www.futuremajority.org

Who Owns Americans' Personal Information and What Is It Worth?[1]

Robert Shapiro and Siddhartha Aneja

Findings

Americans are rightly concerned about mounting evidence that the internet's major platforms and many large companies are systemically gathering, analyzing and selling everyone's personal information. More than nine in ten Americans believe that they should determine who can see their personal information, and nearly nine in ten believe they should be able to direct any website to dispose of their personal data.[2] Yet, these operations – gathering and analyzing as much personal information as possible about every American and selling it in various forms – are an essential part of the online economy's current business model.

The largest search engine and social media platform, Google and Facebook, created this business model as they came to recognize that people use search engines and social media in ways that reveal extraordinary amounts of personal information. What people search for online and what they say and do in online communities often reveal their interests, likes and dislikes, income, debts, politics, sexual orientation, health status, addictions, education and intelligence, as well as their gender, age, marital status, ethnicity, religion, friendships and family background. Both sites also have access to what people write on G-mail and Facebook messages; and the companies' algorithms have created hundreds of millions of comprehensive profiles that advertisers and other businesses pay billions of dollars to access or use.

For a clear sense of how much personal information the large internet platforms collect and analyze, the two authors of this study, a Millennial and an older Baby Boomer, downloaded their personal data files from Google and Facebook. It is unsurprising that Millennials in their 20s and 30s use Google and Facebook much more than Boomers in their 50s and 60s. So, we discovered that Google holds 3.51 gigabytes of personal data on the Millennial, and Facebook holds 631 MB; for the Boomer, Google has 51 MB of personal data, and Facebook has 71 MB. One MB of information is the equivalent of 583 Word pages, and one gigabyte is the equivalent of 583,038 Word pages.[3] So, Google's current files of personal information on the two of us would fill the equivalent of 2,09467,463 Word pages on the Millennial and 29,735 Word pages on the Boomer, and Facebook's personal data files on us would fill the equivalent of 367,897 Word pages on the Millennial and 41,396 Word pages on the Boomer.

The major web platforms are the most prolific and profitable hunter-gatherers of personal information, which they typically analyze and use for considerable fees to attract advertisers by offering to target their digital ads. Led by Google, Facebook, Microsoft, Amazon, Verizon and Twitter, these operations occur on a very large scale. Virtually all of the ads on those platforms

[1] We gratefully acknowledge the financial support for this research and analysis provided by Future Majority.
[2] Zuboff, Shoshanna (2019). *The Age of Surveillance Capitalism.* Public Affairs, Hachette Book Group. New York.
[3] LexisNexis (2019). "How Many Pages in a Gigibyte?"
https://www.lexisnexis.com/applieddiscovery/lawlibrary/whitepapers/adi_fs_pagesinagigabyte.pdf

depend on targeting algorithms based on tens of millions of individual profiles created from the personal information the platforms gather on their users.

This drive to know as much as possible about everyone and profit from it is not limited to the giant search engines, browsers and social media platforms. Most other U.S. companies lack the scale, skills or incentives that would justify creating and operating their own data analysis programs. Instead, thousands of companies gather personal information that their customers or clients provide in the course of doing business with them, and then sell their information to large data brokers, such as credit bureaus. In turn, those data brokers analyze, package and resell the information, often as personal profiles. Their customers range from employers involved in hiring and companies planning marketing campaigns, to banks and mortgage lenders, colleges and universities, political campaigns and charities. In addition, credit card companies and healthcare data firms also routinely gather, analyze and profit from the personal information of their users.

We also examine the gathering, analysis and sale of personal information by credit card companies and healthcare businesses, because most Americans feel strongly that their personal financial and health information is especially sensitive and rely on the government to restrict the dissemination of such information. However, those restrictions apply only to certain types of financial information -- for example, personal bank balances, but not loan repayment data – and the requirements on health care information apply to healthcare providers but not to pharmacies or medical device producers. Moreover, personal financial and health-related information can be gathered, analyzed and sold in anonymized forms, which algorithms can match to most people or simply generate detailed financial and health-related profiles based on the extensive information that internet platforms and data brokers have on everyone.

Finally, the personal data now routinely used for commercial ends are not limited to the information that people reveal through their activities on internet platforms or through the goods and services they purchase. In addition, the Internet of Things has projected personal data gathering into many other aspects of people's lives. For example, smart TVs collect, analyze and sell personal information on who owns them and what they watch. Smart cars and smartphones collect, analyze and sell personal information on who owns them and every place they go. Smart beds and smart fitness bands collect, analyze and sell information on who uses those products and their temperatures, heartrates and respiration. Further, the new generation of wifi-based home devices that respond to people's voice commands – led by Amazon's Alexa, Echo, and Dot, and Google Nest and Google Home – can capture not only personal information about the people who buy and install them, but what they say in the range of those devices.

Since most Americans believe that their personal information is their own property, this analysis estimates the market value that the major internet platforms, data brokers, credit card companies and a leading healthcare data firm derive from capturing, analyzing and selling Americans' personal information. As expected, the large web platforms are the most prolific and profitable gatherers, users and sellers of personal information. Data brokers also derive billions of dollars from buying, analyzing and reselling Americans' personal information, as do credit card companies and the leading healthcare data business.

Table 1, below, summarizes our findings: *The revenues derived from these operations totaled nearly $52.5 billion in 2016, $63.8 billion in 2017 and $76.0 billion in 2018.*

Table 1. The Value of Americans' Personal Information Gathered and Used by Major Internet Platforms, Data Brokers, Credit Card and Healthcare Data Companies 2016-2018 ($ billions)

Platform	2016	2017	2018	Increase
Major Internet Platforms				
Google	$15,303.6	$18,132.4	$21,453.5	40.2%
Facebook	$6,432.4	$9,344.4	$11,882.0	84.7%
Amazon	$582.4	$920.4	$2,397.2	311.6%
Microsoft	$1,736.8	$1,944.8	$2,339.4	34.7%
Oath (Verizon)	$1,830.4	$1,872.0	$1,917.8	4.8%
Twitter	$707.2	$608.4	$728.2	2.9%
Other	$10,951.2	$14,180.4	$17,045.7	55.7%
Subtotal	**$37,544.0**	**$47,002.8**	**$57,763.9**	**53.9%**
Major Data Brokers				
Axciom	$804.0	$824.6	$1,053.0	31.0%
CoreLogic	$1,755.9	$1,664.7	$1,650.5	- 6.0%
Epsilon	$2,062.4	$2,174.3	$2,080.2	0.08%
Equifax	$1,938.7	$2,026.9	$2,066.4	6.6%
Experian	$2,412.5	$2,597.5	$3,070.7	27.3%
FICO	$572.9	$596.6	$681.4	18.9%
Harte-Hanks	$348.6	$330.9	$249.8	28.3%
RELX	$1,910.3	$1,973.6	$2,061.9	7.9%
Transunion	$1,452.2	$1,636.2	$1,934.3	33.2%
Subtotal	**$13,257.6**	**$13,825.3**	**$14,848.2**	**12.0%**
Credit Card Firms				
MasterCard	$1,010.1	$1,185.4	$1,418.1	40.4%
American Express	$238.2	$279.6	$334.5	40.4%
Subtotal	**$1,248.3**	**$1,465.0**	**$1,752.6**	**40.4%**
Healthcare Data Firm				
IQVIA	$443.4	$1,478.1	$1,681.5	379.2%
TOTAL	**$52,493.3**	**$63,771.2**	**$76,046.2**	**44.9%**

This analysis clearly establishes the high value derived by internet platforms, data brokers, leading credit card companies and the major healthcare data firm from gathering, analyzing and selling the personal information of Americans who use their services: It also shows that the revenues derived from these operations jumped 21.5 percent from 2016 to 2017 and another 19.2 percent from 2017 to 2018.

