To: National Telecommunications and Information Administration
ATTN: David J. Redl
From: Cassandra Baloga – NYLS 2019
Date: November 9, 2018
Re: Docket No. 180821780-8780-01 – Developing the Administration's Approach to Consumer Privacy: Request for Public Comments

INTRODUCTION

This comment will address the idea of "personal information," referencing the goals listed in the call for comments. The definition of the term "personal information" is crucial when discussing several of the core privacy outcomes consumers should expect from organizations. Personal information is the crux of these outcomes and what this privacy act is attempting to protect. The only reason to implement them is to help protect a consumer's personal information when it is given to an organization. Without a solid, unambiguous definition and adequate understanding of what "personal information" encompasses, the outcomes listed are not easily attainable. This runs the risk of the privacy act not being uniformly applied across all organizations and further runs the risk that courts will apply the law inconsistently in the event of litigation.

THE CURRENT STATE OF PERSONAL INFORMATION

The first thing to address are the different terms used to describe "personal information." The terms personal information, personal data, and personally identifiable information are all used as terms describing and defining an individual's personal information. Many statutes, regulations, and other documents define these terms differently—meaning the definitions are highly dependent on context.[1]

The National Institute of Standards and Technology (NIST), for example, defines "personally identifiable information" as "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 (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information."[2] As defined by the NIST, personally identifiable information includes information that is already linked to an individual's identity or information that *can be linked* to an individual's identity.

The General Data Protection Regulation ("GDPR") defines "personal data" as "any information relating to an identified or identifiable natural person . . . who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification

[1] David A. Zetoony, *Defining Personal Information*, BRYAN CAVE LEIGHTON PAISNER (Aug. 13, 2018), https://www.bryancave.com/en/thought-leadership/defining-personal-information-1.html.
[2] Erika McCallister, Time Grance, & Karen Scarfone, *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)*, NAT'L INST. OF STANDARDS AND TECH. (April, 2010), https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-122.pdf.

number, location data, an online identifier, or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural, or social identity of that natural person. "[3] It is also important to note that the GDPR states that cookies and IP addresses specifically can be considered personal data.[4]

In stark contrast, the Maryland Personal Information Protection Act defines "personal information" as "[a]n individual's first name or first initial and last name in combination with any one or more of the following data elements, when the name or the data elements are not encrypted, redacted, or otherwise protected by another method that renders the information unreadable or unusable:" (1) Social Security number; (2) Driver's License number; (3) Financial Account number; (4) health information; (5) health insurance policy information; or (6) biometric data.[5] The act further defines what is *not* personal information.[6] Because the Act specifies precisely what information is and is not considered personal information, it is an example of a rigid, non-fluid, and uncomprehensive definition of "personal information."

THE VIDEO PRIVACY PROTECTION ACT AND THE CIRCUIT SPLIT ON WHAT QUALIFIES AS PERSONALLY IDENTIFABLE INFORMATION

As can be seen, the definition of personal information can be ambiguous. In another example, the Video Privacy Protection Act (VPPA) states "[a] video tape service provider who knowingly discloses, to any person, personally identifiable information concerning any consumer . . ." shall be liable.[7] The Act defines "personally identifiable information" as "information which identifies a person has having requested or obtained specific video materials or services from a video tape service provider."[8] There is a circuit split on whether "personally identifiable information" is only information that would allow an "ordinary person" to identify an individual.[9]

In *Eichnberger v. ESPN, Inc.*, users of the ESPN app on a Roku device, brought suit after discovering that Defendant knowingly gave their information to a third-party, namely the Roku serial number and the videos Plaintiff viewed.[10] Defendant gave the information to the third-party knowing the third party would identify the plaintiff by connecting the information with already existing data.[11] Plaintiff's identity was then returned to the defendant which was, in turn, sold to advertisers.[12] The court acknowledged that the term "personally identifiable information" included "more than information than that which, by itself, identifies an individual as having

