CENTER FOR DEMOCRACY & TECHNOLOGY

May 21, 2024

Jessica Herron
Legislative Clerk
Subcommittee on Innovation, Data, and Commerce
House Committee on Energy and Commerce
2125 Rayburn House Office Building
Washington, D.C. 20515-6115

Re: Samir Jain's Responses to Additional Questions for the Record

Dear Ms. Herron:

I want to thank the Subcommittee for inviting me to appear before it on April 17, 2024, to testify at the hearing entitled, "Legislative Solutions to Protect Kids Online and Ensure Americans' Data Privacy Rights."

Pursuant to the Rules of the Committee on Energy and Commerce, I am attaching my answers to additional questions for the record in the required format.

Thank you again for your help, and please let me know if you need any further information.

Sincerely,

Samir Jain

Samir Jain
Vice President of Policy

Mr. Samir Jain

Attachment – Additional Questions for the Record

The Honorable Russ Fulcher

In response to my question about third-party data providers, a.k.a., "data brokers," where I noted tailored marketing messages and advertisements to different customers are commonly used through marketing automation and other tools, I asked about where is the line on the level of tailoring of messaging and advertising, given the American Privacy Rights Act (APRA) strongly keeps the decision ownership with the consumer, the issue of delineating between contextual ads, first-party ads being okay provided they are not behind paywalls, and that ARPA generally strikes the right balance. I want to ensure a company can use a marketing automation tool but is not excessively or intrusively profiling individual customers.

1. **Does this just come down to disaggregating individual customers into categories and the types and amount of tailoring of a message that can be done under ARPA? When it comes to the categories of ads you mentioned – "contextual, first-party not behind a paywall, etc." – can you discuss further what should be allowable versus not, fitting within the requirements of ARPA?**

Answer:

As I noted in my testimony, the APRA discussion draft lacks definitions for several key terms related to advertising, leading to some confusion as to how the data minimization rules would apply to different types of online advertising. My response reflects CDT's interpretation of the sponsors' intent based on the bipartisan compromise that underlies the APRA. CDT has long advocated for strict limits (and in some cases prohibitions) on commercial surveillance systems used to drive targeted advertising, which can produce unfair or discriminatory outcomes. We understand that at least some stakeholders on the industry side continue to advocate for rules that impose few meaningful limits on the collection and processing of personal data. But the result of prior negotiations over the ADPPA and APRA was a compromise position where certain ad-related uses are prohibited, most are subject to an opt-out regime, and a smaller subset are exempt from the opt-out. The charts below provide definitions and concrete examples of use cases in each of those categories.

The definitions and examples in this chart focus on the ways advertisers and advertising-technology (ad-tech) intermediaries use data to match ad content with relevant audiences in some of the main media environments where people encounter digital advertising. We recognize there may be differences in the way these terms are used among advertising professionals, civil society, and policymakers; the use of terms in this document reflects the way these terms are used by the general public and in many state privacy laws.

Mr. Samir Jain

The charts illustrate which practices the discussion draft of APRA would allow, subject to opt-out rights, or prohibit:

- Table 1 explains how six types of ads—*contextual ads*, *self-directed ads*, *in-house ads*, *first-party targeted ads*, *third-party targeted ads*, and *cross-context behavioral ads*—use non-sensitive data to reach specific audiences. It also provides examples for each category.
- Table 2 provides examples for each category to illustrate how the six ad categories apply to different media types (*websites/apps*, *social media*, *video platforms*, and *marketplace sites*).
- Table 3 illustrates how the six ad types intersect with the Sensitive Covered Data (SCD) rule, using examples of ads that would or would not be permissible.

The tables use color-coding to show our understanding of how APRA treats different ad types:

- Green: These types of ads are always allowed so long as they do not rely on sensitive data, and are not subject to the targeted advertising opt-out.
- Yellow: These types of ads are subject to the targeted advertising opt-out. They are allowed for consumers who have not opted out, but prohibited for consumers who have opted out.
- Red: This type of ad is prohibited.

These charts are simplifications that necessarily elide some nuance. For example, in some cases consumer expectations may vary depending on whether they are logged in to a site or platform or not, affecting the types of ads that should be permitted. In a similar vein, these charts do not reflect the discussion draft's inclusion of data collected over time by high-impact social media companies on their platforms as sensitive data.

