FRONTIER TECHNOLOGY QUARTERLY

January
2019

Data Economy: *Radical transformation or dystopia?*

Data is shaping the future of humanity. The production, distribution and consumption of digital data—the data economy—are driving rapid advances in machine learning, artificial intelligence and automation. Individuals and businesses are using data to reduce search and transaction costs and make informed choices. Data is facilitating scientific and medical research, making societies more productive. It is helping to improve the efficacy of public policy, delivery of public services, transparency and accountability. Data is helping us track progress on every sustainable development goal in the 2030 Agenda and deliver broad-based social welfare.

I. Data economy is different

Distinctions between buyers and sellers or consumers and producers are blurred in the data economy. Supply and demand do not necessarily determine price, price is often indeterminate or implicit, and yet, enormous values are created in the data economy. Data is increasingly a critical factor of production, complementing labour and physical capital. But unlike capital or labour, data is non-depletable. The use of data by many does not diminish its quantity or value. On the contrary, the use of the data by many may increase its value. At the same time, data can become less relevant, and less valuable, over time. The value of data, unlike physical capital, also depends on its unique characteristics. An individual data point can carry little value, but its value can multiply manifold when aggregated and analyzed with other relevant data.

Standard economic theories are increasingly deficient to explain the workings of the data economy. Data is also non-rival—millions can use it simultaneously. Yet, data is not necessarily a public good because data owners can exclude people from accessing and using it. The value of data can depend on it being private, determining who can use it and who cannot. Furthermore, data can be stored and transported at very low cost. Individuals, households, businesses are both often producers and consumers in the data economy, with firms extracting, analyzing and intermediating data. A Google search, for example, produces a data point—its algorithm crunching and analyzing millions of underlying data stored in hundreds of different computers across many countries. That search result—a new data point—then instantaneously feeds into the algorithm and becomes a factor input to refine future Google searches.

Given the reasons above, it is hard to precisely value data. How value is generated in the data economy—and how that value is shared among market participants—has important competitive and distributional implications that merit appropriate policy responses. On one hand, the data economy is radically transforming many economic activities and creating new levels of prosperity. On the other, it presents the possibility of a perilous dystopia, where participants in the data economy can face chronic trust deficits and insecurity. People cannot often trust data. They do not know what they actually pay and what they get in return. People do not know whether they are paying more or less than others. They do not know if they are targeted by market researchers or advertisers. Consumer protection is generally weak in the data economy. Furthermore, the collection and use of personal data, designed to influence behavior, carries with it an ever-present potential for abuse. Political interests can access personal data to engage in highly targeted campaigns that appeal to the narrow interests of specific groups rather than societal interests. These efforts can be as effective as they are devastating. A market economy cannot function without trust, and the data economy is no exception. Trust deficits can unravel the data market and undermine social cohesion, stability and peace. Uniform standards, quality controls and regulations enhancing consumer protection can instill trust in the data economy and make it work for all.

World Economic and Social Survey 2018: Frontier Technology for Sustainable Development—a flagship publication of the United Nations Department of Economic and Social Affairs (DESA)—generated considerable interest in new technologies and their development impacts. Inspired by this strong interest, the Economic Analysis and Policy Division has undertaken to produce quarterly reviews on frontier technologies, delving deeper into specific aspects of a new technology. The series will identify challenges and raise many questions—and answer a few—while motivating policy research in UNDESA and beyond. The first edition of the series focuses on data economy. The second edition will discuss genetic technologies. Subsequent quarterly reviews will address technological advances in new materials and blockchain technology. The quarterly reviews will be shared and discussed in development policy seminars and social media platforms to enrich policy discourse on frontier technologies.

Hoi Wai Jackie Cheng, Marcelo LaFleur and Hamid Rashid authored the Quarterly under the supervision of Hamid Rashid, Chief of the Development Research Branch, Economic Analysis and Policy Division of UNDESA. Nazrul Islam, Alex Julca, Hiroshi Kawamura and Mariangela Parra-Lancourt provided useful comments on the draft. Research support was provided by Nicole Hunt. Typesetting was done by Nancy Settecasi. The views and opinions expressed herein are those of the authors' and do not necessarily reflect those of the United Nations Secretariat.

