No. 15-1441

IN THE UNITED STATES COURT OF APPEALS
FOR THE THIRD CIRCUIT

IN RE NICKELODEON CONSUMER PRIVACY LITIGATION

On Appeal from the United States District Court
For the District of New Jersey
Case No. 2:12-cv-07829
The Hon. Stanley R. Chesler

BRIEF OF *AMICUS CURIAE* ELECTRONIC PRIVACY INFORMATION CENTER (EPIC) IN SUPPORT OF APPELLANTS

Marc Rotenberg
Counsel of Record
Alan Butler
Julia Horwitz
John Tran
Electronic Privacy Information Center
1718 Connecticut Avenue, N.W.
Suite 200
Washington, DC 20009
(202) 483-1140

May 4, 2015

CORPORATE DISCLOSURE STATEMENT

Pursuant to Federal Rule of Appellate Procedure 26.1 and 29(c), *amicus curiae* Electronic Privacy Information Center ("EPIC") certifies that it is a District of Columbia corporation with no parent corporation. No publicly held company owns 10% or more of EPIC stock.

TABLE OF CONTENTS

TABLE OF AUTHORITIES

STATUTES

OTHER AUTHORITIES

INTEREST OF THE AMICUS

The Electronic Privacy Information Center ("EPIC") is a public interest research center in Washington, D.C., established in 1994 to focus public attention on emerging civil liberties issues and to protect privacy, the First Amendment, and other Constitutional values.[1] EPIC has written extensively on the privacy implications of the collection, storage, and disclosure of sensitive consumer information.

EPIC routinely participates as *amicus curiae* before federal and state courts in cases concerning consumer privacy rights. *See*, *e.g.*, *First Am. Financial Corp. v. Edwards*, 132 S. Ct. 2536 (2012) (defending consumer standing claims); *Sorrell v. IMS Health Inc.*, 131 S. Ct. 2653 (2011) (defending state prescription privacy law against commercial speech challenge); *Fraley v. Facebook*, No. 13-16918 (9th Cir. Feb. 20, 2014) (defending consumer interests in a class action privacy settlement)*; Joffe v. Google*, 746 F.3d 920 (9th Cir. 2013) (defending Internet users against unlawful interception of private wi-fi communications); *Harris v. Blockbuster, Inc.*, No. 09-10420 (5th Cir. Nov. 9, 2009) (preserving privacy safeguards for video rental records).

[1] In accordance with Rule 29, the undersigned states that no monetary contributions were made for the preparation or submission of this brief. This brief was not authored, in whole or in part, by counsel for a party.

SUMMARY OF THE ARGUMENT

The definition of "personally identifiable information" set out in the Video Privacy Protection Act is purposefully broad to ensure that the underlying intent of the Act—to safeguard personal information against unlawful disclosure— is preserved as technology evolves.

The lower court wrongfully concluded that unique, persistent identifiers and other "transactional information" obtained from a consumer by an entity subject to the Act do not constitute personally identifiable information. This result is directly contrary to the language of the statute, the intent of Congress, and the express statement of the Act's sponsor. If unique Internet identifiers are excluded from the definition of personally identifiable information, then the VPPA will cease to have any meaningful application to Internet providers of video services. The court's conclusion was also fundamentally flawed because it failed to recognize the ability of Google and other marketing firms to identify users based on their browsing data.

ARGUMENT

The Video Privacy Protection Act of 1988 (VPPA), 18 U.S.C. § 2710 (2014), is a quintessential federal privacy statute—the law prohibits the disclosure of "personally identifiable information concerning any customer" of a video service provider (or provide of other "similar audio visual materials"), except under certain limited circumstances, and provides customers with an opportunity to

seek relief if their rights were violated. *Id*. § 2710(b)(1). The statute specifies that "the term 'personally identifiable information' *includes* information which identifies a person as having requested or obtained specific video materials or services." *Id*. § 2710(a)(3) (emphasis added). Congress intended this definition to "establish a minimum, but not exclusive, definition of personally identifiable information." S. Rep. No. 100-599, at 12 (1988), *reprinted in* 1988 U.S.C.C.A.N. 4342. As Senator Patrick J. Leahy, the bill's original sponsor explained in his introductory floor statement, "A person maintains a privacy interest in the transactional information about his or her personal activities. The disclosure of this information should only be permissible under well-defined circumstances." 134 Cong. Rec. 10260 (1988) [hereinafter Leahy Floor Statement]. The data before the lower court is precisely the type of "transactional information" that the VPPA was enacted to protect.

As Professor Jerry Kang explains in his analysis of the collection and use of personally identifiable information by Internet firms, definition of PII is not limited to names and addresses; the term "describes a relationship between the information and a person, namely that the information—whether sensitive or trivial—is somehow identifiable to an individual." Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 Stan. L. Rev. 1193, 1207 (1998). Information can be "identifiable" to a person in one of three ways: (1) authorship, (2) description, or

(3) instrumental mapping. *Id*. Information that an individual creates and claims authorship over is identifiable, as is information that "could describe the individual in some manner" including characteristics like age and sex; and persistent identifiers (like social security numbers, usernames, Internet Protocol addresses, and unique device addresses) that can be used to map an individual's interactions with an institution are also identifiable. *Id*.