This two-year growth rate of 44.9 percent makes the capture and use of Americans' personal information the fastest-growing part in the U.S. economy.

Moreover, if these revenues continue to increase at the same rate over the next two and four years – that is, assuming these operations and the revenues they produce neither slow nor accelerate – the value of the personal information for the same businesses would increase from more than $76.0 billion in 2018 to $115.9 billion in 2020 and $197.7 billion by 2022. These projections are presented in Table 2, below.

Table 2. Projected Value of Americans' Personal Information Gathered and Used by Major Internet Platforms, Data Brokers, Credit Card and Healthcare Data Companies 2020 and 2022 ($ billions)

Platform	2018	2020	2022
Major Internet Platforms			
Google	$21,453.5	$30,077.2	$42,167.3
Facebook	$11,882.0	$21,948.6	$40,543.7
Amazon	$2,397.2	$9,867.1	$40,613.5
Microsoft	$2,339.4	$3,151.9	$4,246.6
Oath (Verizon)	$1,917.8	$2,010.4	$2,107.5
Twitter	$728.2	$749.7	$771.7
Other	$17,045.7	$26,531.7	$41,296.8
Subtotal	**$57,763.9**	**$94,336.5**	**$171,747.1**
Major Data Brokers			
Axciom	$1,053.0	$1,379.1	$1,806.2
CoreLogic	$1,650.5	$1,551.4	$1,458.3
Epsilon	$2,080.2	$2,098.2	$2,116.3
Equifax	$2,066.4	$2,202.5	$2,347.6
Experian	$3,070.7	$3,908.5	$4,974.8
FICO	$681.4	$810.5	$963.9
Harte-Hanks	$249.8	$179.0	$128.3
RELX	$2,061.9	$2,225.5	$2,402.2
Transunion	$1,934.3	$2,576.5	$3,431.8
Subtotal	**$14,848.2**	**$16,933.1**	**$19,629.3**
Credit Card Firms			
MasterCard	$1,418.1	$1,991.0	$2,795.3
American Express	$334.5	$469.6	$659.3
Subtotal	**$1,752.6**	**$2,460.6**	**$3,454.5**
Healthcare Data Firm			
IQVIA	$1,681.5	$2,177.6	$2,820.1
TOTAL	**$76,046.2**	**$115,907.8**	**$197,651.1**

For context and perspective, U.S. agriculture contributed $157.6 billion to U.S. GDP in 2018, so the value of the personal information gathered, analyzed and sold by the companies included here was equivalent to nearly half the total value of all U.S. agricultural output.[4] Moreover, our projections suggest that by 2022, the value of Americans' personal information as used by the companies included here, at $197.65 billion, will surpass the total value of U.S. agricultural output in that year.

Given how valuable and relatively inexpensive these data are to gather and use, it is virtually certain that in coming years, the capture of personal information will spread to more industries and businesses, the analyses of those data will be more detailed and sophisticated, and the commercial, social and political uses of the personal profiles drawn from those data and analyses will proliferate. It also seems likely that technologists will figure out how to sidestep or defeat any conventional regulation aimed at controlling or discouraging these activities, as they have been already with regard to financial healthcare information.

Here, we will consider a different or additional approach. Since gathering, analyzing and selling people's personal information is part of the basic business model for these companies, we propose that the government recognize people's property rights in their personal information and require that these companies do so as well. The companies contribute to the commercial value of the personal profiles they create by gathering, analyzing, and selling people's information. Accordingly, these companies should share the revenues from these operations with the people who provide the essential inputs, and we propose a 50-50 division.

An estimated 312 million people in the United States used the Internet in 2018, or 95.2 percent of a total population of 327.9 million.[5] A 50-50 split of the $76.05 billion in revenues generated from their personal information in 2018 would translate into a personal payment of about $122 per-person on the internet in 2018. By 2020, the population is expected to reach 332.6 million. Assuming 95.2 percent of them are online or 316.7 million people, projected revenues of $115.9 billion from using their personal information and the 50-50 split, the payments to those Americans for the use of their personal information would come to $183 per-person. By 2022, a projected 321.1 million Americans will use the internet and the companies considered here will generate an estimated $197.65 billion from their personal information, so the 50-50 split would come to about $308 per-person.

Alternatively, the same 50-50 split could cover much of America's outstanding infrastructure needs. The World Bank has estimated that in 2020, the United States will spend $309 billion on infrastructure or $116 billion less than the World Bank's $425 billion estimate of U.S. infrastructure requirements in 2020.[6] A 50 percent fee on the revenues derived from using people's personal information would cover half of that difference, or $58 billion. By 2022, The

[4] Bureau of Economic Analysis (2019). "GDP by Industry." https://www.bea.gov/data/gdp/gdp-industry
[5] Statistica (2019). "Internet usage in the United States – Statistics and Facts." https://www.statista.com/topics/2237/internet-usage-in-the-united-states/ ; U.S. Census Bureau (2019). "Population Projections." https://census.gov/programs-surveys/popproj/about/faq.html.
[6] World Bank (2019). Global Infrastructure Outlook." https://outlook.gihub.org/

World Bank estimates that we will need $442 billion in infrastructure investments that year, or $135 billion less than projected spending of $317 billion. The 50 percent fee would cover $98.8 billion of the $135 billion shortfall or about 73 percent of the gap. Or, the 50 percent fee on the projected revenues from using people's personal information could cut the federal budget deficit by $57.9 billion in 2020 and $98.8 billion in 2022.

Evidence and Analysis

U.S. Digital Advertising on Major Internet Platforms

The shift of millions of transactions from brick and mortar stores and service locations to the digital internet has produced millions of gigabytes of personal information about the buyers, sellers and their transactions. As we have seen, monetizing those personal data has become a major source of revenue. Giant web platforms led by Google, Facebook, Microsoft's Bing, Yahoo and LinkedIn mine and analyze the personal data they capture from their users, ranging from a user's age, location, family status and religion to occupation, finances, health status and sexual orientation, and personal social, political and consumer interests and views. Based on thousands of data points for each person, these platforms create personal profiles that, for a fee, advertisers can use to target their ads to those most likely to respond.[7]

Every time a person uses Facebook, Google and the others, he or she exposes personal information that is automatically captured by a platform's servers. The platforms also identify the devices each of us uses to connect to the platforms, so they can also target advertising to those devices even when the user is no longer on the platform. Moreover, many of these platforms are part of larger companies that gather personal information from a variety of websites. For example, Google integrates personal information on anyone using not only the Google search engine, but also the Alphabet-owned YouTube, Gmail, Google Maps, and advertising networks. A more complete list of Google's sources for personal information includes Android Device Configuration Service, Bookmarks, Calendar, Chrome, Classic Sites, Classroom, Contacts, Data Shared For Research, Drive, Fit, G Suite Marketplace, Google Help Communities, Google My Business, Google Pay, Google Photos, Google Playbooks, Google Play Console, Google Play Games Services, GooglePlay Movies & TV, Google Play Music, Google Play Store, Google Shopping, Google+ +1s on websites, Google+ Circles, Google+ Communities, Google+ Stream, Groups, Handsfree, Hangouts, Hangouts on Air, Home App, Input Tools, Keep, Location History, Mail, Maps, My Activity, My Maps, News, Posts on Google, Profile, Purchases & Reservations, Reminders, Saved, Search Contributions, Shopping Lists, Street View, Tasks, Textcube, Voice, and YouTube data. Google's wide reach gives the company an ability to amass and analyze an incredibly broad and diverse dataset of personal information, from demographic characteristics, email messages, location and movement

