Fall Alumni Networking Lecture & Reception 2018: Professors David Lie & Lisa Austin

Farid Najm: So, my name is Farid Najm, I am professor and department chair for the Electrical and Computer Engineering Department. The full name is, of course, who knows to say this with me: The Edward S. Rogers Sr. Department of Electrical & Computer Engineering. It took a few years to be able to say this.

So recently as you will have noticed there has been a lot of exciting innovations in the field of electrical and computer engineering. Things like artificial intelligence, electric vehicles, robotics, quantum computers. This stuff is here to stay and it will have a disruptive influence on the field so it's a very exciting time. Of course what's under the hood with all these areas of technology is exciting to people like you and me. Things like power electronics, deep convolutional neural networks, Lasalle's invariance principle, right Jess? Where's Jess?—it's used in robotics—quantum key distribution. All of these techie innovations are very exciting but of course what's really exciting is to think of their impact. These technologies will have an impact on society whether it's clean energy impacting climate change, robots impacting surgery or AI to predict financial markets. So again these things are going to have a high impact and they're disruptive and these technologies start in the labs and the classrooms in here at ECE. But they don't stay in the labs and the classrooms and our alumni play an important role in this transition. Many of our alumni hire or mentor our students, speak at our alumni events, or make philanthropic gifts to the university or the department, including many of you in the room tonight, here. So thank you for your continued engagement with the university and the department.

Today's talk is on a topic that really affects everyone; it's on privacy, transparency and technology. This is obviously a topic that's interdisciplinary—it's not a purely engineering topic. It's a really unique collaboration that's happening here at U of T in the Information Technology, Transparency and Transformation Lab, also called IT3 Lab. So without further ado, it's my pleasure to introduce the co-founders of IT3 and our featured speakers here tonight; Professor David Lie from ECE and Professor Lisa Austin from the Faculty at Law at U of T.

Lisa Austin is a Professor and Chair in Law and Technology at the University of Toronto Faculty of Law. She's also cross-appointed to ECE currently. In addition to her legal training, Lisa holds a PhD in philosophy from the U of T. Lisa's research focuses on privacy, property and legal theory with an emphasis on the nature of the rule of law and the boundaries between what the law considers private and public. Currently she's interested in exploring issues at the intersection of law and computer science and engineering, which is what she's here to talk about today.

Professor David Lie received his Bachelors of Applied Science from the U of T in 1998 and his Masters and PhD from Stanford in 2001 and 2004. He is currently a professor in the Department of Electrical and Computer Engineering here. He's known for his seminal work on the XOM architecture, which was an early precursor to modern execution processor architectures like ARM TrustZone and IntelSGX. Currently David's interests are in securing mobile platforms, cloud computing security and bridging the divide between technology and policy.

With this, please join me in welcoming Professors Lie and Austin to give their talk.

[Applause]

David Lie: Alright. Thanks for the introduction Farid. So 2007, announcement of a new device, the iPhone. Probably even Steve Jobs at the time didn't know how big of an impact that device would have on our lives. There's a chart there showing the growth and the use of smartphones, as you can see it's going up. By 2021 its projected to have 40% of the world's population will be smartphone users, so a large number of people. And of those, a large percentage are also young people. People who are children and young adults who may not be aware of all the implications of smartphone usage, and some might consider somewhat vulnerable.

What is really happening on a Smartphone? slide

A question I often get is, what's really happening on these devices? And it turns out there's quite a bit. At the bottom layer, the manufacturer, either Apple or Android (put out by Google) has an operating system and that operating system interacts with the sensors on the phone, collecting information, helping to run the phone better but also sharing that information with applications you may have on your device. Things like Yelp or Uber or Facebook. Those apps will use that information to provide you better experience, give you suggestions, so on and so forth. Now, a lot of these applications are free, so they of course need to pay for the development costs and the cost of running these applications in some way. And a lot of them do that through advertising and analytics. These logos may less familiar to you but they are an integral part of how these devices are made available and how they run. Of course, there are features that help you understand and control what's going on. Both iOS and Android devices have a permissions systems so you can grant permissions to these applications and these advertisers to access some of your information and what they can do with it. Of course, not everybody may be aware of that, or know how to use these permissions. In the end, from a legal regulatory perspective, certain countries have privacy laws that require them to have privacy policies that dictate what data they can collect and what they can do with that. But of course nobody really reads these privacy policies. So, a couple of years ago we thought, this is an interesting problem and what can we do to look at that.

