Online Consumer Data Collection

October 31, 2022
Online Consumer Data Collection

Large amounts of consumer data can be collected, processed, and analyzed by operators of websites and mobile applications (apps) and third parties, which are entities other than the website or app primary operator (e.g., data brokers). Operators collect data for multiple purposes, including providing services, selling user data to third parties, or sending targeted ads directed to specific individuals.

The value of consumer data often comes from identifying users and linking their data from various sources to a common identifier. Operators can identify individuals using their personally identifiable information (PII)—such as name, address, or date of birth—and other identifiers, such as those associated with a particular device. Some federal laws prevent entities from collecting or sharing specific types of PII or identifiers in certain circumstances. However, in recent decades, the ubiquity of non-PII (data not directly linked to an individual’s identity, including anonymized or aggregated data) and the emergence of new data collection and tracking tools have made it easier to identify individuals.

Consumer data can be collected using various data collection and tracking tools, such as cookies, pixels, device and browser fingerprinting, application programming interfaces (APIs), and software development kits (SDKs). These tools can continuously collect different types of data, including identifiers, even when the consumer visits a different website or app. Some of these tools are necessary for websites and apps to provide services, and others typically are used for online advertising. Some of these tools can be used to help develop a website or app and offer services provided by other operators, which can increase competition. They also can be used to collect large amounts of data, particularly by third parties, causing some to raise consumer data privacy concerns.

Although some intermediaries—such as web browsers and operating systems—offer anti-tracking tools allowing users to limit the data collected by websites, apps, and third parties, they can also make it difficult for individuals to access these tools. Most web browsers offer “private browsing” or “incognito” modes that allow users to refuse cookies, and some browsers ban third-party cookies. Apple released an iPhone operating system update that requires apps to ask for permission before tracking user activity. Google has restricted access to a consumer identifier in its advertising (ad) network and banned certain ad blocking apps, some of which were reinstated, from its app store. Intermediaries such as these may be able to obtain a competitive advantage if they are able to collect data that others cannot access.

The United States does not have a comprehensive federal data protection law, although multiple federal statutes create data protection obligations for particular types of information or for entities engaged in certain activities. For example, the Children’s Online Privacy Protection Act (COPPA, 15 U.S.C. §§6501-6506) requires online services directed to children under 13 years of age that collect personal information to notify users about the data collection, receive parental consent, and maintain “reasonable procedures” to protect the data. COPPA is enforced by the Federal Trade Commission (FTC), which has brought enforcement actions against companies for their consumer data collection practices under its authority to prevent “unfair or deceptive acts or practices in or affecting commerce.” For example, it has taken action against companies for allegedly handling personal information in a way that contradicts their privacy policies. The FTC is also considering whether it will implement new rules on data collection and security to protect consumers’ data and privacy.

A comprehensive federal data protection law may have differential effects. For example, prohibiting the collection of consumer data would provide the highest level of data security but could also prevent some operators from providing their services or degrade the quality of their services. Prohibiting the transfer or sale of consumer data might provide some data protection, depending on the operators that consumers are willing to share their data with, but could also further entrench incumbents that have already collected large amounts of consumer data. Increasing transparency on the collection and use of consumer data—particularly if doing so would heighten public scrutiny of the operator—might incentivize operators to adjust their current practices but might not significantly alter operators’ behavior if consumers continue to use the website or app.

Some Members of the 117th Congress have introduced bills to create a comprehensive data protection law. If Congress chooses to pursue legislative action, it may consider (1) if the legislation would broadly address consumer data collection or focus on specific types of data; (2) whether to implement requirements for operators or allow consumers to determine which entities can receive their data; (3) if the legislation would preempt state laws; (4) whether to include a private right of action; and (5) potential unintended effects.
Contents

Introduction ................................................................................................................................. 1
How Consumer Data Is Collected Online ............................................................................... 2
   Identifying and Linking Consumer Data ................................................................................. 3
      Personally Identifiable Information ................................................................................. 3
      Persistent Identifiers ........................................................................................................ 4
   Selected Data Collection and Tracking Tools ........................................................................ 5
      Cookies .............................................................................................................................. 7
      Tracking Pixels ................................................................................................................ 7
      Device and Browser Fingerprinting ................................................................................... 9
      Application Programming Interfaces ............................................................................... 9
      Software Development Kits ............................................................................................. 10
   Anti-tracking Tools from Intermediaries ............................................................................. 11
   Federal Trade Commission Enforcement of Consumer Data Collection .......................... 14
Potential Effects of Consumer Data Protection Requirements ............................................. 16
   Prohibit the Collection of Consumer Data ......................................................................... 16
   Prohibit the Transfer or Sale of Consumer Data ................................................................ 17
   Increase Transparency on the Collection and Use of Consumer Data .............................. 18
Policy Considerations for Congress ..................................................................................... 18

Figures

   Figure 1. Overview of Selected Data Collection and Tracking Tools................................. 6

Contacts

   Author Information .............................................................................................................. 22
Introduction

Over the past few decades, developments in information and communication technologies have enabled operators of websites and mobile applications (apps) and third parties to collect, process, and analyze large amounts of consumer data. These developments have led to new products and services that collect and provide real-time information, including navigation services that update users with current traffic conditions and suggest faster routes; health-related services that can track physical activity or help users determine whether they are ovulating; and finance-related services that can help users analyze their spending habits.

Website and app operators may collect user data for multiple purposes, such as providing or improving their services, selling user data to third parties, or sending targeted advertisements. Some data collection may benefit consumers when, for example, it allows consumers to access services at a cheaper price or for free. However, some consumers may be unaware that their data are collected and how that data are used.