[7] Google (2019). "Targeting your Ads." https://support.google.com/google-ads/answer/1704368?hl=en

to travel, browser history, bookmarks, calendar data, news consumed, and videos viewed. An estimated 84 percent of the revenues of Alphabet, Google's holding company, come from these operations[8] and have made Google the world's third largest private company.[9]

Similarly, Facebook combines and integrates personal information captured on Facebook itself as well as Messenger, WhatsApp and Instagram.[10] An estimated 98 percent of Facebook's revenues come from targeted advertising based on these operations[11] and have made Facebook the world's fifth largest private company.[12] Microsoft also launched its own advertising platform in 2018 based on personal information captured by its Bing search engine, Edge/Explorer browser and the range of Microsoft products and services, including LinkedIn, X-box, Skype and Microsoft Office.[13] Microsoft also captures additional personal information through its data partnerships with third parties including Yelp, Twitter, Foursquare, and TripAdvisor.[14]

Many companies also trade personal data with these platforms. For example, Facebook allows some third-party applications to use its platform in exchange for the data app collects, such as extensive dating and sexual information on OKCupid users who sign up for the dating service through Facebook. Companies that integrate their product offering through the Facebook platform also gain access to some of Facebook's data. The *New York Times* reported in 2018 that Facebook has established such data sharing arrangements with 150 companies since 2010, including in some cases access to people's personal messages on Facebook's platform.[15] In one well-known instance, a Russian with access to Facebook's unique user IDs has been accused of sharing those data with the Kremlin. Less sinister but equally troubling, Facebook had partnership agreements with Spotify, Netflix and the Royal Bank of Canada allowing them to read and edit private messages between Facebook users. [16] A public backlash forced Facebook to alter its data sharing practices in ways that make it more difficult for third parties, yet still possible, to gain follow-on access to personal data on Facebook's users.

---

[8] Shaban, Hamza (2018). "Google parent Alphabet reports soaring ad revenue, despite YouTube backlash/" the *Washington Post*. February 1, 2018. "https://www.washingtonpost.com/news/the-switch/wp/2018/02/01/google-parent-alphabet-reports-soaring-ad-revenue-despite-youtube-backlash/?utm_term=.1456c39043b3

[9] Shen, Lucinda (May 21, 2018). "Here Are the Fortune 500's 10 Most Valuable Companies." *Fortune*. http://fortune.com/2018/05/21/fortune-500-most-valuable-companies-2018/

[10] Twitter does own subsidiary platforms but still gathers extensive user data from including, extensive demographic and work history information, the content of tweets posted, the content of tweets read or retweeted, links clicked, and followed., and personal data collected from other websites with embedded Twitter features. Twitter (2019). "Your Twitter Data." https://twitter.com/personalization.

[11]" Facebook's advertising revenue worldwide from 2009 to 2018 (in million U.S. dollars)." Statistica. https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/.

[12] Shen, Lucinda (May 21, 2018). "Here Are the Fortune 500's 10 Most Valuable Companies." *Fortune*. http://fortune.com/2018/05/21/fortune-500-most-valuable-companies-2018/

[13] Raehsler, Lisa (2018). "How Microsoft Audience Ads Work." *SearchEngine Journal*. https://www.searchenginejournal.com/how-microsoft-audience-ads-work/251621/#close.

[14] Davies, Jamie (2018). "Microsoft Launches its own Personalised Advertising Platform." Telecoms.com http://telecoms.com/489563/microsoft-launches-its-own-personalised-advertising-platform /

[15] Romano, Aja (Dec. 19, 2018). "Facebook let Netflix, Spotify, and other Companies read your Private Messages." *Vox*. https://www.vox.com/2018/12/19/18148136/facebook-privacy-violations-nyt-netflix-spotify-amazon-yahoo.

[16] https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html.

In addition, Facebook and other platforms have extensive data-trading and data-sharing arrangements with thousands of applications that use Facebook or the other platforms for logins. Applications have similar login arrangements with Google, Amazon, Twitter and other major platforms, so people can use the applications without creating an additional specific account. These arrangements enable people to access those applications without creating separate user IDs and passwords; and for facilitating this one step access, Google, Amazon, Twitter, Facebook and the others gain access to the personal information people provide when they use the applications. For instance, a person who uses LinkedIn to access WallStreetOasis, a popular website for aspiring finance professionals, in so doing provides Microsoft (LinkedIn's owner) access to whatever the user does on WallStreetOasis. Similarly, people who access Pinterest through Google in so doing share with Google all of their "pins" and responses to other people's pins.

The internet service provider (ISP) Verizon approaches its users' personal information in much the same way as Google, Facebook and Microsoft. Verizon created a corporate division that captures and analyzes its users' personal information, again to help advertisers target their offerings. As yet, Verizon does not attract targeted advertising on the scale of Facebook and Google. However, a 2017 law gives ISPs a right to access their customers' browser histories.[17] Verizon Media can now construct very detailed personal profiles on the customers of Verizon internet service from their browser histories, including use of Verizon-owned AOL, Yahoo, and Tumblr, as well as its mobile search engine Gemini. Verizon can also integrate that personal information with additional data from its millions of mobile customers, including GPS data, cell phone applications, streaming data, and data from its online properties and from third parties.[18] At this time, however, the other major ISPs, AT&T and Comcast, have invested relatively little in the information infrastructure for digital advertising. In fact, AT&T shut down its online and mobile ad network in 2013,[19] although it recently purchased several advertising properties, including AppNexus and created a new advertising subsidiary called "xandr."[20]

Amazon also collects large volumes of personal information, from every visit to its platform, starting with information about every Amazon consumer's browsing and purchases. It also can integrate those data with additional personal information collected from owners of Alexa, users of Amazon Fresh, Amazon Prime and Kindle, and users of other Amazon-owned websites such as Audible, Zappos, PillPack and Twitch. Like AT&T, Amazon has been a

[17] Solon, Olivia (March 28, 2017). "Your browsing history may be up for sale soon. Here's what you need to know." *The Guardian*. https://www.theguardian.com/technology/2017/mar/28/internet-service-providers-sell-browsing-history-house-vote.
[18] Verizon (2019). "Privacy Policy." https://www.verizon.com/about/privacy/full-privacy-policy.
[19] Walsh, Mark (2013). "AT&T AdWorks Dumps Online, Mobile Ad Net." *MediaPost*. https://www.mediapost.com/publications/article/211171/att-adworks-dumps-online-mobile-ad-net.html?edition=&mod=article_inline.
[20] Castillo, Michelle (August 10, 2018). "AT&T's Recent Acquisition Spree is Part of a Plan to Dominate Advertising on Connected TVs and Devices." *CNBC*. https://www.cnbc.com/2018/08/10/att-wants-to-sell-ads-on-connected-tvs-and-devices.html; AT&T (2018). "AT&T Launches New Advertising Company, Xandr." https://about.att.com/story/2018/att_launches_xandr.html.

relatively late adopter of targeted digital advertising – yet from 2016 to 2018, its share of internet advertising revenues grew from 1.6 percent to 4.1 percent, a 168 percent increase.[21]

Companies also gather personal information in other ways. The Internet of Things has introduced hundreds of consumer products that collect and transmit data to the companies that produce them. Most prominently, new virtual assistants such as the Google Home and the Amazon's Alexa record and transmit everything users say after activating the device with a "wake" word such as "OK Google" or "Alexa."[22] Other smart devices also capture large volumes of personal information. Smart TVs gather all viewer information including location and sell that information to cable companies and others to advertise offerings chosen by algorithms. As the chief technology officer of Vizio said recently, "It's not just about data collection. It's about post-purchase monetization of the TV."[23]

Finally, thousands of smaller companies on the internet contract with third-party ad servers, ad networks and ad exchanges that collect end users' personal data for targeted advertising. Ad servers store, maintain, and deliver advertisements to end users' computers based on algorithmic analysis of their interests.[24] Ad networks are middlemen between websites and advertisers, freeing the websites from setting up their own ad sales teams and the advertisers from recruiting and managing multiple websites.[25] Ad exchanges also auction off the right to sell inventory to the highest bidder, based on the number of times an ad is viewed.[26] The giant web platforms also own ad network subsidiaries, especially Google which operates Google Ad Manager and Google Marketing Platform for third party advertisers.

