Who Shares Your Information With Facebook?

Sampling the Surveillance Economy in 2023

BY DON MARTI, FENGYANG LIN,
MATTHEW SCHWARTZ, AND GINNY FAHS

JANUARY 2024

CR Consumer Reports

Table of Contents

Background

This study sets out to investigate elements of the surveillance economy that are largely invisible to consumers, consumer advocates, and researchers.

What is the surveillance economy? One way to understand it is as the subset of consumer marketing in which the data being used is obtained from the surveillance, or covert observation, of ordinary consumer activities such as visiting websites, buying goods or services from an online or physical retailer, using one's credit card, and consuming entertainment content.

The surveillance economy is "cross-contextual," meaning that it uses information about individuals that's been collected in one context—such as a website visit, an action taken in an app, or a visit to a physical location—and applies it to another context to affect how you are advertised to, what prices you see, and how you are otherwise treated.

The surveillance economy is not limited to one kind of business. Although large tech companies are well known for surveillance, and some recent legislative attention has been paid to data brokers—firms that buy and sell surveillance data—most of the companies that you interact with every day are participating in the surveillance economy in some way. Key players include:

- **Internet platform companies,** which sell search-based and social media-based advertising. These platforms, including Meta/Facebook and Google, typically get surveillance data from advertisers, and tend not to share the data they collect directly but instead "rent" it to advertisers by letting them use it indirectly to target paid ads.
- **Data brokers,** which buy and sell consumer data. These companies get data from many sources and make data available for many purposes, from anti-fraud to direct mail and social media ad targeting.
- **Retailers,** which collect large quantities of data from customers. Large-scale retailers increasingly operate their own advertising systems.
- **Marketing agencies and service providers,** which use customer data from brands and/or third-party data from data brokers to place advertising on behalf of client-advertisers.

The Current State of Consumer Privacy

In recent years, consumers have achieved several important wins in privacy law and technology that imposed modest constraints on the surveillance economy. Under a handful of state privacy laws such as the California Consumer Privacy Act (CCPA), many consumers now possess legal backing for their rights to learn how companies are using their personal information and to better control its use.

Many of these new laws also create universal opt-out mechanisms for consumers and require companies to respond to "authorized agents" designated by consumers to exercise their privacy rights.

These provisions enable consumers to exercise their data rights across many company websites at once, eliminating the need to do so sequentially, one company website at a time.

At the same time, web browsers and at least one major mobile platform have technologically limited the ability of sites and apps to track users across contexts. For example, Apple Safari's Intelligent Tracking Prevention features and Firefox's Enhanced Tracking Protection features both make it harder for surveillance companies to follow consumers around the internet using tracking technologies such as cookies.

Surveillance companies have responded to the changing privacy landscape in a variety of ways. While some good actors are extending privacy rights to all of their users (not just those who live in states with privacy laws) or adapting to platform-level restrictions of data access, many have merely turned to workarounds, reallocating their surveillance efforts into areas that are less susceptible to user understanding and control. A common practice is to replace or supplement on-device or in-browser tracking with data transfers that take place server-to-server outside of the consumer's awareness or control. This change presents a challenge to the research community and to consumer organizations seeking to understand which companies and practices represent the most active privacy threats.

The earlier that consumer advocates detect new surveillance technologies, the more likely it is that consumer researchers and advocates will be able to influence the practices of all businesses to use data in ways that benefit consumers. CR's "Who Shares Your Information With Facebook?" study was designed to help researchers better understand the state of the surveillance economy in 2023 and real-world surveillance practices, including hard-to-measure server-to-server data transfers.

Based on the findings of this effort, we have also developed policy proposals.

Approach

The list of data brokers and other companies holding consumer data is extremely long and ever-changing, but many of the companies have one thing in common: They choose to advertise on the Ads Manager platform operated by Meta, the corporate parent of social media platform Facebook. Using Meta Ads Manager, advertisers can target particular audiences by placing ads in their social media feeds across Facebook, Instagram, and Messenger.

Facebook's Download Your Information tool allows users to access some of the information about them that companies have transferred to Facebook while using the Ads Manager platform.[1] This study primarily examines two datasets available through this tool: Custom Audiences and Events.

Custom Audiences are lists of personal identifiers, such as email addresses, postal addresses, phone numbers, and mobile ad IDs transferred by Facebook advertisers to Facebook for the purpose of targeting advertisements. Targeted advertisements are directed to their intended recipients using data collected from the advertiser's own customers or partner businesses, purchased from data brokers, or by licensing data held by another business that the advertiser contracts with to help target advertisements to specific groups. Ads can be targeted either directly to the individuals whose data has been shared or to a "lookalike audience" of other Facebook users who share certain characteristics with those individuals.[2] (Custom Audiences data can also be used to exclude Facebook users with certain characteristics from seeing an ad.) Advertisers or their service providers generally add personal data to the Custom Audiences database without knowing whether the individuals on their lists are Facebook users.

The second dataset that Facebook lets consumer users download is Events, a detailed compilation of actions the user has taken when using the internet or interacting with a company. Typical entries include viewing certain pages of the company's website, the purchase of a product or service, visits to physical retail locations, and even "leveling up" in a video game.

As with Custom Audiences data, Events data are transferred by Facebook advertisers to Facebook, generally for the purpose of targeting advertisements or for measuring whether a Facebook ad resulted in an off-Facebook action such as a sale. Advertisers typically collect Events data using a script, or "tracking pixel," running on their websites; using Facebook Software Development Kit (SDK) code running in a mobile application; or by a server running Facebook Conversions API (CAPI). Events created by a pixel or SDK are visible from a web browser extension or proxy server (and have been the

[1] Meta Accounts Center, https://www.facebook.com/dyi.
[2] "How to format a customer list when creating a custom audience," Meta Business Help Center, https://www.facebook.com/business/help/2082575038703844?id=2469097953376494.

subject of previous consumer research[3]), whereas CAPI events cannot be seen by consumer users or in client-side applications and can be seen and analyzed only by running a study like this one.

For an individual, Custom Audiences and Events data has limited value: They can determine which companies have collected data on them and/or likely targeted them with ads. But they cannot know how many (or few) people were similarly tracked and targeted by the company, or which of the companies tracking and targeting them represent the biggest risk to their privacy. Such insights would require access to and analysis of hundreds of consumer records, and thus remain unattainable to most individuals and research teams.

Consumer Reports, however, is uniquely positioned to conduct such research because of our consumer-first mission, commitment to consumer privacy, and relationships with more than 6 million consumer members. Our research team launched a campaign to pool together the Facebook records of many consumers so that we could search for patterns in the data and begin to identify which companies—and which types of companies—present the greatest threats to consumer privacy.

One analogy we used when conceiving and designing this study is that of a sampling well, a method commonly used to test groundwater for harmful chemicals. Engineers design these wells to tap into low points in the earth where water collects (imagine the bottom of a funnel), enabling them to detect toxins dispersed across a wide area of the surface (the top of the funnel). Facebook Ads Manager is, in effect, a sampling well for the broader surveillance marketing ecosystem: Because so many advertisers leverage Meta Ads Manager to reach consumers, and because Meta makes its advertising system easy to use by companies of all sizes, consumer data is likely to end up with Meta eventually. By collecting a large sample of consumer data from Facebook, we are thus able to study an exceptionally broad—though not statistically representative of American consumers or the U.S. population—cross section of the data moving through the surveillance marketing economy.

