you're getting a lot of the larger questions that the Census will aggregate.

But, for example, when you're looking at representation of incredibly small communities within a state, then you could end up in a situation where they effectively get erased, and so that's one concern.

But as far as big tech companies getting it, that's not the case. This is information that would only be held by the Census for I think it's 70 years or 80 years, or something like that? Yeah.
And you wouldn't have access by private companies. So, Julia, if I have poorly messed up the differential privacy analysis, please, please set the record straight. No, absolutely.

You are absolutely covering it correctly. I think the big point here is that it is never the case that the one size fits all, right? So, it's never quite as simple as all models are wrong, some are useful. Different models will be useful in different contexts. This is like a more general statement of all AI is biased.

So, with differential privacy, I think what will make or break our ability to be comfortable with the use of this technology by the sensors and others is, whether we, as members of the public, are able to understand sufficiently well, how these tradeoffs are set. So, that we can speak up about where we think there should be more precision. So, for example, for redistricting really, we want data as precise as we can get, and where we can tolerate as a society more. Really, so, there's this magic parameter in these differential privacy setups called Epsilon that just tells you how privacy preserving you are, and the more privacy, the more noise, the less utility, right? So, absolutely, we, as the people should be deciding what the rules are, and that's for everything, including also how these are made by the Census, right?

So, at the moment, the Census is making these decisions. They are working very seriously
and thinking very hard about how to help members of the public understand these technical details and how they can really participate in the debate. But I don't think on the other hand that there really is a choice.

I personally think, and many will disagree that the Census, given the current legal requirements, has to add noise in the principled way, and so they have to use differential privacy. Again, there are different readings of this by different people.

But that's my point, is that we need to use it, but we need to enable members of the public to understand what the rules are, so that they're not hurt, but that they benefit. Great. Faine, your mic's on, I didn't know if you wanted to respond to it. That's all. Yeah. Okay.

So one is, kind of, a standing question that's come up in a number of our different panels and it's about this kind of vast marketplace, this kind of gray market of mobile location data that's out there, collected by a host of different apps and different companies, and it's aggregated, sometimes anonymized, and then resold, and that's kind of going on while others have been thinking a lot about the ethics of data re-use, right? We talked about limiting the amount of data, how long we keep the data, only using it for its original purpose.

what are the ethics of data re-use? And this may be a very American situation but,
with all of this different data out there and ethic,

you know,

or ethical guidelines saying,

you know,

anonymize it or aggregate it, once you smash a certain number of different data sets together there's no such thing as anonymity,

right? There's enough opportunity technologically to de-anonymize.

So, given that that's kind of this reality or maybe an emerging reality, or maybe in a reality we want to go away.

Here,

at least in the United States.

How do we think about that when these are data sets that someone can just purchase and drop into their big data system because it's legal, those have been anonymize everybody, followed the original ethical guidelines,

you know.

Well, let me stipulate that they probably, you know, they may have. How do we think about that? Albert? We try to outlaw it, I mean, just be speaking for myself, right now,

I helped author a bill that would outlaw the use of geo-location data by police in New York State, whether they purchase it from a data broker, or whether they get it by use of a court order,
just to say that it is not permissible in a democratic society for the information to be used by police departments.

And I think that we’re, at a moment, especially in the US legal context, where we’ve allowed the fiction of terms of service to govern our discourse for far too long, and we’re going to see a lot more laws at the state and local level, and eventually the federal level that simply say it is not ethical or legal to use this data, even if you had someone check a box, you know, thirteen steps before this in the aggregation process.

Right. Faine, I saw you nodding your head. Yeah, I just wanted to voice my complete agreement with that. It should not be legal. It should be illegal.

We should not be relying upon companies and other users, established to self regulate, which we have done for quite a long time and it’s obviously not working. And again, it’s not ethical. It shouldn’t be allowed to happen.