It is not reasonable to conclude, as the lower court did in this case, that information is not "personally identifiable" because it consists only of "anonymous user IDs, gender and age, or data about a user's computer." *In re Nickelodeon Consumer Privacy Litig.*, No. 12-7829, slip op. at 5 (D.N.J. Jan. 20, 2015) [hereinafter *Nickelodeon II*]. A court must first consider whether the information disclosed is linked to a "specific [video] transaction," such as information about "whether a person patronized a [provider] at a particular time or on a particular date." S. Rep. No. 100-599, at 12. If the court finds that the information disclosed links a particular customer to a specific transaction, *see id*. at 7 (noting that PII is "information that links the customer or patron to particular materials or services"), then that information should be considered PII unless it was sufficiently anonymized to remove all potentially identifying information. But even information that may seem anonymous when it is disclosed, could be potentially linked to a known individual in the future. That is one the reasons why PII has

been routinely defined in federal privacy laws to include information that both identifies or could identify an actual individual. *See*, *e.g.*, California Online Privacy Protection Act, Cal. Bus. & Prof. Code §§ 22575–22579 (2014) (including information that "permits the physical or online contacting of a specific individual"); E-Government Act of 2002, 44 U.S.C. § 3501 *et seq.* (2014) (including both "direct" and "indirect" identifiers); Children's Online Privacy Protection Act of 1998, 15 U.S.C. §§ 6501–6506 (2014) (including "persistent identifiers that can be used to recognize a user over time and across different Web sites or online services").

The VPPA and other federal privacy laws do not limit the definition of PII to names and addresses; the laws "define PII in a much broader way" and "account for the possibility of deductive disclosure . . . and do not lay down a list of informational attributes that constitute PII." Arvind Narayanan & Vitaly Shmatikov, *Privacy and Security: Myths and Fallacies of "Personally Identifiable Information"*, 53 Comm. ACM 24, 24–25 (2010).

The lower court's decision is clearly contrary to the intent of Congress and the purpose of the Act (to safeguard privacy) because the recipient of the customer's data in this case is Google, the single biggest aggregator of personal information in the world. The court, when discussing the potential for users to be identified by Google, appears to be unaware that Google routinely records not only

the search queries of Internet users but also tracks users' activities based on their IP address. Peter Fleischer & Nicole Wong, *Taking Steps to Further Improve Our Privacy Practices*, Google Official Blog (Mar. 14, 2014).[2] The court found that nothing in the allegation supports the inference that "Google can identify the individual Plaintiffs in this case, as opposed to identifying people generally." *Nickelodeon II*, slip op. at 6. The court likens this question to whether the provider has disclosed "a unique identifier and a correlated look-up table." *Id*. (internal citation omitted)).

But the court fails to recognize that Google *is* the "look-up table." Google's entire business model is premised on delivering targeted ads to users based on their browsing history, and this is only possible by *identifying* and tracking their online browsing habits. It is nonsensical to say that Google is unable to identify a user based on a combination of IP address, MAC address, and other browser cookie data; that is precisely what Google does best. It would be like concluding the company that produces the phone book is unable to deduce the identity of an individual based on their telephone number.

I. The Term "Personally Identifiable Information" Is Broadly Construed Under Federal and State Privacy Laws

[2] http://googleblog.blogspot.com/2007/03/taking-steps-to-further-improve-our.html.

The concept of PII is the key to all modern privacy laws, regulations, and industry standards. Indeed, under many privacy regimes, PII is the jurisdictional or substantive trigger. *See, e.g.,* Nat'l Conf. State Legislatures, *Security Breach Notification Laws* (2015) (listing data breach notification laws triggered by breach of PII enacted in forty-seven states, the District of Columbia, Guam, Puerto Rico, and the Virgin Islands).[3] *See also* Christopher Wolf, *Envisioning Privacy in the World of Big Data*, *in Privacy in the Modern Age: The Search for Solutions* 204, 207 (Marc Rotenberg, Julia Horwitz, & Jeramie Scott eds., 2015) ("Personally identifiable information ('PII') is one of the central concepts in information privacy regulation.")

The Government Accountability Office ("GAO") has also provided one of the most detailed and comprehensive definitions of PII in its report evaluating federal government privacy laws. *See* U.S. Gov't Accountability Office, *GAO-08-536, Privacy: Alternatives Exist for Enhancing Protection of Personally Identifiable Information* (May 2008).[4] The GAO adopted a broad, comprehensive definition of PII:

> [PII is] any information about an individual maintained by an agency, including (1) any information that can be used to distinguish or trace an individual's identity, such as name, Social Security number, date and place of birth, mother's maiden name, or biometric records; and

[3] http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx.
[4] http://www.gao.gov/new.items/d08536.pdf.

> (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information.

Id. at 1 n.1 (emphasis added).