---

[3] 2016 O.J. (L 119) Art. 4(1).
[4] 2016 O.J. (L 119) Recital 30.
[5] Md. Code Ann., Com. Law § 14-3501(e)(1)(i) (West 2018).
[6] Md. Code Ann., Com. Law § 14-3501(e)(2) (West 2018).
[7] 18 U.S.C.A. § 2710(b)(1) (West 2013).
[8] 18 U.S.C. § 2710(b)(3) (West 2013).
[9] Heather Egan Sussman et al., *United States: Ninth Circuit Weighs In On Scope of Identifiable Informaton Under VPPA*, ROPES AND GRAY (Dec. 28, 2017), http://www.mondaq.com/unitedstates/x/659440/Data+Protection+Privacy/Ninth+Circuit+Weighs+In+On+Scope+of+Identifiable+Information+under+VPPA.
[10] Eichenberger v. ESPN, Inc., 876 F.3d 979, 981 (9th Cir. 2017).
[11] *Id.*
[12] *Id.*

watched certain videos[,]"[13] but "covers some information that *can be used* to identify an individual."[14] Nevertheless, the court, using the "ordinary observer" test, concluded that because the "information *cannot* identify an individual unless it is combined with other data in [the third-party's] possession . . . an ordinary person could not use the information that Defendant allegedly disclosed to identify an individual."[15]

The "ordinary observer" test came from *Nickelodeon Consumer Privacy Litig*.[16] There, Defendant disclosed, *inter alia*, Plaintiffs' IP addresses, browser settings, and a "unique identifier" through the use of cookies.[17] Plaintiffs' argued that these pieces of information permitted third-parties to track their computers and identify their specific internet usage.[18] The court, however, stated that "[t]o an average person, an IP address or digital code in a cookie file would likely be of little help in trying to identify an actual person."[19] The court held that "personally identifiable information" applies to information that "would readily permit an ordinary person to identify a specific individual's video-watching behavior[,]" and that digital identifiers were not such information.[20]

On the other hand, the court in *Yershov v. Gannett Satellite Info. Network, Inc.*, adopted a broader interpretation of personally identifiable information. There, Defendant sent information to a third-party every time Plaintiff viewed a video without Plaintiff's consent.[21] This information included the video title, the GPS coordinates of the device, and other identifiers such as the unique Android ID.[22] Using this information, the third-party was able to link Plaintiff to an individualized profile they already maintained.[23] The court held this was "personally identifiable information" within the meaning of the VPPA.[24] The court noted that "[w]hile there is certainly a point at which the linkage of information to identity becomes too uncertain, or too dependent on too much yet-to-be-done, or unforeseeable detective work, here the linkage, as plausibly alleged, is both firm and readily foreseeable to [Plaintiff]."[25] The court here did not use the "ordinary person" test, but instead asked whether the information was "*reasonably* and *foreseeably* likely to reveal" the identity of Plaintiff.[26]

The problem this circuit split creates is one where very similar facts result in different outcomes during litigation. All three courts looked at the history of the VPPA and legislative intent behind its enactment while still disagreeing on how "personally identifiable information" should be interpreted.

---

[13] *Id.* at 984.
[14] *Id.* (emphasis in original).
[15] *Id.* at 986. (emphasis in original).
[16] In re Nickelodeon Consumer Privacy Litig., 827 F.3d 262 (3d Cir. 2016), *cert. denied sub nom.* C. A. F. v. Viacom Inc., 137 S. Ct. 624, 196 L. Ed. 2d 516 (2017).
[17] *Id.* at 281–82.
[18] *Id.* at 282.
[19] *Id.* at 283.
[20] *Id.* at 287.
[21] Yershov v. Gannett Satellite Info. Network, Inc., 820 F.3d 482, 484 (1st Cir. 2016).
[22] *Id.*
[23] *Id.* at 485.
[24] *Id.* at 486.
[25] *Id.*
[26] *Id.* (emphasis added).