The Value of the Personal Information Used to Target Digital Advertising

To estimate the value of the use of people's personal information by major platforms including Google, Facebook, Amazon, Microsoft, Verizon's Oath, Amazon, Twitter and others, we begin with their 2016, 2017 and 2018 revenues from digital advertising, drawn principally from their annual reports and SEC filings.[27] (Table 3, below) In 2018, these platforms earned $111.1 billion from U.S. advertisers targeting American consumers. Google and Facebook dominated this area in 2018, accounting respectively for 37.1 percent ($41.3 billion) and 20.6 percent ($22.9 billion) of total digital advertising revenues.

---

[21] See Table 3 below
[22] Simonite, Tom (May 31, 2016). "How Alexa, Siri, and Google Assistant Will Make Money Off You." *MIT Technology Review*. https://www.technologyreview.com/s/601583/how-alexa-siri-and-google-assistant-will-make-money-off-you/
[23] Gilbert, Ben (Jan. 12, 2019). "There's a Simple Reason your New Smart TV was so Affordable: It's Collecting and Selling your Data." *Business Insider*. https://www.businessinsider.com/smart-tv-data-collection-advertising-2019-1.
[24] *Muvi* (2016). "Ad Servers vs Ad Networks, What's the Difference?" https://www.muvi.com/blogs/ad-servers-vs-ad-networks-whats-difference.html
[25] *Ibid.*
[26] *SelfAdvertiser* (2017). "Ad Network vs Ad Exchange." https://blog.selfadvertiser.com/ad-network-vs-ad-exchange
[27] See also Rodriquez, Ashley (October 9, 2018). "Amazon may be the Biggest Threat to Google since Facebook." *Quartz*. https://qz.com/1417775/amazon-ads-may-be-the-biggest-threat-to-google-since-facebook; Twitter (2018). "Form 10-Q." http://d18rn0p25nwr6d.cloudfront.net/CIK-0001418091/e1171c36-99a4-4e08-884c-35f46f987f83.pdf

Table 3: U.S. Digital Advertising Revenues by Platform, 2018

Platform	2016	2017	2018	Share, 2018
Google	$29,430	$34,870	$41,257	37.1%
Facebook	$12,370	$17,970	$22,850	20.6%
Amazon	$1,120	$1,770	$4,610	4.2%
Microsoft	$3,340	$3,740	$4,500	4.1%
Oath (Verizon)	$3,520	$3,600	$3,688	3.3%
Twitter	$1,360	$1,170	$1,400	1.3%
Other	$21,060	$27,270	$32,780	29.5%
Total	**$72,200**	**$90,390**	**$111,084**	**100.0%**

These financial data, however, overstate the value of the personal information, because companies would advertise on the Internet even if the platforms did not target consumers based on personal profiles. We begin, therefore, by calculating the value of people's personal information as the difference between what companies will pay for ads targeted based on personal profiles and what they will pay for untargeted advertising. Several studies have tried to measure the value of personal data for advertising on online platforms. A 2011 study estimated that privacy protections enacted by the European Union in 2003 and 2004 to curtail the use of personal information reduced the effectiveness of online advertising by 65 percent.[28] A 2013 study used structural modeling to simulate how advertisers bid for space on internet platforms and found that a ban on internet tracking would reduce revenues for banner ads 38.5 percent.[29] However, these studies did not have consumer-level data to work with to estimate the value of personal data. Therefore, we rely on a 2018 study that analyzed data from AdChoices, the internet industry's self-regulatory program. The researchers found that the cost of ads uninformed by people's personal data was 52 percent less than the cost of comparable ads informed by those data.[30]

Since virtually all advertising on these platforms is placed based on the personal profiles created by those platforms for their advertisers, we estimate that the value of the personal

[28] Goldfarb, Avi and Catherine Tucker (2011). "Privacy Regulation and Online Advertising." *Management Science*, Vol. 57, No. 1. http://papers.ssrn.com/paper/taf?abstract_id=2333193.

[29] The researchers analyzed survey data from nearly 10,000 advertising campaigns from 2001 to 2008 to estimate the relative effectiveness of advertising by country before and after the EU implemented those privacy protections. Johnson, Garrett (2013). "The Impact of Privacy Policy on the Auction Market for Online Display Advertising." *Simon Business School Working Paper No. FR 13-26*.

[30] Johnson, Garrett, Scott Shriver, and Shaoyin Du (2018). "Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry?" *Simon Business School Working Paper No. FR 17-19*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3020503. The study also examined the frequency with which people opt out of their offsite browsing behavior being captured. However, few consumers were aware of the opt-out option, and the researchers found that 0.23 percent of 62 million sampled ad impression transactions involved opt-out consumers. It is also worth noting that even when a consumer opts out under AdChoices, Google has acknowledged "[a]ds can still be targeted with info like your general location or the content of the website you're visiting." Google (2019). "Block Certain Ads." https://support.google.com/ads/answer/2662922?hl=en.

information accounts for 52 percent of the revenues earned from U.S. digital advertising on those platforms. Therefore, the use of people personal information increased online advertising revenues in 2018 by $57.8 billion (0.52 * 111.08), including $21.5 billion for Google and nearly $11.9 billion for Facebook. (Table 4, below)

Table 4: Value of Personal Information Used in U.S. Digital Advertising, By Platform (2018)

Platform	Revenues
Google	$21,453,495,904
Facebook	$11,882,025,060
Amazon	$2,397,200,000
Microsoft	$2,339,436,145
Oath (Verizon)	$1,917,760,000
Twitter	$728,240,116
Other	$17,045,698,197
Total	**$57,763,855,422**

Table 3, above shows clearly that the revenues drawn from the analysis and sale of personal information for U.S. digital advertising increased sharply over recent years. From 2016 to 2018, those revenues grew 53.9 percent overall, including gains of 40.2 percent by Google, 84.7 percent by Facebook, and 311.6 percent by Amazon.[31] Table 3, below, tracks those increases in the value of the personal information used to target digital advertising.

Table 5: Value of Personal Information in U.S. Digital Advertising by Platform 2016, 2017 and 2018 ($ millions)

Platform	2016	2017	2018	Increase
Google	$15,303.6	$18,132.4	$21,453.5	40.2%
Facebook	$6,432.4	$9,344.4	$11,882.0	84.7%
Amazon	$582.4	$920.4	$2,397.2	311.6%
Microsoft	$1,736.8	$1,944.8	$2,339.4	34.7%
Oath (Verizon)	$1,830.4	$1,872.0	$1,917.8	4.8%
Twitter	$707.2	$608.4	$728.2	2.9%
Other	$10,951.2	$14,180.4	$17,045.7	55.7%
Total	**$37,544.0**	**$47,002.8**	**$57,763.9**	**53.9%**

[31] Rodriquez, Ashley (October 9, 2018). "Amazon may be the Biggest Threat to Google since Facebook." *Quartz*. https://qz.com/1417775/amazon-ads-may-be-the-biggest-threat-to-google-since-facebook; Twitter (2018). "Form 10-Q." http://d18rn0p25nwr6d.cloudfront.net/CIK-0001418091/e1171c36-99a4-4e08-884c-35f46f987f83.pdf; eMarketer (2018). "Data Suggests Surprising Shift: Duopoly Not All-Powerful." https://www.emarketer.com/content/google-and-facebook-s-digital-dominance-fading-as-rivals-share-grows

If this growth rate of online advertising revenues is sustained over the next two and four years – assuming its growth neither accelerates nor slows -- the value of Americans' personal information to online platforms that use that information to target advertising will reach $94.3 billion in 2020 and $171.7 billion in 2022. Based on these calculations, the advertising value of Americans' personal information for Facebook and Amazon will reach$40.5 billion and $40.6 billion, respectively, by 2022, rivaling the projected value of such information for Google in 2022 of $42.2 billion.

