

Your Data Isn't Gold; It's Not Even Yours

DAVID MOSCHELLA | APRIL 2022

Defending Digital Series, No. 4: Claims that Big Tech is making too much money off of “our data” are wrong in two fundamental ways: The data about most individuals isn't worth very much—and when consumers use a business service, the resulting data isn't “theirs.”

We hear it so often that it's easy to assume it must be true. “Our data is gold, and we should be compensated for it.” These two statements basically tell consumers that they are being taken advantage of, even ripped off by Big Tech. Not surprisingly, this has led to resentment and calls for action. But there is just one problem: Both assertions are much more wrong than right. Today's leading technology companies are extraordinarily profitable, but this is far more due to the unique features of information economics than any data ownership or usage abuses.

HOW MUCH IS YOUR DATA WORTH?

In recent years, there have been numerous efforts to determine how much our individual data is worth. Some studies have looked at Big Tech's revenue per user; others focus on market capitalization per user. We think profits per user is the best starting point because any ongoing payments to individual consumers would come directly out of a company's bottom line.

Calculating profits per consumer is much easier to do at a global level because that's where publicly reported financial data is most consistently available. We can further simplify things by just assessing Alphabet (Google) and Meta (Facebook). While Apple, Microsoft, and Amazon all sell data-driven advertising to some extent, it's a very small share of their overall businesses. Similarly, we don't need to include Twitter because it's relatively small, and it lost money in 2021.

Let's begin with Alphabet. In the year ending December 31, 2021, the company had revenues of \$257 billion and a net income of \$76 billion. According to its annual 10K filing, advertising revenues accounted for \$209 billion, or 81 percent of total revenues. As ads are Alphabet's cash cow, let's assume that 90 percent of its profits were also ad-based. This translates into \$68 billion in advertising profits in 2021. Now let's suppose that, for whatever reasons, Alphabet makes the extraordinary gesture of giving half of this (\$34 billion) back to its data sources.

Since Alphabet gets valuable data from both the users of its services as well as Internet, YouTube, and other content providers, let's say that it gives both groups half of the \$34 billion, or \$17 billion each. How should we divide up this \$17 billion? Alphabet says that it has nine businesses with more than a billion users—Search, Chrome, Gmail, Android, YouTube, Maps, Google Play, Google Drive, and Google Photos. Best estimates are that roughly 4 billion people use at least one Alphabet product. Using the same 81 percent share above, let's say that some 3

billion use its ad-based services. If we divide \$17 billion by 3 billion users, we get \$5.60 per consumer during 2021, more like a cup of artisan coffee than a pot of gold.

We can do the same analysis for Meta. The company's 2021 revenues were \$118 billion, with a net income of \$39.4 billion, virtually all of which was driven by advertising. Once again, let's assume Meta keeps half of its profits for itself and splits the other half (\$19.7 billion) between its consumers and content providers. However, as Meta is much less reliant on outside content providers than Alphabet, let's give just one-quarter of this split (\$4.9 billion) to providers and three quarters (\$14.8 billion) to Meta's estimated 2.9 billion consumers. Dividing \$14.8 billion by 2.9 billion customer translates to \$5.10 per user, remarkably similar to Alphabet's \$5.60.

If the Alphabet and Meta payments to consumers are added together, you get \$10.70 per customer per year. If you include half of the advertising-based profits of Microsoft, Amazon, and Apple, you'd get a little more than \$12. But remember: This is making the colossal assumption that all five of these companies give up *half* of their advertising-based profits. If they only gave back 10 percent, then the total annual payment would be about two bucks per user per year.

Although any such annual numbers will continue to grow, they won't add up to much anytime soon. Moreover, the effort needed to calculate and disperse small individual payments would be enormous. Perhaps most fundamentally, who among us thinks that the services these companies provide *for free* aren't worth far more than \$12 per year? By this metric, it's one of the great bargains in economic history, which is why these companies are so successful.

Of course, one could argue that some people's data is worth a lot more than \$12. This is surely true. But the data that is worth the most tends to come from the most affluent among us. Do we really want to argue that Alphabet and/or Meta's most affluent customers should get a lot more than \$12, knowing that this would mean that everyone else would get a lot less? The same is true geographically. Per-user profits in the United States and Europe for both Alphabet and Meta are higher than in the developing world, but do we really want consumers in less-developed nations to get very little, even if they use the services as much as consumers in wealthier countries?

The low payment numbers, high implementation costs, and fairness sensitivities explain why Big Tech will likely remain reluctant to support such schemes in any major way, even though they are really just the digital version of a traditional loyalty program. These operational challenges also explain why the Data Dividend Project and related start-ups, such as Invisibly and UBDI, have yet to catch on. It's entirely possible—perhaps even likely—that, eventually, some combination of blockchains, peer-to-peer architectures, avatars, virtual agents, NFTs, cryptocurrencies, and new data-collection and usage norms will spawn a generation of important data intermediaries, but today such efforts are mostly unfeasible and of dubious value.

IT'S NOT YOUR DATA ANYWAY

The principles of customer data ownership are as familiar as business recordkeeping. Credit card issuers itemize our purchases; cable TV operators track what we watch; telecom companies know what numbers we've called; loyalty-card issuers reward frequent customers; health-care providers store our medical records; schools know what we studied and what our grades were; governments record what real estate we own, what countries we've visited, and so much more. No one doubts

that these organizations own this data. We might be able to see it, challenge its accuracy, or limit its use, but in no way is it “ours.”