Who has your data? slide

The project basically looked at analyzing privacy policies and applications together and was featured recently in the news, there was an article in CBC news about this.

AppTrans: Transparency for Mobile Applications slide

What our project does, which we call AppTrans, is it combines machine learning on the privacy policies to understand what the privacy policy is saying, what the app is saying about collecting, collection of private information. And on the other side we analyze the coding itself to see what the application is actually doing. And then we compare the two, to see, does the privacy policy which dictates what the app should be allowed to do—and that's what a regulatory agency like the Privacy Commissioner of Canada would use to say is this app compliant, is it compliant with the law or not—and compare that against what the code says, which is what the app is actually doing. We do this checking so that basically

you don't have to. You don't have to read the privacy policy and god forbid you shouldn't actually have to look at the code to see what it's doing.

What did we find? slide

So what did we find? Interestingly, a couple of things. At the top there is just the number of privacy policy pairs that we compared, there's about 400 of them. The first row there says non-compliance. Non-compliance means the app does not conform to its privacy policy. The privacy policy says one thing but the app does something else, probably collects more information than was declared in the privacy policy. And it's about 60%, which maybe is alarming. I have to confess we weren't completely surprised, we were expecting this. But what's interesting is, if you go further down, 80%— actually 85%— of those violations are due to third-party code. What is third party code?

Question for the class: What do you think third-party code is in an application?

Audience member: Libraries.

David Lie: Libraries. Libraries that do what?

Audience Member: Doesn't matter, libraries that are included by the app itself.

David Lie: Yeah, that's the definition of third-party code, yes, external code. In this case, these libraries—the app developer didn't write them and they included them from somewhere. And the reason they're including the ones that are particularly causing problems is advertising and analytics. So it's the advertising libraries that are monetizing— allowing these app developers to support themselves— that are doing most of the data collection that isn't declared in the privacy policy. You might ask well, why doesn't the app developer just disclose what the advertising libraries are doing?

How come developers don't disclose? slide

So we looked at this too. And it turns out that's exactly what they're supposed to do. So if you read the privacy policy of—this is AdMob's, AdMob's privacy policy— what they're supposed to do is, it's the responsibility of the app developer to disclose what their application does and if you include our code, well that's in your application so you should disclose that. Turns out that is what they're supposed to do. But for an app developer to do this they actually need to understand AdMob's privacy policy. So here's just a screenshot of the top level and each one of those links leads to a whole bunch of other links and a whole bunch of pages. So can an app developer actually do this? Well, we looked at it and first of all, it's a lot of work, and then on top of that if you actually look at the code in the advertising library this is some of the most heavily armored and obfuscated code you can find. It's basically bordering on what malware is doing. So, in fact, it's just a big tangled mess. Our conclusion is that even if app developers wanted to be compliant, it's going to be hard, maybe even impossible, for them to actually be compliant because they cannot know what the advertising library is doing. So that's kind of the state of things now.

IT3 Lab slide

So to solve these challenges and address other problems like this we founded the IT3 Lab and I'm going to let Lisa talk a little bit about that.

Lisa Austin: Cheers. So, when we were working on this project together we realized there were lots of larger issues that were coming up around— that we sort of categorized around—transparency. By transparency we mean all the ways in which were are all becoming more transparent to corporations, to governments, and yet the technologies and processes that make us transparent are increasingly opaque. They're opaque to individuals, they're opaque to developers, they're opaque to many people and this is a problem in that we decided to set up this lab to kind of close that transparency gap and to do it through research that was truly collaborative and interdisciplinary. So David's talked about the AppTrans project from a developer's point of view— how do possibly comply with these obligations. But from the regulators point of view—and I've been talking about this project with Canadian privacy regulators, and they're very intrigued because they feel overwhelmed, like how do you regulate this world when you have thousands of apps that come on to the market every single day? You can't use the kind of process we've been using in the regulatory space based on individual complaints that are then investigated by humans. You'll never get at this, you need to start doing some automation in order to raise some flags and help regulators understand how to prioritize their scarce resources if they want to move to a more proactive regulatory model. So building these kinds of tools at scale to make the world transparent for regulators is actually really, really important.