The United States does not have a comprehensive federal data protection law, although multiple federal statutes create data protection obligations for particular types of information or for entities engaged in certain activities. Some Members of the 117th Congress have introduced bills to create a comprehensive data protection law, and five states have implemented their own comprehensive consumer data protection laws. In October 2022, the White House Office of Science and Technology Policy (OSTP) released Blueprint for an AI Bill of Rights, which includes recommended data privacy protections for all Americans. Although OSTP does not have regulatory or enforcement authority, it notes, “where law or policy does not already provide guidance, the Blueprint should be used to inform policy-making to fill those gaps.”

---

1 In this report, consumer data includes any data about an individual, including personally identifiable information (PII), demographic data, behavioral data (e.g., search history, buttons clicked on a website), and purchase data. Operator is used to describe the entity—such as a firm, nonprofit organization, or individual—that controls and operates the website or app. Third parties are entities other than the primary operator of the website or app; see text box in “Identifying and Linking Consumer Data.”


4 For example, see the American Data Privacy and Protection Act (H.R. 8152) and the Consumer Online Privacy Rights Act (S. 3195). For comparisons of some of these bills, see CRS Legal Sidebar LSB10776, Overview of the American Data Privacy and Protection Act, H.R. 8152, by Jonathan M. Gaffney, Eric N. Holmes, and Chris D. Linebaugh.

5 The five states are California, Colorado, Connecticut, Utah, and Virginia.


This report provides a brief overview of methods for consumer data collection on websites and mobile apps. It then discusses some efforts to limit consumer data collection and considers the potential effects of a comprehensive federal data protection law on website and app operators, which may be applicable to other types of data collection as well.

How Consumer Data Is Collected Online

Each day, individuals interact with digital technologies that collect vast amounts of their data—including self-reported personal information and online behavioral data—without their knowledge. Website and mobile app operators, as well as third parties (see text box), use these data to make inferences and decisions, inform machine learning models, and build detailed user profiles for targeted advertising or personalized content, among other purposes. This section discusses some of the tools used to collect consumer data and track individuals online.

<table>
<thead>
<tr>
<th>Sources of Data Collection: First vs. Third Party</th>
</tr>
</thead>
</table>
| In this report, first-party data refers to data that the primary operator of a website or app collects directly from its users. Some first-party data may be self-reported, that is, actively provided by the individual creating an account (e.g., name, age) or uploading content. Third-party data is collected by entities other than the primary operator of a website or app. Some data collection tools can blur the distinction between first and third parties. For example, some providers of online advertising (ad) services embed code on websites that use their services; the ad service would be considered a third party in this report because it is not the website's primary operator. First- and third-party data collection can raise different concerns. First-party data may be collected without a user's express consent or knowledge; this is nearly always the case with third-party data unless the user resides in a state or country that requires entities to obtain consumer consent. Additionally, some website and app operators might supplement their first-party data with third-party data. For example, an operator might purchase or exchange user data with third parties; a firm could also acquire another firm to obtain its consumer data. Nearly all of the most popular apps and websites provide data to third parties. One study found that the top one million websites send data to at least 34 third parties on average. Another study found that nearly one million

---

9 This report does not discuss data collected by internet service providers (ISPs) or other internet-connected devices, such as smart devices, that may present similar data privacy concerns. For more information, see FTC, A Look At What ISPs Know About You: Examining the Privacy Practices of Six Major Internet Service Providers. Staff Report, October 21, 2021, at https://www.ftc.gov/reports/look-what-isps-know-about-you-examining-privacy-practices-six-major-internet-service-providers; CRS In Focus IF11239, The Internet of Things (IoT): An Overview, by Patricia Moloney Figliola; and CRS Report R44227, The Internet of Things: Frequently Asked Questions, by Patricia Moloney Figliola.

10 Some data protection regulations use different terms for first-party data collection. For example, the European Union’s (EU’s) General Data Protection Regulation (GDPR) uses the terms controller and processor for data collection methods that would be considered first party in this report. See Article 4(7, 8) of the GDPR, at https://gdpr-info.eu/art-4-gdpr/.

11 The EU’s GDPR defines third party as “a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data.” See Article 4(10) of the GDPR, at https://gdpr-info.eu/art-4-gdpr/.

12 For example, the EU’s GDPR requires the controller to demonstrate that the data subject has consented to processing of his or her personal data. For the definition of consent, see Article 4(11) of the GDPR, at https://gdpr-info.eu/art-4-gdpr/. For the definition of conditions for consent, see Article 7 of the GDPR, at https://gdpr-info.eu/art-7-gdpr/.

Identifying and Linking Consumer Data

The value of consumer data often comes from identifying users and combining their data from various sources. This is possible, in part, through the ubiquity of personally identifiable information (PII) and unique identifiers, as well as identifying individuals from non-PII, such as aggregated or anonymized data. The ability to identify users enables website and app operators to combine data and track user activity across devices.  

Personally Identifiable Information

PII is “information that can be used to distinguish or trace an individual’s identity, either alone or when combined with other information that is linked or linkable to a specific individual.” Examples of PII include names, addresses, dates of birth, telephone numbers, Social Security numbers, financial information, medical records, and employment information. This information can be used to uniquely identify an individual across many different contexts and devices. 


The study on Google’s trackers included its advertising services DoubleClick, Admob, and Adsense, and the study on Facebook’s trackers included Facebook, Liverail, and Lifetreet. See Binns et al., “Third Party Tracking.” 


numbers, precise location information, and biometrics. Multiple federal statutes regulate the use of specific types of PII in certain circumstances.

The proliferation of non-PII and its widespread collection have raised additional privacy concerns. For the purposes of this report, non-PII is all other kinds of data not directly linked to an individual’s identity, including anonymized or aggregated data. The emergence of new data tracking tools, data analysis, and statistical techniques in recent decades has made it possible to identify individuals with a high degree of accuracy from non-PII. For example, one study found that four approximate location data points are sufficient to identify 95% of individuals in a dataset of one and a half million people. The Federal Trade Commission (FTC) has said, “the traditional notion of what constitutes PII versus non-PII is becoming less and less meaningful and should not, by itself, determine the protections provided for consumer data.”