United Nations promotes the use of data for development

The United Nations has an important role in shaping how data will impact our future, ranging from facilitating negotiations on a multilateral framework on data to making sure data is a positive force for peace, development and human rights. Achieving the SDGs requires global action of unprecedented ambition on social, environmental and economic challenges. The data revolution helps in this regard, bringing about a shift in the way governments and the public and private sectors use data and statistics.

Some examples of how the UN System is working to identify and promote the use of data for development include:

- In 2017, the United Nations Development Group produced a guidance note on Big Data for the 2030 Agenda, which aims to establish common principles across UNDG to support the operational use of Big Data for achieving SDGs, to serve as a risk-management tool accounting for human rights, and to set principles for obtaining, retaining, using and controlling the quality of data from the private sector.
- The UN Global Pulse has produced numerous reports on how Big Data can be used for supporting sustainable development.
- The UN Data Revolution Report (2014) focuses on how to fill data gaps, close data divide, improve data quality, and prevent people from data-related abuses.
- The UN World Data Forum 2018 addressed how to leverage data for sustainable development and how to improve migration statistics and the political economy of statistical capacity. The first Forum in 2017 focused on capacity development for better data, innovations and synergies across different data ecosystems, development use of data, and data principles and governance, all with the objective of supporting government programmes and initiatives.

Data is ubiquitous

The digital age is producing a vast amount of data every second. Devices and people actively share their data and leave behind rapid, real-time trails of data. Google, for example, now processes on average over 71,966 search queries every second, which translates to over 6.2 billion searches per day and 2.3 trillion searches per year worldwide. If we were to produce hard copy of every piece of data, we would generate nearly 1 billion 200-page books per second.

Table 1

Data generated in a second

Emails sent	2,763,771
Google searches	71,966
YouTube videos watched	77,134
Tweets sent	8,342
Internet traffic (Gigabytes)	67,023
Total data* (Gigabytes)	289,351

Sources: http://www.internetlivestats.com/one-second/#google-band
* https://www.domo.com/learn/data-never-sleeps-5?aid=ogsm072517_1&sf100871281=1
Note: 1 GB= 677,963 pages of texts

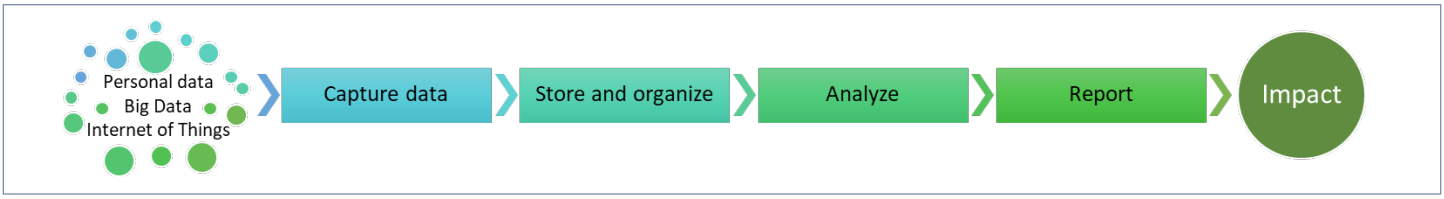
The value chain (Figure 1) in the data economy begins with collecting personal and non-personal data and making them available for storage and eventual analysis. Machine learning algorithms use copious amounts of data to detect patterns and relationships in the data that are otherwise too difficult to detect. Developers of self-driving cars, for example, process large datasets of road and traffic information using machine learning to "train" their software on how to cope with the unpredictable nature of real traffic conditions. The processed data, that is, traffic and road information, is then made available to end-users, who use it to make decisions. A weather app—giving forecasts of temperature, rain or sunshine available on our Smartphone—helps us decide what to wear in the morning. A transit or a ride-sharing app—analyzing commuting times and costs—help us choose the most efficient and cost-effective route to work. Google searches throughout the day allows us to conduct research. Like individuals, businesses also interface and interact with data to reduce costs and maximize profit.