When I mention to people in law that we are trying to do this collaborative work—you know lawyers like to tell people what to do—so they think, oh well this is great, go hang out with some engineers and you can tell them about all the values and ethics and norms that they're supposed to embed in their technology. And I say, well, that might be somebody's project and who knows if it's a good project, but that's not our project. Our project is actually a two-way street, it's a collaborative effort in order to try to figure out that if you're looking at a problem that we care about but from very different professional angles and with very different professional toolkits, how can we figure out how you can open up a kind of solution space in a very new and different way? And that's what sort of creative and interesting for us.

IT3 Goals slide

We have three major pillars or goals for looking at this. Which is around Transparency for Individuals— so we care about if you know how the technology works and if you know what's happening with your data, for example. Transparency for Government— regulators, auditors, law enforcement is certainly part of that and Transparency for Businesses as well. So we're interested in that transparency gap across a range of activities and different people involved in our lab have slightly different interests in relation to that.

So we thought we'd tell you about a few other projects that we've launched under this rubric of the IT3.

Open Data and Sharing slide

One issue that we're interested in is around data sharing, and sometimes in the open data context is one form of data sharing. And so sharing, there's nothing wrong with sharing, we teach our kids to share, sharing is good, there's all sorts of contexts in which we want data sharing to happen in the new data economy. But there's also some problems if that data sharing involves personally identifiable information or information that's easily identifiable. One of the context in which we started thinking about this—one of the contexts that many people in Toronto are thinking about this—is Sidewalk Labs. Sidewalk Labs is saying we're going to make the data collected in this project open by default. Open by default, they have two very important goals behind that. One is to allow other people to get access to the data, to innovate, this is about helping to foster innovation. And the other one is a response to the concerns about data monopolies, right, we don't want all this data to be hoarded by one or two companies, we want it to be a public resource that others can get access to. So all of this sounds right, but there's a concern—and those of you in the technical community know this better than those people in my community—which is, it's very difficult to take large datasets with individual-level data and properly de-identify it so that you can't re-identify.

Data re-identification slide

And there's been many, many prominent examples of re-identification of datasets; AOL, Netflix user-data, there was a big one in Australia around Australian health data. So there's a lot of risks of re-identification when you make this data easily available and publicly accessible, even if you de-identify it. And a lot of the methods for managing that risk of re-identification have a lot of really problematic trade-offs in this context, they reduce the accuracy of the information. So if you goal is innovation and reducing data monopolies and your solution reduces the accuracy of the data, you have a big problem.

Safe Sharing Sites slide

So what we tried to do—we have a piece that we're drafting and workshopping right now—we call it Safe Sharing site just a little bit of a tongue in cheek play on Safe Injection Sites. In my space, in the privacy policy world, I like to joke that folks think that sharing data is a bit like doing drugs. I was workshopping this at NYU last week and my commentator said "No, no, no privacy law people think doing drugs is way more acceptable." I think engineers think differently about this and I certainly don't share that view. So, we had this idea about safe sharing sites where we can actually foster the sharing of data that doesn't have the same kinds of privacy and security risks involved in it. The basic idea, and David can elaborate on the technical side of it, but the basic idea is one organization that has data they want to share, there is another organization who wants to use that data in some way. If you do it through this intermediary of a sharing site, they can use that data without having to get actually access it, get a copy of it, for that data to get out and start proliferating in all these other contexts where you can't control how it's being used and how it might be re-identified in that sort of context. The other part to it is that sharing can happen in lots and lots of different legal contexts, lawyers like to say, well what would it look like in data protection law, what would it look like if you thought about this problem in a litigation context and let's solve this sharing across all these variant legal contexts. What we're trying to do is say instead, what if you solve that sharing problem once, in one particular kind of way and do it properly and then make it work with multiple legal contexts? So we had this idea of creating what we

call a legal interface, so could you have this plug in to multiple legal contexts. And our idea was that if you build in a lot of transparency and auditability it actually can be made to work with multiple different sorts of regimes. And so we're playing with that idea in that project.