The National Institute of Standards and Technology ("NIST") adopted this definition and issued guidance explaining its important terms. *See* Nat'l Inst. of Stds. & Tech., U.S. Dep't of Commerce, *Special Pub. 800-122, Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)* (April 2010).[5] To "distinguish" an individual, "is to identify an individual." *Id.* at 2-1. A "name, passport number, social security number or biometric data," for example, are sufficient to distinguish a person. *Id*. Although a list containing only credit scores without additional identifying information would be insufficient to distinguish a person, if that list were supplemented by data such as age, address, and gender, "it is probable that this additional information would render the individuals identifiable." *Id.* at 2-1 n.18. A person is "traced" if the information is sufficient "to make a determination about a specific aspect" of the person's activities or status. *Id.* Thus, for example, "an audit log containing records of user actions" is sufficient to trace a person's activities. *Id.*

"Linked" information is information "about or related to an individual that is logically associated with other information about the individual." *Id.* For example, two databases containing different PII elements that reside "on the same system or

[5] http://csrc.nist.gov/publications/nistpubs/800-122/sp800-122.pdf.

a closely related system and lacks security controls to effectively segregate the data" are considered linked. *Id.* In contrast, "linkable" information is a broader category of data. It is information "about or related to a person for which there is a *possibility* of logical association" with other data about that person. *Id* (emphasis added). In other words, it is information that may be associated with separately maintained data, such as data contained in public records or data easily obtainable through an online search engine.

These definitions of PII under current federal and state privacy regimes show that the category is broadly defined, and that the primary focus is differentiating between information that distinguishes an individual and can be used to track that individual or otherwise be matched with new information about that individual in the future. In the context of a specific privacy law, like the VPPA, the focus should be on the ability of information to reveal personal details about the customer in connection with the subject matter of the law (in this case, video rental records). In some contexts, it will be useful to identify the types of customer-related information that are *not* personally identifiable, as Congress addressed in the VPPA Committee report. *See* S. Rep. No. 100-599, at 12, *reprinted in* 1988 U.S.C.C.A.N. 4342.

One category of information that has traditionally been excluded from the definitions of PII in federal privacy statutes is aggregate, anonymized (or de-

identified) statistical data. But unlike the lower court's casual use of the term "anonymous user IDs" to describe the data at issue in this case, the concepts of de-identification and anonymization are highly technical and complex, as recent data science research has shown.

II. Internet Protocol Addresses And Other Identifiers Are PII

The lower court concluded, without explanation, that Viacom's disclosure of a customer's "username; IP address; browser setting[s]; [and] 'unique device identifier" to Google along with the "detailed URL requests and video materials requested and obtained" and a code corresponding to the user's gender and age could not "serve to identify an actual, identifiable Plaintiff and what video or videos that Plaintiff watched." *In re Nickelodeon Consumer Privacy Litig.*, No. 12-7829, slip op. at 20 (D.N.J. July. 22, 2014) [hereinafter *Nickelodeon I*]. This conclusion is illogical, and appears to be based on a series of mistaken assumptions about the underlying functions of Internet networks and the business practices of Google and other advertising agencies. These marketing agencies rely on usernames, IP addresses, and other digital to identify and track users across the web, and to deliver targeted ads. *See* Jessica Rich, Dir., Bureau of Consumer Prot., Fed. Trade Comm'n, *Beyond Cookies: Privacy Lessons for Internet Advertising*

(Jan. 21, 2015).[6] These firms are not only capable of identifying and tracking users using this data, it is their entire business model.

A. Personal Data Can Only Be Considered "Anonymized" When It Has Been De-identified to Remove All PII

The "transactional information," *see* Leahy Floor Statement, *supra*, disclosed in this case can not fairly be described as "anonymous" or "anonymized." Those are key terms in privacy law that should not be used without a clear definition and close analysis. Since the publication of the HEW Report in 1973, which outlined the "Fair Information Practices" that make up the core of modern privacy law, the use of aggregate and anonymized data for statistical research purposes has been well established. *See* Sec'y's Advisory Comm. on Automated Personal Data Sys., U.S. Dep't of Health, Educ., and Welfare, *Records, Computers and the Rights of Citizens* 6 (1973). But more recently, scholars have shown that, even in large data sets, it is not sufficient to remove only "explicit identifiers, such as name, address and phone number," because at least "87% (216 million of 248 million) of the population in the United States had reported characteristics that likely made them unique based only on [5-digit ZIP, gender, and date of birth]." Latanya Sweeney, *Simple Demographics Often Identify People*

[6] *Available at* https://www.ftc.gov/system/files/documents/public_statements/620061/150121beyondcookies.pdf.

Uniquely 1 (Carnegie Mellon Univ., Sch. of Computer Sci., Data Privacy Lab., Working Paper No. 3, 2000).

In order to avoid the risk of disclosure of PII and the identification of a data subject, institutions processing aggregate statistical data to be publicly released must go through a process of "de-identification." De-identification " refers to the process of removing or obscuring any personally identifiable information from [records] in a way that minimizes the risk of unintended disclosure of the identity of individuals and information about them." Privacy and Tech. Assistance Ctr., U.S. Dep't of Education, *Data De-identification: An Overview of Basic Terms* 2–3 (2013). *See also* Robert Gellman, *The Deidentification Dilemma: A Legislative and Contractual Proposal*, 21 Fordham Intell. Prop. Media & Ent. L.J. 33 (2010); Latanya Sweeney, *k-anonymity: A Model for Protecting Privacy*, 10(5) Int'l J. on Uncertainty, Fuzziness, & Knowledge-based Sys. 557 (2002).