It appears the Federal Trade Commission (FTC) agrees with the "reasonably foreseeable" test. In 2016, the Director of the FTC stated that the FTC regards "data as 'personally identifiable,' and thus warranting privacy protections when it can be *reasonably linked* to a particular person, computer, or device. In many cases, persistent identifiers such as device identifiers, MAC addresses, static IP addresses, or cookies meet this test."[27]

THE BENEFITS OF PERSONAL INFORMATION COLLECTION

Despite the risk to consumers, gathering personal information has many benefits for companies, consumers, and society. Gathering a user's personal information helps retailers with marketing and sales by controlling costs and tracking inventory.[28] The financial and banking industries benefit from the collection of personal information by identifying credit risks, tailoring loan packages, and detecting fraud.[29] In the medical field, personal information is being used to "predict epidemics, cure disease, improve quality of life and avoid preventable deaths."[30] Researches are also studying Alzheimer's, breast cancer gene therapy, and the prevention of neural tube birth defects.[31]

HOW SHOULD PERSONAL INFORMATION BE DEFINED?

All of this leads us to ask: how should personal information be defined? If the definition is too narrow, it will "fail to protect privacy in light of modern technologies."[32] If it is too broad, "it could encompass too much information, and threaten to transform privacy law into a cumbersome and unworkable regulation of nearly all information."[33] With this in mind, and in order to reconcile the idea that one's personal information should be protected with the idea that sometimes personal data plays an important role in research,[34] personal information should be divided into three categories and be defined and interpreted separately: identified information, personally identifiable information (PII), and non-personally identifiable information (non-PII).[35]

Paul Schwartz and Daniel Solove, law professors Berkley and George Washington, have suggested a method, called PII 2.0, based on a risk identification model.[36] The first category of information, identified information, refers to information that singles out and identifies a specific

---

[27] Jessica Rich, *Keeping Up with the Online Advertising Industry*, FED. TRADE COMM'N (Apr. 21, 2016, 10:30 AM), https://www.ftc.gov/news-events/blogs/business-blog/2016/04/keeping-online-advertising-industry (emphasis in original).
[28] Rich, *supra* note 27.
[29] Rich, *supra* note 27.
[30] Bernard Mar, *How Bit Data Is Changing Healthcare*, FORBES (Apr. 21, 2015, 10:50 AM), https://www.forbes.com/sites/bernardmarr/2015/04/21/how-big-data-is-changing-healthcare/#2d9e89dc2873.
[31] Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and A New Concept of Personally Identifiable Information*, 86 N.Y.U. L. Rev. 1814, 1827 (2011).
[32] *Id.*
[33] *Id.*
[34] *Id.* at 1866.
[35] *Id.* at 1877.
[36] *Id.* at 1877–79.

individual.[37] This concept is generally agreed upon.[38] This category should also include information that "brings a substantial risk of identification of an individual."[39] When there is a substantial possibility that the links necessary to identify a person will be made, this information presents a great risk of identity.[40]

The second category, PII, refers to information that makes "an individual identifiable" when "there is some non-remote possibility of future identification."[41] There is a low to moderate amount of risk associated with this category of information.[42] The third category is non-PII, which has only a remote risk of identification associated with it.[43] Non-PII is data that is anonymous, has no need for encryption, and has only a small risk that misuse would result in harm to an individual.[44] Aggregate data, for example, is not always identifiable—a census of the United States could not necessarily pinpoint one particular individual.

The line between PII and non-PII is a sliding scale, dependent on technology.[45] Information not personally identifiable today may be identifiable tomorrow, or at least have a higher chance of being identifiable, thereby sliding down the scale closer toward the end that carries a higher risk of personal identification.

The breadth of information that already exists about an individual is a factor in determining whether information is identifiable. For example, medical information, even if redacted, can be identifiable to "skillful Googlers," friends, family, or colleagues. [46] Google search queries, on their face, seem to be non-PII, however in reality they may become PII depending on what the search contains or how many search queries by the same user are released.[47] At some point, the specificity, quantity, and character of search queries allows for a person to be identifiable.[48] For instance, when AOL released 20 million search queries for 657,000 of their users, one user in particular was identified through her searches for a landscaper in her town and about her dog, searches looking for single men, and queries of people with the same last name.[49] This user was identified based on information AOL previously thought to be non-PII, and may have been if her particular search queries were not released in such mass quantities.