**Persistent Identifiers**

In the context of online tracking, a persistent identifier—also called unique identifier—links a specific device or user with a unique ID. A persistent identifier can detect a specific device or user across different websites or online services and target content or advertisements to specific users. Companies may also use persistent identifiers to share and combine user data with one another. Some persistent identifiers may be necessary to provide certain online services or functionalities, such as retaining unpurchased items in an online shopping cart. Persistent identifiers are intended to be long-lasting and track users over time. Some may reset or expire after a certain period, while others can identify users over a number of years.

Examples of persistent identifiers include device identifiers, such as Media Access Control (MAC) addresses, and web-based identifiers, such as IP addresses; third parties can also assign their own identifiers to users. Mobile advertising IDs (MAIDs) are popular persistent identifiers.

---

22 While PII typically includes these forms of data, it may include other types of data. Shaun Donovan, *Preparing for and Responding to a Breach of Personally Identifiable Information*, Executive Office of the President, Office of Management and Budget, M-17-12, January 3, 2017, at https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2017/m-17-12_0.pdf.

23 For example, the Gramm-Leach-Bliley Act prohibits financial institutions from sharing “nonpublic personal information,” and the Health Insurance Portability and Accountability Act prohibits “covered entities” from using or sharing “protected health information,” which includes “individually identifiable health information.” For more information about these laws, see CRS Report R45631, *Data Protection Law: An Overview*, by Stephen P. Mulligan and Chris D. Linebaugh.


27 Persistent identifiers may also refer to long-lasting references to digital resources, often used in library science and archival research.

28 A Media Access Control (MAC) address is a long string of numbers and letters that uniquely identifies every device connected to a network and typically does not change.

29 An IP address is a numerical identifier assigned to a computer or device that connects to the Internet. IP addresses can be used for location tracking by linking an individual IP address to a particular device that is associated with a specific individual. See FTC, *Self-Regulatory Principles*.

consisting of a long string of numbers and letters that uniquely identify a user’s mobile device through its operating system. Apple’s iPhone operating system (iOS) uses an “Identifier for Advertisers” (IDFA) as its MAID, and Google’s mobile operating system (Android) uses “Google Advertising ID.” MAIDs enable targeted ads to be sent to a specific device based on data collected about the user. Some data brokers have also used MAIDs to identify and sell user data, linking MAIDs to PII or other persistent identifiers.

Persistent identifiers can constitute PII in certain circumstances. For example, in 2013, the FTC published an amended rule adding “persistent identifier” to the definition of personal information in the Children’s Online Privacy Protection Act (COPPA). In 2016, then-FTC Chair Edith Ramirez stated, “as a result of the increased ability to identify consumers, we now regard data as personally identifiable 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, and retail loyalty card numbers meet this test.”

Selected Data Collection and Tracking Tools

Website and app operators use various tools to collect consumer data. Some of these tools can continuously collect data, even when the user visits a different website or app.

This section discusses selected examples of data collection tools, including cookies, pixels, fingerprinting, application programming interfaces (APIs), and software development kits (SDKs). Figure 1 provides an overview of these tools. This list is illustrative, not exhaustive. Each tool can collect various types of data, including persistent identifiers.

---

31 NIST defines an operating system as “a program that runs on a computer and provides a software platform on which other programs can run.” Karen Kent et al., Guide to Integrating Forensic Techniques into Incident Response, NIST, Special Publication 800-86, August 2006, at https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-86.pdf.

32 Bennett Cyphers and Gennie Gebhart, “Behind the One-Way Mirror: A Deep Dive Into the Technology of Corporate Surveillance,” Electronic Frontier Foundation (EFF), December 2, 2019, at https://www.eff.org/wp/behind-the-one-way-mirror (Cyphers and Gebhart, “Behind the One-Way Mirror”). Google Advertising ID has also been referred to as Android Advertising ID.


36 For example, this report does not discuss data collection through device sensors, such as motion or light sensors.
Figure 1. Overview of Selected Data Collection and Tracking Tools

<table>
<thead>
<tr>
<th>Definitions</th>
<th>Providers</th>
<th>Tracking &amp; Anti-Tracking</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookies</td>
<td>1st-party cookies are created by a website visited by a user</td>
<td>May be banned by certain web browsers, such as FireFox or Safari</td>
<td>Analytics cookies (e.g., Google Analytics cookies used to distinguish visitors for two years)</td>
</tr>
<tr>
<td>Pixels</td>
<td>3rd-party cookies are created by an entity other than the operator of the website a user is visiting</td>
<td>May be used to evade web browser bans on cookies</td>
<td>Advertising cookies (e.g., Cookies used by advertisers to measure performance of ads)</td>
</tr>
<tr>
<td>Fingerprinting</td>
<td>Code is often provided by third-parties with ad services, primarily for targeted advertising</td>
<td>May be used to evade web browser cookie bans</td>
<td>Meta Pixel</td>
</tr>
<tr>
<td>APIs</td>
<td>Some third-parties provide fingerprinting and device identification services to websites</td>
<td>May be used to evade ad blockers</td>
<td>TikTok Pixel</td>
</tr>
<tr>
<td>SDKs</td>
<td>Certain APIs may be provided by large companies or included in SDKs</td>
<td>Certain APIs may be used to evade ad blockers, since they are less impacted than pixels</td>
<td>Fingerprint formed from data such as:</td>
</tr>
<tr>
<td></td>
<td>Complete toolkit of software and code, often including APIs</td>
<td></td>
<td>Screen size &amp; resolution</td>
</tr>
<tr>
<td></td>
<td>Helps develop and integrate mobile apps with other services</td>
<td></td>
<td>Browser information</td>
</tr>
<tr>
<td></td>
<td>Certain SDKs may be provided by large companies for interoperability and data sharing</td>
<td></td>
<td>Operating system details</td>
</tr>
<tr>
<td></td>
<td>Some SDKs may be provided by the manufacturer of a hardware platform, operating system, or programming language</td>
<td></td>
<td>Analytics APIs provide website operators data about their visitors (e.g., Facebook’s Ads Insights APIs)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Feature APIs allow websites and apps to integrate another service’s feature into their own offering (e.g., API for Google search within a separate website)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Analytics or Monetization SDKs (e.g., Google Mobile Ads SDK for Ad Manager; Firebase SDK for Google Analytics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Platform or language-specific SDKs to develop apps (e.g., iOS SDK)</td>
</tr>
</tbody>
</table>

Source: Congressional Research Service.