But even anonymized data should only be released with caution. The danger of re-identification of presumably anonymized datasets sparked an FTC investigation of the online video streaming company Netflix. *See* Letter from Mineesha Mithal, Associate Director, Division of Privacy and Identity Protection, Federal Trade Comm'n, to Reed Freeman, Counsel for Netflix, Inc. (Mar. 12,

2010) [hereinafter FTC Netflix Letter].[7] In 2006, Netflix released six-years' worth of customer viewing data in connection with the company's efforts to improve its video recommendation algorithm. *See* EPIC, *Re-identification: Concerning Re-identification of Consumer Information* (2015).[8] Before releasing the data, Netflix attempted to anonymize the information by replacing customers' names with unique numbers. *Id.* The 2006 dataset also did not contain customers' addresses, telephone numbers, or other direct identifiers. *Id.*

Despite Netflix's attempt to anonymize the 2006 dataset, two university researchers demonstrated that it was possible to re-identify individual Netflix users. Arvind Narayanan & Vitaly Shmatikov, *Robust De-Anonymization of Large Sparse Datasets (How to Break the Anonymity of the Netflix Prize Dataset)*, 2008 IEEE Symp. on Sec. & Privacy 111 (Feb. 5, 2008).[9] Using publicly available movie reviews posted by Netflix users on the website www.imdb.com, the researchers were able to determine a subscriber's complete movie rating history. *Id.* "Releasing the data and just removing the names does nothing for privacy," observed one of the study's authors. "If you know their name and a few records, then you can identify that person in the other [private] database." Robert Lemos,

[7] https://www.ftc.gov/sites/default/files/documents/closing_letters/netflix-inc./100312netflixletter.pdf.
[8] https://epic.org/privacy/reidentification/.
[9] *Available at* http://arxiv.org/abs/cs/0610105.

Researchers Reverse Netflix Anonymization, Security Focus (Dec. 4, 2007).[10] The study prompted the FTC's investigation into Netflix, and eventually led the company to cancel a second release of subscriber data. *See* FTC Netflix Letter, *supra*.

Companies that claim to deal only in "anonymous" data, "do not mean that they have no way to distinguish a specific person," or that "they have no way to recognize [the user] as the same person with whom they have interacted previously." Solon Barocas & Helen Nissenbaum, *Big Data's End Run Around Anonymity and Consent*, *in Privacy, Big Data, and the Public Good* 53 (Julia Lane et al. eds. 2014). Instead, these companies simply mean that they "rely on unique persistent identifiers that differ from those in common and everyday use (i.e. a name and other so-called [PII])." *Id.* The limitations of a name-focused conception of PII is illustrated by the widespread use of the Social Security Number ("SSN"). On its own, an SSN is nothing more than a nine-digit number. Large institutions, however, frequently use SSNs for identification because they are "necessarily more unique than given names, the more common of which (e.g. John Smith) could easily recur multiple times in the same database." *Id.* at 54. This is precisely case with unique persistent identifiers that are routinely swept up by online companies.

[10] http://www.securityfocus.com/news/11497.

Thus, it is axiomatic that any discussion of PII considers the collection, storage, or sharing of unique persistent identifiers beyond names.

B. Persistent Identifiers, Including IP and MAC Addresses, Are Not "Anonymous"

It is well established that Internet Protocol ("IP") addresses and other unique, persistent identifiers constitute personal information that "can be used to recognize a user over time and across different websites or online services." Fed. Trade Comm'n, *Complying with COPPA: Frequently Asked Questions* (2015).[11] IP addresses can be used to identify users and link consumers to digital video rentals. They are akin to Internet versions of consumers' home telephone numbers.

Every computer connected to the Internet receives an IP address that is logged by web servers as the user browses the Internet. These logs allow companies to record a trail of the user's online activity. Companies engage in extensive tracking and data collection about the online activities on consumers. *See generally* Emily Steel & Julia Angwin, *On the Web's Cutting Edge, Anonymity in Name Only*, Wall St. J. (Aug. 4, 2010);[12] Jessica E. Vascellaro, *Google Agonizes on Privacy As Ad World Vaults Ahead*, Wall St. J. (Aug. 10, 2010).[13] Furthermore,

---

[11] *Available at* https://www.ftc.gov/tips-advice/business-center/guidance/complying-coppa-frequently-asked-questions.

[12] http://www.wsj.com/articles/SB10001424052748703294904575385532109190198

[13] http://www.wsj.com/articles/SB10001424052748703309704575413553851854026

user names, which are frequently disclosed in URLs, can be used to personally identify users. Jonathan Mayer, *Tracking the Trackers: Where Everybody Knows Your Username*, Stanford Ctr. For Internet & Soc'y (Oct. 11, 2011).[14] All of these unique, persistent identifiers are used to track the online activities of specific users, and are also used to target advertising and otherwise influence the content delivered to those users by the websites they visit. Laura J. Bowman, *Pulling Back the Curtain: Online Consumer Tracking*, 7 I/S: J.L. & Pol'y for Info. Soc'y 721 (2012).