And that’s really about how complex it needs to be. Hyeong-ah, did you want to chime in on this?

No, I think I'll just pass it. Okay, and Julia? It might be the only question on which I’ll pass. Okay, I want to end with, you know, a little hopeful toward the future but, I have teenagers, many of us have children that are born as digital natives. They have these phones. They think it’s perfectly natural that, you know, they should know exactly where they are all the time. They should be able to see over the horizon to far away places, just by looking down at the palm of their hand.

So the cultural assumptions that the next generation has around location, around privacy,
around the quality of their data are kind of different from what I would have grown up with,
you know,
I didn't grow up with a cell phone in my hand, I didn't have digital maps available to me.
As soon as I walked out of the house,
as long as I was back before sundown,
I didn't grow up with a cell phone in my hand, I didn't have digital maps available to me.
As soon as I walked out of the house,
as long as I was back before sundown,
As soon as I walked out of the house,
as long as I was back before sundown,
So, given that many of these burdens, these challenges are not going to be faced, you know, necessarily by us but by this next generation that's coming up with a different set of cultural assumptions around this technology,
what's your advice to them or what do you think they should be thinking about? And maybe, you know, what should we be educating them about?
So, your experience with your children, are they very aware of the privacy issues, security issues for those teenagers?
Yeah,
I mean,

you know,

they read the full read-it review of the TikTok app and all the surveillance data that's streaming off of it and they keep using TikTok, and so I look at that and I say,

you know,

they get value out of TikTok that, whatever, cultural value

ey get out of it, and they don't see that they have anything to lose, even if they are 100% surveilled. And I just look at that and I say, fast forward 5 or 6 years, you know, when you have a credit score, and when you have assets, and when you have salary, you know, when you have a physical premise that you want to secure and you worry about the safety of your family, will you change your world view?

And I suspect the answer is yes, right, but that's this interesting educational conversation you have with your kids, you go, I know you don't care about it today, but, like, you know, you need to gird your loins, you're going to need to prepare for a different future
when you graduate from being an adolescent to an adult with assets that can be arrested and over-policed in new and interesting ways. So I don't know if I answered your question Hyeong-ah, but yeah. Yeah. I think I asked it to you because I mean, me as the other, the educator, to have a conversation with young people about what could happen in the future down the road. Because right now, I understand they don't really care about those issues because of most of the apps these days, they require that you provide all of the personal information. So I think it's really important for us to be responsible with the young people. Great. Julia, yeah, our other professor, another professor. I have a seven year old, so he doesn't really use this quite as much, but I'm just very naive in that, I think I am a hopeless optimist. So I think technology is nothing special. It's part of society, right? So, when we teach our children, we should just teach them to be good people and to change society in such a way that data is not seen as a liability that your existence in the digital realm was not against someone's sinister experiment watching you, and trying to figure out how to disadvantage you I think that we will have, with the next generation, an even better chance of building a society where things are more harmonious.
And I realize that, I'm saying this, you know, despite the federal climate, despite what we're seeing here with COVID and, kind of, with the very inadequate response the current government is bringing.

But I'm hoping that again, our children will be better than us. And that they will deserve a better society and a better government by being better. And so scaring people is, I don't think it's the right tactic. That's great. Faine?

Well, this is a tricky one for me, because I'm a millennial, so I was in that first generation, I was in one of those generations.

We were the first generation to be called the digital natives, and we grew up hearing about how our Facebook accounts, and our MySpace accounts are going to be used to ruin our lives when we got jobs, which is not entirely untrue.

But I do think people, honestly, I would actually really defend teenagers and other young people on this, and I've actually done some digging on some of the research that exists out here. I think they're pretty darn aware of the privacy violations, but as we saw with your kids, they're aware, they may make different calculations. There's some research that shows the younger people, and I think I would include myself here to some extent, are particularly prone to looking at, kind of, privacy helplessness and the sense that they know they're being surveilled. They know it's problematic. They don't think there's a heck of a lot that they can currently do about it. They don't necessarily think the individual choices that they would make have a huge impact.