Many firms in the data market perform multiple value generating functions: capturing, storing, transporting, analyzing and reporting data outputs. In many cases, data is exchanged without a monetary transaction. Some data firms, for example Facebook and Google, provide a service that users enjoy without paying for it directly. Others such as Airbnb and Uber intermediate the supply of and demand for a service, and collects a fee for this service. Then,

Figure 1

The data market as a value chain: from data sources to economic impact

Source: Author elaboration, based on "European Data Market - SMART 2013/0063 - Final Report".



there are service or content providers that sell digital contents, products or services. At the end of the spectrum, there are firms that develop operating systems, software and hardware that enables capture, storage, process and transmission of data. Data giants like Amazon engage in the entire data value chain, capturing data from consumers, organizing and analyzing this data, extracting useful insights, sharing with third party sellers thus creating new markets.

Table 2
Dominant players in the data economy: a partial snapshot

Niche areas	Dominant firms
Search engines	Google
Social media/Messaging	Facebook, WhatsApp, WeChat
Share economy platforms	Uber, Airbnb
Content and service provider	Netflix, Venmo, Expedia
Retailer	Amazon, eBay, Alibaba
Operating systems	Microsoft, Apple, Google
Data hardware	Apple, Samsung, Cisco

Source: Authors.

How large is the data economy?

The data economy is still very small as a share of GDP. In the European Union, the value of the data market—the aggregate revenue of all firms in the data economy—reached 65 billion euros in 2017, representing only 0.49 per cent of GDP and employing[1] 6.7 million people. In the United States and Japan, the data economy is 1 per cent and 0.8 per cent of GDP, respectively. The size of the data market is much smaller in emerging and developing economies. In the European Union, for example, the total impact of the data market on the region's economy in 2017 was 335.6 billion euros, or 2.4 per cent of total GDP (Figure 2).[2]

The share in GDP, however, belies the real market size and economic influence of the data economy. Twenty-five of the largest technology firms, mostly data firms, had a combined market valuation of nearly $6 trillion in 2016, representing nearly 20 per cent of market capitalization in the United States. The five largest firms in the world—Apple, Amazon, Facebook, Google and Microsoft—are actors in the data economy with a combined market value of nearly $4 trillion in 2018. In 2008, only one among the five largest firms was a data firm.

The disproportionate large share of market capitalization of the data firms represents investors' confidence in their innovation prowess, product niche and monopoly market powers. Their share prices reflect large valuation premiums. Large data firms (Table 3) typically employ fewer people, invest less in physical assets and generate less revenue, yet they are valued significantly higher compared to their traditional brick-and-mortar counterparts of similar size. This phenomenon has some first- and second-order effects. Market over-valuation of data firms can represent an irrational exuberance among investors, with detrimental effect on the regular economy. More importantly, higher market valuation of these firms represents transfers of purchasing power and capacity to invest from households and smaller firms to larger data firms that may not necessarily invest those proceeds. The net effect on the economy can be negative, depressing overall employment, investment and aggregate demand. The higher valuation of these firms

Table 3
Five largest data firms and their brick and mortar counterparts

	Total number of employees	Total revenue (bn, $)	Total market valuation (bn, $)	Price-Earning ratio
Apple	132,000	266.0	868.8	17.20
Amazon	566,000	178.0	560.0	190.16
Facebook	25,105	40.7	508.9	32.74
Google	85,050	89.4	720.8	58.65
Microsoft	134,944	110.0	570.0	57.34
Exxon	69,600	237.0	348.6	17.52
Johnson and Johnson	134,000	76.0	375.0	22.13
Proctor and Gamble	95,000	65.0	226.8	23.93
Royal Dutch Shell	92,000	305.0	271.9	12.51
Walmart	1,500,000	500.3	289.8	22.69

Source: Authors, based on https://www.macrotrends.net/ and other sources.

Figure 2
Components of EU's data economy in 2017, share of total

Source: Author elaboration, based on "European Data Market - SMART 2013/0063 - Final Report".

1 Workforce who collect, store, manage and analyze data as their primary activity, or as a relevant part of their activities.

2 The concept of the data economy is limited to economic impacts and does not include the impact of data-driven innovations on health, safety, recreation, air quality, or other quality-of-life measures.

also translates to larger income and wealth inequality, as wealth is increasingly concentrated among the few owners and investors of large data firms. The five richest individuals in the world today are data entrepreneurs.