Notes: APIs = application programming interfaces; SDKs = software development kits; and VPNs = virtual private networks.
Cookies

A cookie is a small information file stored on a user’s device when a user visits a website. Cookies can store a wide range of information, including a persistent identifier known as a cookie ID. The lifetime of a cookie can vary; some cookies are deleted as soon as the user leaves a website, while others remain on the user’s device for a predetermined amount of time.

First-party cookies—created by a website visited by a user—can be used to personalize a website’s features and may be necessary to provide certain services. For example, a first-party cookie might save a user’s preferred language setting or unpurchased items in a shopping cart.

Data privacy concerns have focused primarily on third-party cookies, also called tracking cookies, which are created by an entity other than the operator of the website a user is visiting. For example, if a user visits a website using a third-party advertising (ad) network, the third party can place cookies on the user’s browser. Third-party cookies enable advertisers and websites to track user behavior and build user profiles in order to serve personalized ads.

Many websites serve ads from different ad networks, meaning multiple third parties can use and compare their cookies by combining data using cookie IDs. This process, known as “cookie syncing,” enables third parties that track users’ online activities to merge data without users’ knowledge. Some studies have found that the majority of such third parties are involved in cookie syncing.

Tracking Pixels

A tracking pixel—also called a pixel, marketing pixel, image tag, or tracking code—is a small piece of code that an operator can add to its website to collect and send data to a third party, typically an ad network that provides the pixel code. The code creates an “image” the size of a single pixel, the smallest unit of a digital image or graphic, that is not visible to users. When a user visits a web page, their actions automatically trigger the pixel to send data to the pixel provider’s servers. Pixels can be embedded in websites, ad banners, and emails.

Typically, Pixels are used for online advertising to collect data on each website user, including user behavior, such as clicks or searches, IP address, and device model and browser information.

41 Cyphers and Gebhart, “Behind the One-Way Mirror.”
42 Steven Englehardt and Arvind Narayanan, “Online Tracking.”
43 Ibid.
Data are sent to the provider of the tracking pixel (e.g., Meta Pixel, TikTok Pixel), which can be used to target ads to specific users and obtain metrics on the efficacy of ads.

Pixels are often used in combination with other data collection and tracking tools, such as first- and third-party cookies. For example, Meta has stated that its pixel relies on Facebook cookies, which enable Meta to match a company’s website visitors to their respective Facebook accounts. Additionally, if Meta Pixel is present on a website, Meta can then place a first-party cookie on the user, expanding its data collection tools. Web browsers cannot disable pixels easily, unlike cookies, because they send data directly to the pixel provider’s servers.

In recent years, pixels have become increasingly popular as an alternative to tracking cookies. In 2018, Meta CEO Mark Zuckerberg testified that there were more than two million of its pixels installed on websites; the number has likely grown since. One investigation found that 30% of the 100,000 most popular websites embed Meta Pixel.

Privacy advocates have criticized the use of pixels for sharing sensitive user information without users’ knowledge and potentially violating data privacy laws. For example, in 2022, the news outlet MarkUp found that Meta collects sensitive user data through Meta Pixels embedded in the websites of hospitals, crisis pregnancy centers, and federal government agencies. Another investigation found that Meta Pixel and TikTok Pixel collect data from thousands of popular websites before a user submits information to the website via an online form, such as signing in or signing up for a service. These pixels collected personal information, including email addresses and passwords, even if the user never submitted the form or provided consent.

---


Device and Browser Fingerprinting

Device and browser fingerprinting is a process to collect and combine a variety of data (e.g., device or browser properties) to uniquely identify and track a particular device, thereby tracking the user. In fingerprinting, a piece of code on a website typically collects user data—such as screen size, screen resolution, browser information, and operating system details—to identify and track users. A fingerprinting algorithm combines the data to create a “fingerprint” or identifier. Commercial fingerprinting companies can provide the code to websites and store the results as a service.

One study found that 80%-90% of desktop fingerprints and about 80% of mobile device fingerprints are unique, suggesting that the majority of users’ browsers are uniquely identifiable. Another investigation found that at least one-third of the 500 most popular websites collected details about user devices for fingerprinting. Even when an individual uses “private browsing mode, a virtual private network (VPN), or clears third-party cookies, these websites were still able to successfully track users.” Unlike cookies or other data collection tools that can be “reset,” fingerprints typically last as long as the user has the same hardware and software.

Application Programming Interfaces

An API is a piece of software that enables one software program to communicate and interact with another, typically to share data and make certain features or tools available. Data communications between apps and operating systems, between apps and websites, and between apps often use APIs. APIs can connect the platforms and share data in a form that the developers of a separate website or app can easily use, sometimes for tracking users as they move across platforms.

Data collection and sharing user data with third parties, including advertisers, may use some APIs. For example, Meta and Google offer analytics APIs to help operators gain information about visitors to their websites based on their Google or Meta account and other data, allowing all the companies involved to build detailed user profiles for targeted advertising. Third parties can take advantage of APIs to extract consumer data deceptively. For example, in 2018, the company

---


55 A fingerprinting algorithm inputs data and produces a shorter string that identifies the data.


59 Ibid.

60 For example, apps use application programming interfaces (APIs) to access and send service requests to the operating system of a mobile device. Thus, all apps need these kinds of APIs to be compatible with an operating system. Cameron Russell et al., “APIs and Your Privacy,” SSRN, February 5, 2019, at http://dx.doi.org/10.2139/ssrn.3328825.
Cambridge Analytica collected user data through Facebook’s API without users’ knowledge or consent to build voter profiles (discussed later in “Federal Trade Commission Enforcement of Consumer Data Collection”).