So, I think it's one more. I don't think it's a matter of ignorance. In many cases. I really don't think it is. I think it's a matter of they're not sure what agency they have. Right. The utility of the technology outweighs the downside.

But, again, I do think that there's also, in terms of age based stuff, remember, we have a lot of evidence that, in terms of online disinformation. In disinformation, older people were much more vulnerable to that of younger people.
So, there’s a lot of technological education that we need to do, but it definitely should not just be directed as telling young people what they need to know, I think there also needs to be concerns about how age works across the differential. What are the risks to seniors and how would they differ from the rest of young people? But, again, the narratives often become that young people aren’t aware of the risks.

Like, they are, many know. I think, you know, before I pass it over to Albert, I'll just say I went to the Lafayette Park protests with my older daughter, and the same girl that doesn't care that TikTok is just sucking everything off her advice, she had the full Instagram, kind of, instructional on how to set the configurations on all your tech to limit police surveillance on US protesters,

so, I mean, she's definitely thinking about it but it's kind of like, you know, transactional and content, you know, related to certain contexts. But anyways, Albert.

Yeah, I think that's exactly to the point I wanted to make, which is that, in the work we've done, we’ve seen that younger users, especially young Zoomers,

they actually have a far more granular sense of privacy with online interactions than a lot of older users, and that. Yes, they may be fine with the sort of data aggregation that you see through TikTok, which to be frank, isn't that different from what you see with Facebook, but as soon as they are told that maybe their parents are going to see all the same data, they become a lot more privacy protective. The same goes for the police, or for other government agencies.

So, I do think there is a very nuanced approach to privacy that may not be as, where you may have a sense of that privacy helplessness when it comes to corporate data collection,
I do think that quite frankly, as someone who tries to pass laws against surveillance, you know, the only reason we've seen the success we have in recent months, is because of the young people who've been taking to the streets, who have been raising up their voices, who have been opposing the sort of draconian surveillance that's gone on in New York for far too long and who are pushing truly transformative changes. So I do think that the next generation will be a powerful force for progress in this area.

Well, that's wonderful. I really appreciate the kind of wisdom you've given on that last question, a kind of, positive ending to the discussion. It's been a real honor and a pleasure for me to have all of you with us today to broaden our horizons and give us more depth on this really complex set of issues.

So I'd like to, kind of, finish by thanking you, thank our panelists Mr. Albert Fox Cahn, Dr. Hyeong-ah Choi, Ms. Faine Greenwood, and Dr. Julia Stoyanovich for joining us today. Thank you so much for taking time out of your busy schedules. I'd also like to thank the Henry Luce foundation for sponsoring this series of Blue-Ribbon Panels and our Location Tech Task Force, without their support we would not have been able to convene such bright lights on these important issues of the day. I'd also like to make sure that those of you in our audience today know about our upcoming events. This is the series of events that we've had. I think this is the fifth of our Task Forces and we've also had a series of Leadership Spotlights.
So, coming up, we have some additional events, one is actually tomorrow on Digital Contact Tracing Tools featuring lead researchers at Harvard Kennedy School's Belfer Center for Science and International Affairs.

That'll be a spotlight hosted by my colleague, AGS President, Professor Marie Price, and then next, I guess, a little more than a week from now, we'll have another spotlight that Marie will be hosting on Technology and LGBT Location Privacy during COVID-19: A Conversation with Mr. Victor Madrigal-Borloz from the United Nations.

So, I encourage all of you to tune in for those, but also look into our previous panels that you may not have known about, share them with your friends, spur some discussion and please, if you have thoughts and feedback, we would love to receive that from you.

So, thanks again for joining us today, we appreciate you being part of this conversation and we look forward to seeing you again out there in cyberspace. Thank you very much. Thank you.