HOW THIS MODEL RELATES TO THE SUGGESTED OUTCOMES AND HOW IT WILL BE AN IMPROVEMENT FROM THE CURRENT REGIME

---

[37] *Id.* at 1877.
[38] *Id.*
[39] *Id.* at 1878.
[40] *Id.*
[41] *Id.*
[42] *Id.*
[43] *Id.*
[44] LATENTVIEW, https://www.latentview.com/non-pii-data (last visited Nov. 9, 2018).
[45] Schwartz, *supra* note 31, at 1846.
[46] Nw. Mem'l Hosp. v. Ashcroft, 362 F.3d 923, 929 (7th Cir. 2004) (quotations removed).
[47] Schwartz, *supra* note 31, at 1847–48.
[48] Schwartz, *supra* note 31, at 1848.
[49] Michael Barbaro & Tom Zeller Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, THE NEW YORK TIMES (Aug. 9, 2006) https://www.nytimes.com/2006/08/09/technology/09aol.html.

By determining the risk of personal identification associated with personal information on a sliding scale, and by making the context of the information a center point of the analysis, this method will move with technology and will not become obsolete or unworkable with technological advancement. "The line between PII and non-PII is not fixed, but depends upon technology. Thus, today's non-PII might be tomorrow's PII."[50] Technology moves quickly, but legislation moves slow. A fluid definition will allow organizations, users, and courts to adapt the definition of personal information as technology continues to grow.

Further, the risk-based approach to defining personal information will encourage privacy by design. Privacy by design's goal is to build privacy principles into a product at the development stage.[51] It is proactive and seeks to embed privacy directly into the design and operation of technological systems.[52] Currently, organizations take a "collect first, ask questions later" approach to the collection of personal information.[53] By setting boundaries within the definition, even on a fluid scale, organizations will be encouraged to adopt the idea of privacy by design. Broadening the definition of personal information in tandem with the other outcomes and goals the NTIA seeks to accomplish will shift the burden of protecting their personal information from users to organizations by encouraging them to make privacy protection part of the default.[54]

PII 2.0 Will Help Advance Many of the Outcomes and Goals that the NTIA Would Like to Achieve

Employ a Risk and Outcome-Based Approach

The NTIA desires to create a model that doesn't create "cumbersome red tape" but instead is based on risk modeling and user-centric outcomes.[55] That is exactly what this definition of "personal information" would achieve. It is based on a scale from identified information to non-identified information, which is correlated to the amount of risk associated with the data. This is a flexible approach that will not be immobile and run the risk that organizations will be disinclined to develop new products, services, or business models while also providing privacy protections to users.

Reasonable Minimization

---

[50] Schwartz, *supra* note 31, at 1846.
[51] Robin Kurzer, *What Is Privacy By Design? A Deeper Dive Into This GDPR Requirement*, MARTECHTODAY (Mar. 20, 2018, 12:57 PM), https://martechtoday.com/privacy-design-deeper-dive-gdpr-requirement-212463.
[52] Deloitte LLP, *Privacy By Design: Setting a New Standard for Privacy Certification*, https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/risk/ca-en-ers-privacy-by-design-brochure.PDF (last visited Nov. 9, 2018).
[53] WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES 6 (Harvard Univ. Press eds., 1stt ed. 2018).
[54] FED TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE 24 (2012), https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.
[55] Developing the Administration's Approach to Consumer Privacy, 86 Fed. Reg. 48600, 48602 (proposed Sept. 26, 2018).

The "reasonable minimization" goal seeks to reduce an organization's collection, storage, use and sharing of personal information.[56] PII 2.0 will encourage organizations to reduce the risk of holding personal information that is readily identifiable. Organizations will have an incentive to collect only the data they need and to only collect PII under necessary circumstances.[57] PII 2.0 is contextual, "it should consider factors such as the lifetime for which information is to be stored, the likelihood of future development of relevant technology, and parties' incentives to link identifiable data to a specific person."[58] If organizations are motivated to invest resources to maintain information in either identifiable or non-identifiable form, the goal of reasonable minimization will "become easier to comply with as they move along [the] continuum away from identified information."[59] The risk that already existing information about an individual will be combined with new information about that same individual—sliding down the scale from identifiable to identified information—will encourage organizations to think long and hard about what information they have stored, what information they would like to collect, and what information is no longer worth the risk to continue to collect or store.

