# Global Distribution of Economic Values of Data and Data Flows

**Wendy C. Y. Li**
Moon Economics Institute

Working Paper

2023

**MOON ECONOMICS INSTITUTE**
https://moonecon.org/

# Global Distribution of Economic Values of Data and Data Flows

Wendy C.Y. Li[1][2]

Moon Economics Institute

Revision Date: January 5, 2023

## Abstract

Data is a key asset for firms to innovate, produce, and compete, and the scale of data flows affects the international division of labor. The distribution of the values of data and cross-border data flows is an important topic for global data governance, trade, investment, development, and tax policies. Policymakers around the world, however, have been negotiating agreements on data flows while estimation of the economic values pertained in the cross-border data flows and how they vary with time is unavailable. Using the internet traffic data from ITU and TeleGeography, I present the first estimate of the economic values of international data flows at the global, regional, and national levels. I find that: First, because the value of data depreciates, the value of data flows grows at a slower rate than that of data flows. Second, ten leading countries “control” over two thirds of international data flows. Though adopting a strict data policy, the value of China’s international data flows outpaced the US counterpart in 2014, and the gap has been widening. Third, the value of China’s international data flows is the highest in the world: It was worth US $201 billion with an annual growth rate of 22.5% in 2017, when the U.S. counterpart was valued at US $82 billion and continued to decline. Fourth, the value of cross-region data flows distributes very unevenly across the globe. The U.S./Canada-Latin America has the highest average value US $75.2 billion, in 2021, with an average annual growth rate of 12.9% during 2018-2021. Lastly, Taiwan experienced the world’s highest growth in the value of international data flows during 2012-2017, when the average annual growth rate was 154%, indicating that smaller countries can increase the growth rates of international data flows, their economic values, and related investments by adopting an open data policy.

JEL Codes: O3



[2] E-mail: wendyli@moonecon.org

## 1. Introduction

Data are crucial for firms to innovate, produce and compete in the digital era (Li and Chi, 2021). Firms using data to organize production enjoy higher productivity and market valuation. Among them, Big Tech possesses a tremendous value of data, and data can increase Big Tech's profitability significantly (Li, 2022). The core of Big Tech relies on artificial intelligence (AI) algorithms and data. As AI algorithms become more affordable and adaptable, data becomes the key that determines the accuracy of the algorithms. In other words, AI is a tool of finding patterns in data, and data determines the performance of AI. As shown later, although adopting a strict data policy, China outpaced the U.S. in terms of international data flows in 2014, and the data gap has been widening since then. China's international data flow has the world's highest value with a double-digit growth during 2012 to 2017; however, the U.S.'s value has less than one half of China's and has been declining. According to Stanford University's 2021 AI Index Report, China's share of journal citations in AI has surpassed that of the U.S. and was ranked the top of the world in 2020 (Stanford HAI, 2021).

Data is not only a strategic asset, but its cross-border flows are important for firms' daily operations and for the development of developing countries. Big Tech and online platform companies, typically multinational firms, can concentrate their operations in one or a handful of countries while providing digital services worldwide, and their daily operations inevitably involve tremendous cross-border data flows. Their data-driven business models allow them to serve the global market without facing traditional physical constraints, to easily scale up their businesses, and to enter adjacent industries and markets, as long as they can access the data needed. Moreover, firms in developing countries can also operate domestically and access to overseas markets through online platforms (Coyle and Li, 2021). All of these activities rely on cross-border data flows.

Because data is important for countries to build their AI supply chain, many countries have also started building data fortress to ensure their firms' competitiveness in the market. A main reason that EU is establishing a series of laws and regulations on data and data flows is to help European firms compete with Big Tech (Cory et al., 2022). Another extreme example is China, where the government sees data as a critical element in developing its digital economy and is aggressively formulating a strict data governance policy, building platforms for data sharing, and constructing data exchange markets (Coyle and Li, 2021). As countries are engaging in talks and agreements on how to govern data flows, however, there is a void of knowledge about the values of data and cross-border data flows and how they may change with time. Not having such critical information but knowing that data is a strategic asset, policymakers may simply control or block the data flows, a situation that can result in rising trade costs, restricting data sharing, limiting the potential of data, and hindering productivity growth. Examples of countries that have taken or are considering this approach include Bangladesh, Indonesia, Pakistan, and Vietnam (Cory et al., 2022). The lack of knowledge about the value of data associated with cross-border data flows is also an immediate problem for governments to identify effective measures to address new issues introduced by firms that are data-intensive and have data-driven business models.