[3] Angie Waller and Colin Lecher, “Help Us Investigate Facebook Pixel Tracking,” *The Markup,* January 21, 2022, https://themarkup.org/pixel-hunt/2022/01/21/help-us-investigate-facebook-pixel-tracking.

Methodology

Consumer Reports is a 6-million-plus-member, nonprofit organization with a track record of conducting participatory research that shines a light on problems in the consumer marketplace. We have completed participatory studies of home tap water quality[4], errors on credit reports[5], price variation on broadband bills[6], the usability of new data rights[7, 8], and more. We decided to apply a similar research model to the surveillance economy by asking volunteers to enroll in the Facebook Surveillance Study, download a portion of their Facebook data, and donate that data to CR for analysis.

Enrollment

We asked consumers to sign up for the study via our Community Reports volunteer site, where they were prompted to provide first name, last name, and email address, and to agree to CR's privacy policy and user agreement (Appendix A). Upon submitting the enrollment form, volunteers received an email with instructions for downloading their data from Facebook (Appendix B) and were also taken to a data donation form (Appendix C), where they were prompted to provide their email address and review an informed consent agreement. After consenting to participate in the study and avowing that they were over 18 years old, participants again received instructions for downloading their data from Facebook (Appendix D).

Data Donation

In its Download Your Information tool, Facebook enables consumer users to download both the Custom Audiences data and Events data associated with their Facebook accounts. CR's instructions guided volunteers through the process of downloading these two datasets using either a desktop or mobile interface. We did not ask volunteers to donate any photos, posts, or other personal information.

After volunteers uploaded a zip file of the pertinent Facebook data, we asked for their ZIP code and whether they had ever used Global Privacy Control (GPC), a browser setting that signals a user's intent

[4] Ryan Felton, Lisa Gill, and Lewis Kendall, "How Safe Is Our Drinking Water?" *Consumer Reports,* March 31, 2021, https://www.consumerreports.org/water-quality/how-safe-is-our-drinking-water-a0101771201.

[5] Lisa L. Gill, "More Than a Third of Volunteers in a Consumer Reports Study Found Errors in Their Credit Reports," *Consumer Reports,* June 11, 2021, https://www.consumerreports.org/credit-scores-reports/consumers-found-errors-in-their-credit-reports-a6996937910.

[6] James K. Willcox, "You May Be Paying Too Much for Your Internet," *Consumer Reports,* November 17, 2022, https://www.consumerreports.org/electronics-computers/telecom-services/you-may-be-paying-too-much-for-your-internet-a7157329937.

[7] Kaveh Waddell, "California's New Privacy Rights Are Hard to Use, Consumer Reports Study Finds," *Consumer Reports,* March 16, 2021, https://www.consumerreports.org/electronics-computers/privacy/californias-new-privacy-rights-are-tough-to-use-a1497188573.

[8] Kaveh Waddell, "Why It's Tough to Get Help Opting Out of Data Sharing," *Consumer Reports,* March 16, 2021, https://www.consumerreports.org/electronics/privacy/why-its-tough-to-get-help-opting-out-of-data-sharing-a7758781076.

to opt out of the sale and sharing of their personal info to each website they visit. Both questions were optional. We asked about ZIP code in order to understand which geographic areas our responses came from and about GPC to help us project the extent to which future GPC adoption could limit surveillance.

Volunteer Recruitment

As noted above, the success of our “sampling well” approach depends on collecting data from a large number of volunteers. Furthermore, the multistep process of finding and downloading one’s personal data in the Facebook database is mined with opportunities for error and participant dropout. Considering this, our research team aimed to recruit widely for the study.

We started with a base of active CR members, emailing more than 50,000 consumers who had previously participated in at least one of our community research projects and were therefore familiar with the notion of donating personal data to the organization. We hosted a live webinar for interested volunteers to explain the purpose of the project and demonstrate the data donation process in real time.

In order to maximize our chances of getting enough usable data, we also recruited participants more widely, using CR social media channels and outreach to interested media, researchers, and peer organizations. The Markup, another nonprofit journalism organization with a track record of investigating the surveillance economy, wrote an article outlining our study plan and asking for participants from among its readership.[9]

Volunteer Participation

Overall, more than 2,600 people signed up to participate in the Facebook Surveillance Study between June 28 and September 28, 2023. Some volunteers were not able to complete their data donation, because they could not access their Facebook accounts, had never had a Facebook account, struggled to find the data files we were seeking, or did not complete the donation for some other reason.

Ultimately, more than 1,000 people contributed their Facebook data to CR, and after removing the duplicate files, more than 700 of those files were able to be cleaned, processed, and included in our analysis. Among the files we were unable to parse and analyze, some included information other than the advertising data we asked for and others contained no data.

Sign-ups to participate in study	2,643
Facebook data files donated	1,043

[9] Sisi Wei and Maria Puertas, “Help Us Investigate Surveillance Marketing Using Facebook Data,” *The Markup*, August 2, 2023, https://themarkup.org/pixel-hunt/2023/08/02/help-us-investigate-surveillance-marketing-using-facebook-data.

Facebook data files included in statistical analysis	709

Sample Representation

We did not ask our data donors for any demographic information, and thus cannot make any claims about how representative this sample is of the U.S. population as a whole. Because we asked participants to provide their ZIP codes, we know that a disproportionate number of volunteers lived in California, New York, and Texas. Individuals who care about privacy are surely overrepresented in this data, because we recruited them from mailing lists of individuals who had participated in privacy research studies in the past, through privacy-concerned circles on social media, and through a peer organization that publishes regularly about privacy and surveillance. And while we don’t have exact numbers, the timing of volunteers’ participation suggests that the vast majority of participants were CR members.

Furthermore, all participants were tech-savvy enough to use Facebook, download their data from Facebook, and upload their data to a form—meaning they possessed a relatively high level of digital literacy. We don’t know precisely why more than 1,500 people signed up but never donated their data, but we suspect that many did not have a Facebook profile, did not feel safe uploading their data using the donation system we set up, or lacked enough digital literacy to complete the steps.

All of which is to say, again, that our dataset is not representative of the U.S. population as a whole.

Nevertheless, we believe this study will both help advance researchers’ understanding of the surveillance economy and suggest more ways that the surveillance economy should be further analyzed in the future.

Statistical Analysis

Once CR received donated data from volunteers, our research team transferred the data to a secure environment for cleaning and processing. To ensure that consumers’ personal information was not used in our analysis, we set up a privacy-first analytical environment that stripped donated files of any user-specific information and selected only the advertising information we needed for analysis.

Filtering and aggregation was the first processing task: Our code walked through the directory of uploaded zip files, opened and read only the specific JSON-formatted files we had requested from users, and extracted specific data points from the JSON file. The code we installed was incapable of reading any information other than what was requested.

