ENGINEERING TRUST: PRIVACY @SXSW

By James A. Baril, VOX Global Senior Vice President

EXECUTIVE SUMMARY

Over the course of five days in March, key stakeholders descended on Austin, Texas, for SXSW Interactive 2016 to discuss technology trends and innovations. How those trends impact the privacy debate going forward was one of the underlying themes of the conference.

While there were a number of innovations and technologies discussed — virtual reality, drones, machine learning, big data, etc. — there were two key technologies that much of the privacy discussions focused on:

1. Internet of Things (IoT) – how do you manage privacy when your fridge is talking to your electric panel, which is talking to the grid – and with none of these managed by the same company? How do consumers navigate who owns their data for their connected devices, and how is that data managed when they sell their device?
2. Machine Learning/Algorithmic Inferences – Companies are now learning more about consumers from algorithmic inferences versus the data they collect. What are the privacy implications if those companies assume certain information – like race or sexual identity? What if the algorithms are wrong, and what are the privacy risks?

The following report goes into detail about the overall privacy conversation at SXSW this year among these stakeholders, including: chief privacy officers, other privacy professionals, 3rd party organizations, reporters and policymakers – including President Obama.

While the format of SXSW resembles a number of smaller conversations (in the forms of panels and speeches) on various topics, the following outlines the broader summary of the collective conversation. It is not intended to capture every word of every panel, but to provide the holistic top-line insights.

For many companies, this information may not be new – or, for some, may not be completely relevant. However, if unable to read much else, the following are the key considerations for privacy organizations:

- Consent and control have always been foundational to most privacy programs. How you manage these programs with these new technologies will become more difficult, and require creativity.
- “Biased algorithms,” “planned obsolescence,” “downstream security” and “data portability” are topics that will get even more attention from 3rd parties, regulators and the media in the near future. How will your products and services fare if put to scrutiny against these issues?

- A significant number of questions posed were more ethical than technical in nature. As business cases make their way through privacy teams, those teams should be sure to ask both: “should we do this?” as well as “how should we enable this, while still being a responsible organization?”
- Policymakers appear willing to give (most) companies a bit of latitude to innovate, as long as they use “reasonable security” and act responsibly (see here for the FTC’s *Guide for Business*). But, they warn that one incident could force their hand.

OVERVIEW

In September 2014, during a visit to an arts festival in Brooklyn, several hundred New Yorkers voluntarily gave their social security number, birthdate and other sensitive information to a complete stranger … in exchange for a cookie.

According to several panel discussions at SXSW, this social experiment demonstrated three common themes about the privacy discussions in recent several years:

1. Consumers are willing to give up a bit of privacy in exchange for something that they see as having value – although, they may not understand the true value of that information.

2. Often, it is the simple act that brings the greater threat to privacy, not the orchestrated malicious attack (whether that be an action the consumer takes, an employee takes, etc.).

3. Privacy is contextual. There are cases that are more likely to make the consumer more willing to give up a bit of their privacy than others.

Just a year and a half ago, it was easy enough to simplify the issue of privacy to such an amusing narrative. However, over the course of five days at SXSW, that privacy narrative began to shift during numerous panels on privacy policy, big data, IoT, drones, encryption, corporate reputation and other topics.

If the previous narrative was about the data a consumer voluntarily consented to provide to a company, the next era of privacy will be about the data gathered based on behaviors of those around them, inferences from algorithms and, in many cases, collected without the direct consent of those tracked.

These issues will all be reviewed in the report below, based on the following categories:

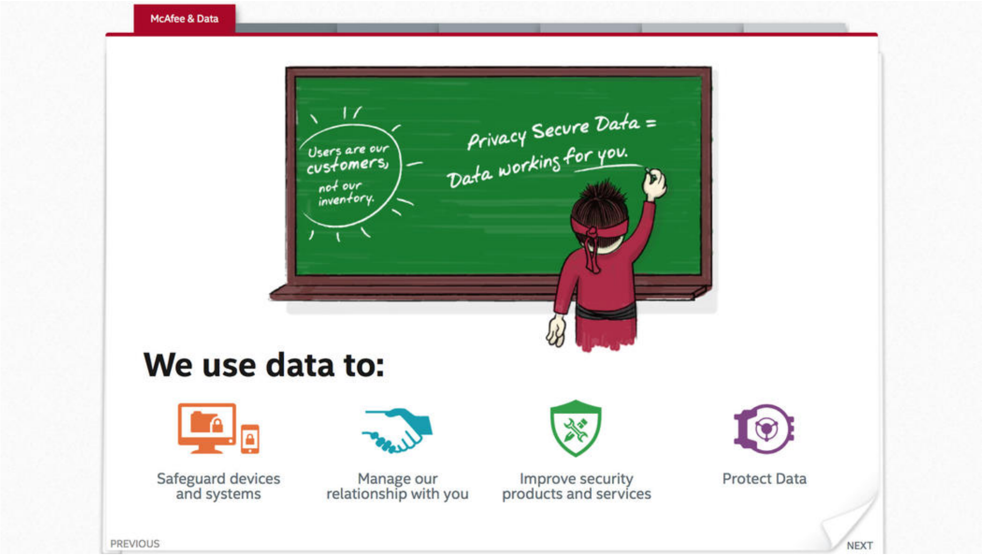
❖ Consent, Transparency and Control
❖ Device and Data Lifecycle
❖ Regulations and Ethics

CONSENT, TRANSPARENCY AND CONTROL

CONSENT

Consent is the foundation of any privacy program. As with most discussions of improving privacy programs, the discussion around consent begins with the language used in the privacy notices. I.E., “don’t write your policy just to protect your company, write it in plain English, so that your users understand how you are using and aggressively protecting their information.”

While this was a common refrain among third party organizations and regulators, several corporate executives noted that many companies have been making this transition to ***better-crafted privacy policies***. An example of going beyond and becoming more creative, is what Michelle Dennedy, chief privacy expert at Cisco, and Ruby Zefo, vice president of Law & Policy Group at Intel Corporation, created when they worked together at Intel Security (formerly McAfee): an edutainment model of reviewing the privacy policy.


McAfee & Data
Users are our customers, not our inventory.
Privacy Secure Data = Data working for you.
We use data to:
Safeguard devices and systems
Manage our relationship with you
Improve security products and services
Protect Data
PREVIOUS
NEXT

However, in other sessions, companies acknowledged that one area where they are having a challenge with consent is for their massive social experiments — where data scientist are able to use data to test a variety of social models, in many cases without the users’ knowledge that they are a part of the experiment.

In fact, the greater challenge, according to those executives, as well as other panelists and attendees, is ***consent fatigue***.

We’ve all been there — you download the app or go to the website and the first thing you see is the check box to consent to the terms and conditions. The routine has become rote: check the box without reading anything, and get straight to the content you are trying to get to, without worrying about how your data is used.

Stated another way, the longer it takes to read through that language, the longer before the user gets their cookie in exchange for their data. So, it doesn't really matter how well written the privacy policy or terms and conditions are, if the end user doesn't ever bother to read them.

Like many issues below, this is an area where there were more questions or complaints about the problem, and few solutions were put forward. Two ideas, though, were to: 1) use the principles of behavioral science (e.g., consistency, social proof, etc.) and 2) use the principles of automation design to build in better methods of communication (e.g., use Siri to communicate that terms have changed).

The above two issues — better crafted notices and consent fatigue — are issues that most companies, regardless of their industry, have been dealing with for a while and likely will continue to grapple with for the foreseeable future. However, looking ahead are two newer challenges beginning to arise for many companies in how they manage consent.

First, ***how to deal with algorithmic inferences***. The issue is becoming more common with technology companies and is best exemplified by the following article that highlighted this practice at SXSW: Facebook Ad Platform Now Guesses at Your Race Based on Your Behavior.

The practice of inferring personal information (race, ethnicity, gender, sexual orientation) based on inputted data or actions taken is raising a number of questions for companies to consider:

- If a company begins to append personal information (race, ethnicity, gender, sexual orientation, etc.) from your data, without you ever having provided it, are there special consent issues required to review?
- Are there special protections to ensure there is no bias inserted into how the company treats you?