II. Why idiosyncrasies of the data economy matter

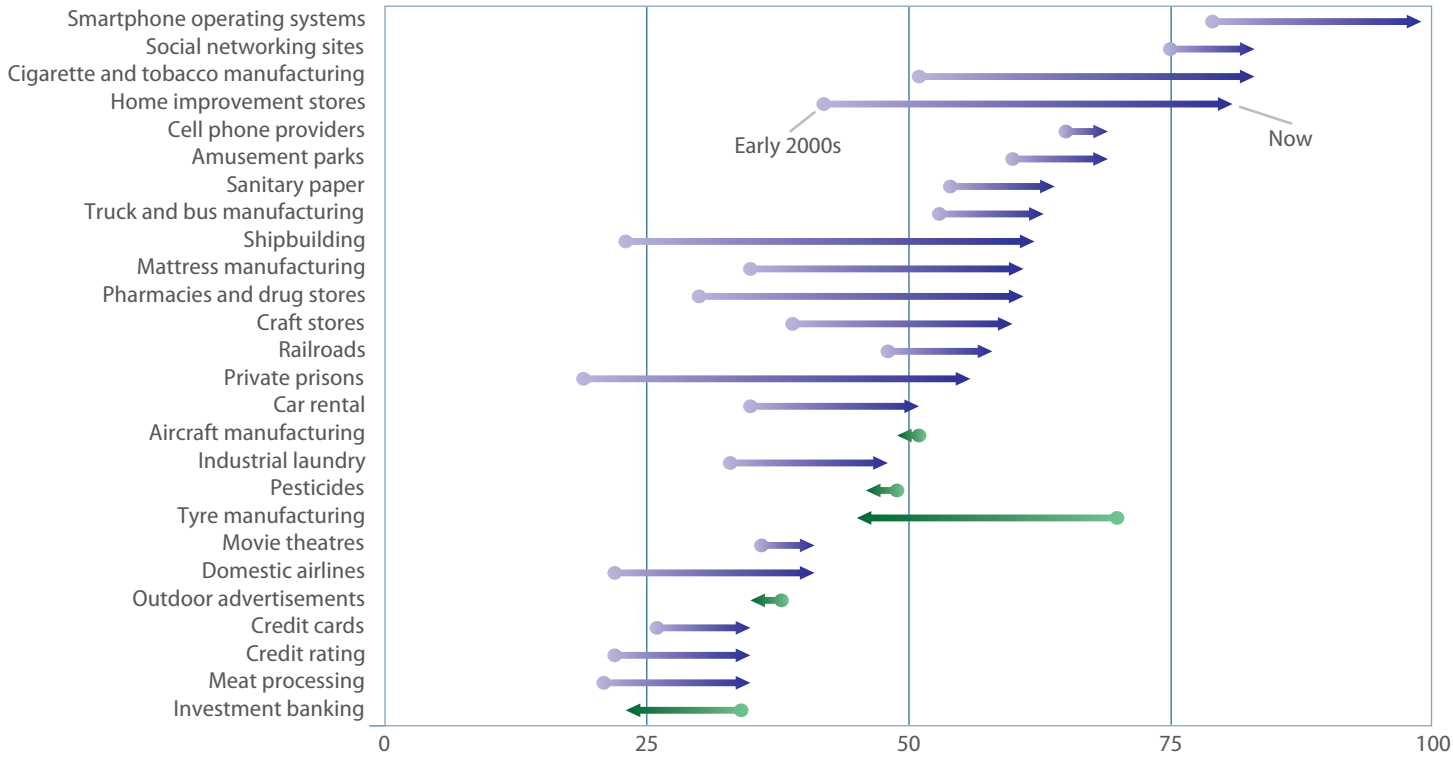
The unique characteristics of data has allowed the rise of data monopolies, where each large data firm has carved out a niche with highly differentiated and specialized data products. Absent national and global regulations, large data firms increasingly dictate the terms and conditions of data availability and use. Absent market prices of data, large data firms are capturing all surplus values created by data, amassing unprecedented wealth and exacerbating income and wealth inequality. Public policies at national and international levels will need to address the competitive market structure, in the data economy and the lack thereof, taking into account the unique characteristics of data as a factor of production and the challenges in assessing the fair market value of data.