David Lie: Yeah so as Lisa was saying as she was explaining these various problems to me it really struck me how they were similar. You know as engineers we like it when we can build one thing and then tweak it a bit and re-use it for something else, and tweak it a bit and re-use is for something else. And this struck me as a good way of solving these various problems. And what do we mean by various problems?

Other Sharing Contexts slide

Well, one of the problems is that, remember the FBI suing Apple to get access to an iPhone, and that was a huge technological and legal quagmire because there was no easy way to give the FBI access to that one particular phone without giving them access to all phones. But, again, a safe sharing site where two sides could share data and one side could access the data that they need and not access all the other data. And on the other side, you would also not know what they are doing with the data but they just get the data that they need and the information out of that dataset that they need, would solve that problem.

Another area is online advertising. So in online advertising typically ads are done via a bidding process now. So actually when you look at an ad there is a millisecond auction happening where information about you as an individual is being sent to a whole bunch of bidders and whoever bids the highest for that slot gets to show that ad to you. And that all happens as you're accessing that webpage or as you're opening that app. But of course, for them to bid on the app they have to know something about you and so all this information is being spread out throughout the internet about you so that they can bid on that ad. Now this is a problem, in fact there is a complaint under GDPR raised against this, against all online advertisers. But the problem there is that they need this information to do the bid but there is no way to ensure— even though the regulations says after you do that bid you've got to throw away that information— we know there is no way to guarantee that information has been destroyed once its sent out. You cannot prove to me that you've forgotten some information. Again, a safe sharing site could solve this because you could run your bidding algorithm inside the site where it could see the data but when it comes out it's just whether you won the auction or not, you never actually see the data.

Finally on a smartphone, there's going to be a lot of things on the smartphone, for example health monitoring. One of the things I have up there is that a lot of insurance companies are now pushing these apps where you load it on your phone, it kind of monitors how you drive and scores you. And then that score is sent to the insurance company and it's going to affect your rates. Now there's a lot of problems with that, but one of the problems with that is also that that data is being collected and sent to the insurance company and some of it is probably extraneous, your insurance company probably doesn't really need to know this. Again, a safe sharing site would allow the computation to be done on the phone and then just the score to be sent to the insurance company. And the insurance company could also know, because the safe sharing site protects itself, that you haven't tampered with that

computation—that you aren't lying to the phone or causing the phone to send fake results to the insurance company.

So all these things involve this third trusted party that both sides trust. You trust the safe sharing site not to leak your data and the insurance company, or law enforcement, or the online advertiser trusts the safe sharing site not to lie and say 'oh, I've shown an ad when they haven't or I'm a great driver but I'm not.' It's almost like that adage in computer science, all things can be solved with a level of indirection. We think a lot of things here can be solved with a third party that both sides trust.

Implementing Safe Sharing Sites slide

Now where do we implement these? Well, we can implement them in data centres, large data trusts like what we could do for Sidewalk Labs. If all the data goes in there then people could do searches on that data but then we can guarantee that privacy isn't leaked. There's lots of instances where you might want to do it on the phone; online advertising, online apps that are profiling you for insurance or whatnot. And a third growing area is IoT devices, so smartphones. These are going to collect lots of information about you. Some of that is useful to… for example your power utility to just general information about what's going on at homes, or if you have a security monitoring company. But you don't want to share everything, you want to make sure that sharing is controlled. So again we want to look at implementing these in the huge case in the data centre, down to individual devices like a thermostat.

Finally, we're going to talk a little bit about a third project.