In general, there appears to be some sense that the legalese within the privacy policy will be enough to protect the company, if drafted properly from a legal perspective. However, over time, it is anticipated that these types of practices will become more commonplace, and the inferences will be as important to providing the consumer with a services as the data they generate themselves.

- If so, are there greater risks of liability if the inferred data is used improperly, used without the customer's consent or knowledge, and/or if the data is compromised?
- What is the role of transparency? (more on that below)

Second, there are ***often no user interfaces for connected devices in the Internet of Things***, so how do consumers consent to certain practices if as Mike Hintze, chief privacy counsel at Microsoft Corporation, noted, "We won't have a privacy policy taped to a light switch."

While there were not a lot of solutions put forward, there were some ideas mentioned (the fact that these could be considered far-fetched speaks to the challenge of solving this):

- **Phone App** — program your phone to have your desired privacy settings, and as you walk through your house, a department store on the train, etc., the connected devices in those areas will read your preferences via RFID or related technology and react accordingly.
- **Home Center** — similar to your thermostat at home, have your privacy preferences for a variety of actions and the connected devices would act accordingly (e.g., like the iPhone settings interface). This would, of course, not solve for the connected devices outside of your home.
- **Register Your Device** — after the point of sale, register your device (toaster, light bulb, etc.) online, consenting to the agreements, before the device is enabled.

Finally, there's the question of whether this requires a solution or if the fact that you are purchasing these connected devices is consent in itself, and the user can go to the website if they want to learn more. The counter to this argument was that if that's the minimum bar, then transparency is important.

TRANSPARENCY AND CONTROL

According to many panelists, privacy notices are a foundation within the privacy world for a number of reasons, high among those reasons is that this is the easiest place for consumers to learn most clearly how their data is being used, stored, removed, etc., I.E., the privacy notice is the bedrock of transparency, which is why well-written privacy policies are important.

If a company were only storing information in a database and doing nothing more, then transparency for some companies could be considered to begin and end with the privacy notice.

However, the IoT, predictive/self-learning algorithms, massive online experiments and other newer trends in technology are creating greater pressure for more transparency of data privacy and security issues.

Some examples of recommendations made include:

- Companies would release their obsolescence schedule as a part of their packaging so that, for example, if a consumer is choosing between two smart TVs, and one is $50 cheaper but will no longer be supported for security patches in one year, then that fact may be enough to persuade someone's decision.
- Provide users with the ability to determine if they were a part of a company's research (e.g., online massive experiments), if the data the company has on them includes any algorithmic assumptions, etc.

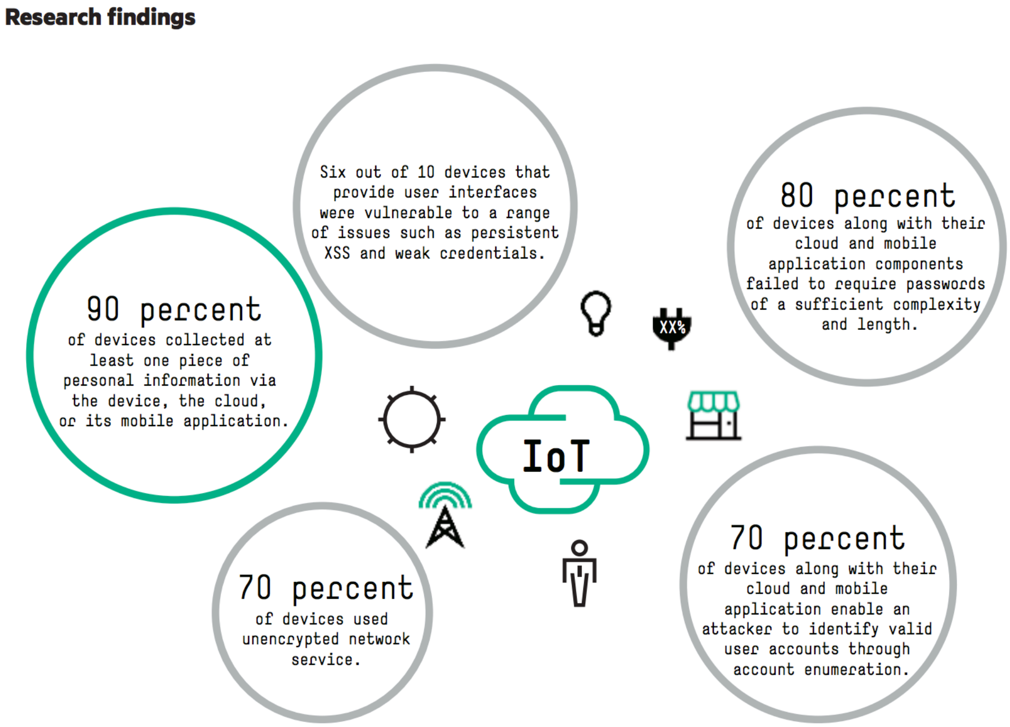
- Black box systems and algorithms for independent internal or external third parties to conduct a review of their systems and data to determine whether there are any issues of privacy, unintentional bias, etc. While not fully transparent, this was seen as a compromise between opening your algorithms to the world (and your competitors) and simply saying "trust us, nothing to see here, we're not doing anything wrong."
- Enable users to identify issues with you, and assume that their issues may be credible (e.g., if they think the data is discriminating or has errors).