Mr. Samir Jain

Table 1: Simplified description of the main online advertising categories

Ad types	Explanation	Examples
Contextual advertising	Ads are placed exclusively based on the content of the website, app or publication and/or based on a user's general location (not precise geolocation).	- Hoka advertises running shoes in a fitness app. - American Airlines advertises flight deals on a website about the best travel destinations. - A new restaurant advertises to people whose devices are located within a 15-mile radius on the results page for a search query about restaurants.
Self-directed advertising	This is an example of an innovation ad tech companies can make to adapt to privacy laws. Ads would be based on information voluntarily provided by the user (which can be changed any time) for the purpose of informing ad targeting.	- While logged in to Instagram, I select "the great outdoors" as one of several types of ads that I am interested in receiving. REI buys ad inventory to reach me and other users who have made the same selection.
In-house advertising	The company whose website/app you're viewing advertises its own products/services based on your activity (as long as it does not reveal or use sensitive data) on that company's website/app.	- The New York Times advertises a NYT Cooking subscription while I'm reading news on the NYT app. - Old Navy shows me an ad on its own website or sends me an email advertising a blazer based on the fact that I bought the matching pants.
First-party targeted advertising	Targeting is based on your activity on the website/app you're viewing (as long as it does not use sensitive data).	- I read several articles about travel on the Washington Post website. The Washington Post then uses this data to decide to show me ads for Costco vacation packages.
Third-party targeted advertising	Targeting is based on the non-sensitive data of a single third-party advertiser, such as retargeting.	- Last week I visited the L.L.Bean website but didn't buy anything. L.L.Bean placed a cookie in my browser, and now every website I visit shows me L.L.Bean ads. - Nordstrom gives Meta the email addresses of 10,000 previous customers and directs Meta to show those customers ads for an upcoming sale.
Cross-context behavioral advertising	Ads are based on my browsing activity over time and across websites (using cookies/device fingerprinting/etc.) and/or digital dossiers that ad-tech vendors or data brokers compile about me on the basis of cross-site tracking data (sensitive data that cannot be used for advertising).	- Based on a profile compiled about me using my online activity, Acxiom infers that I am a sports fan who likes to gamble. FanDuel uses Acxiom's profile to show me ads for its sports-betting feature on every site I visit. - Over time I visit a series of websites to shop for Nike shoes and an Osprey backpack. I also do my taxes on a tax filing website. All of these sites use the Meta Pixel and transfer my browsing data to Meta, who profiles me as a high-income outdoor enthusiast. REI places an ad across Meta's social media services to target users who are interested in purchasing expensive outdoor goods, and my online activities are used to target me with the ad.

Table 2: Examples of the different environments where consumers encounter online ads

	Websites/apps	Social media	Video platforms	Marketplace sites
Contextual advertising	REI advertises hiking gear on a website about local hiking trails.	While viewing content in a Facebook Group about dog ownership, I see ads for pet supplies.	While watching episodes of Cupcake Wars on a Netflix plan with ads, I see ads for KitchenAids.	I view baby formula on Target's website and Target shows an ad for diapers on the page.
Self-directed advertising	While logged in to the NYT, I select "nature" as one of several types of ads I am interested in. Patagonia buys ads to reach me and others who have made the same selection.	On my Instagram, I select "travel" as one of several types of ads that I am interested in. Marriott buys ads to reach me and others who have made the same selection.	While logged in to Hulu, I select "beauty" as one of several types of ads that I am interested in. Ulta buys ads to reach me and other users who have made the same selection.	While logged in to my Amazon account, through which I indicated that I am interested in ads for books, I am shown ads for upcoming new release books.
In-house advertising	While reading the NYT, I see ads for a NYT Cooking subscription.	While scrolling through my Facebook feed, I see ads for other Meta products like Instagram.	While watching videos on Amazon Prime, I see ads for Whole Foods or another Prime show.	Based on my Amazon browsing, Amazon shows me ads for hats from Amazon Basics.
First-party targeted advertising	I read several articles about travel on the Washington Post site. The Washington Post uses that data to decide to show me ads for vacation packages from travel agents.	I consistently engage with (like, comment, etc.) Instagram content related to health and fitness. Meta uses that data to show me ads for Planet Fitness memberships.	I watch sports content daily on ESPN. ESPN decides, based in part on this activity, to show me ads for Nike.	Based on my Amazon browsing and purchase history, Amazon shows me ads for Duracell batteries that fit products I have previously purchased.
Third-party targeted advertising	Last week, I visited the L.L.Bean site but didn't buy anything. L.L.Bean placed a cookie in my browser, and now every website I visit shows me L.L.Bean ads.	Last week, I attended a Broadway show. The venue uses my email address to pay Meta to show me ads for upcoming Broadway shows on Facebook.	While logged in to HBO Max on a plan with ads, I am shown ads for Nike running shoes based on Nike's data that I buy running shoes every spring.	Brita gives Amazon the email addresses of 10,000 customers who have bought pitchers and directs Amazon to show those customers ads for Brita filters.
Cross-context behavioral advertising	LiveRamp aggregated my browsing history into a profile that indicates I might have a toddler. I now see ads for toddler toys that are targeted to me based on my LiveRamp profile through Real-Time Bidding/ad exchanges.	Meta has been tracking nearly every website I visit for years and now has a detailed profile on me, which it uses to target me with ads on Facebook, Instagram, and across the web.	With help from data brokers, Disney has been tracking my online activity for years and now has a detailed profile for me, which it uses to target me with ads on Disney-owned platforms and across the web.	Amazon has been tracking my online activity for years and now has a detailed profile for me, which it uses to target me with ads in its marketplace site, during Amazon Prime content, and across the web.

