Privacy: Front and Center

Ann Cavoukian | Information & Privacy Commissioner of Ontario, Canada
Alan Davidson | MIT Technology and Policy Program
Ed Felten | US Federal Trade Commission
Marit Hansen | Deputy Privacy and Information Commissioner of Land Schleswig-Holstein, Germany
Susan Landau | privacyink.org
Anna Slomovic | Equifax

In the 10 years since *IEEE Security & Privacy*'s initial launch, privacy has moved from being a side story occasionally in the newspaper to a central issue of our times. Through the rise of online social networks, tracking technologies such as cookies and Web beacons, and the sharing of data with third parties, as well as with the government's increasing use of surveillance mechanisms, almost everyone experiences far less privacy than they did just a decade ago. At the same time, governments and industry are taking much more of an interest in privacy protection than they did when *IEEE Security & Privacy* first appeared. In this roundtable, we discuss some recent concerns with five privacy leaders: Ann Cavoukian, Ontario's privacy commissioner; Alan Davidson, a recent head of Google's US public policy office; Ed Felten, who recently served a term as chief technologist at the US Federal Trade Commission; Marit Hansen, deputy privacy and information commissioner of Land Schleswig-Holstein, Germany; and Anna Slomovic, chief privacy officer at Equifax.

—Susan Landau, *IEEE Security & Privacy* editorial board member

Susan Landau: How easy is it for average users to protect their privacy in day-to-day use of the Internet and other communication technologies?

Ed Felten: This is a real challenge for users. The technical choices offered to them are very complicated, and their effects are subtle. I think we're starting to see users get more help from intermediary technologies, but we need to move to an approach where users make their general preferences known, and then technology helps them manage the landscape of technical choices. That's part of the problem—how do you operate the technical controls available?—but another part is understanding the practices of the different parties to whom you give data. Where will the data go, what will happen to it, and what are the implications for the information flows that happen? This is all quite complicated. We'll need to have more technologies, more institutions to help users make these decisions because if a user has to work all of this out himself from first principles, that user is going to be in big trouble.

Ann Cavoukian: I totally agree with Ed. The ability of users to not only anticipate but stay up to date with technology in any kind of nuanced manner is unrealistic. Users are like all of us—busy. Most people don't have the time or the inclination to read privacy policies or understand the nuanced uses of their data. I think openness, transparency, and an increasing abundance of attention on the part of data organizers and collectors to be clear and straightforward with what they will do with the data is crucial—identifying the purpose of data collection to users, the reason why.

Anna Slomovic: I think the answer to this question very much depends on what you mean by privacy. If you mean keep all activities between me and the people I do business with private, I completely agree with Ed and Ann—if you're technologically savvy or have help, it's possible to retain some measure of privacy. If you mean doing something without being observed or recorded, we're beyond that. Anytime you go online, your activities are recorded and analyzed. The issue is more about how your data is being used more than preventing it from being collected in the first place.

Marit Hansen: More encryption would address this last issue, so that even if some communication is monitored all the time, there's the possibility of having something said that is at least not easily decipherable. In principle, the technologies are available but not for the average user. Where we can, we should have more encryption by default.

Cavoukian: I just want to give one example in agreeing with Marit. The government runs the casinos and gaming operation in Ontario; each of the 27 casinos has video surveillance cameras at the front that match to a facial recognition program for a very limited purpose. It sounds like a lot of data will be collected, and certainly that would be anyone's assumption, but it's not. The programs we have in place involve biometric encryption, so the limited amount of data collected can't be accessed without the user's presence: the user's face or finger is what decrypts the information. No question, there's more recording and data collection taking place, but it doesn't necessarily happen in a way that makes the data widely accessible.

Alan Davidson: I think we all agree that it's not as easy as it should be for people to protect their privacy. But at the same time, we're seeing a tremendous amount of innovation that offers more tools for users to protect themselves. We're certainly seeing this on the industry side. People understand that consumers are not going to use services if they don't trust them, and building in strong privacy tools is essential to that. The conversation about encryption gets at an important question: What is the threat model that most people face? There are really powerful tools for people who want to understand what information they share with other users or companies. Where it gets much more difficult is when you deal with very sophisticated collectors of information or adversaries. For example, we've only just touched on this question of how to make consumers feel protected in the face of government access to information. That's a big issue unto itself, and an area where government needs to get its own house in order.