And, more broadly, in terms of control, Apple's user interface on their iPhone was mentioned a number of times as being a best practice for companies to emulate. In that case, the user can go to their settings, and with the swipe of a button, turn their location data on/off, etc. Several recommendations were made for companies to consider having similar user controls on their website, within their apps, or somehow with their connected devices, etc.

DEVICE AND DATA LIFECYCLE

The data lifecycle for many companies has historically been focused on the collection, storage, use, access and disposal of data. The IoT is adding additional layers to this lifecycle.

One of the key challenges cited with these new layers to the data lifecycle issues are the vulnerability to attack with IoT devices. As FTC Commissioner Julie Brill cited, an HP study shows that while 90 percent of IoT devices are handling sensitive information, 70 percent lack encryption, and even higher percentages are at risk due to failing to require passwords or complexity.



Research findings

Six out of 10 devices that provide user interfaces were vulnerable to a range of issues such as persistent XSS and weak credentials.

80 percent of devices along with their cloud and mobile application components failed to require passwords of a sufficient complexity and length.

90 percent of devices collected at least one piece of personal information via the device, the cloud, or its mobile application.

70 percent of devices used unencrypted network service.

70 percent of devices along with their cloud and mobile application enable an attacker to identify valid user accounts through account enumeration.

In addition to this key component of data security, a number of privacy issues were raised related to data/device lifecycle. Primarily the following:

1. Obsolescence.
2. Downstream Issues.
3. Data Ownership and Portability.

Planned obsolescence was one of the most common issues raised. What happens when the connected refrigerator will only be supported with patches and other support for five years, but consumers generally own them for 13 years on average? If there is a security flaw exposed after the fact, is there an obligation to patch the issue anyway or does the obligation rest with the consumer to buy a new refrigerator after it has lost its support? What if the consumer is buying the refrigerator in year four of its lifecycle? Is there a greater obligation for transparency, so that they are aware that it may have security issues and require a new purchase shortly?

Even with devices where obsolescence has been an issue already (e.g., smart phones, software, etc.), most consumers aren't really aware of the impact of having dated equipment and software. When they are aware and plan to make a change, the cost to the consumer is generally less expensive. However, if their connected car has a lifecycle of five years, and the average consumer owns their car for six years, then the extra year would either be unprotected or would require a change in behavior of some sort. And, if that change is to buy a new car, then it's an expensive "fix."

Beyond the cost associated with the fix, other issues of consumer issues were raised to ensure customers are aware of this new issue when they buy their TVs, etc. The demographic group most often cited as being at risk are seniors, who will be more likely to be on a fixed income. For them, more frequent purchases may create problems, and they are less likely to have experience dealing with these obsolescence type of issues.

As Commissioner Brill stated, "when the connected pen becomes unsupported, then it becomes a security threat in your home."