Measuring economic values of cross-border data flows is difficult (UNCTAD, 2021). Through our series of studies on the value of data and data flows, we develop a methodology to measure the economic values of data and cross-border data flows around the world (Li et al., 2019, Li and Chi, 2021, Li, 2022). To our knowledge, this is the only method that can measure the value of cross-border data flows. Using this method, Li (2022) finds that the economic value of cross-region data flows in 2020 is up to several hundred billion dollars, and the estimated global value of data is several trillion dollars (Li, 2022). The economic values of data flows, however, are distributed very unevenly due to the inhomogeneity of cross-region data flows. As the world economy becomes more digitized and the rapid adoption of AI technologies in production and innovation across industries, data and data flows have increased dramatically within and across regions and countries. A rich understanding of the location of value creation and distribution from the use of data can also benefit the analysis of trade and the development of future data trade policies (Coyle and Li, 2022).

In this paper, I use the methodology as described in Section 2 and the internet traffic data from the International Telecommunication Union (ITU) and TeleGeography (see Section 3) to investigate the answers to the following questions: How are international data flows distributed around the world and how do they grow with time? What are the economic values associated with each country's international data flows? How are cross-region data flows distributed and what economic values are associated with those cross-region data flows? What policy implications can we derive from the research results?

This paper presents the first estimate of the economic values of international data flows at the global, regional, and national levels. I find that: First, because the value of data depreciates, the value of data flows grows at a slower rate than that of data flows. Second, the four leading countries, including China, the U.S., the U.K., and Taiwan, "control" over 50% of international data flows. Additionally, top ten countries "control" over two thirds of international data flows. Although China adopts a strict data policy, its value of international data flows outpaced the US counterpart in 2014, and the gap has been widening. Third, the value of China's international data flows is the highest in the world: It was worth US $201 billion with an annual growth rate of 22.5% in 2017, when the U.S. counterpart was valued at US $82 billion and with a negative growth. Fourth, the value of cross-region data flows distributes very unevenly across the globe. The

U.S./Canada-Latin America has the highest average value US$ 75.17 billion, in 2021, with an average annual growth rate, 12.9% during 2018-2021. Last but not least, Taiwan experienced the world's highest growth in the value of international data flows during 2012-2017, when the average annual growth rate was 154%. This finding indicates that an open data policy can help smaller countries not only raise the growth rates of international data flows but also their economic values and related investments.

The rest of paper proceeds as follows. Section 2 introduces the empirical methodology. Section 3 describes the data sets used. Section 4 shows the empirical analysis results. Section 5 concludes.

## 2. Methodology

### Measuring Economic Value of Data by Organizational Capital

The methodology of estimating the value of data and data flows builds on the principles developed and the empirical relations discovered in previous studies by the author (Li and Chi, 2021; Li, 2022). Li and Chi (2021) provides a theoretical foundation of why and how a firm's organizational capital represents its economic value of data. As defined by Prescott and Visscher (1980), organizational capital is the accumulated information of the firm. The value of data lies in the firm's utilization of information derived from data to drive its productivity and competitive advantage. In the measurement of organizational capital, this study follows the practice well established by previous studies [e.g., Lev and Radhakrishnan, 2005] and uses the selling, general, and administrative (SG&A) expense minus R&D expense as the proxy for the firm's investments in organizational capital.

Li and Chi (2021) also demonstrates the use of a proven method developed by Li and Hall (2020) in estimating the depreciation rate of organizational capital from firm-level investment data. As explained by Li and Chi (2021), the depreciation rate is the key to calculating the stock of organization capital, and other studies in the past have only assumed a fixed value for this quantity rather than truly estimated it. The depreciation rate of organizational capital should vary across firms and across industries. It can also change with time: When an online platform enters the industry, the existing incumbents with a lower degree of digital transformation will suffer from a higher depreciation rate of organizational capital, a phenomenon called "online platforms' creative disruption in organizational capital" discovered by Li and Chi (2021).

