

MEMORANDUM

TO: ESPC

FROM: D. Reed Freeman

DATE: December 10, 2012 FILE: 68223-0000001

RE: FTC Workshop: The Big Picture: Comprehensive Data Collection

We write to provide you with a summary of the Federal Trade Commission (“FTC” or “Commission”) workshop called “The Big Picture: Comprehensive Data Collection” held December 6, 2012. “Comprehensive data collection” is the collection of all or most of the online activities of users across multiple locations. The FTC indicated in its March 2012 privacy report that it would hold this workshop to explore **the privacy implications of the data collection practices of large platform providers** such as ISPs, operating systems, browsers, search engines, and social media platforms, particularly with respect to consumer data, and their current and expected uses of such data.

The workshop consisted of a presentation on the technological landscape of data collection and three panel discussions. FTC staff moderated the three panel discussions. Panel members consisted of industry representatives, privacy advocacy groups, and academics.

EXECUTIVE SUMMARY

In her introductory remarks, Commissioner Brill:

- ◆ Noted that data collection can provide important benefits to consumers, but not in all contexts;
- ◆ Indicated her concerns that:
 - Companies are gathering enough data about consumers to create “databases of ruin”—e.g., comprehensive profiles that include sensitive medical, financial, and other types of information; and
 - The “consistent with the context” framework might not be effective, particularly if data collection is taking place in the background.

In remarks prefacing the afternoon sessions, Commissioner Ohlhausen:

- ◆ Emphasized balancing consumer privacy with innovative uses of technology and data so consumers benefit without being harmed;
- ◆ Suggested that context shaped by user expectations will be an effective guide to the use of data; and

- ◆ Embraced self-regulation to avoid stifling innovation and creativity in a dynamic system in which collecting consumer information is valid and legitimate.

In closing remarks, associate director of the FTC’s Division of Privacy and Identity Protection Maneesha Mithal:

- ◆ Stated that Facebook and Google were invited to participate but chose not to.

These were the only substantive pronouncements by the FTC at the conference. However, questions asked by FTC staff moderators suggested they think comprehensive data collection could be inherently harmful, and that a “consistent with the context” framework may not work for comprehensive data collection. No questions were answered by FTC staff.

In all three discussions, panelists disagreed about:

- ◆ ***The scope and nature of harms resulting from comprehensive data collection.***
 - Whether data collection itself is a harm, or whether harm only flows from misuse of the data.
- ◆ ***The lack of consumer understanding of data collection.***
 - Do consumers understand or care how much data is being collected?
 - Can the market provide consumers with choices about the privacy of their data?
- ◆ ***Whether companies are willing and able to compete over privacy, or if legislative or regulatory intervention is necessary.***

However, there was general agreement that:

- ◆ ***Data collection policies should be technology neutral.***
- ◆ ***There are real, tangible benefits resulting from the large-scale collection of data from online users*** including free online content and services.
- ◆ ***Certain uses of data are improper***, such as discrimination and targeting individuals based on sensitive information.
- ◆ ***Competition and innovation in privacy should be embraced and pursued.***

The FTC is accepting comments on the workshop but has not yet indicated a deadline for comments. The FTC will use this information to determine if its guidance on comprehensive data collection needs to be revised, but has not indicated what form those revisions may take.

PRESENTATION: THE TECHNOLOGICAL LANDSCAPE OF COMPREHENSIVE DATA COLLECTION

Professor Dan Wallach of Rice University began the workshop with an overview of how consumers’ activities can be tracked and the limits of that technology. In addition to the basic mechanics of first party and third party tracking, Professor Wallach focused on these key issues:

1. *Tracking is ubiquitous.*
2. *Data collection is part of a tradeoff between consumers and business.*
 - a. We “pay” for a lot of what we do on the internet by giving away a little bit of privacy.
3. *There are multiple uses of data.*
 - a. Including forensics, advertising, and marketing.
 - b. ISPs use Deep Packet Inspection (DPI) and active engagement such as blacklisting websites and controlling data traffic flows.
4. *Regulation should be focused on information and thus technology-neutral.*
5. *Relying on individual websites to engage consumers to take actions about tracking and cookies does not work.*

PANEL 1: BENEFITS AND RISKS OF COMPREHENSIVE DATA COLLECTION

This panel considered both the benefits to consumers from the technologies that allow comprehensive data collection as well as the various privacy concerns associated with the ability to track all, or virtually all, of a consumer’s online activities. The panel focused in particular on whether comprehensive data collection is inherently harmful or harmful only if misused.