Table 6: Value of Personal Information in U.S. Digital Advertising by Platform, 2018 and Estimates for 2020 and 2022 ($ million)

Platform	2018	2020	2022
Google	$21,453.5	$30,077.2	$42,167.3
Facebook	$11,882.0	$21,948.6	$40,543.7
Amazon	$2,397.2	$9,867.1	$40,613.5
Microsoft	$2,339.4	$3,151.9	$4,246.6
Oath (Verizon)	$1,917.8	$2,010.4	$2,107.5
Twitter	$728.2	$749.7	$771.7
Other	$17,045.7	$26,531.7	$41,296.8
Total	**$57,763.9**	**$94,336.5**	**$171,747.1**

Credit Bureaus and Other Data Brokers

Beyond the major platforms' digital advertising programs, many other companies specialize in purchasing, analyzing, aggregating and selling personal information collected by a broad range of third parties. Data brokers including credit bureaus collect people's personal information from such sources as public records and social media platforms, phonebooks, public records of arrests and police complaints, records of real estate sales, as well as information from retailers, online surveys and cookies that track people's internet behavior.[32] After aggregating hundreds or thousands of records, data brokers use data mining and data analysis to build profiles on tens of millions of people, including demographic, income, employment, consumption, tax and health characteristics. One leader broker, Acxiom, reported as far back as 2013 that it had personal data on 10 percent of the world's population, averaging 1,500 files per individual.[33] Data brokers sell their personal profiles and analytics to companies planning marketing campaigns or considering hiring decisions, insurers, educational institutions, airports, public agencies and even giant internet platform such as Facebook.[34] In this way, for example, data

[32] Privacy International (2018.) "How do Companies get our Data?" https://privacyinternational.org/feature/2048/how-do-data-companies-get-our-data.
[33] Bachman, Katy (March 25, 2014). "Confessions of a Data Broker." *AdWeek*. https://www.adweek.com/digital/confessions-data-broker-156437/.
[34] Grauer, Yael (March 27, 2018). "What Are 'Data Brokers,' and Why Are They Scooping Up Information About You?" *Vice*. https://motherboard.vice.com/en_us/article/bjpx3w/what-are-data-brokers-and-how-to-stop-my-private-data-collection.

brokers compiled and sold lists of people based on their HIV status, records of alcoholism, and issues with erectile dysfunction.[35]

Credit bureaus are perhaps the most prominent form of data broker. Credit bureaus are companies that collect consumer data specifically on people's personal financial habits. Credit bureaus and other data brokers can freely gather, purchase, analyze and sell very detailed personal financial information, because relevant federal statutes only fully bar sales of certain, relatively narrowly-defined financial data. Bank account balances, account numbers, and access codes are supposed to remain confidential, but other personal information such as a person's answers in a credit card application or records of loan payments can be traded or sold if the bank informs the customer. As a result, companies such as Equifax, Experian, and TransUnion, under no legal obligation to anonymize the financial information they collect and sell, gather detailed personal information from an estimated 10,000 to 30,000 sources to build people's credit ratings. They often start with public records, from car and home ownership to bankruptcy and divorce proceedings. Mortgage lenders sell data brokers the information on people's loan amounts, loan balances, and payment histories; and while the FTC has barred data brokers from selling personal identifying information about people with late mortgage payments, Equifax did precisely that from 2008 to 2010.[36] Equifax has acknowledged selling employment data to debt collectors, financial institutions, and others; and according to one report, Equifax also sold people's salary records to those companies. [37]

Credit bureaus also purchase detailed records of people's non-mortgage payments and delinquencies from bill collection agencies.[38] The credit bureaus and other data brokers create financial profiles from all of these sources and sell hundreds of millions of such reports to employers, banks and other financial institutions. The Fair Credit Reporting Act technically requires that employers secure a person's written permission to access their credit reports; but most people assume that declining to provide permission could preclude their being hired or cost them their jobs.[39] Credit bureaus and other data brokers also sell their stores of personal financial information to thousands of businesses for other purposes. For example, TransUnion works with hospitals and healthcare systems to help them recover funds from patients and

[35] Kashmir, Hill (December 19, 2013). "Data Broker Was Selling Lists Of Rape Victims, Alcoholics, and 'Erectile Dysfunction Sufferers.'" *Forbes*. https://www.forbes.com/sites/kashmirhill/2013/12/19/data-broker-was-selling-lists-of-rape-alcoholism-and-erectile-dysfunction-sufferers/#7ec6cdee1d53.
[36] Federal Trade Commission (2012). "FTC Settlements Require Equifax to Forfeit Money Made by Allegedly Improperly Selling Information about Millions of Consumers Who Were Late on Their Mortgages." https://www.ftc.gov/news-events/press-releases/2012/10/ftc-settlements-require-equifax-forfeit-money-made-allegedly.
[37] Sullivan, Bob (January 30, 2013). https://www.nbcnews.com/technolog/exclusive-your-employer-may-share-your-salary-equifax-might-sell-1B8173066.
[38] Consumer Financial Protection Bureau (2012). "Key Dimensions and Processes in the U.S. Credit Reporting System." https://files.consumerfinance.gov/f/201212_cfpb_credit-reporting-white-paper.pdf.
[39] Sweet, Ken (Oct. 6, 2017). "Equifax Collects Your Data, and Then Sells It." *Inc*. https://www.inc.com/associated-press/equifax-data-money.html.

insurance companies, Equifax helps lenders assess a person's default risk, and Experian sells its personal information to help businesses confirm their customers' identities.[40]

The FTC's definition of data brokers as "companies that collect consumers' personal information and resell or share that information with others" also covers many other types of businesses that sell important people's personal information to third parties. For example, mobile phone companies including AT&T, Sprint, Verizon, and T-Mobile sell their customers' real-time location information to data brokers such as Zumigo and LocationSmart, which resell those data to almost anyone who will pay for them.[41] On occasion, this practice exceeds the public's tolerance: The major cellphone service providers sold GPS location data to a prison technology company in a form that allowed it to track the locations not only of inmates, but any of those providers' customers.[42] Once the arrangement was reported publicly in mid-2018, it ended. Political campaigns and non-profit organization also act as data brokers when they sell personal data on their supporters, including telephone numbers, email addresses, physical addresses and sometimes passport numbers.[43] In 2016, Marco Rubio's presidential campaign collected more than $500,000 from selling personal information on its donors, and Donald Trump's presidential apparatus sold the use of data on his supporters and his donors to GOP candidates running in the 2018 midterm elections.[44]

Finally, many companies that do not operate as data brokers *per se* maintain extensive data partnerships with other companies, which they hope will enhance the value of their own products. Wells Fargo, JP Morgan Chase and other major banks allow Intuit, whose products include QuickBooks and TurboTax, to access some of their banking data, so both parties to the arrangement can better integrate their products and enhance cross-selling opportunities.[45] A person who uses the QuickBooks accounting software package links his Intuit account to his bank account by using the Intuit portal to log into his bank. The user can then import data from

[40] TransUnion (2019). "Solution: Customer Acquisition Strategies." https://www.transunion.com/solution/customer-acquisition; Equifax (2019). "Solution: Commercial Insight Suite." https://www.equifax.com/business/commercial-insight-suite/; Experian (2019). "Business-to-Business." https://www.experianplc.com/about-us/our-business-activities/business-to-business/

[41] Cox, Joseph (January 8, 2019.) "I Gave a Bounty Hunter $300. Then He Located Our Phone." *Vice*. https://motherboard.vice.com/en_us/article/nepxbz/i-gave-a-bounty-hunter-300-dollars-located-phone-microbilt-zumigo-tmobile.