## Law of the Value of Data

Also known as "Li's law of the value of data," the empirical relationship between the global data flows with Big Tech's organizational capital can provide a useful means of estimating the value of data flows (Li and Chi, 2021). Specifically, the Big Tech in the analysis includes Alphabet, Amazon, Microsoft, Apple, Meta (formerly known as Facebook), Alibaba, and Tencent, and the depreciation rate of the organizational capital for each company is estimated by using the Li and Hall (2020) method to properly calculate the combined stock of organization capital. Figure 1 shows the clear relationship in Li's law that the Big Tech's value of data doubles when the global data volume increases by five folds.

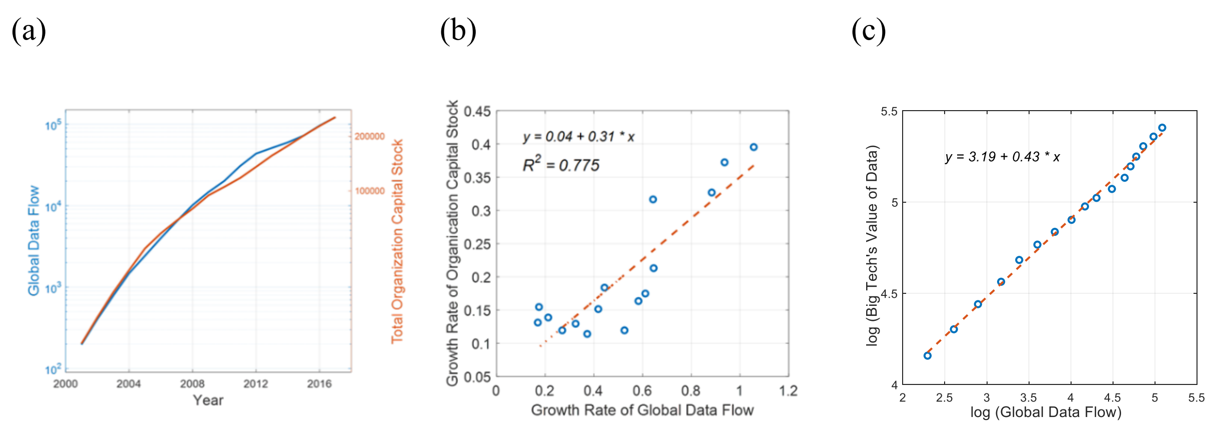


**Figure 1.** Big Tech's combined organizational capital (in million US dollars) vs. global data flow (in PB/month). Taken from Li and Chi (2021).

Li's law provides very useful mapping connecting organization capital (the value of data) to the Internet traffic (data flow). It also has a profound implication that, on the macroscopic scale, the economic value of data in the business sector can be estimated by the amount of data flows, which is essential to the methodology of estimating the value of data flows.

## Global Value of Data

In order to estimate the economic value of each individual data flow, Li (2022) develops the method of using the global value of data as an intermediate step. This approach avoids the intrinsic difficulties in directly estimating the value of a data flow, such as the need of dissecting the types of data in the data flow and the value in each category. When the global value of data is a known quantity, estimating the value of every individual data flow can be straightforward through finding the share of the data flow in the global Internet traffic.

As also explained by Li (2022), it is far more feasible to estimate the global value of data through the use of the organization capital investments of major Big Tech companies and the market capitalization data. While it is conceptually feasible to calculate the global value of data by summing the organization capital investments of all companies in the world, the task will be a formidable challenge in practice, and estimating the organizational capital investments by private firms that do not publish SG&A data requires large-scale surveys in all the countries involved. Due to these difficulties, I use the combined organizational capital and the share of market capitalization of seven largest Big Tech companies to estimate the global value of data, because it has been well demonstrated that organization capital of a firm is positively correlated with its market capitalization (Eisfeldt and Papanikolaou, 2013).

## Economic Values of Data Flows

When the global value of data is a known quantity, Li's law provides a natural basis for estimating the value of a large-scale data flow. Because the economic value of data is correlated with Internet traffic, the value associated with each subset of Internet traffic, including data flows within borders and cross-border data flows, can be estimated by the share of the subset in the global Internet traffic.

While measuring the value of data is a relatively young and growing research topic where several approaches have been proposed (see the recent review by Coyle and Manley (2022)), it is worth noting that the methodology developed by the author as described above has several critical advantages that cannot be matched by any other proposed method:

1. The only existing method that can estimate the depreciation rate without using an ad hoc approach.
2. The only existing method that can connect an economic quantity (organizational capital) to the volume of data (with "The Li's law").
3. The only existing method that can estimate the value of data flows.