IP addresses are the "housing addresses" of networked devices. Christopher Parsons, *IPv6 and the Future of Privacy* (2010).[15] Generally, each device is assigned a unique number, which directs all packets of information going to and from the device. *Id*. Like a housing (or business) address, however, multiple devices can share the same IP address. This is possible through the use of routers, which assign separate, sub-addresses to individual devices in local networks. Riva Richmond, *We Know Where You Are*, Wall Street J. (Sep. 29, 2008).[16] But a family's sharing of a router (and public IP address) is ultimately no different than their sharing of a Blockbuster or Netflix account—the address can be used to track

---

14 http://web.archive.org/web/20140415113817/https://cyberlaw.stanford.edu/blog/2011/10/tracking-trackers-where-everybody-knows-your-username.

15 http://www.christopher-parsons.com/ipv6-and-the-future-of-privacy.

16 *Available at* http://online.wsj.com/article/SB122227759888771725.html.

and identify all of the activities of the subscriber(s), across the Internet. In this sense, public IP addresses are both persistent (because they are linked to all of a given subscriber's Internet activities) and they are unique (because they identify traffic from a specific subscriber account).

In addition to the IP address, each device with a network connection has "its own unique MAC address" for each "distinct point of attachment." IEEE Computer Society, *802—IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture* 22 (2002).[17] This means that a single laptop could have more than one MAC address (if it uses both a wireless and Ethernet connection for example), but no two devices share the same MAC address. *But see* A. Deepak Gupta et al., *MAC Spoofing and Its Countermeasures*, 2(4) Int. J. of Recent Trends in Eng'g & Tech. (2009). These MAC addresses were specifically designed to be unique, persistent identifiers that would be used to identify and communicate with specific devices over the Internet and other networks.

While IP addresses and MAC addresses have been used to identify users' computers and other devices since the development of modern Internet protocols, new tools have emerged that will allow even more extensive tracking of users across different networks and even different devices.

[17] *Available at* http://www.ieee802.org/secmail/pdfocSP2xXA6d.pdf.

III. Google and Other Internet Advertising Firms Identify Internet Users With Advanced Tracking Techniques

Internet users are already subject to targeted advertising and tracking from Google and other major firms, but the extent of that tracking is currently expanding at an incredible pace. *See* Rich, *supra*. But the lower court fundamentally misunderstood the dynamic, finding paradoxically that "information about a computer used to access" Internet videos was not "information about the Plaintiff himself." *Nickelodeon I*, slip op. at 20–21. This is equivalent to saying that a video rental store could share rental records linked with license plate numbers, because those numbers are "information about a [car] used to access" the store, rather than information "about the Plaintiff himself." The devices that customers use to access videos are directly linked to them, and tracking those devices is functionally identical to tracking the customers themselves.

A. Soon Every Internet-Connected Device May Be Assigned a Unique, Persistent IPv6 Address

Several new Internet tracking tools and techniques are posed to greatly expand the ability of Google and other advertising firms to track users across different web services, and even across different web-enabled devices. These new developments include: the assignment of unique, device-specific IP addresses; the evolution of tracking cookies that can identify users across services and across devices; and the use of digital fingerprinting.

A key change that is poised to allow even more granular identification of individual internet-connected devices is the rollout of a new Internet address protocol, IPv6. EPIC, *Search Engine Privacy* (2015).[18] *See also* Sangam Racherla & Jason Daniel, IBM, *IPv6 Introduction and Configuration* 2 (2012).[19] The current protocol, IPv4, only has the capacity to assign 4.3 billion unique IP addresses (2^{32} addresses). Laurie J. Flynn, *Drumming Up More Addresses on the Internet*, N.Y. Times (Feb. 14, 2011).[20] However, the rapid increase in the number of Internet-connected devices has led to an exhaustion of unique IP addresses. *Id*. Having anticipated this issue, network engineers developed a new protocol and began implementing it as early as 2008. Christopher Parsons, *IPv6 and the Future of Privacy* 2 (2010). IPv6 assigns 128-bit addresses, which means that it has the capacity to assign 2^{128}, or approximately 340 trillion trillion trillion addresses. Racherla & Daniel, *supra*, at 10.

As a result of this shift to IPv6, there will be an essentially limitless supply of IP addresses in the future, which means that every device with an IPv6 address will be uniquely identifiable based on its IP address. Early IPv6 implementations used an addressing scheme tied to the embedded network hardware (MAC) address used by each device. This mechanism has the effect of creating an unchangeable,

---

[18] https://epic.org/privacy/search_engine/.
[19] *Available at* http://www.redbooks.ibm.com/redpapers/pdfs/redp4776.pdf.
[20] http://www.nytimes.com/2011/02/15/technology/15internet.html?_r=0.

unique identifier that could be used to correlate unrelated activity and to allow a user to be tracked across multiple networks. See Comments of EPIC, Request for Comments on Deployment of Internet Protocol, Version 6, NTIA Docket No. 040107006-4006-01 (Mar. 8, 2004).[21]

B. The Consolidation of User Data Among a Few Firms and the Expansion of Internet-enabled Devices Makes Users More Traceable and Identifiable Based on Their Browsing Data

As internet-connected devices proliferate in the "Internet of Things," users will become increasingly identifiable based solely on the devices they use. The "Internet of Things" refers to the capability of everyday devices to connect to other devices and people through the existing Internet infrastructure. Devices connect and communicate in many ways. Examples of this are smartphones that interact with other smartphones, vehicle-to-vehicle communication, connected video cameras, connected medical devices, and televisions that track the movies and television shows that users watch. EPIC, *Internet of Things (IoT)* (2015).[22] They are able to communicate with consumers, collect and transmit data to companies, and compile large amounts of data for third parties. The FTC estimates that this year, over 25 billion devices will be connected to the Internet. Edith Ramirez, Remarks at the Federal Trade Commission Internet of Things Workshop (Nov. 19,

[21] *Available at* https://epic.org/privacy/internet/IPv6_comments.pdf.
[22] https://epic.org/privacy/internet/iot/.