[42] Fung, Brian (June 19, 2018). "Verizon, AT&T, T-Mobile and Sprint suspend selling of customer location data after prison officials were caught misusing it." *Washington Post*. https://www.washingtonpost.com/news/the-switch/wp/2018/06/19/verizon-will-suspend-sales-of-customer-location-data-after-a-prison-phone-company-was-caught-misusing-it/?noredirect=on&utm_term=.c3f12073f67b.

[43] Maass, Dave (2016). "Voter Privacy: What You Need to Know About Your Digital Trail During the 2016 Election." *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2016/02/voter-privacy-what-you-need-know-about-your-digital-trail-during-2016-election.

[44] Pagliery, Jose (July 7, 2016). "Here's how Presidential Candidates Sell your Personal Information." *CNN*. https://money.cnn.com/2016/07/07/news/presidential-candidate-sell-donor-data/index.html; https://www.seattletimes.com/nation-world/trump-campaign-selling-email-and-phone-lists-for-millions-of-supporters/

[45] Crosman, Penny (2017). "Wells Fargo Latest to Share Customer Data with Intuit via API." *American Banker*. https://www.americanbanker.com/news/wells-fargo-latest-to-share-customer-data-with-intuit-via-api.

his bank account to his QuickBooks easily and efficiently – and in the process, his bank Fargo and Intuit gain access to the data on him held by the other.[46] Similarly, PayPal has data sharing agreements with more than 600 companies, including payment processors, fraud detection companies, financial product developers, commercial enterprises, and marketing firms. Paypal captures data on its partner's customers and in exchange provides those partners with Paypal's data on its customers' names, addresses, telephone numbers, email addresses, birth dates, along with how they pay and details of their transactions.[47]

The Value of the Personal Information Sold by Data Brokers

Our next task involves estimating the value of the personal information sold by data brokers. While data-brokering has become a sprawling industry, here we will focus on the nine companies widely recognized to be leading or major data brokers/credit bureaus: Axciom, CoreLogic, Epsilon, Equifax, Experian, Fair Issac Corp., Harte-Hanks, Transunion, and RELX. Our analysis is based principally on the 2016, 2017, and 2018 annual reports issued by those companies.[48] Their own data show that the nine companies earned revenues totaling $13.3 billion in 2016, $13.8 billion in 2017 and $14.8 billion in 2018 from brokering the personal information of Americans, led by Experian, Epsilon, Equifax and Transunion. (Table 7, below)

Table 7: Major Data Brokers' U.S. Revenues from Reselling Personal Information, 2016, 2017 and 2018 ($ millions)

Company	2016	2017	2018	Increase
Axciom	$804.0	$824.6	$1,053.0	31.0%
CoreLogic	$1,755.9	$1,664.7	$1,650.5	- 6.0%
Epsilon	$2,062.4	$2,174.3	$2,080.2	0.08%
Equifax	$1,938.7	$2,026.9	$2,066.4	6.6%
Experian	$2,412.5	$2,597.5	$3,070.7	27.3%
FICO	$572.9	$596.6	$681.4	18.9%
Harte-Hanks	$348.6	$330.9	$249.8	- 28.3%
RELX	$1,910.3	$1,973.6	$2,061.9	7.9%

[46] Intuit (2017). "Intuit Signs New Data-Exchange Agreement with Wells Fargo." https://www.intuit.com/company/press-room/press-releases/2017/INTUIT-SIGNS-NEW-DATA-EXCHANGE-AGREEMENT-WITH-WELLS-FARGO/

[47] Paypal (2019). "List of Third Parties (other than PayPal Customers) with Whom Personal Information May be Shared." https://www.paypal.com/ie/webapps/mpp/ua/third-parties-list.

[48] One of the companies, RELX, includes a number of businesses such as Lexis Nexus that earn revenues from operations largely unrelated to data brokering. RELX reports four distinct types of operations – "Scientific, Technical, and Medical," "Risk and Business Analytics," "Legal," and "Exhibitions." All four segments draw on RELX's large personal data resources, manner, but the Risk and Business Analytics segment operates in a manner very similar to the eight other leading data brokers. Therefore, we use only those revenues derived from that segment. For some other data brokers, we also adjusted their reported revenues to account for varying fiscal years and to isolate their American revenues.

Transunion	$1,452.2	$1,636.2	$1,934.3	33.2%
Total	**$13,257.6**	**$13,825.3**	**$14,848.2**	**12.0%**

Data brokering has been expanding, but at a much lower rate (12 percent from 2016 to 2018) than sales of personal information by the major internet platforms (59.6 percent over the two years) If the nine major data brokers/credit bureaus sustain their recent growth rate over the next two and four years, the value of the personal information on Americans that they resell will reach $16.9 billion in 2020 and $19.6 billion in 2022. (Table 8, below)

Table 8: Major Data Brokers' Estimated U.S. Revenues for 2020 and 2022 ($ millions)

Companies	2018	2020	2022
Axciom	$1,053.0	$1,379.1	$1,806.2
CoreLogic	$1,650.5	$1,551.4	$1,458.3
Epsilon	$2,080.2	$2,098.2	$2,116.3
Equifax	$2,066.4	$2,202.5	$2,347.6
Experian	$3,070.7	$3,908.5	$4,974.8
FICO	$681.4	$810.5	$963.9
Harte-Hanks	$249.8	$179.0	$128.3
RELX	$2,061.9	$2,225.5	$2,402.2
Transunion	$1,934.3	$2,576.5	$3,431.8
Total	**$14,848.2**	**$16,933.1**	**$19,629.3**

U.S. Credit Card and Healthcare Data Companies

As noted earlier, most Americans do not consider their personal financial and health-related information to be market goods that can be bought and sold. Yet, in addition to credit bureaus and other data brokers, two major U.S. credit card companies also sell extensive personal information about Americans who use their services. Further, medical and other personal healthcare information is also bought and sold, much of it the leading company in the healthcare data business.

Credit Card Companies

The annual reports and SEC filings of the large credit card companies contain little information about whether they sell their users' personal records or, if they do, how they go about it and to whom they sell it. According to *Forbes Magazine,* Mastercard in particular "has bragged about the growth in its business of selling data to retailers, banks and governments on spending patterns."[49] To be sure, the Financial Services Modernization Act (FSMA) restricts the

[49] Cohan, Peter (2018). "Mastercard, AmEx And Envestnet Profit From $400M Business Of Selling Transaction Data" https://www.forbes.com/sites/petercohan/2018/07/22/mastercard-amex-and-envestnet-profit-from-400m-business-of-selling-transaction-data/#3fbe8c8e7722.

personal financial information that can be disclosed by financial institutions, including credit card companies.[50] Consequently, much of the personal financial data collected and analyzed by both banks and credit card companies are aggregated and/or anonymized before being sold. Hedge funds and institutional investors, for example, regularly use aggregated financial data purchased from credit card companies.[51] In practice, however, these legal limitations are at best problematic. To begin, companies with an interest in piercing the anonymity of financial data from covered institutions can do so easily with only a small amount of identifiable information about a person. A recent study in *Science* reported that given the time and place data on four transactions by an anonymized person, transactions drawn from a dataset of transactions by 1.1 million credit card users over three months, algorithms can de-anonymize and identify correctly more than 90 percent of the individuals in that dataset.[52] Once an individual is so identified, a purchaser of such anonymized data could access that individual's entire purchasing history.