The seminal work of Romer (1990) explains how the non-rivalrous and partially excludable characteristics of data can make research and development investments highly profitable. Non-rivalry of data implies some positive externalities that can benefit entities other than the original data collectors. Amazon, Facebook or Google collects data on consumer choices that may not be valuable to them but extremely valuable to sellers and advertisers. Moreover, the ability to exclude some, but not all, gives the data collectors kind of of monopoly power to profit from data collection. Facebook or Google would lose monopoly powers if the data they collect were available to all interested parties at the same time. More importantly, their monopoly power increases if more people use their services. These characteristics offer data firms increasing returns to scale—doubling up inputs resulting in more than doubling of outputs—to their investments in data. The increasing return to scale allows data firms to create larger values relative to firms that receive constant returns to scale. It is often the source of rapid growth and the rise of a monopoly. Many large data firms enjoy increasing returns to scale, which makes a strong case for regulating and taxing data firms differently than their counterparts that receive constant return to scale.

No free lunch

It is hard to value data in monetary terms. The *World Economic and Social Survey 2018* highlighted the difficulties in valuing data, taking into account privacy, equity, and distributional concerns. The data valuation problem fundamentally affects competition in the data economy. The largest data firms like Facebook and Google adopted business models to offer their services for free but earn advertisement revenues from third parties. Typically, data firms and data users engage in barter trade, in which personal data are exchanged for "free" digital services. This created a natural barrier for competitors to enter the market, allowing incumbent firms to extract monopoly rents from third party users of data. Facing little

Figure 3
Market share of the two largest firms in the United States by sector, early 2000s and now



Source: Leonhardt, David. "The Monopolization of America", The New York Times 25 Nov. 2018. Available at https://www.nytimes.com/2018/11/25/opinion/monopolies-in-the-us.html

or no competition, they are able to expand their customer base and increase market power. Individuals and businesses are increasingly beholden to few large firms to meet their data demand, while concentrated market power lends a few firms the ability to influence social and consumer behaviour. Absent fair market price, both data firms and data users underprice the consequences of privacy and security breaches. There is no free lunch after all.

Asymmetries in the data economy

The data economy is highly uneven, manifesting asymmetric bargaining powers, among firms, between firms and consumers and between firms and the State. Asymmetries in power certainly are not unique to data economy, but the unique properties of data make it difficult to reduce these asymmetries. There is a high level of market concentration in the data market, which is likely to impede market competition, innovation and productivity growth.

Dominant market positions of the leading firms—initially attained with breakthrough innovation—could make it difficult for smaller firms to effectively compete, allowing dominant firms to maintain or even expand their market shares without necessarily being more innovative. The true concern therefore is not market dominance per se, but its possible unjustified permanence. From a political economy perspective, a higher concentration of market power also increases the likelihood of "regulatory capture"—a situation where policymakers or enforcement agencies are in a constant state of being influenced by powerful firms (Hempling, 2014).

Data on consumer choices also gives firms undue pricing advantages, enabling them to charge each customer different prices. In pure economic terms, data firms can take all consumer surplus, making consumers worse off and potentially reducing aggregate demand in the economy for other goods and services. Price discrimination, price gouging and monopoly prices—increasingly possible in the data market—will undermine free market practices and consumer welfare.

III. How do we make the data economy work for all?

Data is changing the competitive landscape, concentrating market power and the economic gains in fewer, larger companies. For individuals, data is raising concerns about trust, privacy and security, as well as the equitable distribution of gains from data. For governments, data is creating new social and political dynamics and changing the relationship with civil society. Governments are tasked with finding a balance between providing incentives to innovation, supporting a healthy competition in the data economy, and defending the rights and interests of individuals and consumers.

Given the global nature of data economy and the cross-border capabilities that exist to collect and trade data products and services, regional and multilateral institutions also have a central role to play in understanding and addressing many of these questions. These institutions must support rules-based regimes for data that promote national and collective interests, for example, by establishing the appropriate boundaries of acceptable use of private data, or data for medical research.

Levelling the playing field

Appropriately valuing data transactions, enhancing privacy and security standards and finding innovative ways of taxing values generated can make the data market more competitive and level the playing field. Authorities need to refine their analytical tools so that they can properly define markets, measure market power—especially in the case of multi-sided markets—and define conduct to be considered anticompetitive. There is also a clear need to strike the appropriate balance between supporting innovation in the data economy and protecting individual rights and customer interests.