Operators of websites and apps have various incentives to use or provide certain APIs. Small operators might use APIs offered by large operators to improve their website or app and offer special features, and large operators might offer APIs to expand their consumer data collection and integrate their services with others. For example, one of YouTube’s APIs allows other developers to embed YouTube videos on their websites and apps. Large companies may collect data and track users across different contexts through various APIs. Facebook, Google, and Apple offer login APIs that allow users to use their accounts with the respective companies to log in to a different website or app. These APIs may present privacy and cybersecurity concerns, as each company may be able to control user data and access to hundreds of other websites and apps. APIs can also evade certain anti-tracking software. For example, Meta’s Conversions API enables advertisers to share user data and online activity, even if a user has an ad blocker to block third-party cookie tracking.

Through sharing data and functionality, APIs can enable interoperable and interactive online services, which may benefit competition but create privacy risks. Some APIs can increase competition by helping companies innovate and interoperate with other services. Some large companies have been accused of stifling competition by revoking competitors’ access to their APIs. For example, Facebook disabled Vine’s (then owned by Twitter) access to a friend-finding API, which allowed users to find their Facebook friends on Vine; it also disabled Twitter’s access to a cross-posting API, which allowed users to post on Facebook and Twitter simultaneously. According to internal company documents released by the UK Parliament in 2018, Facebook executives, including CEO Mark Zuckerberg, intentionally severed Twitter’s access to Facebook’s friend-finding API. Some researchers have argued that dominant companies that are vertically integrated, operating as a “supplier” and as a direct competitor, no longer have an incentive to provide potential competitors access to their APIs.

**Software Development Kits**

An SDK is a set of software tools and code—often containing APIs—that entities can use to develop and integrate websites and apps with other services. This section focuses solely on

---

61 Ibid.
62 Ibid.
67 Chao and Schulman, *Promoting Platform Interoperability*.
68 For example, one of Meta’s software development kits (SDKs) includes login APIs and cross-posting APIs.
69 U.K. CMA, “Appendix F: The Role of Data in Digital Advertising,” *Online Platforms and Digital Advertising*; and
SDKs found in mobile devices due to their increasing adoption to track users in mobile ecosystems. Google and Meta are two of the most popular SDK providers on both Android and iOS, according to a 2019 report. Google had SDKs in over 85% of the most popular apps on Google Play Store, and Facebook had the second highest with SDKs in over 40% of these apps.

Some SDKs are required to develop an app for a particular operating system. These SDKs are necessary for mobile ecosystems to function and not commonly used for mobile tracking. Other SDKs specifically track and collect user data across mobile apps, including users who have never accessed the SDK providers’ consumer-facing products. Consumers are often unaware of the integration of third-party SDKs in apps or that they collect their data. When users interact with a mobile app, they may unknowingly share many kinds of data with all the third-party SDKs integrated in the app.

SDKs are extremely common, in part because they make it faster and easier to develop apps by reducing the amount of original code needed. Developers can also use ad network SDKs to obtain revenue. For example, apps that use Meta’s advertising SDK can send targeted ads to users based not only on the user’s interaction with the app but also on data that Meta collects from other sources. This benefit may incentivize developers to use Meta’s SDKs while simultaneously expanding Meta’s own data collection capabilities. Some app developers may be unaware of the data that third-party SDKs collect, particularly if the data are not essential for the app’s functionality. For example, a 2020 investigation found that Zoom’s iOS mobile app was sharing users’ data with Meta through one of the Facebook SDKs, even if the user did not have a Facebook account. Zoom removed the Facebook SDK after learning that it was collecting device information that was unnecessary for Zoom’s purposes.

Anti-tracking Tools from Intermediaries

Some intermediaries—such as web browsers and operating systems—permit users to limit the data collected by websites, apps, and third parties. Most web browsers offer “private browsing” or “incognito” modes that allow users to refuse cookies. However, other tools can still track users, such as pixels, fingerprinting, and APIs. Some operating systems on mobile devices provide an “opt-out” setting that limits tracking on the device, but most users do not use this setting. Nevertheless, some studies have found that most individuals would prefer to opt out of third-party cookies. One study found that when users are given the option to opt out of third-party cookies,


Cyphers and Gebhart, “Behind the One-Way Mirror.”


Warzel, “The Loophole.”


less than 0.1% of respondents agreed to be tracked (i.e., 99.9% opt out) for personalization and advertising.77

Some web browsers, including Firefox and Safari, have banned third-party tracking cookies. In 2019, Google announced plans to “develop new standards that advance privacy, while continuing to support free access to content” by replacing tracking tools, such as third-party cookies and fingerprinting, with a new set of tools for targeted advertising—collectively known as the Privacy Sandbox.78 This proposal raised competition concerns from antitrust enforcers in the United States and Europe, who argued the Privacy Sandbox could disadvantage Google’s ad competitors by denying them access to user data that Google still collects.79 Google has delayed implementing the Privacy Sandbox until 2024.80 Some commentators have argued that Google is delaying implementing its cookie ban because it could harm Google’s advertising platform.81

Apple released an iOS update in April 2021 that required apps to ask for permission before tracking user activity.82 Specifically, if users choose the option “Ask App Not to Track,” their Apple IDFA would be withheld. An app analytics firm estimated that during the first month after the release of the update, the percentage of U.S. daily users who allowed apps to continue tracking them ranged from 2% to 6%.83 Some commentators noted, however, that the update allows app providers to collect aggregated and “anonymized” data, such as a device’s IP address,84 and that Apple cannot prevent developers from tracking users with other identifiers, such as their email address or usage data.85 Furthermore, some companies reportedly use other