Control, Access & Correction

Obligations to give users full notice, access, and correction rights to their personal data "would decrease rather than increase privacy by requiring that all such data be associated with a specific person."[60] To allow an individual to exercise her rights of notice, access, and correction, "the law would create a vicious circle" that would turn non-identifiable or identifiable information into identified information.[61] The "control" outcome seeks to allow users to exercise control over the collection, use, storage, and disclosure of the personal information provided to organizations.[62] This becomes somewhat difficult under the PII 2.0 model because user control over their personal information would simultaneously transform that information from non-PII or PII to identified information.

This does not mean users cannot exercise any control over their personal information. Instead of approving what specific information an organization collects about them, users can instead approve sets or types of information. After being informed of what risks are associated with identified, identifiable, and non-identifiable information and given an explanation about how each set of information can affect their risk of personal identity, a user can make an informed choice as to what type of information they wish to share. Further, once a user is informed of what they are sharing and the risks associated with it, a user will also be informed enough to be able to withdraw the consent or limit the information they allow to be collected.

Accountability

---

[56] Developing the Administration's Approach to Consumer Privacy, 86 Fed. Reg. at 48601.
[57] Schwartz, *supra* note 32 at 1888.
[58] *Id.* at 1878.
[59] *Id.* at 1883.
[60] *Id.* at 1880.
[61] *Id.*
[62] Developing the Administration's Approach to Consumer Privacy, 86 Fed. Reg. 48601–02 (proposed Sept. 26, 2018).

Under the PII 2.0 model, organizations will not be able to use the repetitious defense that the information they collected is not PII.[63] The test for whether information is identifiable or non-identifiable is extremely contextual and is determined by the circumstances and the facts surrounding a particular situation and, therefore, will not allow organizations to default to the excuse of non-PII.

This will hold organizations accountable because there will be no stark line between PII and non-PII. The fact-specific inquiry that determines how much risk is associated with a user's personal information will force organizations to collect data in ways that minimize risk and give users the information they need to make informed decisions. If expectations are clear, an organization has no excuse to say it didn't know the information being collected was PII—which will hold them responsible for what happens to it. If personal information continues to have an ambiguous definition, organizations can skirt the edge and deny accountability.

Transparency

PII 2.0 will also encourage organizations to be more transparent. A thorough definition of personal information will allow both organizations and users to know exactly what type of information can be collected and the risks that information holds. The risks for an organization and a user are different. A user should be informed about what information is being collected and the risk that it can personally identify them. An organization should be aware of what kind of information fits into each one of the personal information categories and of what secondary information can be used to identify an individual. In other words, a fluid definition of personal information will allow organizations to assess the risks of data collection with better precision, thereby allowing them to be more specific—and therefore, transparent—in informing its users about what information it is collecting and why. This would benefit both users and organizations by allowing users to assess the risks associated with using the service and decreasing an organization's risk of liability.

Legal Clarity While Maintaining the Flexibility to Innovate

This model will protect users without stifling innovation. Privacy is linked with innovation. Offering more data privacy through a regulatory action will produce potential consequences for emerging technologies.[64] Protecting both privacy and innovation requires the balancing of strong consumer protections and legal clarity with flexibility.[65] PII 2.0 is a flexible model that does not draw a harsh line between what organizations can and cannot do. Because this definition is robust, but not rigid, it does not mandate best practices. Instead, it allows organizations to develop their own practices and make decisions on what information to collect, how long to store it, and in what ways to use it. By defining personal information in a way that is essentially open-ended, organizations will not be forced to stifle their creativity or innovation and can continue to use data collection in ways that benefit them, users, and society.

---

[63] Schwartz, *supra* note 31` at 1890.
[64] Avi Goldfarb & Catherine Tucker, *Privacy and Innovation*, NBER WORKING PAPER SERIES (June, 2011), https://www.nber.org/papers/w17124.pdf.
[65] Developing the Administration's Approach to Consumer Privacy, 86 Fed. Reg. at 48602.

CONCLUSION

Although the definition of personal information is one small part of an overall privacy scheme, the rest of the scheme is dependent upon how personal information is defined. The proposed definition will help achieve the NTIA's outcomes and goals. This fluid definition will allow users to be informed of the risks associated with personal information and will incentivize organizations to collect and store data with care.