## 3. Data

As mentioned earlier, I use the selling, general, and administrative (SG&A) expense minus R&D expenses as the proxy for a firm's investment in organizational capital (OC), and for each firm I estimate the OC depreciation rate to calculate its stock of organization capital. Firms report their SG&A expense in their annual income statements, and this expense includes most of the expenditures that generate organizational capital, such as employee training costs, brand enhancement activities, consulting fees, and the installation and management costs of supply chains. Specifically, I used the data for top Big Tech companies, including Microsoft, Amazon, Apple, Alphabet, Meta, Alibaba, and Tencent, up to 2017. The calculation of the OC depreciation rate also requires sales data reported by these firms and the GDP deflation factor.

The estimation of the global value of data also requires data for market capitalization of companies. While the top seven Big Tech companies have public information of their market values, the total market capitalization of all publicly traded companies in the world through 2020 are provided by the World Bank.

This study uses several data sources for the Internet traffic and bandwidths. The Cisco System provides the data for the global Internet traffic from 2002 to 2017. The published International Telecommunication Union (ITU) data includes international Internet bandwidths of 236 reported economies between 2007 and 2021, and the data for global and regional international data flows during 2017 and 2021 come from TeleGeography.

## 4. Empirical Results

This section presents the economic value of international data flows using the methodology and data described above. The main results include those at the global and regional levels during 2017-2021 and the national level during 2012-2017 for economies with highest international data flows.

### 4.1 Economic Value of Global International Data Flow

As described in the methodology, the economic value of every large-scale data flow can be derived from the global value of data and the share of the data flow of interest. Li (2022) estimates the global value of data for the first time and found the value to be approximately five trillion dollars during 2014-2017. Li (2022) also shows that the growth in the global value of data can be affected by the degree of diversity in accessing the data and does not necessarily follow the growth in Internet traffic. According to UNCTAD, the global data flow has been undergoing explosive growth after 2018, particularly after the start of Covid-19 pandemic in 2020, and the data growth has further increased due to the accelerated digitization of the world's economy (UNCTAD, 2021). Nonetheless, even if the global internet traffic continues to grow, the global value of data may saturate, or possibly decline, if Big Tech continues to gain a higher share of global data at a rapid pace. That is, the growth of data and data flows may not result in direct growth of the value of data and data flows, it also depends on the degree of concentration of data access.

The direct comparison between the global value of data and the Internet traffic volume provides a mapping between the value of data and the data flow. The mapping can then provide a way of estimating the values of subsets of the global Internet traffic, such as international data flows. The mapping parameter is assessed every year. It is found that, for years of 2012-2017, the mapped values of a fixed amount of data amount gradually depreciated roughly following an exponential fashion with a depreciation rate of 23%.

With additional statistics from TeleGeography, this study extends Li (2022) and provides a more precise estimate of the global and regional international data flows in 2012-2017. Based on the statistics by TeleGeography, the International Telecommunication Union (ITU), and Cisco Systems, the international data flows account for approximately one tenth of the global Internet traffic during these years. Figure 2 shows the economic value of the global international data flows estimated through mapping. The value grew from US $380 billion in 2017 (which is 8% of the world's value of data, see Li (2022)) to US $825 billion in 2021. The substantial rise in the global value of international data flows in 2020 can be attributed to the dramatical increase in the Internet activity due to the COVID-19 pandemic.