85 Thorin Klosowski, “We Checked 250 iPhone Apps—This Is How They’re Tracking You,” New York Times, May 6,
tools for data collection, such as device fingerprints, to circumvent the tracking permission requirement.\textsuperscript{86}

In 2018, Google restricted ad buyer and ad agency access to its persistent identifier in its ad network DoubleClick ID, indicating that the identifier could be tied to sensitive user information.\textsuperscript{87} However, restricting access to DoubleClick ID made it impossible to compare metrics from ads purchased in Google’s ad network with those purchased from other intermediaries. In response, the U.K. Competition and Markets Authority (CMA) proposed mandating the creation of secure common user IDs to address competition concerns within digital advertising and to increase interoperability by ensuring adherence to common standards.\textsuperscript{88} According to the CMA, market participants could assign a common digital ID to their own data for targeting and attribution purposes. The CMA also recognized potential privacy concerns that the proposal could raise, such as sharing data collected under the common user ID with third parties.

Some intermediaries can make it difficult for individuals to access anti-tracking tools. For example, Google has banned certain ad blocking apps and tracker blockers, some of which were reinstated,\textsuperscript{89} from its app store for violating its policy prohibiting developers from interfering with, disrupting, or damaging devices and other apps.\textsuperscript{90} In July 2022, Google announced updates to its app store policy, including allowing only VPN apps that use its proprietary VpnService interface and have VPN as its core functionality.\textsuperscript{91} The updates prohibit VPN apps from collecting “personal and sensitive user data without prominent disclosure and consent”; redirecting or manipulating “user traffic from other apps on a device for monetization purposes (for example, redirecting ad traffic through a country different than that of the user)”; or manipulating “ads that can impact apps monetization.”\textsuperscript{92} These changes could increase user privacy but might also prevent some ad-blocking VPN apps.\textsuperscript{93}

\begin{footnotesize}
\begin{enumerate}
\item[88] U.K. CMA, Online Platforms and Digital Advertising, last updated July 1, 2020.
\item[90] Google’s Developer Distribution Agreement states, “You will not engage in any activity with Google Play, including making Your Products available via Google Play, that interferes with, disrupts, damages, or accesses in an unauthorized manner the devices, servers, networks, or other properties or services of any third party including, but not limited to, Google or any Authorized Provider.” See Google Play, “Google Play Developer Distribution Agreement,” effective as of November 17, 2020, at https://play.google.com/about/developer-distribution-agreement.html; and Google Play Console Help, “Device and Network Abuse,” at https://support.google.com/googleplay/android-developer/answer/9888379?hl=en.
\item[92] Ibid.
\end{enumerate}
\end{footnotesize}
Intermediaries may be able to obtain a competitive advantage if they can collect data that others cannot access. Some commentators have noted that a third-party cookie ban may have detrimental effects on competition in digital advertising by limiting the availability of data and potential revenue from targeted ads for smaller competitors. Smaller digital advertising platforms may need third-party cookies, while larger incumbents—such as Google and Facebook—may be able to rely on first-party data.

Federal Trade Commission Enforcement of Consumer Data Collection

Although the United States does not have a comprehensive federal data protection law, the FTC has brought enforcement actions against companies for their consumer data collection practices under its authority to prevent “unfair or deceptive acts or practices in or affecting commerce.” The FTC’s complaints have included allegations that companies handled personal information in ways that contradict their privacy policies; failed to adequately protect personal information from unauthorized access despite promises that they would do so; and implemented default privacy settings that are difficult to change. The FTC also enforces some laws related to data collection, such as COPPA.

The FTC has taken enforcement actions against companies for allegedly using some of the data collection tools discussed in the “Anti-tracking Tools from Intermediaries” section. On August 29, 2022, it filed a complaint against the company Kochava, a data broker, for acquiring and selling consumers’ precise geolocation data, allowing “entities to track consumers’ movements to and from sensitive locations.” Other examples of complaints filed by the FTC that resulted in settlements include the following:

- In 2011, the FTC reached a settlement with ScanScout, an online advertiser, which allegedly misrepresented its data collection practices and consumers’ privacy settings that are difficult to change.

---

98 For more information, see CRS Report R45631, Data Protection Law: An Overview, by Stephen P. Mulligan and Chris D. Linebaugh.
100 See “Selected Data Collection and Tracking Tools.”
Online Consumer Data Collection

ability to control collection of their data by deleting some cookies but continuing the use of others to track consumers.\(^{102}\)

- In 2017, the FTC reached a settlement with Turn, a digital advertising company, for allegedly misrepresenting the extent to which it continued to track consumers online and through mobile apps after consumers opted out of such tracking.\(^{103}\)

- In 2019, the FTC reached a settlement with data analytics company Cambridge Analytica, its then-CEO Alexander Nix, and app developer Aleksandr Kogan, who worked with Analytica. In its complaint, the FTC alleged that one of Facebook’s APIs was used to deceptively collect data from millions of Facebook users without their permission.\(^{104}\) The FTC also imposed a $5 billion penalty and additional requirements on Facebook for violating a settlement reached in 2012 that required Facebook obtain consumers’ consent before sharing their data beyond its privacy settings.\(^{105}\)

- In 2021, the FTC reached a settlement with Flo App, a women’s health app, for disclosing millions of users’ health information to various third parties using integrated SDKs.\(^{106}\)

The FTC has also made statements and released reports regarding companies’ uses of consumer data.\(^{107}\) For example, in July 2022, after the U.S. Supreme Court issued its opinion in Dobbs v. Jackson Women’s Health Organization,\(^{108}\) the acting associate director of the FTC Division of Privacy and Identity Protection released a statement that location and information about consumers’ health are among the most sensitive categories of data.\(^{109}\) The statement warned companies that those “that make false claims about anonymization can expect to hear from the FTC” and that “the FTC does not tolerate companies that over-collect, indefinitely retain, or misuse consumer data.”