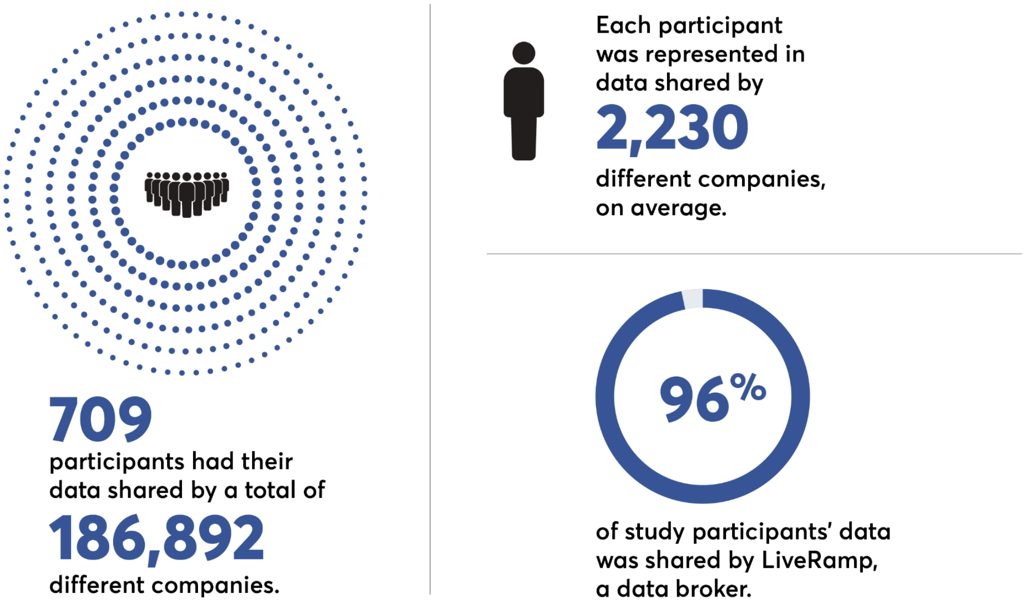
The results of this filtering and aggregation step were two dataframes. The first contained only company information with counts of the number of volunteers affected. The second listed events alongside the company’s name, event type, and a time stamp. Only these two dataframes were used in subsequent processing steps.

To provide a summary of the data we collected, we reported the basic statistics of these variables, including the mean, median, and quantile values. We also calculated the 95% confidence interval for variables about an average company and an average user.

Findings and Discussion

What we discovered in consumers' data files was striking. The overall scope of data sharing and targeted advertising that occurs on Facebook is immense.

Study Highlights



709 participants had their data shared by a total of 186,892 different companies.

Each participant was represented in data shared by 2,230 different companies, on average.

96% of study participants' data was shared by LiveRamp, a data broker.

- We found slightly more than 186,000 different companies represented in the data of 709 participants. Of them, 682 participants have their Events data being sent to Facebook, and 693 of them are included in the Custom Audiences data.
- Each of these 186,000-plus companies shared data on an average of eight participants in our study.
- The average participant in our study was identified in the data by 2,230 different companies; some were identified by more than 7,000 companies.
- The company that shared data on the largest number of participants was LiveRamp, a data broker, which shared data on 679, or about 96%, of study participants.[10]
- The 100 companies that most frequently appeared in our sample each shared or directed their service provider to share data on more than half of the 709 volunteers. (Those companies are listed in Appendix E.) Of those companies, 39 are retailers or

[10] Consumer Reports has a business relationship with LiveRamp and another data broker, Acxiom. Consumer Reports shares data with each of these companies in order to help support its mission.

consumer brands, 28 are agencies or services providers, 19 are data brokers, 4 are political services firms, and 10 are best classified as “other.”

It is worth noting here that in many cases, consumers would have to do their own research to confirm the identity of a company present in the data. This is because Facebook does not have clear rules for how companies are identified to consumers who download their data.

The companies listed in the Facebook data we received from our contributors were identified in a number of different ways:

- Domain names (such as “Amazon.com”).
- Names of reasonably identifiable companies (such as those where one company matching the name comes up as an obvious search result, as with “The Home Depot”).
- Names that are human-comprehensible but can’t be matched to a specific company (such as common words that are used in the names of multiple companies, as with “Viking”).
- Not reasonably identifiable companies (such as strings of digits or meaningless strings of Unicode characters, as with “100130874778177”).

Only 34% of the companies present in the sample data provided a URL linking to the company’s website. The remainder of the companies were listed in free-form text, and in many cases those names did not clearly correspond to a specific or well-known corporate entity. And over 7,000 companies had completely unidentifiable names that would be impossible to associate with a particular business. We address this phenomenon in more detail below.

The data does not reveal each company’s role, or whether the consumer has a relationship with them. Some companies in the data appear to be agencies or service providers acting on behalf of a client, or data brokers that don’t have direct consumer relationships, but this is not visible from the data provided. However, by looking at the data from a large number of consumers, we can detect patterns, such as commonly seen company names that are likely to be large data brokers.

We also found that 52% of the companies shared data on only a single volunteer in our sample, suggesting that these companies were using Custom Audiences or Events data for “microtargeting,” which is also discussed in more detail below.

Retailers and Brands

Aside from data brokers, and among the minority of companies we could identify, the most common types of businesses that showed up in our volunteers’ data were individual brands (such as Heineken), retailers (such as Macy's), and direct-to-consumer brands (where the brand and retailer are the same, such as SmileDirectClub). Home Depot, a major Facebook advertiser, was the retailer that appeared most frequently in the data. Other well-known national retailers, including Amazon and Walmart, were also very frequently seen transferring our volunteers’ data.

The brand reported to the consumer is not necessarily the source of the consumer's information. For example, a company might have rented access to a list of email addresses from a data broker or other third party. When we refer to a company "sharing" consumer data in this report, we use that as a catch-all term to encompass instances where the company directly shared data it collected about consumers with Facebook, as well as when the company made some kind of arrangement to use consumer data from another source.

Interestingly, many of the advertisers that targeted the largest percentages of study participants do not have a national footprint. For example, the Illinois Lottery shared data on nearly 70% of our volunteers. Local auto dealerships were also surprisingly well represented in the sample data, suggesting that they often have access to large marketing lists drawn from national sources. One car dealer in San Benito, Texas (pop. 24,665), for example, was responsible for sending information on approximately 10% of our study volunteers, though only 6.6% of study volunteers reside in the entire state of Texas. Several other local auto dealerships—including, for example, a small-town Porsche dealership—also leveraged contact info on around 10% of our volunteers.

The prevalence of these businesses in our data suggests just how easy it is for even relatively small businesses to collect or leverage large amounts of consumer information—and how attractive it can be for businesses to partner with Facebook to lure prospective customers using that information. In the case of the car dealerships, we suspect that one or more data brokers sold or sold access to large customer lists to a large number of auto dealerships, which then used these lists to serve advertisements on Facebook.

Service Providers, Data Brokers, and Political Services Firms

Many of the companies that appeared in our sample data were advertising agencies or other kinds of service providers, which act on behalf of another company to serve ads to consumers. Many of these companies are also data brokers, which collect and sell data about consumers with whom they often have no direct relationship. Prominent data brokers include LiveRamp, Oracle, and Acxiom.