Canada's Approach to Communications Data slide

Lisa Austin: So we have another project where we're looking at what we call in law, lawful access. So these are types of conditions under which police can get access to certain kinds of data. And there's a debate in Canada about access to communications data that I'll explain t you a bit about because the debate just hasn't moved forward in about 15 years and so we're trying to propose to folks that there's a way forward, it's just not the one they think. So the way that the debate has played out through the last many, many years is that we take a kind of hierarchical and categorical approach to information. We say there are things in communications data like content, the content of your email, and for that it's clear, it's clear—what we call in constitutional law—a reasonable expectation of privacy. What does that mean? If you're found to have a reasonable expectation of privacy in something, you get constitutional protection for it. That usually means if the police want to get access to it they need to get a warrant on what we call reasonable and probable grounds which means that there's a high probability that if they get access to the content of your email they'll have evidence of offence. Those are kind of thresholds and protections that we put around something that we consider really private. So content, high, right? Super private.

But then there's those various forms of communications metadata that the Canadian government has said 'no, actually, that's somewhere in between. It's not as private. It's sort of private but it's not as private and we're going to protect it on a different level. We'll say what you need to show is, you get a

warrant on suspicion. Suspicion that is going to assist in an investigation, so, I suspect it's helpful.' It's not a very high bar at all. And that's for the communications metadata.

And then there's this other category; subscriber information. So this is just the basic, who's the account holder for the telecom. So let's say the police knew someone was doing something from a particular IP address, they would go to the telecom provider and say I want to know who that person is. Give me a name, give me an address. The Canadian government used to call this phonebook information. My students always say, what's a phonebook? **[laughter]** But on that they have this theory that… no reasonable expectation of privacy, it's just not private, its phonebook information and we should be able to get it without a warrant, we should even be able to get it on demand. And this shaped law reform for many, many, many years.

The 2014 SCC Spencer Decision slide

And then a decision in 2014 blew it up. So the 2014 Spencer Decision said 'well actually, for that low level phonebook information, that subscriber information, guess what? Reasonable expectation of privacy, get a warrant on reasonable and probable grounds.' So this completely blew up what the government's theory about how to think about this was. What the court did was it looked at the use context of the subscriber information, saying what you want to do is actually identify what someone is doing online and pierce their online anonymity and we're willing to protect online anonymity as an aspect of privacy under the Canadian constitution. So, blew up their various theories and called in to question this whole edifice.

Legislating around Spencer? slide

It didn't change the practice though because soon after there were all sorts of proposals, can we legislate around this decision. Now you may think, as non-lawyers, can you legislate around a unanimous Supreme Court decision? Actually you can, lawyers are very good at making work for ourselves. There are many ways to do this and they are exploring these different ways. And I've been at some of the stakeholder consultations for some of these proposals and this is what they look like.

WHERE ARE WE NOW? slide

And it's a fight, it's a fight between 'is this data really private, is it not private? What does Spencer mean? Is it private, is it not private? How do we interpret this case, how do we not interpret this case?' And people getting very, very, very heated.

So what a bunch of us, David and I and a number of other people, have tried to do in this project is say 'what if we just stop having that fight? We don't have to agree, just stop the shouting and open up a different solution space. Are there other ways in which we can actually provide tools to police to do the investigations they want to do that don't require such huge Constitutional compromises? It turns out in engineering and computer science you guys have a whole toolkit that we're not thinking of that's very exciting. But I'm going to make David talk about it.

PPLATs as a way Forward slide

David Lie: So one of the examples that I think is quite striking is the problem where because the bad guys also carry cell phones, the police will often request what's called a cell tower dump. So they have the cell phone IMEIs, the identifiers, of the bad guys and they just want to know if they are in a certain area at a certain time. And so they can make a request to the cell phone company to say, give us all of the identifiers of these base stations in this period. But of course that contains a lot of other stuff that's not relevant. And this is a story that plays over and over again, the whole idea of a data trawl to find out some specific piece of information. Well it turns out there's a mechanism that lets both sides get what they want. The cell provider there on the right doesn't want to reveal all the subscriber information and we also don't want them to do that, so that's us and the cell provider. But at the same time, law enforcement doesn't want to reveal the IDs that they're looking for because they're just suspects at this stage and they don't, for various reasons, want to reveal who exactly they are trying to monitor. So there's something called private set intersection which allows you to compute the intersections of two sets without each revealing the contents of that set. So what comes out of that is just the intersection of that set. So that would solve this problem because that's really what both sides are looking for. The cell providers don't want to give the full set to law enforcement. Law enforcement doesn't want to tell exactly what they're looking for. But again if you have a trusted third party that does that computation then you can get that without compromising the requirements of either side. So that's just an example of where breaking apart the legal problem and really understanding it means that there is a better technological solution that might be able to break what is otherwise a deadlock.