MasterCard reports those information sales as "Other Revenues," which account for about 15.6 percent of the company's total revenues. By contrast, Visa does not tout its business or revenues from selling personal information about its "members" to nearly the same extent as MasterCard, and our research failed to find public reports of Visa selling its members' information. Beyond that, while MasterCard acknowledged in August 2018 that under a new partnership with Google, the two companies would work together in tracking people's offline retail purchases, Visa has announced no comparable agreement with Google or other platforms.[53] Visa also reports "Other Revenues" that amount to 3.4 percent of its total revenues, presumably without selling data provided by and about its own users. We will assume here that the difference in the dimensions or scale of the "Other Revenues" reported by the two companies corresponds roughly to MasterCard's earnings from selling its users' personal information. On this basis, we estimate that personal data sales accounted for $1.01 billion of MasterCard's $1.3 billion in "Other Revenues" in 2016, $1.185 billion of MasterCard's $1.52 billion in Other Revenues in 2017, and $1.42 billion of $1.82 billion in other revenues in 2018.

American Express (Amex) also analyzes its' cardholders' personal transactions and sells the results, mainly to advertisers targeting Amex members. However, Amex does not report "Other Revenues" or any other revenue classification associated with information sales. To estimate those revenues, we note that MasterCard earned $1.185 billion from selling information about its 212 million U.S. cardholders, or an average of $5.59 per-cardholder. If we assume that Amex earned a comparable amount per cardholder selling information about its 50 million U.S.

[50] Jolly, Ieuan (2018). "Data Protection in the United States: Overview." *Thomson Reuters*. https://content.next.westlaw.com /6-502-0467?transitionType=Default&contextData=(sc.Default)&firstPage=true&bhcp=1.
[51] Cohan, Peter (July 22, 2018). "Mastercard, AmEx and Envestnet Profit From $400M Business Of Selling Transaction Data." *Forbes*. https://www.forbes.com/sites/petercohan/2018/07/22/mastercard-amex-and-envestnet-profit-from-400m-business-of-selling-transaction-data/#3ee0bb327722; Jones, Rupert (June 24, 2013). "Barclays to sell Customer Data." *The Guardian*. https://www.theguardian.com/business/2013/jun/24/barclays-bank-sell-customer-data..
[52] de Montjoye, Yves-Alexandre, Laura Radaelli, et al (2015). "Unique in the Shopping Mall: On the Re-identifiability of Credit Card Metadata." *Science*, Vol. 347, Issue 6221, pp. 536-539.
[53] Since 2012, press accounts have noted that MasterCard uses its personalized data to generate revenues, and in 2014 a MasterCard executive stated that monetizing transaction data was "an incredibly fast-growing area for us."

members in 2017, American Express earned an estimated $279.6 million in that year selling personal information about member/cardholders.[54] This approach may well understate Amex's earnings from such operations: Since Amex cardholders on average spend more than MasterCard cardholders, advertisers should be willing to pay more for the personal information of Amex members than MasterCard members. With that caveat, we estimate that Amex also earned $238.2 million in 2016 and $334.5 million in 2018 selling personal information. Table 9 below presents the earnings of both credit card companies from selling members' personal information.

Table 9: Major Credit Card Companies' Estimated Revenues from Selling Cardholders' Personal Information, 2016-2018 ($ millions)

Company	2016	2017	2018	Growth
MasterCard	$1,010.1	$1,185.4	$1,418.1	40.4%
American Express	$238.2	$279.6	$334.5	40.4%
Total	**$1,248.3**	**$1,465.0**	**$1,752.6**	**40.4%**

Further, if the growth rate of these revenues remains steady over the next two and four years, the value of the personal information sold by MasterCard and American Express reach $2.46 billion in 2020 and $3.45 billion in 2022 (Table 10).

Table 10: Major Credit Card Companies' Estimated Revenues from Selling Cardholders' Personal Information, 2020 and 2022 ($ millions)

Company	2018	2020	2022
MasterCard	$1,418.1	$1,991.0	$2,795.3
American Express	$334.5	$469.6	$659.3
Total	**$1,752.6**	**$2,460.6**	**$3,454.5**

Healthcare Information

Americans' personal healthcare data also are routinely gathered, analyzed, and sold. The confidentiality of people's personal healthcare information is legally protected under the Health Insurance Portability and Accountability Act (HIPAA), but those protection apply only under certain circumstances. In particular, HIPAA's criminal and civil penalties for breaching confidentiality apply only to "covered entities," principally healthcare plans, healthcare providers, and clearinghouses that process healthcare claims. The restrictions also apply to business associates of covered healthcare plans, providers and clearinghouses, such as third

[54] "Most Popular Credit Card By Country: Visa, Mastercard Or Amex." February 7, 2017. https://merchantmachine.co.uk/visa-mastercard-amex/. For alternate numbers, see Statista (2019). "Number of credit cards in the United States from 2000 to 2017." https://www.statista.com/statistics/245385/number-of-credit-cards-by-credit-card-type-in-the-united-states/

parties who analyze data or process claims for a hospital or physician's office.[55] HIPAA restrictions do not affect most internet-based companies. For example, HIPAA does not reach the personal healthcare information collected by mobile health technologies ("mHealth") such as Glucosio, a mobile application that helps diabetes patients monitor their glucose, or activity trackers such as Fitbit that record heartbeat, exertion and other body data and transmit them back to the producers.[56] Dating websites that prompt members to answer health-related questions for their personal profiles also can ignore HIPAA. Even when the dating app for gay men, Grindr, shared identifying information about its members' HIV status with some of its partner companies, it faced no legal consequences for doing so.[57] Similarly, the personal profiles created and sold by retailers such as Target and Wal-Mart include purchases that the companies' algorithms can quickly align with medical conditions such as depression and obesity

HIPPA's most serious failure to protect the confidentiality of people's personal health information involves the major internet platforms. For example, Facebook captures and sells the personal information people post on Facebook groups for people dealing with cancer, depression, or other conditions, or Facebook discussion groups for people dealing with alcoholism, AIDS, anorexia, breast cancer, heart disease, migraines or chronic pain. HIPAA also does not restrict sites such as WebMD and KidsHealth from gathering, analyzing and selling personal information provided by their users, even inadvertently by the pages they peruse on such websites. More broadly, Facebook, Twitter and other social media capture the health-related information that people provide in their messages and personal posts on the platforms. Similarly, Google and other search engines record every visit to health-related sites for later analysis and sale, as well as health-related information contained in Google Gmail.