Felten: The possible data uses and controls on uses are some key issues. What Ann was describing in the casino situation was a collection of information limited to certain uses and designed technologically so that only the desired uses are feasible. But a lot of collection done today happens with an eye to the collector finding new uses for the information later, which means that it tends to be collected in a relatively uncontrolled way. In some respects, that's a problem you might face image collection out on the street as opposed to in this limited setting where care has been taken. There's no doubt that we have a lot of technical capabilities in terms of designing privacy-preserving technologies, but many of these technologies don't get deployed because [data] collectors don't want to give up the possibility of finding new uses.

Landau: That brings us back to Alan's question about getting the government house in order.

Slomovic: The government in the US, when it collects data, has to have authority to do so. You might not read the *Federal Register*, but the stuff the government is required to publish in the System of Records Notices and other documents, so they are out there for people who want to read them. A different question is what happens when the private sector collects this data and then the government goes and buys it. In many cases, the private sector doesn't have to provide notice, and the government is perfectly happy to go and buy the data if it's available.

Landau: Where might the presssure point be for controlling the use of data in an uncontrolled situation, as opposed to a situation like a controlled Ontario casino.

Cavoukian: That's a tough one. Part of [the demand] is going to come from the public as consumers demand additional trust in terms of the organizations they're willing to work with, but I'm 50/50 on that because of Facebook. Facebook, of course, has introduced face recognition technology to the pictures you upload to it, and we've been urging the company for some time to do something with the program to make it more difficult to enable facial recognition. There is very little interest on the part of Facebook in doing this—in making it more complex or looking at biometric encryption or taking any additional measures—which is why I get very concerned in terms of the future. The public clearly wants its privacy and is interested in pursuing methods to ensure it, but generally speaking, people don't have the breadth of scope to insist on these kinds of things until something bad happens.

Davidson: I think we shouldn't limit our perspective here to the privacy of individuals; we should also look at companies. Industry wants to protect the security of employees and their communications because otherwise it opens a door to company secrets. When US companies use US services subject to US surveillance laws, this might be less of a problem. But think of employees using a Chinese Facebook or some other service, operating where surveillance laws allow more government access. I guess that's when users will need to think about behaving differently.

Landau: For decades, European privacy commissioners have used laws based on fair practices to regulate industry—think of Microsoft Passport or Google Street View, or forcing opt in on cookies. Meanwhile, in the past 15 years or so, the US has been engaged in a sort of race-to-the-top approach with the Federal Trade Commission's prosecution of companies that don't live up to their privacy policies. Which system works better, and what can the communities learn from one another?

> **The issue is more about how your data is being used more than preventing it from being collected in the first place. —Anna Slomovic**

Hansen: I come from a culture with many laws and constitutional guarantees on privacy protection for individuals. But even laws can't guarantee that requirements are being fulfilled. Authorities throughout Europe can only check a very small fraction of data controllers, so noncompliant data processing can go unnoticed. In addition, the economic incentives to be compliant with privacy law aren't very high. Even if a data protection authority can prove an incident happened, the fines aren't very high, or it's difficult to get them enforced. Instead, we need clear criteria and the ability to debate on these criteria. Whether they're imposed by law or some other means, we need a system of incentives, sanctions, and enforcements, with people who can monitor independently, not abstract sanction threats that we can't enforce. I think in Europe and the US, the technological expertise of data protection authorities or the FTC or other institutions responsible for doing the enforcement isn't very high. Computer scientists don't work in these supervisory authorities.

Felten: It's true that the US doesn't have a broad, comprehensive privacy law. But I think that privacy enforcement in the US has been reasonably effective and provides benefits to consumers worldwide. It's an interesting contrast: Europe has more comprehensive laws, but the US has a more aggressive enforcement of the laws. Another point is the technological capabilities of enforcement agencies. In the US, there has been a concerted effort to build up the FTC's and other agencies' capabilities. There's a new trend to have more computer scientists working in these positions, including my own work at the FTC.