Until recently, data brokers mostly functioned as a distinct node in the advertising ecosystem, but as part of a broad consolidation trend in the industry, large agency holding companies have been building or acquiring their own in-house data brokers. Holding company Dentsu Aegis acquired data broker Merkle, for example; Interpublic Group acquired Acxiom LLC; Publicis Groupe acquired Epsilon; and WPP (which owns storied ad agency Ogilvy) combined the data divisions of GroupM and Wunderman Thompson in a new entity called Choreograph. Holding companies and their subsidiaries are well represented in our data sample, and will likely continue to play a large and growing role in the surveillance economy as such firms continue to consolidate and expand their service offerings.

Another sizable subset of the companies that appeared in our sample data are political services firms, which provide advertising, fundraising, and other services to political campaigns.

In most of the cases involving service providers, data brokers, and political services firms, we were unable to determine what brands, products, or candidates these entities were actually advertising. These entities are typically just the middlemen who facilitate the purchase of advertisements on Facebook (or elsewhere), and it is nearly impossible for consumers to understand where the data used to target them originated from.

Unidentifiable Company Names

Many of the supposed company “names” in the data could not be definitively connected to a particular corporate entity. In some cases, the name listed consisted of an indecipherable string of letters and numbers, such as “Bm 5 100tkqc nlm,” which we conjecture may be a company’s internal designation for a particular target audience. In other cases, the names provided were trademarks used by different companies in different markets, such as “Viking,” which could refer to a range of possible businesses.

Overall, only 34% of the companies present in the sample data provided a direct link to the company’s website. Over 7,000 (4%) of companies in the sample had incomprehensible names that made it extremely difficult to identify the company. Ninety-nine percent of study participants were identified by at least one company with an unidentifiable name. And an average of 113 unidentifiable companies shared consumer data on each study participant.

In these instances, it is virtually impossible for even motivated consumers to understand who is tracking and targeting them, and for consumers to exercise their rights under existing state privacy laws. We propose ways to address this issue in the Policy Recommendations section, below.

Microtargeting

A large percentage of the approximately 186,000 companies that appeared in our data appeared to be either small retailers or non-national brands (or were unidentifiable by name). Nearly 48,000 different companies were found in the data of a single volunteer, probably someone with unusual app usage habits or possibly an exceptionally ripe and appealing candidate for microtargeted advertising.

In addition, 96,000 of the companies (52%) were targeting only one of our 709 volunteers. This likely reflects the ease with which even small companies with limited marketing resources can experiment with Meta Ad Manager. Meta provides small businesses with easy implementations of advanced surveillance advertising technologies. The small business owner has only to set a budget, provide personal info on customers, and report customer buying behavior to Meta—Meta’s software does the rest.

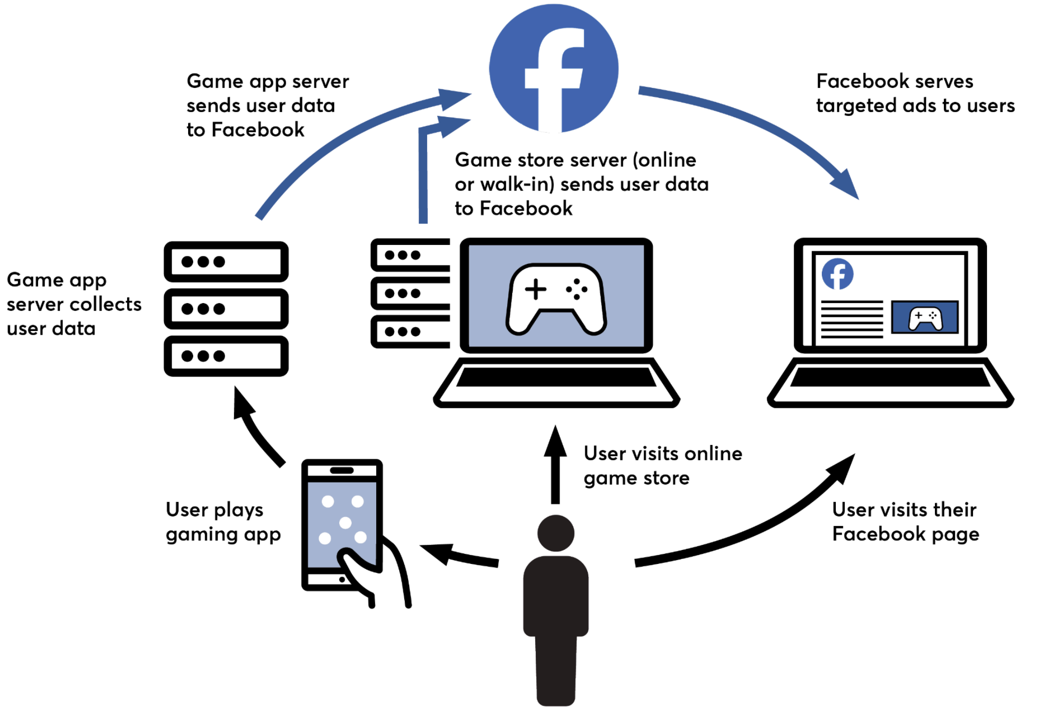
The prevalence of microtargeting in our sample data, combined with the fact that a majority of the company names listed were unidentifiable, also raises an important digital security concern.

Microtargeting has reportedly been used by scammers to target vulnerable people on social media[11,12]. The ability of Ad Manager advertisers to supply an identifier other than a valid company name or domain name, which hides their identity from consumers, makes it easier for this practice to flourish without detection.

Visualizing two targeting scenarios

SCENARIO 1

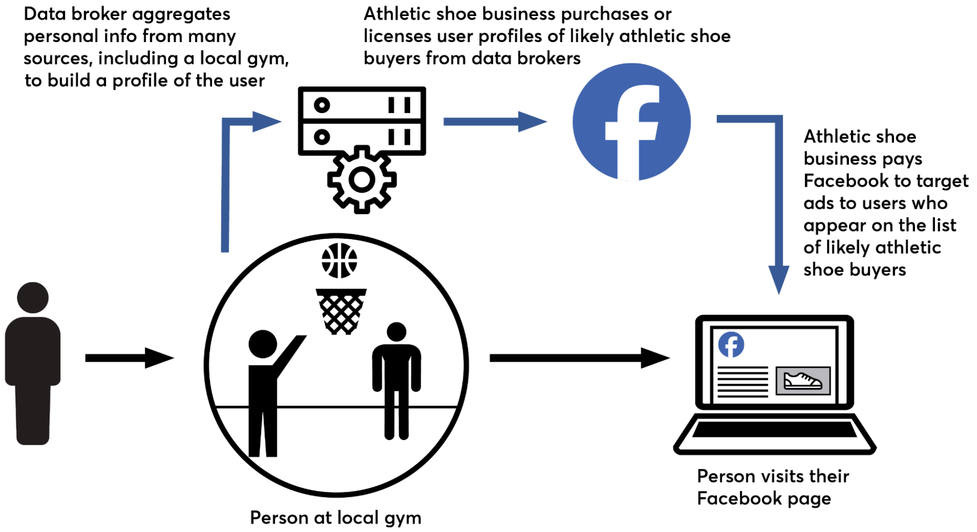
Apps and Stores Help Facebook Track You



[11] Jeremy B. Merrill, “How Facebook fueled a precious-metal scheme targeting older conservatives,” *Yahoo Finance,* November 19, 2019, https://finance.yahoo.com/news/facebook-fueled-precious-metal-scheme-110044886.html.
[12] Emma Fletcher, “Social media a gold mine for scammers in 2021,” *Federal Trade Commission,* January 25, 2022, https://www.ftc.gov/news-events/data-visualizations/data-spotlight/2022/01/social-media-gold-mine-scammers-2021.