Downstream issues were one of the more complex concerns raised. It's one thing for the toaster to connect to the refrigerator to connect to the smart phone. However, what if there's a security patch required for one device in the chain that requires a patch for the other devices connected? And, if those devices are all manufactured by different companies, then how do you coordinate? Finally, what if one of the suppliers refuses to patch the flaw for any reason?

In addition, from a privacy perspective, how does one ensure that when one device is hacked or "goes haywire" it won't impact the other devices connected to it (either from a physical safety or information security perspective)?

For companies in the mobile industry, this is not a new issue. An example that was used several times is that if there's an issue with Android software, it may often require not only Google to patch the problem, but also the phone manufacturer (e.g., Samsung) and/or the network provider (e.g., AT&T). Some third party organizations noted, though, that there have been cases where the phone that has the issue is no longer being supported, so the manufacturer and/or the network provider may decide not to make the patch even if Google does, which may still leave the vulnerability exposed. This is the type of challenge that many manufacturers and technology companies in the IoT community may not be used to dealing with, yet.

Finally, ***data ownership and portability*** continues to be a hot — but, sensitive — topic. Many consumers have become more aware of these issues as they buy connected devices that they have historically been used to owning outright.

For example, a number of farmers, who bought the John Deere smart tractors, learned that not only do they not "own" the data generated on the tractor, but due to the nature of the new type of tractor, they are unable to "tinker" and "fix" things on the tractor, as they used to do with their previous tractors.

This issue is not expected to raise many concerns in the short term, because consumers will have options between connected devices and non-connected devices (e.g., buying a light bulb at Home Depot gives you an option between connected and non-connected, with the latter being more common presently). However, there will be a point in the not-too-distant future in which the only options will likely be connected.

From a privacy perspective, the concern for the end user may be the data that a company or organization may be able to get about a person's habits, schedules, etc., which could be even more problematic if the data were stolen. While most privacy experts advise to collect only the data you really need to provide a service, several companies noted that their data team will often say something to the effect of "we don't know why we'll need the data … yet. So, let's just go ahead and collect it."

Similarly, there are questions about what happens if a company releases data in a way that gives others outside of the company a view of the consumer's private life, which could have a negative impact on that user (e.g., outing a teen unintentionally with their parents, classmates or the public).

An example given in one panel was the data from drones over farm land (in this case in Africa), in which the company thought that the best thing to do was to release the data identified by the drone about crops, irrigation, soil, etc., publicly – with the positive intent of allowing data scientist to find a way to help farmers improve their methods. However, if the data includes information about mineral deposits that

the current owner of the land is not aware of, then someone who reads through the open data would be able to take advantage of the situation (i.e., offer the owner bottom dollar given the deposits, but an amount that would seem fair if the owner were not aware of the deposits).

In terms of portability, there are a number of issues associated with this. The most commonly cited issues at SXSW were related to what happens when the consumer sells their connected device.

For example, if you're selling your car, how do you ensure that the data the car manufacturer and/or any relevant technology companies have been collecting about you, during your time owning the car (locations travelled, average speed, etc.), are not released to the new owner.

Similarly, from a customer intuitive perspective, will there be a way for the data to be transferred to the new car they have purchased, so that their settings and data will carry over seamlessly? And, for closing on the sale of the home, does this add new layers of paperwork? What about cases where the device (house, car, etc.) are being rented?

REGULATION AND ETHICS

There were a number of policymakers (federal agencies, Congress, state and local governments, etc.) and their staff on the ground in Austin for SXSW this year. Most notable, of course, was President Obama who delivered the first keynote at SXSW by a sitting president.

Whether conservative or liberal, nearly all appeared to agree that government regulations to prescribe how companies should deal with most of these issues of privacy would likely do more to stifle innovation than protect consumers.

With that said, consumer protection was a key theme, and there were a number policy issues discussed:

- Tort as a solution
- Baseline of protections
- FCRA-lite for data brokers
- Omnibus privacy bill
- Reasonable security
- Right to be forgotten
- And, of course, encryption

Tort was an idea that came up several times in different panels — *and by policymakers on both sides of the aisle*. Essentially, the idea they floated was that instead of coming up with regulations that will define how companies should behave in advance, Congress should consider a policy to increase the damages and risk of tort liability, so that companies are punished for bad behavior after the fact.