---


\(^{108}\) For more information about the case, see CRS Legal Sidebar LSB10768, Supreme Court Rules No Constitutional Right to Abortion in Dobbs v. Jackson Women’s Health Organization, by Jon O. Shimabukuro.

On August 11, 2022, the FTC announced that it is “exploring rules to crack down on harmful commercial surveillance and lax data security.”\(^{10}\) It published an advance notice of proposed rulemaking and held a virtual meeting seeking public input, using that information to help determine whether new rules are necessary, the practices that might be addressed, and actions it might take to deter those practices.\(^{111}\)

### Potential Effects of Consumer Data Protection Requirements

The potential effects of a comprehensive federal data protection law would depend on how entities and consumers respond. For example, some bills introduced in the 117\(^{th}\) Congress would require certain entities to inform individuals that their data are being collected and require entities to delete individuals’ data if requested;\(^{112}\) the effect of these requirements would depend on the number of users who request their data to be deleted.

This section discusses the potential effects of data protection requirements on website and app operators, specifically examining three possible scenarios: (1) prohibiting the collection of consumer data; (2) prohibiting the transfer, sale, or certain uses of consumer data; and (3) increasing transparency on the collection and use of consumer data.

### Prohibit the Collection of Consumer Data

Prohibiting the collection of consumer data would provide the highest level of consumer data protection but could also prevent some operators from providing their services or degrade the quality of their services. For example, if data protection laws prohibit the collection of certain necessary location data,\(^{113}\) navigation apps may no longer be able to provide real-time traffic information (e.g., notifying users that an accident has increased travel time and resulted in a faster alternate route). However, prohibiting the collection of nonessential data by navigation apps may have a minimal impact on navigation services.

Operators of existing navigation apps might be able to rely on historical data and analysis to continue providing some information—such as estimates for the time a route would take—that new entrants might be unable to provide. Thus, while prohibiting the collection of consumer data could encourage some operators to explore and develop other tools and technologies to obtain similar information from other sources, it also highlights the competitive advantage incumbents may have if consumer data is an integral component of developing their services.

---


112 For example, see the Online Privacy Act of 2021 (H.R. 6027) and the H.R. 8152 American Data Privacy and Protection Act ().

113 For example, GPS and other sensor data from a device may be necessary to determine a user’s location.
Restricting data collection could have a significant effect on advertising. Some websites and apps collect consumer data to provide targeted ads, which often serve as a main source of revenue. Providers of targeted advertising services might experience a loss of revenue if they cannot offer other means of targeting ads. For example, the chief financial officer of Meta Platforms reportedly stated that the company expects to lose more than $10 billion in sales revenue because of Apple’s “do-not-track” iOS update.114

Operators that rely on targeted advertising as their primary source of revenue might explore other options, such as charging users to access the app or using other forms of advertising. The potential effect on their revenue is unclear. An advertising technology company reportedly estimated that Facebook and Snap would lose about 13.2% of their revenue because of Apple’s update, while YouTube and Twitter would lose 7.7% and 7.4%, respectively.115 The estimated loss in revenue that Apple’s update may cause might be partially due to advertisers potentially switching to platforms that would allow them to continue providing targeted ads; implementing a federal ban on the collection of consumer data might not have the same effect.

Some entities that stopped providing targeted ads in the European Union (EU) in response to the General Data Privacy Regulation (GDPR)116—the EU’s comprehensive data protection law—did not experience a loss in revenue. For example, the New York Times reportedly started using contextual and geographical targeting,117 rather than behavioral targeting,118 and did not see a decline in revenue.119 One study found that contextual advertising might be as cost-effective as behavioral targeted advertising.120 Increased reliance on other forms of advertising, however, could increase competition to advertise on popular websites and apps, potentially making it difficult for smaller advertisers to gain visibility.

**Prohibit the Transfer or Sale of Consumer Data**

Prohibiting the transfer or sale of consumer data would provide users with greater control over which operators can access their data. It could discourage operators from collecting consumer data if they would not need it to provide their services. It might also provide greater data protection if it were to result in fewer operators having access to the data, although this would

---


116 For more information, see CRS In Focus IF10896, EU Data Protection Rules and U.S. Implications, by Rachel F. Fefer and Kristin Archick; and CRS Legal Sidebar LSB10846, The EU-U.S. Data Privacy Framework: Background, Implementation, and Next Steps, by Eric N. Holmes.

117 Contextual targeted advertising describes ads targeted to an individual based on the content provided on a website or app.

118 Behavioral targeted advertising describes ads targeted to an individual based on data about the individual’s online behavior, such as the amount of time the individual was watching a video.


depend on various factors, such as the number of operators that consumers are willing to share their data with and the operators’ level of data security.

Implementing restrictions on the transfer and sale of consumer data without restricting its collection and use could further entrench incumbents that have already collected large amounts of consumer data. It could encourage some of these incumbents to expand the markets they provide services in, potentially allowing these operators to obtain more data and expand the types of data they are able to collect. Furthermore, some consumers may be more willing to continue providing their data to a website or app for services they already use, rather than providing it to a new or unfamiliar one. This could make it difficult for nascent websites and apps, as well as operators that currently do not have access to large amounts of data, to compete.

Increase Transparency on the Collection and Use of Consumer Data

While some websites and apps provide general overviews of their data collection practices in privacy policies and other resources, detailed information—such as with whom data is shared or sold to—is typically not publicly available. Requiring operators and third parties to provide users with detailed information on their data collection and transfers—such as with a pop-up notification or a website that provides such information—could increase transparency about the collection and use of consumer data. Increased transparency could incentivize operators to adjust their current practices, particularly if it would heighten public scrutiny of the operators or discourage consumers from using the website or app.