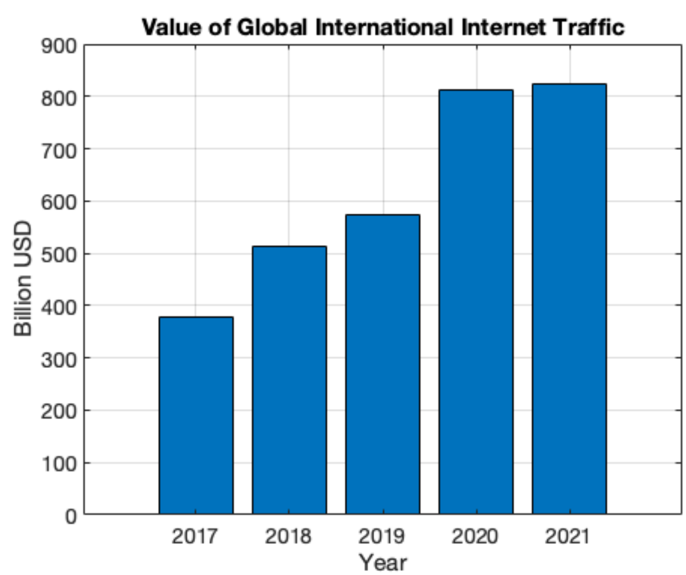


**Figure 2.** The global value of international data flows during 2017-2021.

### 4.2 Economic Values of Regional International Data Flows and Cross-Region Data Flows

TeleGeography provides the statistics of the international Internet traffic for seven regions of the world – Africa, Asia, Europe, Latin America, Middle East, Oceania, and U.S. and Canada – as well as the Internet traffic between these regions. Using the same methodology in estimating the value of each data flow as in Figure 2, Figure 3 shows the value of the annual international data flow for seven regions of the world. It should be noted that the associated internet data flows are international: they include cross-region data flows and cross-border data flows within the same region, but exclude domestic data flows in each nation. Europe has the highest value of international data flows exceeding US $500 billion since 2021, and Asia's value of international data surpassed that of the U.S. and Canada in 2019.

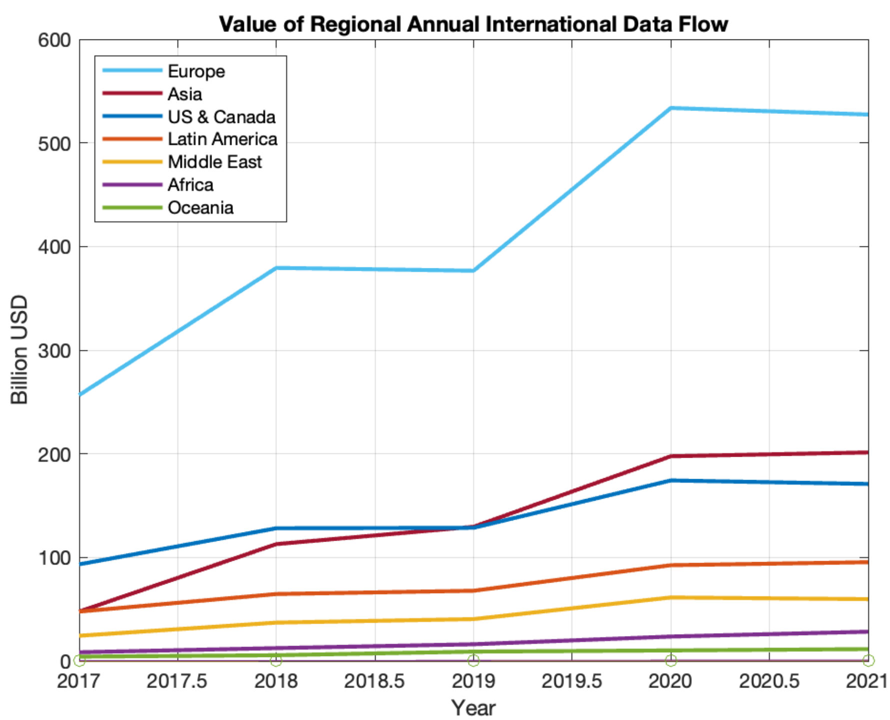


**Figure 3.** Economic value of regional international data flows

In the last two decades, cross-region data flows have experienced not only a growing trend across the board but also a substantial shift in regional activity (see TeleGeography's Global Internet Maps for 2001, 2009, 2018, 2022). In the 2000s and the early 2010s, the Internet traffic between Europe and the U.S. & Canada was the largest cross-region data flow in the world. By the end of the 2010s, the top cross-region Internet traffic was replaced by the data flow between Latin America and the U.S. & Canada. In 2021, the Europe – US & Canada route dropped further to the fifth largest cross-region data flow in the world, and the value of this cross-region data flow had stopped growing. All the top five cross-region data flows are valued at tens of billion dollars a year, and the values for two cross-region data flows from Europe – to Middle East and to Asia – are particularly growing strong (see Table 1).