The panelists debated several topics:

Panelists were divided about comprehensive data collection

- ◆ Lee Tien of the Electronic Frontier Foundation believes companies should not collect large amounts of customer data without consumers knowing and understanding what is being done.
 - Ashkan Soltani agrees, but suggests that companies can—and therefore should—get the information they want from consumers by asking for it openly.
- ◆ Others suggested the focus should be on harms from inappropriate uses of data.
 - Panelists including Michael Altschul of CTIA—The Wireless Association and Professor Howard Beales of George Washington University cited benefits.
 - Data can be used to prevent or identify credit card fraud, identify potential flu outbreaks, plan traffic routing, etc.
 - Large scale data collection also makes information “visible”—analysts can see patterns and trends that nobody knew existed.

- Some benefits therefore aren't known until after data is collected.
 - Data is essential to innovation, new products and improved performance.
 - Mr. Soltani agreed that the focus should be on harms, including the risk of breach.
 - Professor Beales stated that marketing is not an improper use of data.
- ◆ Professor Neil Richards of Washington University in St. Louis noted that data collection is dangerous if misused.
 - He identified three types of problems with comprehensive data collection:
 - **Information asymmetry**, e.g., the fact that companies know a tremendous amount about their customers and customers know relatively little;
 - **Intellectual privacy**, e.g., being able to read, think, ask questions, etc., without being tracked;
 - Private sector data collection can **enable large-scale government surveillance**.
- ◆ David Lincicum, the FTC moderator, asked specifically about the continued relevance of distinguishing first-party and third-party relationships in the context of comprehensive data collection, noting that large platform providers such as ISPs know *a lot* about consumers.
 - Panelists disagreed about whether there is something inherently different about comprehensive data collection.

Data collection policy should be technology-neutral

- ◆ Technology changes too fast and there are too many ways to accumulate information.

How consumers go online has changed

- ◆ Professor Beales presented research showing that consumers use multiple devices, multiple networks, multiple locations, and multiple browsers and services.

Consumer choice

- ◆ Mr. Altschul, along with Markham Erickson, General Counsel of The Internet Association, emphasized the progress of privacy by design.
- ◆ Mr. Altschul also noted that industry now offers more choices.
 - For instance, the Firefox and Chrome browsers provide different browsing experiences.
- ◆ Professor Richards stated that the opacity of consumer data collection makes it difficult for consumers to make meaningful choices.
 - Mr. Tien and Mr. Soltani also doubt the market can provide choice.
 - Consumers don't understand how data is being collected.
- ◆ Mr. Altschul disputed customer indifference, noting consumers respond when they are bothered by particular uses of data.
 - When Facebook changed its Terms of Service they faced a backlash.

- ◆ Professor Beales suggested that if privacy advocates are not happy with the current regime, they should try to change consumer behavior.

PANEL 2: CONSUMER ATTITUDES ABOUT AND CHOICE WITH RESPECT TO COMPREHENSIVE DATA COLLECTION

This panel examined consumer attitudes and knowledge about comprehensive data collection and the role of consumer choice and transparency. The panelists primarily focused on their own attitudes about comprehensive data collection, leading to overlap with the themes from the first panel. Two panelists summarized research about consumer attitudes.

The discussion emphasized several themes surrounding the general topic of comprehensive data collection and consumer behavior:

Whether notice and choice is a viable model

- ◆ Panelists generally agreed that the “consistent with the context” standard is not useful.
 - Moderator Paul Ohm of the FTC suggested that the problem with the standard is that comprehensive data collection takes place in the background.
 - No one-on-one relationship between consumer and provider.
 - Microsoft Associate GC Michael Hintze suggested that there is value in the standard, but it can’t be the only answer.
 - Stu Ingis, Counsel for the Digital Advertising Alliance, suggested a focus on uses and harms, not context.
- ◆ Panelists disputed whether industry efforts to inform consumers are working.
 - Mr. Ohm seemed to doubt that notice can work.
 - Professor Lorrie Faith Cranor of Carnegie Mellon University stated it is inadequate unless consumers have actual choices.
 - Mr. Hintze of Microsoft observed that it is difficult to design policies that will be useful for consumers because there is so much information and data collected.
- ◆ Some panelists believe that consumers are not making meaningful choices.
 - Consumer choice cannot work because there cannot be a market for choice in privacy if consumers don’t know to demand it.
 - Mr. Ingis conceded that there needs to be an increase in transparency and dialogue, and that industry is moving in that direction.
- ◆ Others noted that consumers react and make it clear when data uses are not consistent with their expectations.
 - Mr. Hintze stated that there is a gap between what Microsoft could collect and what it does collect.
 - That gap is driven by consumer expectations, which suggests that consumers effectively are making meaningful choices.
 - Mr. Ingis stated this indicates that the market takes care of ensuring that collection and notice is adequate.