Increasing transparency might not significantly alter consumer behavior, even if consumers must provide their consent before their data are collected. Most consumers do not read terms of service and privacy policies, in which operators provide general information about their data practices. Some consumers may consider the services offered on a website or app valuable enough that they would continue sharing their data, even if they would prefer not to. Data privacy laws have led to a proliferation of notifications on websites providing users the opportunity to accept or reject cookies, causing some users to experience “digital resignation.” Even if operators provide additional information, such as with whom they share or sell a user’s data, it might be difficult to track every entity that has access to that user’s data. The aggregation and bundling of data, as well as the distribution of data to third parties, may present additional challenges for tracking user data.

Policy Considerations for Congress

Multiple federal laws create data protection obligations for specific types of information and entities. However, technological developments have enabled data collection that may provide similar information but are not be covered by existing laws. For example, operators of health-

---


122 Ibid.


related websites and apps that are not associated with health care providers, health plans, or health care clearinghouses may not be subject to the data protection obligations required by the Health Insurance Portability and Accountability Act (HIPAA).\(^{125}\) Similarly, operators of finance-related websites and apps that are not associated with banks may not be subject to the data protection obligations required by the Gramm–Leach–Bliley Act.\(^{126}\)

Some data protection regulations implemented by states and foreign countries govern the use of some data collection tools discussed in this report. For example, the California Consumer Privacy Act (CCPA) requires companies that conduct business in California and serve California residents to disclose the use of certain cookies in their privacy policies and, under certain circumstances, to allow users to opt out of the sale of their cookie-related data to third parties.\(^{127}\) The CCPA does not mandate cookie disclosure banners or require user consent prior to the use of cookies.\(^{128}\) The CCPA also specifies cookies and pixels in its definition of unique identifier, which is included in the definition of personal information.\(^{129}\) The EU’s GDPR and ePrivacy Directive include requirements for websites and apps that serve EU residents to notify users if they use cookies, allow users to opt out of receiving some or all cookies, and allow users to use the bulk of an online service without receiving cookies.\(^{130}\) The results of studies examining the effect of GDPR on the use of cookies are mixed. Some studies find a decline in the use of cookies, while others find that few websites meet the minimum requirements.\(^{131}\)

If Congress chooses not to pursue legislative action, then ongoing efforts by states, federal agencies, and intermediaries may determine standards for data protection in the United States, including the following:

- Five states—California, Colorado, Connecticut, Utah, and Virginia—have passed consumer data protection laws.\(^{132}\)


\(^{129}\) Ibid.


• The FTC is considering whether it will implement new rules on data collection and security to protect consumers’ data and privacy.\

Federal Communications Commission Chair Jessica Rosenworcel sent letters to the top 15 mobile providers requesting information about their data retention and data privacy policies and shared their responses.\

The Consumer Financial Protection Bureau is considering rulemaking proposals to strengthen consumers’ access to and control over their financial data.

Some intermediaries are implementing policies that limit data other entities can access.

If Congress chooses to pursue legislative action, it may consider the following:

• Would legislation broadly address consumer data collection or focus on specific types of data? Some bills introduced in the 117th Congress include additional requirements on data collection and use in sensitive categories, such as biometric data and PII. Depending on the restrictions imposed, it may be possible to infer this information using other data that are seemingly less sensitive by using data collection and tracking tools.

• Would legislation implement requirements for operators—such as restricting or increasing transparency of data collection, or prohibiting the transfer or sale of data—or allow consumers to determine which entities can receive their data? Policymakers could create a privacy-by-default contractual regime that operators would need to follow—except in cases where consumers explicitly choose to provide their data—as suggested by the Biden Administration and other commentators. Alternatively, operators could be required to allow consumers to “opt out” of data collection. If a limited number of consumers choose not to provide their data, they may benefit from operators improving their services using other consumers’ information.

• Would federal legislation preempt state laws and, if so, which laws and to what extent? Without preemption, it may be difficult and costly for operators, particularly smaller and nascent ones, to comply, especially if some laws would create inconsistent regulations. Several U.S. federal statutes related to privacy


136 For example, the American Data Privacy and Protection Act (H.R. 8152) includes biometric information, genetic information, and precise geolocation data under “sensitive data.” Biometric information includes fingerprints, iris or retina scans, and facial or hand mapping.


do not preempt state laws, and preemption could prevent states from implementing stricter laws and enforcement of existing privacy-related state laws.\textsuperscript{139}

- **Would legislation include a private right of action?** Individuals filing suit could increase enforcement of the law, particularly for individual harms, since it may be difficult for a single federal agency to address every case in which an individual is harmed.\textsuperscript{140} Including a private right of action could also lead to frivolous lawsuits, which may be particularly burdensome for small and nascent operators.\textsuperscript{141} Some commentators have recommended a compromise, such as including a private right of action but incorporating a review of potential cases and limiting or setting tiers for damages, depending on the level of harm.\textsuperscript{142}

- **What are the potential unintended effects?** One study estimates that the EU’s GDPR led to over one-third of available apps on Google Play Store exiting and a 47.2% reduction in the rate of entry of new apps.\textsuperscript{143} Another study found that consumers made use of the opt-out capabilities provided by GDPR but that the consumers who chose to opt out were primarily those who used other data protection measures offered by browsers, such as cookie blockers and private browsing.\textsuperscript{144}

---


\textsuperscript{141} Becky Chao, Eric Null, and Claire Park, *Enforcing a New Privacy Law: A Private Right of Action Is Key to Ensuring that Consumers Have Their Own Avenue for Redress*, New America, Open Technology Institute, last updated November 20, 2019, at https://www.newamerica.org/oti/reports/enforcing-new-privacy-law/a-private-right-of-action-is-key-to-ensuring-that-consumers-have-their-own-avenue-for-redress/.