Table 1. Economic values of top five cross-region data flows during 2018-2021

| Cross-region | Value of International Data Flow in 2018 (Billion USD) | Value of International Data Flow in 2021 (Billion USD) | Average Annual Growth Rate during 2018 – 2021 |
|---|---|---|---|
| Latin America – US & Canada | 52.30 | 75.17 | 12.9% |
| Asia – US & Canada | 30.53 | 43.91 | 12.9% |
| Europe – US & Canada | 30.43 | 29.71 | –0.8% |
| Europe – Middle East | 29.46 | 50.11 | 19.4% |
| Asia – Europe | 19.71 | 34.36 | 20.4% |

Figure 4 depicts the distribution of the value of cross-region data flows in 2021. The combined value of cross-region data flows amounts to US $270 billion, which accounts for approximately one third of the global value of international data flows.

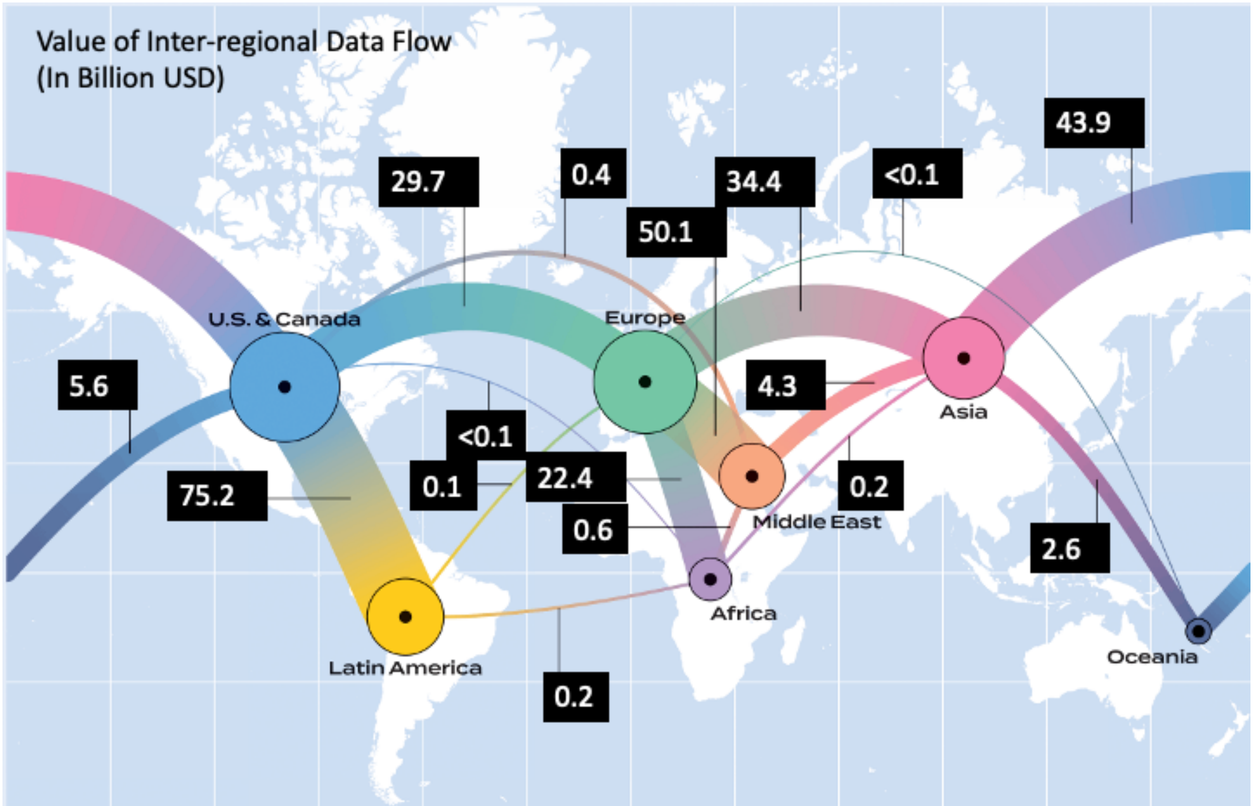


**Figure 4.** Values of cross-region data flows in 2021. The map is adapted from TeleGeography's 2022 Global Internet Map. The economic values are provided by this study.