These loopholes also can be used as work-arounds for entities covered by HIPAA. For example, Optum, a company owned by the giant healthcare insurer and manager UnitedHealth Group, has filed for a patent for software that can scrape data from Facebook and Twitter to extract people's personal medical and healthcare payment information.[58] Finally, personal healthcare information held by companies covered by HIPAA can be captured, analyzed and sold if the information is provided in "anonymous formats." As a result, patient data from pharmacies, insurers, laboratories and electronic medical records are often transmitted to data mining companies in anonymized forms. For example, QuintilesIMS collects data covering most prescription drug sales in the United States, and while the pharmacies and hospitals that sell those data withhold patients' names, they do not withhold their age, doctor's name, partial zip

[55]Centers of Medicare & Medicaid Services (2019). "Are You a Covered Entity?" https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/HIPAA-ACA/AreYouaCoveredEntity.html.
[56] US Department of Health and Human Services (2016). "Examining Oversight of the Privacy & Security of Health Data Collected by Entities Not Regulated by HIPAA." https://www.healthit.gov/sites/default/files/non-covered_entities_report_june_17_2016.pdf.
[57] Neuman, Scott and Camila Domonoske (April 3, 2018). "Grindr Admits it Shared HIV Status of Users." *NPR*. https://www.npr.org/sections/thetwo-way/2018/04/03/599069424/grindr-admits-it-shared-hiv-status-of-users.
[58] Volz, Justin (2018). "Health Insurers are Vacuuming up Details about You – and it could raise your rates" ProPublica. https://www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates.

code, and many other personal details. QuintilesIMS also purchases nationwide healthcare data on people's medical claims and procedures, again without the patients' names but with other more general identifying information. Moreover, in place of the patient's name, the anonymization protocol assigns each person a unique numeric ID that all healthcare providers use for that person. Given the extensive information associated with each unique numeric ID, data mining and analytics can quickly de-anonymize anyone's personal healthcare information.

Let's consider an example of a person with a unique numeric ID who buys one prescription drug from a Duane Reade in Brooklyn, New York, another from a Rite-Aid in Fairfield, Connecticut, and undergoes a medical procedure at the Morristown Medical Center in northern New Jersey. The anonymized data for that person may enable analysts and academics to assess trends on health care utilization. But algorithms can easily identify those data with a specific named individual.[59] While we cannot say at what scale such deanonymization occurs – for example, to help drug companies, medical device producers, or private hospitals target potential customers – one expert in this field noted recently that "It's convenient to pretend it's hard to re-identify people, but it's easy. The kinds of things we did are the kinds of things that any first-year data science student could do."[60]

The company IQVIA dominates the market for anonymized personal pharmaceutical data. IQVIA, created from the merger of the medical data firm IMS Health and drug trial manager Quintiles, collects and analyzes large volumes of health-related information, including the sales of hundreds of pharmaceuticals, the diagnosis and claims data on millions of American patients, and other data on the characteristics and treatment of those millions of patients. One of IQVIA's three main divisions, "commercial solutions," focuses on purchasing and reselling anonymized patient and provider data, and all three divisions draw on the company's vast data resources.

Based on the company's annual reports, IQVIA's revenues from its medical data division jumped 233.4 percent from 2016 to 2017, followed by 13.8 percent growth from 2017 to 2018.

Table 11: IQVIA's Estimated Earnings from Selling Anonymized Personal Medical Information, 2016-2018, ($ millions)

2016	2017	2018
$443.4	$1,478.1	$1,681.5

[59] Tanner, Adam (2017). "Strengthening Protection of Patient Medical Data." *The Century Foundation*. https://tcf.org/content/report/strengthening-protection-patient-medical-data/

[60] Solon, Olivia (2018). "'Data is a Fingerprint': Why You Aren't as Anonymous as You Think Online." *The Guardian*. https://www.theguardian.com/world/2018/jul/13/anonymous-browsing-data-medical-records-identity-privacy. See also Tanner, Adam (2017). "The Hidden Trade in our Medical Data: Why We Should Worry." *Scientific American*. https://www.scientificamerican.com/article/the-hidden-trade-in-our-medical-data-why-we-should-worry/

If we apply the very conservative assumption that IQVIA will sustain its most recent 13.8 percent annual growth over the next two and four years, IQVIA's revenues from selling anonymized personal medical information will reach nearly $2.2 billion in 2020 and more than $2.8 billion in 2022.

Table12: IQVIA's Projected Earnings from Selling Anonymized Personal Medical Information, 2018, 2020 and 2022 ($ millions)

2018	2020	2022
$1,681.5	$2,177.5	$2,820.0

Conclusions

We have established that the gathering, analysis and sale of people's personal information is a very big business. Here, we examined the activities and revenues associated with monetizing Americans' personal information in four areas – the major internet search engines, browsers and social media platforms; large data brokers, credit card companies and the healthcare data business. In 2018, those revenues totaled more than$78 billion. Moreover, if the recent growth rates of those revenues persist in those four areas, the total will reach nearly $116 billion in 2020 and almost $198 billion by 2022.

If we agree with most Americans that each of us owns his or her own personal information, then Americans should receive a substantial share of those revenues, through direct payments or to fund broad public purposes. With a 50-50 division of earnings derived in just these four areas from people's personal information, every American who uses the internet would receive $183 in 2020 and $308 in 2022. Alternatively, a 50-50 split could be used to fund infrastructure, advance universal healthcare coverage, cut payroll taxes, or substantially reduce the budget deficit. Expanding this analysis to cover, for example, major retailers, large hotel and supermarket chains, and U.S. manufacturing could substantially increase the dimensions of the fees or public policies that the revenues could cover. At a minimum, Congress and the President should formally acknowledge that all Americans have clear property rights to their own personal information.

* * *

About the Authors

Robert J. Shapiro is the chairman of Sonecon, LLC, a private firm that provides economic and security-related analysis and advice to senior officials of the U.S. and foreign governments and senior executives of American businesses and non-profit organizations. He is also a Senior Policy Fellow of the Georgetown University McDonough School of Business, a board director of Medici Ventures, and an Advisory Board member of Cote Capital and Gilead Sciences. Dr. Shapiro has advised, among others, President Bill Clinton, Vice President Al Gore, Jr., British Prime Minister Tony Blair, Treasury Secretaries Timothy Geithner and Robert Rubin, British Foreign Secretary David Miliband, and many U.S. Senators and Representatives. He also has advised senior executives of global companies including AT&T, Exxon-Mobil, Amgen, Gilead Science, Google, Elliot Management and Fujitsu, as well as non-profit organizations such as the International Monetary Fund, the Center for American Progress and PhRMA. Before establishing Sonecon, Dr. Shapiro was the Under Secretary of Commerce for Economic Affairs. Prior to that position, he was co-founder and Vice President of the Progressive Policy Institute and, before that, Legislative Director and Economic Counsel for Senator Daniel Patrick Moynihan. Dr. Shapiro also served as the principal economic advisor to Bill Clinton in his 1991-1992 campaign, as a senior economic advisor to Hillary Clinton in 2015-2016, and as economic advisor to the presidential campaigns of Barack Obama, John Kerry and Al Gore. He has been a Fellow of Harvard University, the Brookings Institution, and the National Bureau of Economic Research. Dr. Shapiro holds a Ph.D. and M.A. from Harvard University, a M.Sc. from the London School of Economics and Political Science, and an A.B. from the University of Chicago.

Siddhartha Aneja is a Senior Analyst and Director of Sonecon, LLC, where he has conducted extensive quantitative analysis of the internet, educational outcomes, tax policies, health care costs, and other economic matters. Prior to joining Sonecon, he was a research associate at the Institute for Education and Social Policy at New York University (NYU), where he conducted extended research on links between childhood health, employment, and educational outcomes and on other issues related to urban education. Mr. Aneja's research has been published in peer-reviewed journals including the *JAMA Pediatrics* and the *Journal of School Health*. He also served as an Americorps Volunteer for City Year Little Rock. Mr. Aneja holds a B.A. in Mathematics-Economics from Wesleyan University and a M.P.A. from the NYU Robert F. Wagner Graduate School of Public Service, and he is currently studying law at the Georgetown University School of Law.