The number and size of data breaches worldwide has been increasing at alarming rates (see Figure 4), where costs of such breaches are disproportionately borne by data users. The largest data breaches have each exposed over 1 billion personal records. In 2017, for example, a spam operator exposed 1.4 billion emails and other personal information. There have also been cases of sharing of personal data between companies without users being fully informed. The case of Facebook-Cambridge Analytica, for example, demonstrates the need for clear regulations and standards of acceptable data sharing by companies. The data users are often unable to enforce their legal rights because one-sided agreements typically favor the data firms. Data users typically lack time and expertise to understand the fine prints—and their consequences—in those agreements (Tirole, 2017).

Generally, developing countries face similar but deeper challenges in dealing with the data economy. Key policy priorities include, for example, the need to protect the data of individuals, foster open-data policies, create standards for inter-operability of data functions and develop skills relevant for the data economy. In addition, governments can address existing and emerging barriers to the growth of their domestic data markets and help firms develop strategies to extract and exploit their data. Governments can also address increased market concentration and dominance. Enhancing consumer protection and managing cross-border flow of data will be critical for developing countries. As economic activities shift to the digital space, protection and privacy standards of personal data will largely determine the cost of data and comparative advantages of developing countries in the product market. If a data firm collects, compiles and analyzes consumer preferences and purchases and makes that data available to competing producers in another country, the producers in the host country will likely lose its market share to the foreign competitor.

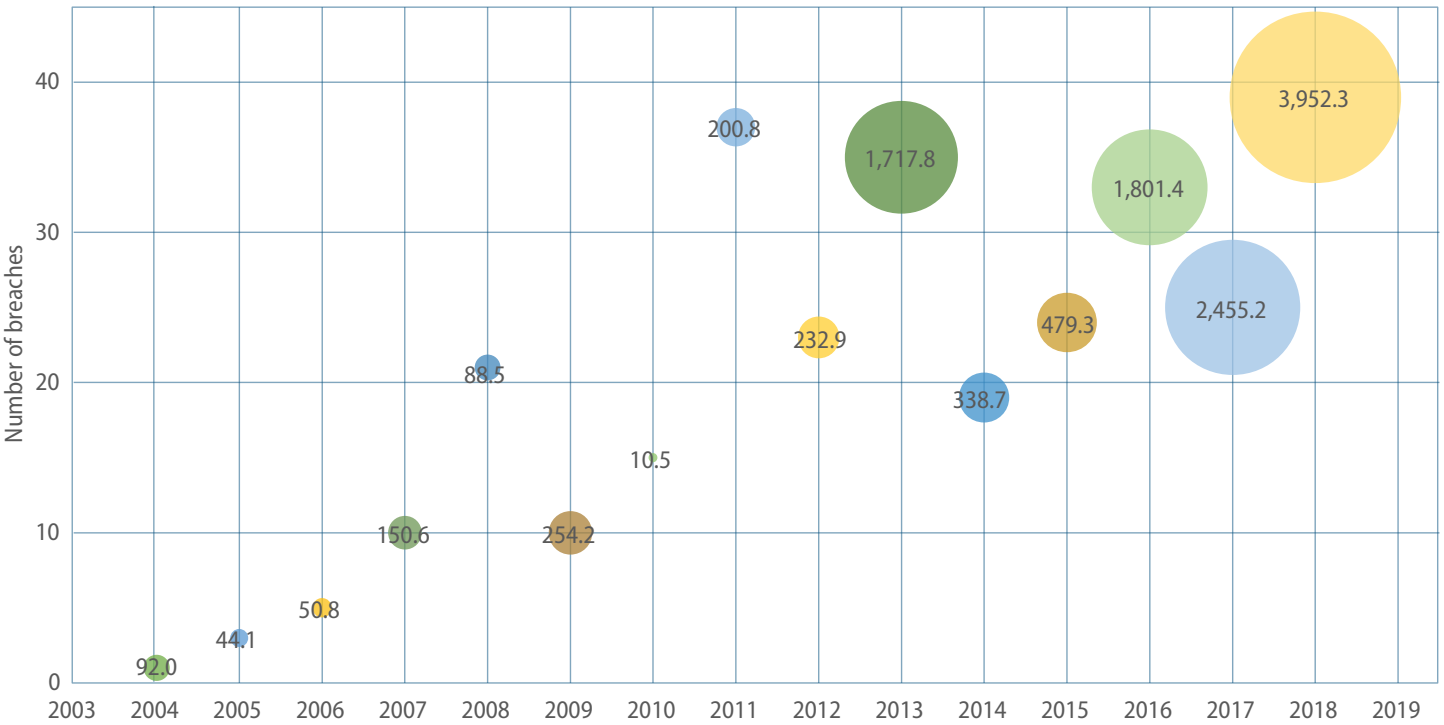
And we need to do more...