### 4.3 Economic Values of International Data Flows for Individual Countries

While the global and regional results can provide top-level views of the value of data flows, the analysis at the national level is particularly useful because of its immediate implications on data governance, trade, and tax policies. Using the ITU database for international bandwidths, this study includes an initial analysis on the value of international data flows at the national level.

The ITU dataset includes the international bandwidths (in Mbits/second) for 236 countries in 2007 – 2021. It is necessary to note that there are gaps in this dataset where the international bandwidth for the country (or economy) was not reported to or estimated by the ITU. For years the ITU estimated the international bandwidths for many major economies, including France, Germany, Japan, the United Kingdom, and the United States, but these estimates are unavailable since 2018[3]. Because of these reasons, the analysis here focuses on the ITU data during 2012-2017.

The methodology for estimating the value of data flow is the same. Figure 5 shows the economic values of 20 countries that had most values in the year of 2017. The Internet traffic is assumed to be 10% of the Internet bandwidth (for international data flows), which is the usage level in 2017 reported by TeleGeography. Because the ITU dataset uses a single value to represent downlink and uplink bandwidths, it is also assumed here that the Internet traffic is equal in both directions. Different from the results in global and regional statistics, the value of international data flows for a country did not always increase with time. The 2017 values and average growth rates are listed in Table 2.

Table 2. Values of international data flows and average growth rates for 20 countries leading in the international bandwidth

| Rank in 2017 | Country / Economy | Value of International Data Flow in 2017 (Billion USD) | Annual Growth Rate of Value of International Data Flow (average value during 2012-2017) |
|---|---|---:|---:|
| 1 | China + Hong Kong | 201 | 22.5% |
| 2 | United States | 82 | –8.1% |
| 3 | United Kingdom | 70.8 | –11.3% |
| 4 | Taiwan | 69.4 | 154.2% |
| 5 | Singapore | 28.4 | 16.1% |
| 6 | India | 26.2 | 52.5% |
| 7 | Russia | 20.2 | 0.9% |
| 8 | Vietnam | 17.4 | 78.1% |
| 9 | Saudi Arabia | 13.6 | 110.2% |
| 10 | Luxembourg | 12.8 | –7.6% |
| 11 | Colombia | 12.8 | 41.0% |
| 12 | Turkey | 11.8 | 6.5% |
| 13 | Thailand | 11.6 | 28.9% |

[3] Most African and Asian countries including China still report international bandwidths to ITU after 2017.

| | | | |
|---|---|---|---|
| 14 | Brazil | 11 | –15.1% |
| 15 | Germany | 10 | –14.6% |
| 16 | Kenya | 9.2 | 35.7% |
| 17 | South Korea | 9 | –3.2% |
| 18 | Mexico | 8 | 1.8% |
| 19 | Japan | 7.8 | –7.9% |
| 20 | France | 7.6 | –16.6% |