Table 3: Examples of ads that would/would not be subject to the SCD rule

Ad type	Permitted	Prohibited under the SCD rule
Contextual advertising **Permitted regardless of SCD rule because this type of advertising does not use covered data**		
Self-directed advertising	- While logged in to Instagram, I select "the great outdoors" as one of several types of ads that I am interested in. REI buys ad inventory to reach me and other users who have made the same selection. I can change my ad preferences at any time.	- A social media platform gives me the option to self-select into an audience category for people with depression and anxiety.
In-house advertising	- The New York Times advertises a NYT Cooking subscription while I'm reading news on the NYT app. - Old Navy shows me an ad on its own website or sends me an email advertising a blazer based on the fact that I bought the matching pants.	- Based on changes in my purchasing habits, a retailer infers that I may have diabetes and starts showing me ads for its selection of glucose monitors.
First-party targeted advertising	- I read several articles about travel on CNN's website. CNN then uses this data to decide to show me ads for vacation packages. - I engage with (like, comment, etc.) sports content on the New York Times app. At the NFL's direction, the NYT uses that information to show me ads for tickets to upcoming NFL games in my city.	- On the basis of my activity within their site, a retailer infers that I may be pregnant and sends me direct mail inviting me to create a baby registry.
Third-party targeted advertising	- Last week I visited the L.L.Bean website but didn't buy anything. L.L.Bean placed a cookie in my browser, and now every website I visit shows me L.L.Bean ads. - Nordstrom gives Facebook the email addresses of 10,000 customers who have previously purchased from Nordstrom and directs Facebook to show those customers ads for Nordstrom's upcoming sale.	- Several of my relatives use the services of a commercial genetic testing company, and as a result learn that they are carriers for a specific condition. The company uses public records to infer that I am related to them, and instructs a social media platform to target me with alarming ads suggesting that I may have the condition and should get tested. - An LGBTQ+ dating app discloses users' precise location data to ad networks for use in Real-Time Bidding. That information is aggregated and sold to data brokers, who combine that data with other information to identify individuals who frequent gay bars and target them with ads.
Cross-context behavioral advertising **Prohibited generally**		

Mr. Samir Jain

The Honorable Lisa Blunt Rochester

I appreciate your attention to civil rights in the context of artificial intelligence. As you mentioned, these tools leverage vast amounts of data and could be utilized in extremely consequential decisions involving employment, housing, and access to credit. It is paramount that AI tools being brought to market do not deepen existing inequality and do not perpetuate discrimination.

1. **Can you expand on how robust data privacy protections could safeguard civil rights?**

Answer:

Privacy protections, particularly those included in the American Privacy Rights Act (APRA), would safeguard civil rights in several ways.

First, APRA defines and then limits use of sensitive data. Included in the definition are a few categories that could have particular impact on civil rights: (1) health information, including disability status; (2) information revealing sexual behavior of an individual; (3) information revealing race, ethnicity, national origin, religion, or sex of an individual; and (4) any data collected for the purpose of identifying any of the above sensitive categories. For a company to collect this information about an individual, it must be "necessary, proportionate, and limited" to the service the individual is requesting, or for a specific allowable purpose. Thus, the amount of data about individuals related to these categories that companies can collect will be significantly lower, and the obligation to prove that the company needs that data will be on the company itself.

With less data, and restrictions on the uses to which such data can be put, there will be less opportunity to discriminate. For example, sensitive data may not be used in either contextual or targeted advertising. This limit means that even if a company otherwise has knowledge that an individual is Black, or is a woman, or has a disability, that company may not use that data to target ads to that person for that reason. Thus, there will be less harmful discrimination in advertising.