Check us out! slide

So we have a website and a twitter account and all that cool stuff. So check us out! We have this and other projects as well as reports and stuff for you to read on our website.

[Applause]

Farid Najm: Thank you guys this was great. We now have some time for questions from the audience.

Audience Member #1: I have a question, I understand that in some countries, in order to access your information that you know who's doing it because you get a little memo that they're looking at your healthcare, or they're looking at your financials. And then you say no thank you very much, or you say ok go ahead. Have we got that here?

Lisa Austin: In some contexts they have to tell you, but not in other sorts of contexts and its certainly an area of law reform that people are promoting with transparency reports and should you have to report to people when you're looking at their data and certainly an area of active research. It's not part of our research project right now but definitely is something that various people are looking at.

Audience Member #2: I'm sure you're aware of Statistics Canada, the latest endeavor. I'd like your comments on, is this constitutional, and are there third parties that can intervene on it? What do you think about that?

Lisa Austin: …legal advice! I've seen good accounts that they have the authority to do what they want to do, it may be ill-advised or they could have done better public relations. It's true there have been reports saying well the banks could stop this and take legal action. Lawyers are very good at finding arguments on many sides to fight that out. My initial reaction to it was that they probably could do this but that they botched the public relations around it. And that you can do it doesn't mean it's a good idea, but we'll see what happens.

Audience Member #2: Are they getting personal information? I mean it's the banks giving the statistical data. Are they being able, by this request, to get the individual who owns that data or is that outside…?

Lisa Austin: My understanding is that they were asking for the individual level and then they were going to —because they wanted to match it with demographic information they already had and then de-identify it later. They have various powers to compel.

Audience Member #2: So CRA could have a look at that then?

Lisa Austin: No, they wouldn't be allowed to do that. Statistics Canada has one of the most restrictive data regimes around. They have their own Act. They have very strict confidentiality rules and they are one of the only employees in Canada subject to criminal penalties for misuse of data. So they're actually very trustworthy and they have a very strict legal regime around what they do with things but they do have powers to compel access to data. Whether they should use them or not is a policy question.

Audience Member #2: Well if they're that trusted why is there so much concern around it?

Lisa Austin: Yeah, yeah.

Audience Member #2: We know if we get the wrong leader, like a Trump, everything could change overnight.

Farid Najm: Alright, thank you. Another question?

Audience Member #3: So these themes of trusted third parties immutability, privacy, security are consistent with blockchain technology, so I was just curious is there a potential for a permission block-chain to help with some of these initiatives?

David Lie: Permission blockchain. So, blockchain is useful for some contexts, we get this question a lot. So my answer is that blockchain is a good technology for audits because it kind of distributes the information among a bunch of parties so even if one…some subset of them lie or somehow corrupt it—it's kind of like the crowd, you have to get a whole bunch of people to corrupt or lie all at once and the chances of that happening are lower. But blockchain isn't necessarily a good fit for things where you want privacy because the way it works is that that information is distributed and replicated across a bunch of nodes and so information is, by its nature, spread. That's how you get the reliability of blockchain. For digital currency, something like bitcoin, you don't actually get privacy, you get anonymity because the actors who are doing the transactions are just private public key-payers, it

doesn't tie you back to a person. But you don't actually get privacy because actually all the transactions done by those key-payers are open for everybody to see. So that's an important distinction.