2013).[23]

Congress anticipated the rise of the interactive television, in particular, as early as the 1980s. *See* William J. Broad, *U.S. Counts on Computer Edge in Race for Advanced TV*, N.Y. Times (Nov. 28, 1989) ("Finally, scientists say, the advent of digital television will aid the merging of computers and television, with the prospect of a rush of combined uses.").[24] Privacy advocates were aware that televisions would enable a wide range of functions in the home, including "home banking, instant voting, storage of personal information, home shopping, instant-response study courses, automatic regulation of utility use, a selection from almost 1,000 data bases of specialized information, and security services which can monitor for fire, home intrusion and medical emergency." David A. Bode, *Interactive Cable Television: Privacy Legislation*, 19 Gonz. L. Rev. 709, 710 (1984).

Cable providers first deployed interactive television through "set-top" boxes, transmitting user data to the service provider. *See* Rachel Powell, *Tech Notes; Televised Give and Take*, N.Y. Times (Apr. 25, 1993) ("Such capabilities require microprocessors atop the television set and high-capacity fiber-optic lines that link

---

[23] *Available at* http://www.ftc.gov/sites/default/files/documents/public_events/internet-things-privacy-security-connected-world/final_transcript.pdf.

[24] http://www.nytimes.com/1989/11/28/science/us-counts-o-computer-edge-in-the-race-for-advanced- tv.html.

the TV with the cable company -- equipment that is far more sophisticated than the set-top converter boxes and copper cable widely used today . . . .").[25] Privacy scholars and policy makers recognized the risk that interactive television would threaten the privacy of users if safeguards were not established. These risks included the "danger similar to wiretapping," of "misuse and interception of 'private' information" during transmission to the central servers, as well as the insecurity of data once it arrived at the central servers. The Cable Communications Policy Act was enacted in 1984 to combat these risks. Stephen Console, *Cable Television Privacy Act: Protecting Privacy Interests From Emerging Cable TV Technology,* 35 Fed. Com. L.J. 71, 79 (1983). The CCPA ensures that cable operators collect only the user data needed to operate the service, keep the data secure while it is in use, and delete the data once it has served its purpose.

The principles codified in the CCPA will become increasingly relevant to the Internet of Things not only as more devices connect to the Internet, but also as the data aggregation firms consolidate their database of information about their users. This is particularly evident in Google's consolidation of user information across services it owns. Google has continued to expand the tracking and profiling of Internet users.

[25] http://www.nytimes.com/1993/04/25/business/tech-notes-televised-give-and-take.html.

Today, Google continues to aggregate data from companies it owns, even when its privacy policy seems to state otherwise. For example, the Google privacy page discussing Google's purchase of the ad tracker DoubleClick states, "We will not combine DoubleClick cookie information with personally identifiable information unless we have your opt-in consent." Google, *Privacy and Terms* (2015).[26] However, Vincent Toubiana, an information technology expert working for the French data protection authority, explained:

> [Y]our Double-Click cookie will not be linked to your personally identifiable information. So Google can not put your name in front of the list of interests they inferred from your browsing behavior and will not put your name (or any other PII) in the ads you see. Because your Web Search history is likely to be unique, it identifies you and therefore can not be combined to your DoubleClick profile. But your search profile (i.e. the list of interests inferred from your search history) is unlikely to be unique and therefore does not identify you so Google can combine it with your DoubleClick cookie information.

Vincent Toubiana, *Google's Ad Targeting Under the New Privacy Policy*, Unsearcher (Feb. 24, 2012).[27]

Similarly, he explained, "your age, gender and interests expressed during Gtalk and Gmail discussions (or any other interest that Google could infer but that you would not be the only one to express) could be associated to your DoubleClick cookie." *Id*.

---

[26] http://www.google.com/intl/en/policies/privacy/.
[27] http://unsearcher.org/google-ad-targeting-under-the-new-privacy-policy.

C. Firms Are Now Deploying Browser "Cookies" That Cannot Be Deleted By the User

A cookie is a piece of code that downloads from a website to a user's browser when the user visits the website. Internet Engineering Task Force, *HTTP State Management Mechanism: Overview* (2011).[28] Every time the user reloads the website that the cookie came from, the cookie sends the website information about the user's computer activity. *See*, *e.g.*, Microsoft, *Windows Internet Explorer 8 Privacy Statement* (2015) ("A cookie is often used to personalize your visit to a website or to save you time. For example, to facilitate a purchase the cookie could contain shopping cart information such as your current selection, as well as contact information such as your name or e-mail address. To help websites track individual visitors, cookies often contain a unique identifier. It is up to the website that created the cookie to disclose to you what information is stored in the cookie and how that information is used.").[29]

Some cookies are privacy-neutral in themselves but may reveal detailed profiles of Internet users in the aggregate. In online shopping, for example, every time the user clicks on a new part of the website, a cookie stores the information about the contents of the user's "shopping cart." David Whalen, *The Unofficial*

[28] *Available at* http://tools.ietf.org/html/rfc6265#section-3.
[29] http://windows.microsoft.com/en-US/Internet-explorer/products/ie-8/privacy-statement.