Baseline protections were an area where, as expected, there was little agreement among panelists and attendees. An anecdote shared by Andrea Matwyshyn, professor of law at Northeastern University, was that there was a time when a landlord would turn off the heat for a tenant in the middle of winter, and so local governments stepped in and passed a minimum standard of services that they would have to provide at all times. Similarly, she and others there proposed having some baseline level of protections.

One area that Ashkan Soltani, former chief technologist of the FTC, suggested that there appears to be some consensus on protections based on race, gender and age — for data practices that impact housing, credit and employment.

Sarah Goyette, senior attorney for Privacy & Security at Intel Corporation, noted that organizations within the broader industry are beginning to develop many of these standards. An example she provided is that AAA is working on developing a "bill of rights for consumer data being collected in cars."

Data brokers continued to come up as one of the challenges to privacy going forward — often cited by regulators and consumer advocacy groups as bad actors in this space. Commissioner Brill recommended considering an ***FCRA-lite for data brokers***, to give consumers a chance to know what is in their databases and the opportunity to delete, when appropriate.

Several individuals called for a ***broader privacy bill*** to address how to treat all data in the new age of technology — noting that the last real legislation to deal with these issues came out before the Internet. The challenges to this proposal were the fact that, in many cases, they wanted to remove barriers on certain data to loosen privacy restrictions (namely HIPAA with healthcare data), while wanting to increase it significantly on other areas where those industries and companies would likely be opposed. For these reasons, it is anticipated that any broad privacy bill would have difficulty getting through Congress.

In a couple of sessions, there was some discussion about the European Union's GDPR and, specifically, the ***right to be forgotten***. As both Keith Enright, legal director for Privacy at Google, and Mike Hintz explained through some history about the cases: the previous rule was better described as the "right to be harder to find," where the new rules are calling for something closer to the "right to be deleted." Beyond the difficulties in accomplishing this through current technology, there are also questions of ethics and public benefit involved (e.g., what if the data being requested deleted is important for public safety). While this rule will create challenges for companies where the EU's data rules apply, there does not appear to be much discussion about these rules being adopted more broadly outside of the EU.

And, of course, the hot topic of SXSW was the matter of ***encryption***. President Obama had quite a bit to say about the matter of encryption, including the following:

> ***"I am not interested in overthrowing the values that have made us a great nation for expediency."***
>
> ***"The question we now have to ask is if technologically it is possible to make an impenetrable device or system where the encryption is so strong that there's no key or no door at all. How do we apprehend the child pornographer? How do we solve or disrupt a terrorist plot?"***
>
> ***"This notion that somehow our data is different and can be walled off from those other trade-offs we make, I believe, is incorrect."***

The president's remarks on the opening day kicked off the privacy discussion at SXSW. For most of the other panelists and attendees, while his remarks were received very positively. However, on the specific issue of encryption, most were of the "agree to disagree" mindset. Regardless of the outcome of the Apple vs. the FBI, most agree that this will be a prolonged fight toward enabling greater encryption.

As Erin Egan, chief privacy officer for Facebook, summarized, "people are demanding (unbreakable encryption). If we are asking U.S. companies to have a lower level security, then customers will go to international companies."

ETHICS

> ***"Ethics is the next big discussion in privacy — not just about compliance, but also should we do certain things? Some harms may not be as obvious. We need philosophers on staff nowadays."***
>
> ***-- Mike Hintze, Chief Privacy Counsel at Microsoft Corporation.***

"Ethics" was the underlying trend of many privacy discussions at SXSW. As was apparent from many of the above issues, many questions raised from panelists and attendees required more considerations for the ethicist than the technologist.