Audience Member #4: Your safe sharing site solution was kind of described mostly in technical terms. I've kind of thought about this stuff for a long time now and I came to the conclusion with some of the ideas I had way back when about some of this stuff was that ultimately it had to be organizational, capital structure solution, almost like a co-op. Ultimately because just this week we heard about Deep Mind betraying its user base, so to speak, by being incorporated in to the larger Google organization which has really terrified the people whose information is in there. Actually I was in the health care field a while ago and I was thinking of about something about privacy-related sharing of information. And I kind of thought at all about the different barriers you could put up technically. And I kind of concluded that ultimately, just by definition, none of them will be sufficient because, as this gentleman pointed out, once you get a Trump or something like that, from the human side of things, from the organizational side of things, who undermines the entire notion of privacy or what have you, or accountability even, all is lost. So ultimately, do you guys anticipate addressing in your lab organizational forms like cooperatives or collectives or other kinds of legal instruments or structuring that can put firewalls at multiple domains around this?

Lisa Austin: The short answer is yes, the longer answer is the safe-sharing site, we were making it with the understanding that it was going to intersect with multiple types of legal regimes. So it doesn't itself solve the accountability question…

Audience Member #4: Yeah, it wasn't a criticism…

Lisa Austin: No, no I'm just saying it's meant to connect with multiple types of regimes. You might have a one-off kind of thing. In litigation, for example, recently—I mean tobacco litigation isn't popular to talk about—but there's a case that went to the Supreme Court of Canada around the tobacco litigation and one of the issues there was should the tobacco get access to the health care data that the province of BC was using to base its assessment of causation and costs. It's worth billions of dollars to the company involved and they wanted to do their own analysis of the data. They went to the Supreme Court of Canada on a very technical statutory interpretation question which I won't bore you with, and the answer was not at this time, not in this way because of this statutory interpretation issue. But it left open the question that there's all sorts of litigation contexts where you might actually need one party to have access to that data. That would be a kind of a one-off context that would have to manage in the context of rules of discovery and court supervision that wouldn't necessarily raise all of those questions that you're saying. But say like a Sidewalk Labs context would definitely. So we have a project we're going to start, but we're just talking to other people that are going to be involved in it, looking at the concept of data trusts, which is more of a legal concept, how you could make it work with our idea of safe-sharing sites and what some of these governance models are legally and how to make them play well with the technical realities that they're meant to intersect with. So we are actually starting to think about some of those other pieces. This is just kind of one piece of the puzzle, not meant to be the whole view of it.

Audience Member #5: I think this is a slightly difficult problem, especially with a lot of closed-source software, where you can't exactly see what's going on. And maybe in a phone environment because it's somewhat isolated you can do this, but, for example on Facebook, I can't see what they're doing with my data, the source code isn't openly available. What would you propose to … Would the onus be on the user to clean up our data or would it be possible to enforce this on larger organizations such as Facebook?

David Lie: Yes, that's a good question. How you kind of get to a world where right now all the data is being held in organizations like Facebook and Google and the code operating of that is also in the same organization. But ultimately what we want to look at is a way to carve out a piece of that. And enforce that, possibly with regulation or the threat of it, that says that some part of this has to be done in an auditable way. So, the comparison I like to make is that Facebook and Google they also have their own bank accounts that they have control over, but the IRS or the CRA here, doesn't worry—I mean they do worry, but not that much—there are ways of ensuring that they're not cheating on their taxes. That's because they're required to keep an auditable trail and if there's something suspect there are legal remedies—legal ways of compelling them to reveal information. So I think technology is part of the solution, but there's a reason we're having this collaboration. Technology—it's impossible well I think it's mostly impossible to build a technology that cannot be corrupted by a human or by its owner. So there has to be a way of regulating and making sure that it's used in the correct way. And I see that at least in Western democracies or such like that, laws are the way to do that. Even someone like Trump, he's still constrained by whatever laws that are placed on him. It is important that law and technology work together to make sure we achieve what the goal is.