Cookie FAQ (2002).[30] However, since cookies provide "a methodology for tracking your Internet usage across various web sites[, t]his information, while virtually useless in the context of a single web site, provides a detailed picture of an individual's web surfing habits in the aggregate, including the potential of being tied back to an actual person's name and address." Terry W. Posey, Jr., *Tony Soprano's Privacy Rights: Internet Cookies, Wiretapping Statutes, and Federal Computer Crimes After In Re Doubleclick*, 29 U. Dayton L. Rev. 109 (2003). These cookies can collect "user-profiling information, IP numbers, shopping cart contents, user IDs, user-selected preferences, serial numbers, frequencies of contact with companies, demographics, purchasing histories, credit-worthiness . . . social security numbers and other personal identifiers, credit card numbers, phone numbers, and addresses." Hal Berghel, *Caustic Cookies*, 44 Comms. ACM 20 (May 2001).

Many cookies, however, are explicitly "tracking cookies," or pieces of code that store information about a user's browsing history for targeted advertising purposes. Omer Tene & Jules Polonetsky, *To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioral Advertising*, 13 Minn. J.L. Sci. & Tech. 281, 292–93 (2012). These cookies are created for the purpose of collecting as much information as possible to create "detailed consumer

[30] http://www.cookiecentral.com/faq/.

profiles that reflect the online practices, preferences and other personal characteristics of each individual who surfs the Web." Michael R. Siebecker, *Cookies and the Common Law: Are Internet Advertisers Trespassing on Our Computers?,* 76 S. Cal. L. Rev. 893, 893 (2003).

There is a lucrative market for firms that use tracking cookies to create extensive user profiles. Companies such as DoubleClick, AdTech, and AdExchange are able to provide websites with details about their visitors using cookie technology. Even though some cookies do not contain identifying information about the user, companies that "specifically provide the service of linking cookies to users' personal information" form an integral part of the web-tracking market. Peter Segrist, *How the Rise of Big Data and Predictive Analytics Are Changing the Attorney's Duty of Competence*, 16 N.C. J. L. & Tech. 527, 540–41 (2015).

Many users are aware of the tracking capabilities of cookies, and take active steps to "clear" or otherwise block websites from using traditional cookies. In response, companies have developed so-called "supercookies" that are very difficult to detect and, if removed, may secretly reinstall themselves. Ashkan Soltani et al., *Flash Cookies and Privacy 2: Now with HTML5 and ETag*

Respawning (2011).[31] Supercookies use a number of different methods designed to evade cookie-blocking features in web browsers and browser ad-ons.

For example, a "Flash cookie" is transmitted through third party software—Adobe Flash—rather than through a normal web browser connection. Flash cookies are not only easier to hide, "they are more persistent than [traditional] cookies" and can store twenty-five times as much data. *Id*. Flash cookies are stored in a way that is not browser-specific, meaning that even if a user switches browsers, Flash cookies enable the user to be tracked." *Id*. Thus, "erasing [traditional] cookies, clearing history, erasing the cache," or even using a browser's "Private Browsing" mode will not prevent Flash cookies from transmitting details about your browsing history. Ashkan Soltani et al., *Flash Cookies and Privacy 1* (2009).[32]

Advertisers use these advanced tracking techniques so that they can indefinitely monitor the browsing data of Internet users, without their knowledge or consent, in order to sell targeted advertising services. The idea behind this tracking is to plant two cookies on the user's machine—a standard cookie that the consumer may erase, and a second Flash cookie that the user will not likely remove because the existence of Flash cookies is not well known. EPIC, *Local Shared*

[31] *Available at* http://ssrn.com/abstract=1898390.
[32] *Available* at http://ssrn.com/abstract=1446862.

Objects—Flash Cookies (2005).[33] Flash cookies transmit data about the user back to the cookies' owners, and "those tidbits of information are added to the user's profile. The more information stored, the more valuable for the gathering source. Even where cookies are disabled, some companies, like Google Chrome, track consumers' entire viewing history by seeking electronic permission to install a seemingly innocuous software that allows for detailed tracking." Alexander Tsesis, *The Right to Erasure: Privacy, Data Brokers, and the Indefinite Retention of Data*, 49 Wake Forest L. Rev. 433, 438 (2014).

