



Let's Build A BETTER DATA ECONOMY

Our digital footprint generates enormous value, but too much of it ends up in Big Tech silos

Yan Carrière-Swallow and Vikram Haksar

Humanity has never been so comprehensively recorded. Smartwatches capture our pulse in real time for a distant artificial intelligence (AI) to ponder the risks of heart disease. Bluetooth and GPS keep track of whether some of us shop at gourmet stores and linger in the candy aisle. Our likes and browsing hours on social media are harvested to predict our credit risk. Our search queries on shopping platforms are run through natural language processors to generate uniquely targeted ads whose unseen tethers subtly remold our tastes and habits.

The generation and collection of data on individual human beings has become a big part of the modern economy. And it generates enormous value. Big data and AI analytics are used in productivity-enhancing research and development. They can strengthen financial inclusion. During the pandemic, data on real-time movements of entire populations have informed policymakers about the impact of lockdowns. Contact tracing apps have notified individuals who have been in potentially dangerous proximity to people infected with COVID-19.

But just as data have helped us monitor, adapt, and respond to COVID-19, the pandemic has brought into focus two fundamental problems with how it flows in the global economy (Carrière-Swallow and Haksar 2019). First, the

data economy is opaque and doesn't always respect individual privacy. Second, data are kept in private silos, reducing its value as a public good to society.

Whose data anyway?

Once the GPS, microphones, and accelerometers in the smart devices located in every pocket and on every bedside table and kitchen counter begin monitoring our behavior and environment, where do the data go? In most countries, they are collected, processed, and resold by whoever can obtain them. User consent is all too often granted by checking a box below lengthy legalistic fine print—hardly a means to serious informed consent. Analysis based on such granular data is a gateway to influencing behavior and has tremendous commercial value. To be sure, this is not a one-way street: consumers get many nice data-driven features for no direct financial cost in exchange. But are they getting enough?

Most transactions involving personal data are unbeknownst to users, who likely aren't even aware that they have taken place, let alone that they have given permission. This gives rise to what is known in economics as an externality: the cost of privacy loss is not fully considered when an exchange of data is undertaken. The consequence is that the market's opacity probably leads to *too much data*

Why are people willing to hand over their location data in exchange for a weather forecast, but not to share it to protect their health?

being collected, with *too little of the value being shared* with individuals.

By agreeing to install a weather application and allowing it to automatically detect its current city, people might unwittingly allow an app designer to continuously track their precise location. Users who sign up for a weather forecast with a sleek interface agree to share their location data, believing it's just to enable the app's full functionality. What they are providing, in fact, is a data trail about their daily routine, travel itinerary, and social activity. The weather forecaster may never get any better at predicting rain but could end up with a better prediction of the user's creditworthiness than the scores compiled by traditional credit bureaus (Berg and others 2020).

Privacy paradoxes

Do we care about our privacy or not? Researchers have documented what is known as a “privacy paradox.” When asked to value their privacy in surveys, people frequently rank it as a very high priority. However, in their daily lives, these same people are often willing to give away highly sensitive personal data for little in exchange.

This paradox should have heralded good news for contact tracing apps, which rely on widespread usage to be effective (Cantú and others 2020). Unfortunately, in many countries where use of these tools is voluntary, take-up has been very low. Why are people willing to hand over their location data in exchange for a weather forecast, but not to share it to protect their health while helping fight a global pandemic that has killed over 2 million people? One reason may be that—unlike the weather app makers—public health agencies have designed their contact tracing apps to transparently announce how they will be collecting and using data, and this triggers concerns about privacy. Another reason is that authorizing governments to combine location information with data on a disease diagnosis may be seen as particularly sensitive. After all, knowledge of someone's preexisting condition could lead

to their exclusion from insurance markets in the future or open the door to other forms of stigma or discrimination.

How to use responsibly

The data generated by our smart devices are essentially a private good held by Big Tech companies that dominate social media, online sales, and search tools. Given how valuable these data are, it is not surprising that companies tend to keep them to themselves (Jones and Tonetti 2020). As more data beget better analysis, which in turn attracts more usage, more data, and more profits, these swollen data war chests fortify their platform networks and potentially stifle competition.

This finders keepers model tends to lead to too much data being collected, but the data are also *insufficiently utilized* exactly when they could be most helpful, kept in private silos while public needs remain unmet. Data sharing can support the development of new technologies, including in the life sciences. Consider how epidemiological research can benefit from scaling up big data analytics. A single researcher analyzing the experience of patients in their home country may be a good start, but it cannot rival the work of many researchers working together and drawing on the experience of many more patients from around the world—the key to the success of a number of cross-border collaborations.