SCENARIO 2

Business Works With Data Broker & Facebook to Track You



Data broker aggregates personal info from many sources, including a local gym, to build a profile of the user

Athletic shoe business purchases or licenses user profiles of likely athletic shoe buyers from data brokers

Athletic shoe business pays Facebook to target ads to users who appear on the list of likely athletic shoe buyers

Person at local gym

Person visits their Facebook page

Policy Recommendations

Many consumers will rightly be concerned about the extent to which their activity is tracked by Facebook and other companies, and may want to take action to counteract consistent surveillance. Based on our analysis of the sample data, consumers need interventions that will:

- Reduce the overall amount of tracking.
- Improve the ability for consumers to take advantage of their right to opt out under state privacy laws.
- Empower social media platform users and researchers to review who and what exactly is being advertised on Facebook.
- Improve the transparency of Facebook's existing tools.

We recommend the following policy interventions to address these needs:

Institute Data Minimization Provisions in Privacy Laws

On average, more than 2,000 different companies targeted each of the study volunteers, many of whom never directly interacted with many of the companies that shared their data. This kind of persistent, cross-contextual tracking would be far less common if privacy laws included the type of strong data minimization provisions Consumer Reports calls for in our Model Bill and in advocacy to regulators like the Federal Trade Commission. Data minimization would prohibit companies from collecting or processing consumer data beyond what is necessary to provide the service requested by the consumer. (For example, putting Cocoa Puffs in your shopping cart at Target.com should not result in the label "sugary cereal lover" being digitally appended to your identity and routinely shared with data brokers.)

Strong data minimization mandates in privacy laws are arguably the most important thing lawmakers can do to reinvigorate consumer privacy, because they would dramatically reduce the amount of data available for advertisers, data brokers, and others to collect.

Give Authorized Agents the Ability to Effectuate Rights Requests

Though data minimization provisions would address many consumer privacy issues, states have so far chosen instead to enact laws that institute an opt-out structure—meaning consumers must affirmatively instruct businesses not to sell or share their information.

With thousands of advertisers per consumer on Facebook alone, many of which were not easily identifiable, it is impractical for consumers who wish to remove themselves from the surveillance advertising ecosystem to contact, one-by-one, every company that has their data. In 2020 Consumer

Reports tested the usability of the opt-out framework and found the process to be both confusing and prohibitively laborious.[13] In some cases, consumers gave up out of frustration.

Since then, CR has advocated that opt-out laws at least include provisions that allow “authorized agents” to effectuate rights requests on consumers’ behalf. An authorized agent is a party designated by consumers to send opt-out requests on their behalf. Authorized agents can dramatically speed up the process of opting out by automating requests to businesses and managing the tedious paperwork that many opt-out processes require. We were first successful in convincing policymakers to embed authorized agent provisions in the California Consumer Privacy Act (CCPA), which has since been enhanced and replaced by the California Privacy Rights Act (CPRA). Other states have subsequently replicated the concept.

CR recently launched an authorized agent service called Permission Slip, which as of October 12, 2023, had sent more than a million opt-out requests on behalf of users. Policymakers should look to empower consumer-friendly actors, like authorized agents, whenever considering rights-based privacy laws.

Institute DSA-Style Ad Archive Mandates

A substantial portion of the companies in our sample data targeted only a single volunteer, and Facebook’s microtargeting abilities may allow fraudsters to target vulnerable individuals with advertisements for scams or illegitimate products while making the ads difficult to discover by researchers and regulators. In the U.S., researchers and regulators have no meaningful ability to review the specific ads that were served to consumers, leaving no opportunity to hold these entities accountable.

To address this, U.S. policymakers could learn from ongoing efforts in Europe to improve transparency in the advertising ecosystem. Through its recently approved Digital Services Act (DSA), the European Union requires large online platforms to preserve all advertisements shown to users in a searchable, publicly available archive for at least one year after the advertisement was presented to users for the last time.[14] This is a crucial oversight mechanism that allows consumers and researchers to conduct a historical analysis of the ad archives to search for ads that facilitate fraud, are discriminatory, or otherwise cause consumer harm.

American consumers deserve the same level of insight into the ad market as our European peers, instead of the mysterious black box we have today.

[13] Maureen Mahoney, “California Consumer Privacy Act: Are Consumers’ Digital Rights Protected?” *Consumer Reports,* October 1, 2020, https://advocacy.consumerreports.org/wp-content/uploads/2020/09/CR_CCPA-Are-Consumers-Digital-Rights-Protected_092020_vf.pdf.

[14] Digital Services Act, Article 39, *Official Journal of the European Union* 65, L 277 (2022), https://husovec.eu/wp-content/uploads/2022/10/Official-Version-OJ_L_2022_277_FULL_EN_TXT.pdf.

Bonus Recommendation: Facebook Should Ensure Transparency Measures Are Usable

Facebook's data download feature is a distinctly positive step toward providing consumers insight into and control over how their data is being used, and the company deserves credit for going beyond both its legal obligations and most of its competitors in this respect.

But this study shows that even this tool is inadequate. A significant portion of companies were allowed to list their name as either an indecipherable series of letters and numbers or in vague enough terms that it could not be easily associated with a corporate entity. Going forward, Facebook should ensure that data quality is standardized and intelligible to consumers. And ads, events, and consumer data like Custom Audiences should be easily traceable to their source via a standardized, readable format such as an https:// URL of a company home page or privacy policy.

Furthermore, one of the most troubling findings of this study is the prevalence of advertising service providers represented in the Facebook data. Service providers are entities that act on behalf of advertisers to deliver ads to consumers. The problem is that state privacy laws typically require only advertisers to honor privacy requests from consumers; advertising service providers might not be required to comply or to pass such requests to their client-advertisers when they receive requests directly from consumers. Thus, consumers are largely unable to use their Facebook data to ask advertisers to delete their data, opt them out of data sales and targeting, or provide access to their data (though in some cases, neither the advertiser nor the service provider may actually have direct access to the data; they may have leveraged data from yet another entity, such as a data broker).

On an even more basic level, the listing of service providers reduces transparency, because consumers will often be unable to determine what company paid for advertisements targeted at them. And adding a layer of complexity, some companies in our data sample are owned by global marketing holding companies that provide advertising services under many different agency names and are likely unfamiliar to many consumers.

Platforms like Facebook should ensure that their advertising transparency tools work in tandem with privacy laws to give consumers actionable choices regarding their data. For example, service providers should be required to list the name and contact information of client-advertisers so that consumers can direct rights requests to the correct business. (Alternatively, privacy laws could clarify that service providers must pass such requests to their client-advertisers or to reveal their client-advertisers upon request; but these options would add technical complexity, such as the need to tag each ad served with an identifier.)