Some supercookies have the additional function of "respawning," or restoring, HTTP cookies that have already been deleted. Tene & Polonetsky, *supra*, at 292–93. These supercookies are programmed to recognize and recreate certain HTTP cookies, which will become the "zombie cookies." First, the supercookie downloads to the user's browser using the Adobe Flash script. Once downloaded, the supercookie assigns a unique identifying number to the host computer and stores that number in the Adobe Flash storage bin. Then the supercookie checks the browser for the presence of the HTTP cookies that it can respawn. Jacqui Chen, *Zombie Cookie Wars: Evil Tracking API Meant To "Raise Awareness,"* Ars Technica (Sep. 22, 2010).[34] If the user deletes the browser's

[33] https://epic.org/privacy/cookies/flash.html.
[34] http://arstechnica.com/business/2010/09/22/evercookie-escalates-the-zombie-cookie-war-by-raising-awareness/.

cookies, the supercookie will regenerate a new HTTP cookie and label it with the unique tracking number. In this way, the "zombified" HTTP cookie is never effectively deleted, and can continue to track the user even after the user's cookies have been cleared. Bruce Schneier, *Evercookies*, Schneier on Security (Sep. 23, 2010).[35]

D. Marketers Can Also Identify and Track Users Based On Their Digital "Fingerprints"

Faced with users' increased understanding of the cookies on their computers and how to erase them, advertisers have begun to use an identification technique called "fingerprinting." Adam Tanner, *The Web Cookie Is Dying. Here's The Creepier Technology That Comes Next*, Forbes (Jun. 17, 2013).[36] By transmitting some code to a user's browser when the user visits a given website, advertisers can make a detailed profile of that user's computer. Erik Larkin, *Browser Fingerprinting Can ID You Without Cookies*, PC World (Jan. 29, 2010).[37] This process involves identifying Internet users based on "unique characteristics of the individual computers people use" under the "assumption that each user operates his or her own hardware, identifying a device is tantamount to identifying the person

---

[35] https://www.schneier.com/blog/archives/2010/09/evercookies.html.
[36] *Available at* http://www.forbes.com/sites/adamtanner/2013/06/17/the-web-cookie-is-dying-heres-the-creepier-technology-that-comes-next/.
[37] *Available at* http://www.pcworld.com/article/188161/browser_fingerprinting_can_id_you_without_cookies.html.

behind it." Nick Nikiforakis & Günes Acar, *Browser Fingerprinting and the Online-Tracking Arms Race*, IEEE Spectrum (July 25, 2014).[38] For example, a fingerprint can identify a user based on their operating system's version and configuration, the type and version of the web browser installed, the plugins attached to the browser, wireless settings, TCP/IP configuration, and the hardware clock skew. Tadayoshi Kohno, Andre Broido & K. C. Claffy, *Remote Physical Device Fingerprinting*, Inst. of Electrical & Electronic Eng'rs Symposium on Sec. & Privacy 1, 211 (2005).[39]

These fingerprinting techniques are evolving to the point where they are impossible for users to block. Julia Angwin, *Meet the Online Tracking Device That Is Virtually Impossible to Block*, ProPublica (July 21, 2014) (describing a new technique called "canvas fingerprinting" that identifies a user's device based on the unique way it draws a hidden image).[40] As a result of these developments, as well as the use of tracking cookies and other identifying techniques, it is no longer possible to consider Internet browsing data to be "anonymous."

---

[38] http://spectrum.ieee.org/computing/software/browser-fingerprinting-and-the-onlinetracking-arms-race.

[39] *Available at* http://homes.cs.washington.edu/~yoshi/papers/PDF/KoBrCl2005PDF-Extended-lowres.pdf. "Clock skew" refers to an imperfection in hardware circuitry that prevents the nodes in the network of memory transistors from synchronizing perfectly. Clock skew can vary, often uniquely, from device to device.

[40] http://www.propublica.org/article/meet-the-online-tracking-device-that-is-virtually-impossible-to-block.

CONCLUSION

Amicus respectfully request this Court reverse the lower court's order granting Appellee's motion to dismiss.

Respectfully submitted,

/s/ Marc Rotenberg
Marc Rotenberg
Counsel of Record
Alan Butler
Julia Horwitz
John Tran
Electronic Privacy Information Center
1718 Connecticut Ave. NW, Suite 200
Washington, DC 20009
(202) 483-1140

CERTIFICATE OF COMPLIANCE WITH FEDERAL RULES

This brief complies with the type-volume limitation of Fed. R. App. P. 29(d) and Fed. R. App. P. 32(a)(7)(B) because it contains 6,430 words, excluding the parts of the brief exempted by Fed. R. App. P. 32(a)(7)(B)(iii). This brief also complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type style requirements of Fed. R. App. P. 32(a)(6) because this brief has been prepared in a proportionally spaced typeface using Microsoft Office Word for Mac 2011 in 14 point Times New Roman.

Dated: May 4, 2015

/s/ Marc Rotenberg
Marc Rotenberg

CERTIFICATE OF COMPLIANCE WITH LOCAL RULES

I certify that I have complied with LAR 31.1(c) because this file was scanned by the most current version of Scan This, https://scanthis.net, and no virus was detected. I also certify that I am a member of the bar of this court, and that the text of this electronically filed brief is identical to the text of the 10 paper copies mailed to the court.

Dated: May 4, 2015

/s/ Marc Rotenberg
Marc Rotenberg

CERTIFICATE OF SERVICE

I hereby certify that on May 4, 2015, I electronically filed the foregoing Brief of *Amici Curiae* Electronic Privacy Information Center Support of Appellant with the Clerk of the United States Court of Appeals for the Third Circuit using the CM/ECF system. All parties are to this case will be served via the CM/ECF system.

Dated: May 4, 2015

/s/ Marc Rotenberg
Marc Rotenberg