Economics tells us monopolies are neither fair nor efficient. Data monopolies are unlikely to be different. New and effective policies protecting data ownership, portability, privacy and security issues will go a long way to fairly distribute surpluses captured by data monopolies. These measures will also facilitate new entrants and make the data market more competitive. In addition, enforcement of privacy and protection standards will increase the cost of data as a factor input and limit the future growth of data monopolies. Furthermore, effective taxation of data trade will allow redistribution of the gains generated in the data economy, which will be critical for reducing income and wealth inequality.

Governments must specify rules for data privacy, data ownership, and security. The General Data Protection Regulation (GDPR)

Figure 4

Number of data breaches and number of records exposed (*indicated within the bubbles in millions*)**, by year**



Source: McCandless, David, Tom Evans, Paul Barton, and Stephanie Tomasevic. 2018. "World's Biggest Data Breaches and Hacks." Information Is Beautiful. December 5, 2018. Available at https://informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/.

in the European Union is the first step in the right direction. The regulation specifies the rights of individuals and the obligations placed on firms covered by the regulation. These include allowing people to have easier access to their data held by companies; a new fines regime with severe penalties; and requiring consent by individuals before organizations can collect their data. The European model of data rights of individuals is an important contribution to the conversation of how the multilateral system can create global standards for data privacy and rights. Most importantly, the GDPR has normalized many of the rules and standards for the data market across the members of the EU, opening the door for an EU-wide data market.

Governments must also consider the interlinkages of different regulatory measures and the possibility of unintended consequences. Changes to privacy laws, for example, are necessary for protecting consumers, but such changes may disadvantage smaller firms. For example, the high costs of compliance with the GDPR privacy laws may disproportionately burden smaller firms with fewer resources, potentially hurting competition. The right to data portability, among other things, imposes interoperability requirements that intend to allow users to easily switch between platforms. While it is a key element in protecting the rights of individuals in the data economy, it could be very costly for smaller firms to comply and undermine consumer welfare (Swire and Lagos, 2013).

The future of humanity hangs on a delicate balance. How we shape the data economy will shape the future of the global economy, as economic activities are increasingly digitalized. The global community must not shy away from asking difficult questions: i) How do we price data fairly, taking into account the cost of privacy and security breaches, to equitably share the gains of the data economy? ii) How the price of data affect countries' comparative advantages in data and product markets? iii) How does the rise of monopolies in the data economy affect income and wealth inequality? iv) How does the rise of monopolies affect investment in other sectors of the real economy? v) Should firms in the data economy be taxed differently and how? Answers to such questions, though not easy to derive, will allow us to ensure that the data economy works for all, not just for few innovators and investors that capture all gains and delivers sustainable development outcomes.

References:

Hempling, Scott (2014). "Regulatory capture": sources and solutions. Emory Corporate Governance and Accountability Review, vol. 1, No. 1. Available at http://law.emory.edu/ecgar/content/volume-1/issue-1/essays/regulatory-capture.html.

ICD, and Open Evidence. (2017). "European Data Market - SMART 2013/0063 - Final Report." http://datalandscape.eu/study-reports/european-data-market-study-final-report.

Romer, Paul M. (1990). Endogenous technological change. *Journal of political Economy*, *98* (5, Part 2), S71-S102.

Swire, P., & Lagos, Y. (2013). Why the right to data portability likely reduces consumer welfare: antitrust and privacy critique. *Md. L. Rev.*, *72*, 335.

Tirole, Jean (2017). *Economics for the common good*. Princeton University Press.