Research on consumer understanding and behavior

- ◆ Professor Cranor summarized her research:
 - Consumers do not understand what data is collected beyond the information they knowingly give to websites.
 - When researchers explain the *scope* of data collection to consumers, they think it is “creepy.”
 - But when researchers explain the *value* of data collection, consumers understand that.
- ◆ Professor Alessandro Acquisti of Carnegie Mellon explained that his research suggests that the more control consumers have over their privacy online, the more inclined they are to disclose sensitive information.
- ◆ Mr. Ingis stated that the right question is whether consumers benefit, not whether they understand.

Whether comprehensive data collection is harmful

- ◆ Mr. Ingis thinks consumers get a good deal.
 - Businesses can give things away for free by collecting anonymous information about consumers.
- ◆ Christopher Calabrese of the ACLU thinks companies will discriminate and consumers will never know, so they will be unable to prevent it.

Legislation versus self-regulation

- ◆ Professor Cranor and Mr. Calabrese were both skeptical of guidelines and believe that there needs to be legal weight behind standards.
- ◆ Mr. Ingis suggested the market should pick winners and losers, not the regulatory regime.
- ◆ Mr. Hintze emphasized that competition has a role and that Microsoft is trying to compete on privacy.

PANEL 3: THE FUTURE OF COMPREHENSIVE DATA COLLECTION

This panel partially discussed potential next steps for industry and policy makers in the area of comprehensive data collection, but primarily engaged in a debate similar to the previous panels about the relationship between data collection and harm. They also focused on the potential for the market to supply real choice for consumers. Christopher N. Olsen of the FTC co-moderated. Through the questions he posed, he indicated a general skepticism about large scale data collection.

Data collection and harm

- ◆ Some panelists argued that information asymmetry can result in harm.
 - Price and other kinds of discrimination.
- ◆ Mr. Olsen asked whether comprehensive data collection should be allowed at all, and suggested that it could be a harm in its own right.

- Panelists including Alissa Cooper of the Center for Democracy and Technology agreed, because big data “surveillance” is *different* from sector-specific activities like OBA.
 - Invokes notions of risk and creepiness.
- Chris Hoofnagle, Director of Information Privacy Programs at Berkeley Law, stated that giving away content and services “free” masks the fact that privacy is part of the price.
- ◆ Jim Halpert, General Counsel to Internet Commerce, does not think “comprehensive” data exists.
 - Too many different platforms and mechanisms of accessing online content for one entity to collect comprehensive information.
 - For purposes of policy making, it’s too difficult to try to define what amount of data collection is “comprehensive.”
- ◆ Panelists largely agreed that there is some data we want to keep private.
 - Mr. Halpert suggested a focus on avoiding unseemly discrimination and behavior.
 - For secondary uses, transparency and control principles should apply.
- ◆ Professor Randal Picker of the University of Chicago warned that regulation could be more harmful than current data collection policies.
 - Tom Lenard of the Technology Policy Institute emphasized that the collection of information is central to how the internet economy works today.

Online versus physical world

- ◆ Mr. Halpert emphasized that traditional invasions of privacy—intrusions upon seclusion—are different.
 - Online, consumers get free content and are exposed to data collection and tracking in what are essentially public areas.
- ◆ Other panelists suggested that models of physical interaction do not map online well.
 - Computers know significantly more about us than people we interact with on the street.

Consumers don’t understand how data is collected and used

- ◆ Panelists largely agreed, but disagreed about implications.
 - Sid Stamm of Mozilla stated that notice doesn’t work because users are focused on one thing at a time (e.g. shopping, reading the news, etc.)
 - Users don’t understand a comprehensive data collection notice.
 - Mr. Halpert suggested that companies think their consumers understand and care about their privacy.
 - They don’t abuse data collection power because they don’t want to lose customers.
 - Consumers are not powerless.