Future Research Directions

Though the findings of this study are not based on a nationally representative data sample, we believe they do reflect consumer tracking behaviors that are broadly typical of and commonplace across the surveillance economy. Researchers studying the surveillance economy should continue to leverage Facebook's data. In particular, we hope the following research ideas will be explored:

Effectiveness of Global Privacy Control

Aside from authorized agents such as the Permission Slip by CR mobile app, the only practical way for consumers to opt out of data sales with retailers and brands they do business with is Global Privacy Control (GPC), a browser preference or extension that broadcasts a consumer's opt-out intent to every website they visit. These types of universal opt-out signals are legally enforceable under a growing number of state laws.[15]

Researchers should measure the extent to which consumer opt-outs broadcast via GPC are actually taking effect. Such a study might involve recruiting volunteers, confirming that volunteers have installed a browser or extension that sends GPC, inviting volunteers to visit the sites of companies that they do business with, then collecting Facebook data and comparing the time stamps of events sent to Facebook to the date on which GPC was installed. This process would enable researchers to determine the extent and efficacy of GPC compliance.

A Nationally Representative Study of Facebook Data

A promising direction for future research would be a study similar to this one with a set of consumers more representative of the U.S. population as a whole. Among other potential benefits of such an effort would be the opportunity to analyze how the surveillance economy affects different populations differently.

Additional Platforms

To build a more comprehensive view of the surveillance economy, further research could leverage other platforms' data access tools. Such research would help detect new surveillance technologies and enable consumer research and advocacy organizations to better influence the practices of advertising and publishing businesses to the benefit of consumers.

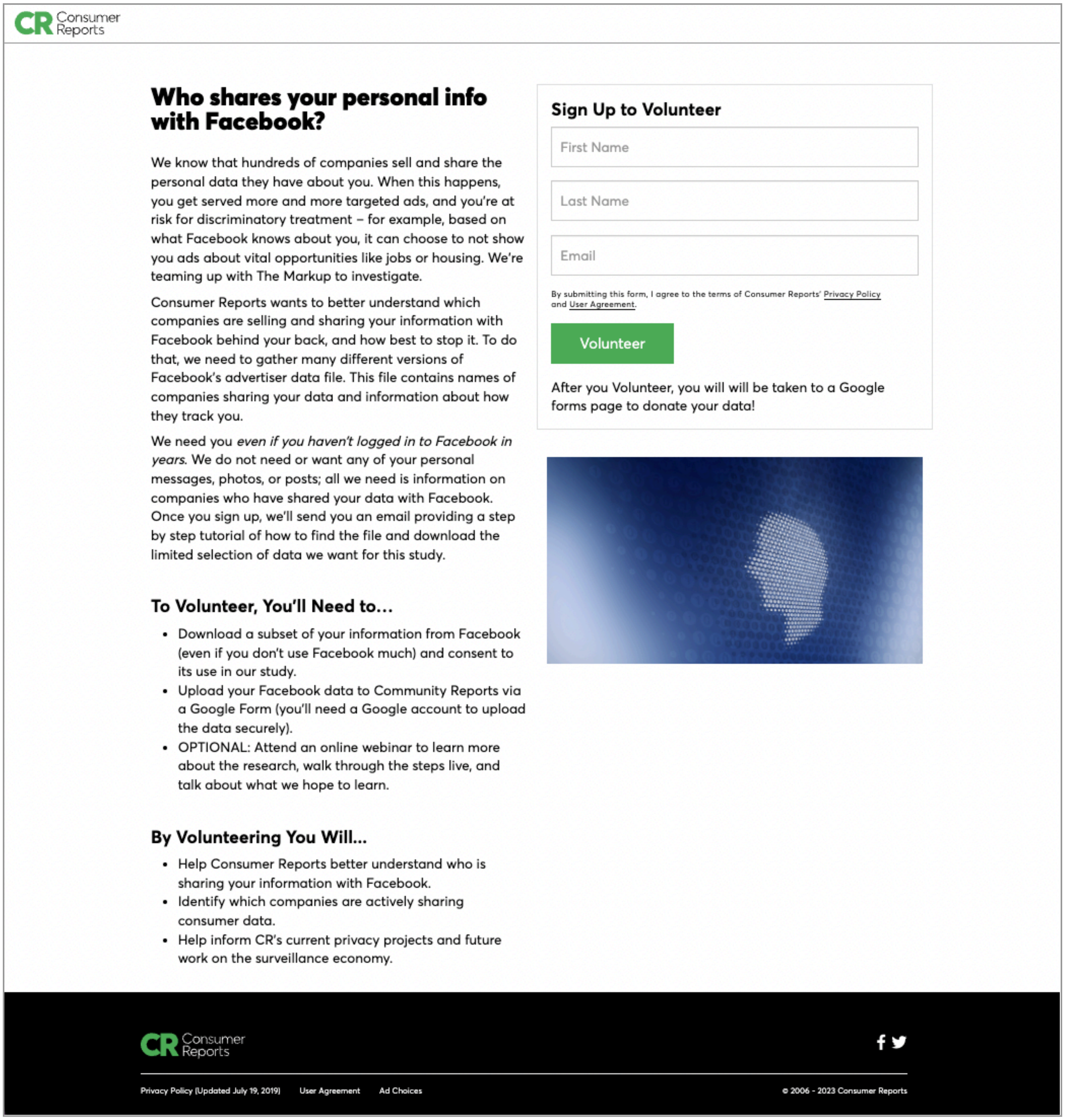
[15] As of October 10, 2023, state privacy laws that require businesses to honor universal opt-out signals like those sent by GPC include: California, Colorado, Connecticut, Delaware, Montana, Oregon, and Texas.

Acknowledgements

The Facebook Surveillance Study team would like to thank Alan Smith and Daniella Raposo, who significantly contributed to the study and this report. We would also like to acknowledge Sisi Wei and colleagues at The Markup for their collaboration, the Cornell Tech PiTech Siegel PhD Impact Fellowship Program for placing Fengyang Lin with CR, and Cristobal Cheyre from Cornell Bowers CIS for useful discussion and comments.

Appendix

A. Facebook Surveillance Study Enrollment Page on CR's Volunteer Website



Who shares your personal info with Facebook?

We know that hundreds of companies sell and share the personal data they have about you. When this happens, you get served more and more targeted ads, and you're at risk for discriminatory treatment – for example, based on what Facebook knows about you, it can choose to not show you ads about vital opportunities like jobs or housing. We're teaming up with The Markup to investigate.

Consumer Reports wants to better understand which companies are selling and sharing your information with Facebook behind your back, and how best to stop it. To do that, we need to gather many different versions of Facebook's advertiser data file. This file contains names of companies sharing your data and information about how they track you.

We need you *even if you haven't logged in to Facebook in years.* We do not need or want any of your personal messages, photos, or posts; all we need is information on companies who have shared your data with Facebook. Once you sign up, we'll send you an email providing a step by step tutorial of how to find the file and download the limited selection of data we want for this study.