Audience Member #5: And I guess there would be the other problem, just like tax evasion, for example, people do have offshore accounts to circumvent a lot of tax laws and I think the other issue is it would be a very difficult issue to enforce these things on a global perspective, right. We can have great laws here but offshore it might not be so great…

Lisa Austin: And that's a great research problem in law around global governance and how do we regulate in a global economy. And there's lots of people interested in that and how that works in the data economy. You know we're slowly finding each other and creating networks. Certainly our lab is with folks doing other pieces of this research. But if you want a more optimistic note than that, it always makes good work for lawyers and engineers. You'll be employed!

David Lie: Lisa can say more to this but there are some examples of, what do they call it? Extra…extra-territorial, right. So like when the US said we won't do business with you unless you reveal the information of all US citizens, like financial data, to us. Or GDPR, like its all European citizens but you can't actually identify who's actually a European citizen. So effectively it's imposing a policy on the entire world. So these are the mechanisms we need to employ along with technology to make things happen.

Audience Member #6: What would the governance model look like for such a system, like the safe-sharing site? Who enforces it, who's running it, are they censoring things? You've kind of moved the

problem another step along the way where now this central sharing site has all of our data and if it gets compromised it now has access to Google's dataset, to Amazon's dataset, to Huawei's dataset. So how do you solve that?

David Lie: We should hire this guy. I think actually the real answer is we're figuring these out. To speak to one thing is that I don't think centralization is necessarily fatal because in some ways a lot of things are already centralized right now, right? Like Google and Facebook and even the Canadian government has a lot of information on you. We expect them to keep it safe. So I think the security problem can be solved as well as it is now, if not better because you'll have more resources if it's centralized. But ultimately the questions about how these entities would be governed, that's something that we don't have the answers right now.

Farid Najm: Ok thanks. One more? Let's have someone who hasn't asked a question yet.

Audience member #7: I find it fascinating we're talking about who is going to do it and talking about the intersectionality of engineering and law and there's an elephant in the room that no one seems to talk about and that's PEO. They're ostensibly here to regulate technology and have always been the intersectionality. Personally I'm an EIT, I will hopefully be a P. Eng. in two weeks, two months. But you talk to a lot of those guys and their philosophy is, the bridges are staying up, we're good. Some of the disturbers seem more interested in the cyberworld but in my mind it's a no-brainer. Can you talk to that? And, I'm sorry, I see frequently more lawyers interested in these issues than P.Engs, so I find it fascinating.

David Lie: Well, yeah I mean. PEO is our Professional Organization. I think ultimately we have to find the right partner for this. It may be a PEO, it might not be. But I think PEO has a job. They need to make sure that people who are claiming they have engineering expertise aren't … sorry people who don't have engineering expertise aren't claiming to be engineers and designing houses and bridges that might fall down and hurt people. When it comes to technology, especially information technology, I feel like PEO kind of steps back a bit. It's not as regulated, they don't necessarily know what their role is. That could be something we define and it could be them or it could be someone else but I think that's a good point but we don't really have the answer right now.

Farid Najm: That's great. We're going to have to stop here, we can continue the questions in the networking session. Thank you David, thank you Lisa.

[Applause]

So as I said, we still have time, we can enjoy some more networking and more food and drinks. As to networking you will see this card on your seat. So this is our attempt at doing good networking with alumni. The ECE CONNECT under the engineering CONNECT platform. You can sign up for this very easily, go to the website on the card. This is a website for networking with other alumni of all ages in various locations around the world. Some of them are very new, junior alumni so you'll have the opportunity to mentor younger alumni. And younger alumni, you have an opportunity to benefit from the experience of older, more experienced alumni. So it's like a curated LinkedIn. We approve people to

join this and you can upload or download or whatever you call it, your data from LinkedIn to CONNECT. I think this is a side-load. So it doesn't take a lot of work to get set up. I think you'll find it fun, we have what almost 1000 ECE alums at this point? More? 1300 ECE alumni on the platform and we only started it a year ago so people are finding it useful and we are discovering new alumni that we didn't know about who join the platform and we get to invite them to these events. So thank you for coming. We still have time for networking and food and drink so we can get to that now. I'm trying to make a joke about a lawyer and an engineer going up to a bar, but I don't know where it goes from there. But they will both be here to answer questions. Thank you and enjoy the rest of this.