Davidson: Having worked at a company that was on the receiving end of an FTC enforcement action, I can say that it's a powerful regulatory tool! Enforcement actions are taken very seriously, particularly by larger consumer-facing companies. Major players such as Twitter, Facebook, and Google have all entered into consent decrees that will govern their behavior for a long time to come. That said, I suspect most people in industry don't view this as an either/or [between principles and prosecutions]. There's a benefit in having a broad set of principles that provides a little bit more certainty, and that's why we've seen people coalesce behind the idea of some sort of baseline privacy law in the US. Whether you have a baseline rule or not, we need to invest in enforcement, and not just against the big names consumers recognize. We need to monitor activity across the board. It's not necessarily the most glamorous work, but it's important.

Landau: Can practices from other parts of the world help here?

Cavoukian: Whether it's the EU, Canada, or the US, the number of data breaches that actually comes to our attention as regulators is getting smaller and smaller. My concern with the growth of privacy-related infractions is that, given the volume, we need something more than enforcement. I applaud what the FTC does—its rulings resonate. But I don't know the effect on smaller players across the board. In my jurisdiction, we go after everybody who comes to our attention. We have order-making power, which means I can order government departments that fall under my jurisdiction to comply with the law and change their practices, and I can order the cessation of collection practices that are in breach of the law. But the scope of harm from infractions is growing so wide with online social media, Wi-Fi, and mobile devices. I just think it's too much to expect any regulator to be able to stay on top of this.

Hansen: The feeling in Europe is that the system right now is just too complicated. Even experts don't get it. The newly proposed European Data

Protection Regulation is going to be the one legal instrument backed by the full European Economic Area and is a more simplified approach. On the other hand, financial incentives might work—for example, if the government wants to purchase software or services, it could prefer products or providers with certifications that prove you comply with data protection law, or such vendors or providers might profit from tax breaks.

Felten: I want to echo others about the sheer number of potential violations versus the resources available for enforcement. What can be done to close that gap somewhat? First, we need to make sure more resources are allocated to enforcement, especially on the technical side. We also need to make better use of what the technical and research communities know about how to design, test, and monitor system behavior to more efficiently identify potential problems. Moreover, it would be helpful to think carefully about where the lines are drawn legally to try to understand how to provide meaningful protection for consumers while simultaneously providing roles that are easier to follow and monitor.

Davidson: In terms of how you influence the broader industry, one example is the Electronic Frontier Foundation list of "who's got your back." It's a list of major companies and with a somewhat subjective assessment of how well they do at protecting privacy versus meeting government requests for data. People pay attention to this stuff, and industry understands that. So there's a huge role for civil society and other actors in terms of monitoring what's going on and talking about it.

Landau: Privacy by design sounds good in theory, but industry wants data from customers, and law enforcement and national security want data from suspects (and sometimes, potential suspects). How does privacy by design actually work in practice?

Cavoukian: The beauty of privacy by design is [that it's] a very comprehensive, holistic, creative way of addressing privacy problems and solutions. Instead of just looking to a regulatory legal system that allows the harm to arise and offers a system of redress after the fact, it prevents the harm from arising in the first place. It seeks to advance privacy interests, but there's room for multiple functionalities to coexist in the same space, allowing for identity management. Privacy by design is focused on use limitation: you have to very clearly identify the use of information to the user, and then you restrict your use to that specific purpose, which becomes embedded by default—users don't have to scour privacy policies and figure out what they mean. They're assured of privacy because it's embedded in the system's design as a default setting.

Slomovic: I'm not really sure how privacy by design would work in two sets of circumstances. First is that, given the fact that we engage in many more transactions that parties need to independently verify, there's an actual need for third parties to independently collect data and provide claim verification, whether it's credit worthiness or identity services or something else. So, if you're designing systems in a way that doesn't permit this, you actually limit the positives you can get out of online interaction. Second, in the private sector, to do anything, you have to provide a business case. In a data-driven economy, very often a business case very much depends on what you can do with data in terms of reusing and repurposing it. So as good as privacy by design sounds, I'm not sure I understand how it would be implemented in practice in at least these two cases.

Cavoukian: The notion of repurposing data, the value of saving it for some other purpose—I'm going to suggest that although it was strong in the past, it has become more of a dated view. In many circumstances, keeping the data just because you might have some better use for it later doesn't address the attendant costs of keeping it, such as a data breach.