Second, the civil rights section in APRA explicitly prevents processing data in ways that discriminate against certain protected classes, closing many gaps in federal law. Our current federal civil rights landscape is mostly a patchwork of protections in various areas (such as housing and employment) related to certain protected classes (in particular, key civil rights laws like Sections 1981, 1982, and 1985 apply only to racial discrimination). APRA would clarify that using data in a discriminatory way against a variety of protected classes would be illegal, and would provide for at least injunctive relief to prevent those practices.

Third, the algorithmic impact assessments would force large companies to consider bias in the development and deployment of their algorithms. APRA would require large companies that develop algorithms to be transparent about the algorithm's design process, purpose and uses, training datasets, and outputs, as well as the necessity and proportionality of the algorithm in relation to its stated purpose. These companies would also need to disclose steps they have taken to prevent civil rights-related harms, such as their role in facilitating advertising for, determining access to, or restrictions on housing, education, employment, health care, insurance, or credit opportunities. APRA also would require any developer of a covered algorithm, prior to deployment, to analyze their algorithms for their impact on these same harms and reduce the risk of those harms.

Fourth, the ability to opt out of algorithms that help make consequential decisions provides a safeguard for people to protect themselves against discrimination that still exists, particularly for disabled people, who may need to request an accommodation. Even after all the up-front protections for civil rights, there may still be algorithms that discriminate based on a variety of factors, or an individual may simply feel as though they would prefer to have a human review their situation rather than an algorithm. In those cases, section 14 of APRA would allow people to opt-out of those automated decisions.

2. Where could APRA, KOSA, and COPPA 2.0 be strengthened in this regard?

Answer:

APRA's protections against discrimination could be strengthened in at least three specific ways. First, Section 13(a)(2)(C) should be removed. That section creates an exception to the strong civil rights protections that allows for "advertising, marketing, or soliciting economic opportunities or benefits to underrepresented populations or members of protected classes as described in paragraph (1)." Despite positive intentions, this exception would allow for advertising toward any "members of protected classes" in harmful ways, such as "white-men-only" advertising for housing. It could also allow "Christians-only" or even "non-Christians only" advertising. These are the types of discriminatory advertising we're trying to *avoid*.

Second, APRA could be improved by clarifying evaluations and impact assessment requirements for algorithms involved in consequential decisions. APRA could be clearer that any covered entity developing or deploying an algorithm involved in a consequential decision should conduct pre-deployment evaluations and post-deployment impact assessments. Pre-deployment evaluations should focus on how the algorithm was developed, what its intended uses and outputs are, what data was used to train the algorithm, anticipated harms including actions taken to avoid those harms, and how it was tested for bias. Post-deployment impact assessments should consider similar issues, including whether the actual use of the algorithm has resulted in

disparate impact or other discrimination. The post-deployment impact assessments should be completed by an independent third party auditor.

Third, APRA could be improved by applying the private right of action to sensitive data collection and processing. Given all the benefits I described above regarding the sensitive data protections and how they help address discrimination harms, it would seem reasonable to apply the private right of action to the collection and processing of sensitive data, as the American Data Privacy and Protection Act did in 2022. APRA as currently drafted applies the private right of action only to transfers of sensitive data, which addresses only one type of harm associated with sensitive data misuse.

While COPPA 2.0 has some improvements over the original COPPA from 1998, it is fundamentally a different type of privacy law than APRA. The focus of COPPA since its passage has primarily been verifiable parental consent. The purpose of APRA is to move us beyond a notice-and-consent regime to one based on data minimization. And while COPPA 2.0 has some data minimization language, it is not as strong as APRA. Thus, there might be some civil rights benefits from that language, but overall APRA's data minimization standard would provide greater protections. Thus, a preferable approach would be for the substantive protections from COPPA 2.0 (bans on targeted advertising for those under 17, for instance) to be incorporated into APRA, so all issues are considered in a single comprehensive law.

Finally, as noted in my testimony, KOSA, while well-intentioned, raises concerns, especially to the extent it would restrict access to certain types of content that the government would deem harmful to children. Although bills to address online child safety pursue an important goal, content-based restrictions can hurt young people, particularly teenagers, who need to access important information, including those in marginalized communities that otherwise face discrimination. Children who grow up in highly restricted environments or face parental or domestic abuse in particular have a strong need for access to information and private communications channels to ensure their safety and mental health, which may be jeopardized by legislation that empowers government officials to sue companies for enabling access to information that they deem "harmful" to young people. Government attempts to protect minors by restricting access to content also raise significant constitutional concerns as they can infringe on the rights of users, including children and teenagers, to access constitutionally-protected information. Thus, KOSA could be improved by ensuring it did not authorize or incentivize these types of content restrictions.