The rules governing the use of this type of data are based on either explicit or implied understandings between consumers and providers, with each industry having its own set of norms, obligations, and limitations. Although consumers can decide whether to keep a record of the products and services they use, most of us don’t bother. Providers, however, have no such choice. Detailed data collection is typically required for legal, accounting, billing, customer service, and many other purposes. The only question is how this data is or isn’t to be used. It’s the same question with Big Tech.

If anything, the rights of providers are even stronger when services are supported by advertising and thus provided *for free*. As the old saying accurately goes, if you are not paying for a product, you are the product, as advertisers are paying for your attention. Additionally, technology consumers often have more choices than in many of the industries listed above. Most people don’t have to use Facebook; it’s easy to switch from Google search to DuckDuckGo, or use one of the many alternatives to Gmail. These Big Tech alternatives will likely continue to gain momentum going forward. Consider the stunning growth of TikTok, or the way Amazon’s Alexa uses Microsoft’s Bing as its search engine.

None of this is meant to argue that consumers and policymakers can’t or shouldn’t try to improve the terms of their data-usage deals. Today’s end-user license agreements, privacy settings, anonymity implementations, customer-profiling services, portability support, and many other practices are byzantine in nature, and even sophisticated consumers often can’t be bothered to understand or tweak them. But even the most worthwhile efforts to improve consumer data protection won’t change the fundamental fact that providers own the data, and you don’t.

ECONOMICS TRUMPS ANALOGIES

Some readers might counter that although all of the above is true enough, it still leaves us with the problem of Big Tech companies making “too much money.” The combined profits of the big five companies reached an astonishing \$350 billion in 2021, and they are still rising. (But keep in mind that these same high profits enabled these companies to invest \$136 billion in R&D in 2020, and surely more in 2021.)

The reason today’s tech leaders are so profitable has much more to do with information economics than anything to do with data ownership. Although information economics is a term used in a variety of ways, we use it to encompass the unique features of digital technology—nearly infinite economies of scale, powerful network effects, zero marginal costs, perfect reproducibility, exponential increases in volume, and winner-take-all market-share tendencies. Google performs over five billion searches every day, and 300 million photos are uploaded to Meta per day. It is this combination of massive scale and high utility—not advertising abuses or unfair data ownership—that best explains the vast riches of today’s tech giants.

This is why analogies with gold or oil—as well as less frequent comparisons to electricity, water, air, and exhaust—typically yield more shade than light. Data is not like gold, which is a scarce commodity with a largely fixed supply and a narrow set of potential uses. The analogy between data and oil is better, as both can power businesses, drive innovation, and create great wealth. However, like gold, oil doesn’t have any of the information economics characteristics listed

above. Perhaps most fundamentally, data and software are non-rivalrous goods, meaning that my use of a product doesn't prevent you from using it, too. No material goods have this property. These differences explain why efforts to understand the extraordinary success of Big Tech should rely much more on economics than analogies.

In most industries, unusually high profit margins tend to get whittled away over time by competition, market maturation, new business models, technological disruption, expensive new ventures, and/or the bloated costs companies' typically take on once they become rich. Odds are that similar economic forces will eventually rein in Big Tech as well. A future installment in this "Defending Digital" series will address the many new forms of competition today's leaders will likely face in the 2020s.

But as of now, data, information, and knowledge comprise a unique hierarchy of value that historical analogies cannot speak to. One of the wonders of the information age is how data that is worth very little at an individual level becomes extraordinarily valuable when collected at massive scale. It's like a modern form of alchemy. Your data isn't gold; it's not even yours. But when everyone's data is collected and leveraged, the value to companies—and to society—is worth more than all the gold on planet Earth.¹

About This Series

ITIF’s “Defending Digital” series examines popular criticisms, complaints, and policy indictments against the tech industry to assess their validity, correct factual errors, and debunk outright myths. Our goal in this series is not to defend tech reflexively or categorically, but to scrutinize widely echoed claims that are driving the most consequential debates in tech policy. Before enacting new laws and regulations, it’s important to ask: Do these claims hold water?

About the Author

David Moschella is a non-resident senior fellow at ITIF. Previously, he was head of research at the Leading Edge Forum, where he explored the global impact of digital technologies, with a particular focus on disruptive business models, industry restructuring and machine intelligence. Before that, David was the worldwide research director for IDC, the largest market analysis firm in the information technology industry. His books include *Seeing Digital – A Visual Guide to the Industries, Organizations, and Careers of the 2020s* (DXC, 2018), *Customer-Driven IT* (Harvard Business School Press, 2003), and *Waves of Power* (Amacom, 1997).

About ITIF

The Information Technology and Innovation Foundation (ITIF) is an independent, nonprofit, nonpartisan research and educational institute focusing on the intersection of technological innovation and public policy. Recognized by its peers in the think tank community as the global center of excellence for science and technology policy, ITIF’s mission is to formulate and promote policy solutions that accelerate innovation and boost productivity to spur growth, opportunity, and progress. For more information, visit itif.org.

ENDNOTE

1. Best estimates are that all the known holdings of gold in the world are worth about \$10 trillion; the current market cap of just Apple, Alphabet, Microsoft, Amazon and Facebook is roughly \$9 trillion.