Some examples of the types of questions being posed include:

- Is it price discrimination if a retail store is able to increase the price for a customer if they live further from a competitor's store?
- Would it be OK for Apple or Google to change the map directions to take you by one of their advertiser's instead of via the most efficient route?
- If I am building the algorithm for a search result where the user asks, "who is the greatest president ever?," then there may be no truly agnostic answer, so should the answer be geared to the psychographic profile of the person asking (assuming our data has that), be tied to the zip code where they are asking from (i.e., a Republican area or Democrat area), etc.?
- If we are calling for more transparency for body cam data from police, in which cases could there be too much transparency? E.G., if a police officer responds to a domestic disturbance, and their camera catches the faces of the abused spouse and/or kids, should that be made publicly available?
- Is there a risk of "mind control" or "mood control" from massive online experiments, such as the famous example from Facebook?

However, the most pressing current ethical issue was related to "***biased algorithms***." These are cases where the algorithms being designed have some form of bias written into the code – more often than not, inadvertently.

Some examples included:

- According to a study from Carnegie Mellon, ads on Google for jobs where the income was $200k+ were more often shown to men than women.
- An autocomplete feature (powered through machine learning) on some search engines completed phrases for the beginning of certain search terms with highly negative and biased results, e.g., "Are Transgenders" would autocomplete with "going to hell."
- During the 2012 election, people doing a deeper search on candidates' positions on the issues for Obama (e.g., healthcare) were shown more information about Obama's positions tied to those issues, but Romney searches were not returning results related to his positions. According to Google, this was because of user preferences (i.e., Obama voters wanted more information, but Romney voters didn't).

These, and other examples, raised key ethical questions:

- Who gets to decide what makes up functions of the algorithm (e.g., when IBM created a "terrorist score" to determine the likelihood that someone was a terrorist for a client, who was in the room to determine what should be included in that type of algorithm)?

- What level of error do we accept when it comes to these predictive algorithms (i.e. if an algorithm guesses incorrectly that you're a criminal)?
- Search engines are supposed to be acting as a neutral party, but do they have a role to identify cases where their ad services are demonstrating bias and correct for it – even if the bias may be on the ad purchaser end?

As mentioned above, one of the key recommendations for resolving this issue is to "blackbox" the algorithms, enabling independent third parties to come in and test it.

Julia Angwin, senior reporter for ProPublica, also best summarized another ethical privacy issue that arose during numerous sessions: "***privacy is becoming a luxury good***."

The criticism is that the cheaper option is often the one that comes at the cost of more data from the consumer. So, it will become more common that if you want to opt out of giving away your data then it will cost increasingly more over time.

A number of individuals noted that there may be more of a need for ***industry-wide ethical guidelines for data scientist*** – although, as some noted, if there was any industry that would rebel against guidelines, it would be the data industry.

Michelle Dennedy, chief privacy expert at Cisco, did note that while the IEEE is working on ethics for data, it's up to the big companies to come up with the tools and standards for the start-ups and smaller companies, who don't have the luxury of thinking about these issues.

CONCLUSION

Building a consensus around the solutions for many of the above challenges will be difficult. Trevor Hughes, president and CEO of International Association of Privacy Professionals, concluded the session he moderated with the following two predictions for the future:

1. These issues are places of ferocious competition. Companies will continue to compete on reputation and trust.
2. With growing complexity in technology, we will be relying on the privacy staffs of the organizations.

Our recommendations to organizations are:

- Don't wait to consider these issues. Implementing Privacy by Design going forward will mean thinking about these types of issues at the onset of any project.

- Implementing Security by Design is critical to ensure that as your organization gathers more sensitive information, you are protecting it to the fullest extent possible – including information that the consumer may not be aware that you have about them.
- Be prepared to handle the breach. As Corey Ealons, senior vice president at VOX Global, and Sterling Miller, senior counsel at Hilgers Graben, outlined during their "You've Been Hacked, Now What?" session, the best way to handle the breach is to have your plan in place in advance, and completed numerous tabletop exercises that include all parties that will be around the table (executive team, data forensics team, IS/IT, HR…and, of course, communications and legal teams).

At the end of the day, the privacy organization is increasingly responsible for that company's reputation. Building and maintaining that reputation takes a strong and well-coordinated team who is focused on executing their privacy mission statement every day. As Keith Enright, legal director of Privacy at Google, closed in one session:

> ***"We are trying to engineer trust."***