Davidson: Privacy by design is a powerful product development philosophy. But we shouldn't think of it as an absolute bar on data collection or a magic talisman that we can use to solve all known privacy problems, which is how people sometimes tend to see it. It's a useful tool for identifying real risks early on in the development cycle, and trying to either mitigate or remove them entirely. We shouldn't expect it to be something that we can invoke and then suddenly erase all data collection or privacy concerns. My favorite example of privacy by design is the Google Chat "off the record" button. Initially there was this notion that Google Chat might collect every chat and log it, which would be very useful for some people. But then the developers realized early in the design cycle that this could create major privacy problems, so they created a way for users to control whether chats would be logged or "off the record." Not to say that data collection doesn't happen sometimes, but it should happen with user control. The idea of getting in early, understanding the issues, and mitigating them is what's so powerful about this approach.

Slomovic: But that only works in certain kinds of applications. It doesn't actually work in other kinds of applications.

Davidson: Absolutely.

Slomovic: Think about risk management, right? If you want to have some kind of a nonrepudiation capability or a capability to demonstrate that somebody actually engaged in a chat when they claimed they didn't, you can't not record it. And that's an issue in the identity space—if you want to have nonrepudiation in the identity space, you can't not allow identity verification to happen, and you can't not record.

Landau: This leads to the issue of de-identifying data by removing names, addresses, birth dates, and so on, which isn't working as well—we see increasing amounts of data stored about users, and the search algorithms to match data have improved tremendously. Yet, using anonymized data is crucial for many types of research. How do you handle this?

Slomovic: The healthcare space offers some insights. First, on the institutional side, we have the concept of a data use agreement, which is specifically designed with data that isn't fully de-identified and deals with human subjects. In human subject research, we also have institutional review boards that evaluate research projects for potential benefits and harms to data subjects. Many social science projects also require a dispensation from the institutional review board before they can proceed. Researchers aren't entirely thrilled because they claim this bureaucracy slows them down, but in fact, slowing down and articulating harms, benefits, and how the project will go forward is the goal. Second, we have an emerging concept of an agreement between institutions and individuals in which individuals are data donors, so the individual consciously donates data for research purposes. It also keeps individuals more connected to institutions and more involved in the research process. I think these kinds of approaches—new models that don't rely on anonymization or technical means alone—could provide us with solutions to many of our privacy issues.

Hansen: I like the option that people are involved if they want to be involved. Even then, I think it isn't necessary that the participants give their full identity in all cases.

Felten: One of the most important issues is making sure we're honestly recognizing the tension between privacy and utility of these datasets. In the policy world, there's a tendency to wish this problem away or to assume or hope that some basic technical mechanism will provide a level of privacy protection that it simply doesn't. We need to either think in a sophisticated, technical way about addressing these problems, or we need to open up the other tools at our disposal, including agreements, limitations on use, the role of institutional review boards, and so on. My concern as a technical person involved in the policy process is that solutions that are too easy and don't actually work will be too attractive to policy makers who aren't fully engaged in understanding the nature of the problem. There isn't going to be a data-scrubbing panacea that will solve these problems for us, although there are interesting advances coming out of the research community that could be one part of our strategy in dealing with this.

Cavoukian: I don't want to be the contrarian in this, but in Canada, the government is the healthcare provider, so the health insurer is the government. This means that a lot of data on individuals is accessible by the government, and, of course, researchers also want access to it. We have a system in Ontario: an organization called ICES [Institute for Clinical and Evaluative Study] that gets all the data in identifiable form on the population of Ontario, but maybe two people see it in this format. Before it goes anywhere, and before it goes into the hands of researchers, all personal data identifiers are removed, algorithms generate unique code used in association with the data, and that's what goes out to researchers who then use the data essentially in nonindentifiable form. If they need to work back to link to an identifier for future purposes, they can, but in a very limited way, so creating algorithms that create encryption tools that create unique identifiers that can be used in association with the data is a very strong possibility. It preserves data utility while holding policy makers absolutely responsible for the data's future.