To Volunteer, You'll Need to…

- Download a subset of your information from Facebook (even if you don't use Facebook much) and consent to its use in our study.
- Upload your Facebook data to Community Reports via a Google Form (you'll need a Google account to upload the data securely).
- OPTIONAL: Attend an online webinar to learn more about the research, walk through the steps live, and talk about what we hope to learn.

By Volunteering You Will...

- Help Consumer Reports better understand who is sharing your information with Facebook.
- Identify which companies are actively sharing consumer data.
- Help inform CR's current privacy projects and future work on the surveillance economy.

Sign Up to Volunteer

First Name

Last Name

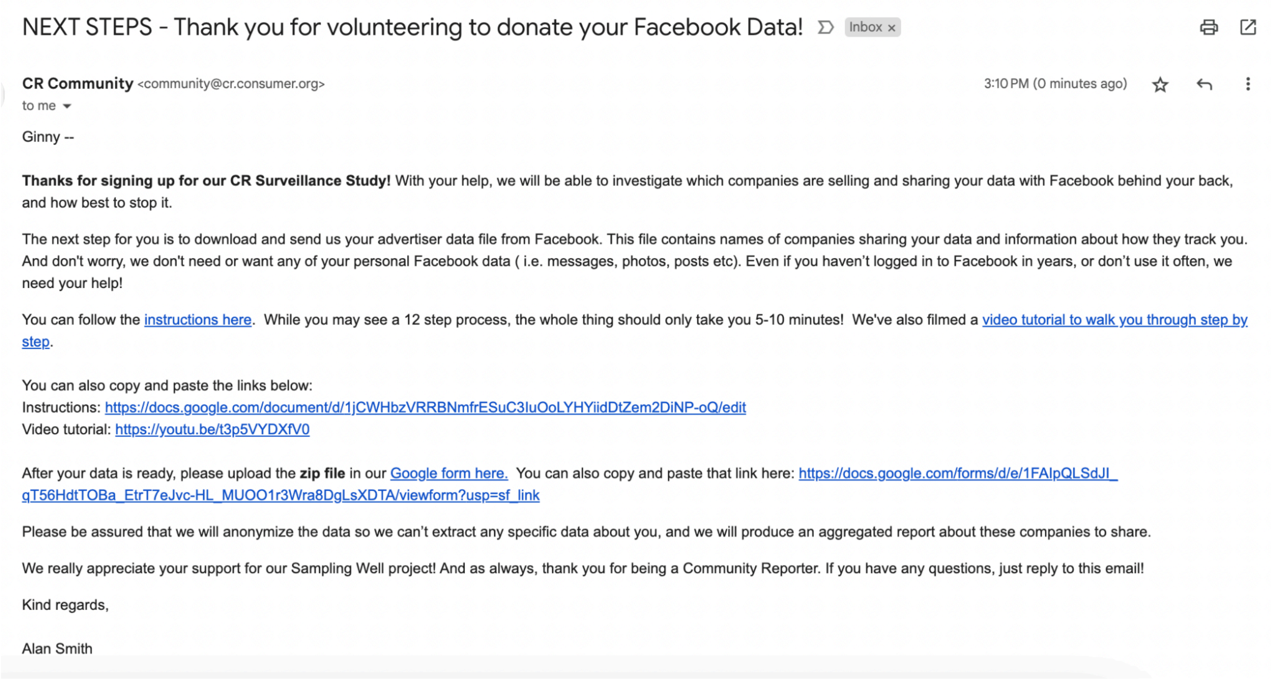
Email

By submitting this form, I agree to the terms of Consumer Reports' Privacy Policy and User Agreement.

After you Volunteer, you will will be taken to a Google forms page to donate your data!

CR Consumer Reports
Privacy Policy (Updated July 19, 2019)
User Agreement
Ad Choices
© 2006 - 2023 Consumer Reports

B. Facebook Surveillance Study - Enrollment Confirmation Email



NEXT STEPS - Thank you for volunteering to donate your Facebook Data!

CR Community <community@cr.consumer.org>
to me

Ginny --

Thanks for signing up for our CR Surveillance Study! With your help, we will be able to investigate which companies are selling and sharing your data with Facebook behind your back, and how best to stop it.

The next step for you is to download and send us your advertiser data file from Facebook. This file contains names of companies sharing your data and information about how they track you. And don't worry, we don't need or want any of your personal Facebook data (i.e. messages, photos, posts etc). Even if you haven't logged in to Facebook in years, or don't use it often, we need your help!

You can follow the instructions here. While you may see a 12 step process, the whole thing should only take you 5-10 minutes! We've also filmed a video tutorial to walk you through step by step.

You can also copy and paste the links below:
Instructions: https://docs.google.com/document/d/1jCWHbzVRRBNmfrESuC3IuOoLYHYiidDtZem2DiNP-oQ/edit
Video tutorial: https://youtu.be/t3p5VYDXfV0

After your data is ready, please upload the **zip file** in our Google form here. You can also copy and paste that link here: https://docs.google.com/forms/d/e/1FAIpQLSdJI_qT56HdtTOBa_EtrT7eJvc-HL_MUOO1r3Wra8DgLsXDTA/viewform?usp=sf_link

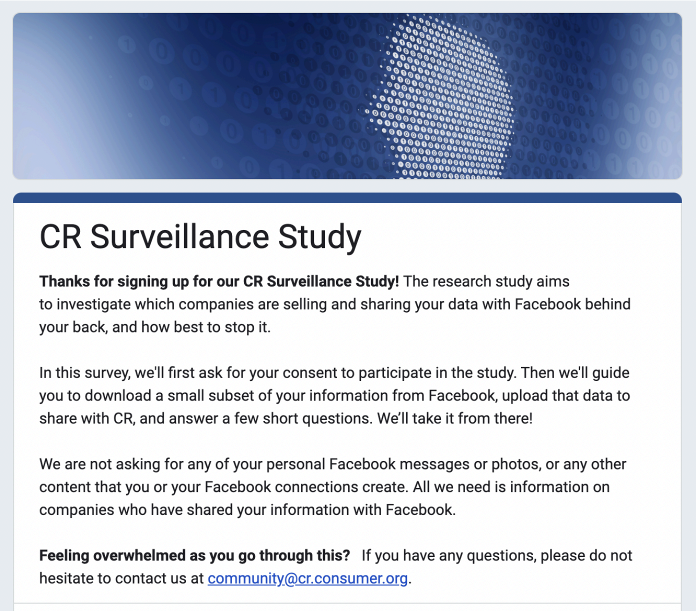
Please be assured that we will anonymize the data so we can't extract any specific data about you, and we will produce an aggregated report about these companies to share.

We really appreciate your support for our Sampling Well project! And as always, thank you for being a Community Reporter. If you have any questions, just reply to this email!

Kind regards,

Alan Smith

C. Facebook Surveillance Study - Data Donation Form



CR Surveillance Study

Thanks for signing up for our CR Surveillance Study! The research study aims to investigate which companies are selling and sharing your data with Facebook behind your back, and how best to stop it.

In this survey, we'll first ask for your consent to participate in the study. Then we'll guide you to download a small subset of your information from Facebook, upload that data to share with CR, and answer a few short questions. We'll take it from there!