How can data be made more of a public good? Commercial interests and incentives for innovation must be balanced with the need to build public trust through protection of privacy and integrity. Clarifying the rules of the data economy is a good place to start. Significant advances have resulted, for example, from the 2018 implementation in Europe of the General Data Protection Regulation (GDPR), which clarified a number of rights and obligations governing the data economy. EU residents now have the right to access their data and to limit how it is processed, and these rights are being enforced with increasingly heavy fines. But even as researchers

have started to see the impact of the GDPR on the digital economy, there are still concerns about how to operationalize these rights and keep them from being simply a box-checking exercise.

People should have more agency over their individual data. There could be a case to consider the creation of public data utilities—perhaps as an outgrowth of credit registries—that could balance public needs with individual rights. Imagine an independent agency tasked with collecting and anonymizing certain classes of individual data, which could then be made available for analysis, subject to the consent of interested parties. Uses could include contact tracing to fight pandemics, better macroeconomic forecasting, and combating money laundering and terrorism financing.

Policies can also help consumers avoid becoming hostage within individual ecosystems, thus contributing to market contestability and competition. The European Union's late-2020 proposals for the Digital Markets Act and the Digital Services Act have many new features. These include third-party interoperability requirements for Big Tech “gatekeepers”—including social media and online marketplaces—in certain situations and efforts to make it easier for their customers to port their data to different platforms.

Policies also have a role to play in keeping data secure from cyberattacks. An individual company does not fully internalize the harm to public trust in the entire system when its customers' data are breached, and may thus invest less in cybersecurity than what would be in the public interest. This concern has special resonance in the financial system, where maintaining public confidence is crucial. This is why secure infrastructure, cybersecurity standards, and regulation are essential pillars of the open banking policies many countries have adopted to facilitate interoperability in sensitive financial data.

Global approach

Many countries have been developing policies aimed at a clearer, fairer, and more dynamic data economy. But they are taking different approaches, risking greater fragmentation of the global digital economy. These risks arise in many data-intensive sectors, ranging from trade in goods to cross-border financial flows. In the context of the pandemic, differing privacy protection standards make it harder to collaborate

on crucial medical research across borders—true even before the pandemic—because of the difficulty of sharing individual results of biomedical trials (Peloquin and others 2020).

Global coordination is always a challenge, especially in an area as complex as data policy, where there is a multitude of interests and regulators even within individual countries, let alone across borders. Dealing with the fallout of the pandemic has spurred a new opportunity to ask hard questions about the need for common minimum global principles for sharing data internationally while protecting individual rights and national security prerogatives.

The current moment also affords an opportunity to explore innovative technological solutions. Consider whether jump-starting the recovery in international travel could be facilitated by a global vaccine registry. This could leverage old-fashioned paper-based international health cards but would call for development of standards and an interoperable data management system for reporting and consulting on people's vaccination status—potentially linked to digital identity—as well as agreements on protection of individual privacy and barriers to access for other purposes.

There is a strong case for international cooperation to ensure that the benefits of the global data economy can build a more resilient, healthier, and fairer global society. To find a way forward together, we can start by asking the right questions. **FD**

YAN CARRIÈRE-SWALLOW is an economist in the IMF's Strategy, Policy, and Review Department. **VIKRAM HAKSAR** is an assistant director in the IMF's Monetary and Capital Markets Department.

References:

- Berg, Tobias, Valentin Burg, Ana Gombovic, and Manju Puri. 2020. “On the Rise of FinTechs: Credit Scoring Using Digital Footprints.” *Review of Financial Studies* 33:2845–97.
- Cantú, Carlos, Gong Cheng, Sebastian Doerr, Jon Frost, and Leonardo Gambacorta. 2020. “On Health and Privacy: Technology to Combat the Pandemic.” *BIS Bulletin* 17 (May).
- Carrière-Swallow, Yan, and Vikram Haksar. 2019. “The Economics and Implications of Data: An Integrated Perspective.” Departmental Paper 19/16, International Monetary Fund, Washington, DC.
- Jones, Charles I., and Christopher Tonetti. 2020. “Nonrivalry and the Economics of Data.” *American Economic Review* 110 (9): 2819–58.
- Peloquin, David, Michael DiMaio, Barbara Bierer, and Mark Barnes. 2020. “Disruptive and Avoidable: GDPR Challenges to Secondary Research Uses of Data.” *European Journal of Human Genetics* 28:697–705.