Slomovic: Ann, I don't think anybody is talking about stopping de-identification, at least, I'm not talking about stopping efforts to anonymize or de-identify data. But so much data is available now that what constitutes de-identification today might not constitute it tomorrow—look at what Alessandro Acquisti has been doing at Carnegie Mellon. What I think you're describing is actually a combination of de-identification and administrative and legal controls, which is pretty much what I was talking about. If you can't be sure that you're going to have data permanently and completely de-identified and still have it be useful, then what you do is bring in other mechanisms to make sure that whoever is handling it is not re-identifying it.

Landau: I seem to be hearing from you all the idea that privacy protection is going to be a toolbox that includes technology, but also policy and law, and all of these pieces need to work together. Shifting the topic slightly, we've grown accustomed to free search, free Facebook, free "you name it" Internet service, but as many have observed, if you're getting a service for free, you aren't the customer, you're the product. Does this mean that the services we've grown to rely on—search on Google, sharing social information via Facebook or LinkedIn, scheduling meetings via Doodle—will need to become paid services for those who want privacy?

Davidson: I question the premise that advertising-supported models are incompatible with privacy. They can be, and arguably, we've seen some quite privacy-friendly ad-supported free services. One good example is the success of context-driven advertising that some would argue minimizes the amount of data collected about users. We should also recognize that free ad-supported products provide enormous benefits to consumers. We're living in a wonderful age when consumers have access to a huge amount of free content online, and it's driven by advertising. As long as [consumers] are fully informed and have good choices, we shouldn't take those choices away from them. All that said, I suspect that we'll see other models evolve. Services like Dropbox or Doodle prove that consumers are willing to pay for valuable services online. Sometimes, they're even willing to pay to avoid advertising. An interesting question for this group is whether we think that subscription-paid services are actually more privacy friendly than advertising-supported models. It's not at all clear that it's better for a person to be followed around the Internet by their credit card number as opposed to an advertising cookie.

Slomovic: In traditional publishing, you had both advertising and subscription. What I'm not seeing online is the ability to pay to have greater privacy—for example, on LinkedIn, you can pay to see more data about other people, but you can't pay to have others see less about you.

Cavoukian: I'm actually sympathetic to companies trying to offer these amazing services that we're all getting very, very accustomed to as users. We love them, we get them for free, and we're getting spoiled. There has to be some exchange. I think we'll eventually see a trade-off that will benefit customers and that will preserve the advertising model because it's what fuels the online world. I should also point out that just because advertising is directed to me in targeted way, if as a customer I know that my identifiable data isn't in anybody's hands, it gives me great solace.

Felten: I think it's reasonably likely that in the long run, effective targeting methods that don't rely on extensive data collection will emerge and be successful. But I think we also need to look beyond advertising because information about consumers' behavior online gets collected for all kinds of purposes. It's a red herring in this discussion. Even if effective targeted advertising doesn't require information collection, some entities out there will want to collect the information for other purposes, and we'll still be having this debate.

Hansen: It's not sufficient to not show ads to people who pay for privacy. I know of models where the full data are analyzed and processed; the ads aren't shown, but the information is there and even sometimes used. Even if we take the principle of information privacy seriously, that doesn't mean you can abandon your rights—for example, in several areas in Europe, your personal communication must not be monitored [by private companies], and that can't be easily waived.

Davidson: It's actually surprising that we haven't seen more models put forward to consumers, whether they're based on subscriptions, ads, or other kinds of collection. A big key is going to be making sure that people really understand the model they have chosen. ■

Ann Cavoukian is the Information & Privacy Commissioner of Ontario, Canada. Contact her at commissioner.ipc@ipc.on.ca.

Alan Davidson is a visiting scholar at MIT. Contact him at alanbdavidson@gmail.com.

Ed Felten is chief technologist at the US Federal Trade Commission. Contact him at felten@cs.princeton.edu.

Marit Hansen is Deputy Privacy and Information Commissioner of Land Schleswig-Holstein, Germany. Contact her at marit.hansen@privacyresearch.eu.

Susan Landau works in cybersecurity, privacy, and public policy. Contact her at susan.landau@privacyink.org.

Anna Slomovic is chief privacy officer at Equifax. Contact her at Anna.Slomovic@equifax.com.