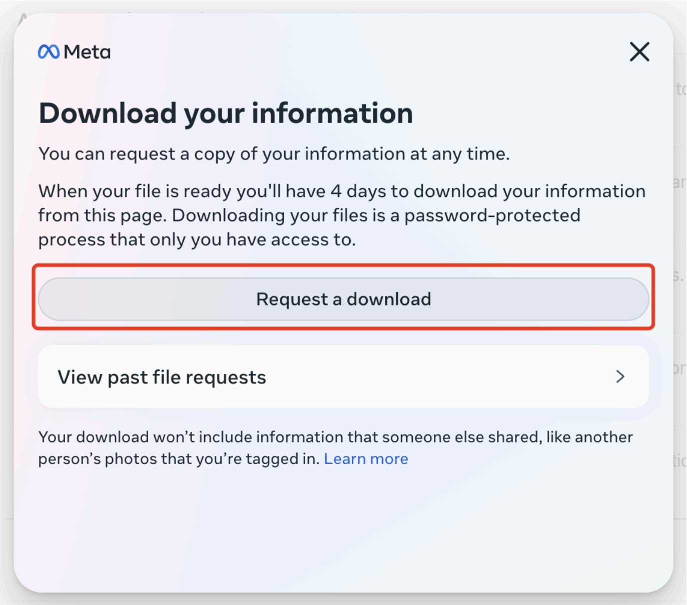
We are not asking for any of your personal Facebook messages or photos, or any other content that you or your Facebook connections create. All we need is information on companies who have shared your information with Facebook.

Feeling overwhelmed as you go through this? If you have any questions, please do not hesitate to contact us at community@cr.consumer.org.

D. Facebook Surveillance Study - Instructions for Downloading Your Data From Facebook

Thank you for your work in this study! In all likelihood, the process of logging in to Facebook and navigating to our first step will look like Version A, below. On the off chance that you have a different flow and Version A is not appearing like you expect, check out Version B, further down.
You can either watch a video walk-through here or follow the step-by-step instructions below. If you have any questions about the steps, please don't hesitate to contact Fengyang Lin.

1. Using a desktop browser, go to http://facebook.com/dyi. Log in to your Facebook account to continue.
2. Click the button "Request a download."



Meta

Download your information

You can request a copy of your information at any time.

When your file is ready you'll have 4 days to download your information from this page. Downloading your files is a password-protected process that only you have access to.

Request a download

View past file requests

Your download won't include information that someone else shared, like another person's photos that you're tagged in. Learn more

If you are not seeing this image, you should go to version B.

3. Under "Select information," go to "Select types of information."
4. Scroll down to the bottom and click the two boxes for "Apps and websites off of Facebook" and "Ads information." Click "Next" to proceed to the next page.
5. Under "Select file options" go to "Date range (required)" and select "last 3 years". Click "Save" to proceed.
6. Under "Select file options" go to "format" and select "JSON". Click "Save" to proceed.
7. Click the button "Submit request"
8. Scroll up to the top of the page. You should see an entry for your request in progress that looks like this:

Requested Jun 12, 2023
Your account name
Facebook
under “Pending Download.” At this point you should also get an email notification stating “Thanks for requesting a copy of your Facebook information. Once we've finished creating your file, we'll send you another email letting you know it's complete and ready to be downloaded.”

9. After about 2-10 minutes (or within a day), your data will be ready. You can either:
    a. Refresh the page (http://facebook.com/dyi). When your “pending” turns to “download” you are good to go OR
    b. Wait to receive another email from Meta that will let you know that your data is ready. You can then follow the link in the email, which should take you to: https://www.facebook.com/dyi/?tab=all_archives
10. Click “Download” and re-enter your password if prompted.
11. Click “Confirm” if prompted.
12. There will be a new file ending in “.zip” in your Downloads folder. This is the file that you will need to upload. You can upload your zip file in the Google form here (You will need a Google account to submit your data securely).
    a. If your zip file is unzipped automatically, you can compress it by right clicking the folder and select “compress [your folder]”, then upload the zip file.

<<If you are interested in seeing what is in the file, you can unzip it and check it out yourself.>>

If you can’t get to “Download Your Information”

1. Using a desktop browser, visit facebook.com. If you are not already logged in, log in.
2. Click on your profile picture on the upper right.
3. When the menu appears, select “Settings & Privacy”
4. Select “Privacy Center”
5. Scroll down to “Information: Manage information across your activity and accounts” and click “Get started”
6. Click “Download your information”

Then continue with the steps above.

E. Facebook Surveillance Study - 100 Most Frequently Occuring Companies

1. LiveRamp
2. Acxiom
3. Experian Marketing Services - Audiences
4. Hearts & Science
5. ODC CA
6. Epsilon Audience Data Provider
7. The Home Depot
8. OMD USA
9. 4C
10. Amazon.com
11. Neustar FB Syndication
12. Starcom USA
13. Walmart.com
14. Basis Technologies
15. TargetSmart
16. Merkle Incorporated
17. UM NY
18. PayPal
19. 360i
20. ADARA
21. Macy's
22. iProspect Detroit
23. Nordstrom
24. Spark Foundry
25. Zeta
26. Ovative Group
27. Decoded Advertising
28. Harris Teeter
29. Foursquare City Guide
30. Illinois Lottery
31. Massage Envy
32. Bubly
33. Vizeum US
34. Barkley
35. HongKong Zoom Interactive Network Marketing Technology Limited
36. Endless Pools
37. MullenLowe U.S.
38. Spark Foundry USA
39. Hopper
40. Wavemaker USA
41. Bully Pulpit Interactive
42. J3
43. Bayer Consumer Health Canada
44. iProspect NY
45. Kepler
46. GolfNow.com
47. Iron Store
48. Predictive Media Analytics, LLC
49. Test Page
50. Pep Boys
51. Gap Inc
52. WBCI
53. GCommerce
54. SmileDirectClub
55. AdParlor
56. Consumer Reports[16]
57. Datonics
58. Drive Toyota
59. Heineken USA: Social Media
60. Drury Hotels
61. Rezonate Media
62. Dentsu X North America
63. Faraday
64. Initiative
65. Hapulico
66. Digitas North America
67. Deep Root Analytics
68. Mediacom USA
69. Semcasting, Inc.
70. Coupons.com
71. Etsy
72. Haleon
73. Discover
74. Launchpad Ignite
75. Planet Partnership
76. Meetsocial HK Digital Marketing Co.
77. New Ad Accounts for Real Chemistry
78. Tuyen Bus
79. Uber
80. Lowe's Home Improvement
81. M+R
82. I360 LLC
83. Nordstrom Rack
84. Essence
85. Windfall Data
86. Real Chemistry
87. Bed Bath & Beyond
88. Rising Tide Interactive
89. truth
90. Ford Motor Company
91. Huntington National Bank
92. Merkle Data Partner
93. eBay.co.uk
94. HMI Social
95. Sojern
96. Walt Disney World
97. Conill Advertising
98. Change Research
99. Old Navy
100. Aero

[16] CR appears in this list because, as noted on page 9, we believe study participants were disproportionately CR members, to whom CR markets through these platforms. Our full privacy policy is available at CR.org/privacy.