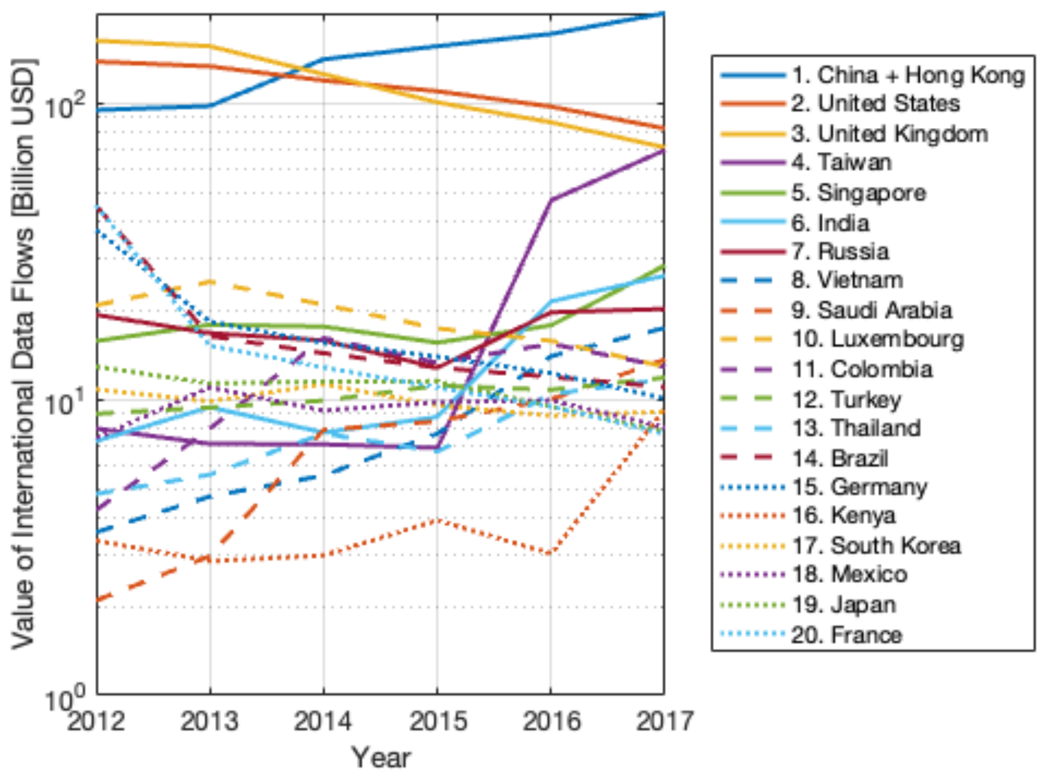


**Figure 5.** Economic value of international data flow for the top 20 countries

To visualize the large range of the economic values of international data flows among countries and how it changes with time, the shares of the 10 leading countries for the years of 2007, 2012, and 2017 are plotted in Figure 6, where the values for China and Hong Kong are combined to aid the comparison. While the United States and the United Kingdom held the top spot in 2007 and 2012, respectively, China (including Hong Kong) rapidly moved up from the fifth place to the first in ten years. It is also revealing that some of the small economies such as Taiwan, Singapore, Vietnam, and Luxembourg have been actively expanding the capacity for international data flows recently and climbed to the top ten list in 2017. Evidently, the majority of the economic values of international data flows are tied to a small number of countries. To our surprise, as the largest economic power in EU, Germany does not enter the top 10 countries controlling international data flows. This may reflect the situation where Germany is lagging behind other countries in digitalization. Table 3 lists the proportions of the global values associated with the leading countries.

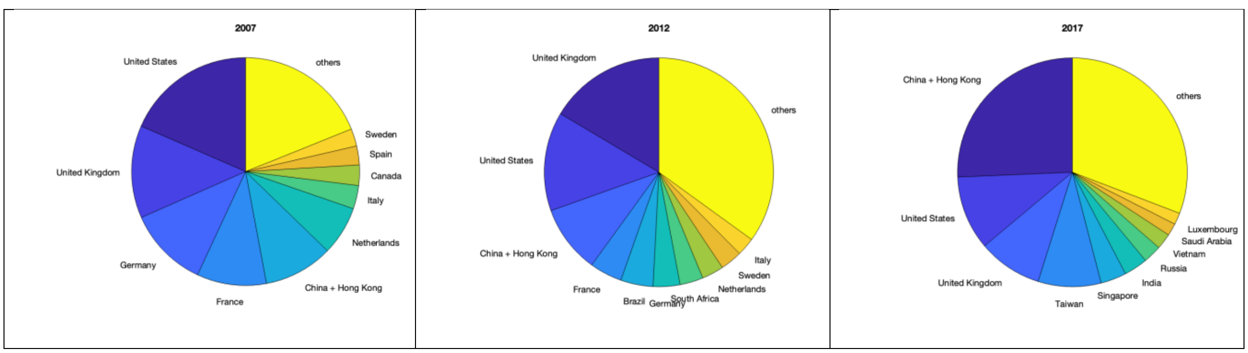


**Figure 6**. Shares of 10 leading countries in the value of international data flows in 2007, 2012, and 2017

**Table 3.** Shares of leading countries in the global value of international data flows

| | Share in global value of international data flows | Countries |
|---|---|---|
| Top 5 | > 50% | China (including Hong Kong), United States, United Kingdom, Taiwan, and Singapore |
| Top 10 | > two thirds (2/3) | The above and India, Russia, Vietnam, Saudi Arabia, and Luxembourg |
| Top 20 | > 80% | The above and Colombia, Turkey, Thailand, Brazil, Germany, Kenya, South Korea, Mexico, Japan, and France |

These results on the national level present many interesting facts and implications. First, China (including Hong Kong) has the largest international data flows since 2014. Given that China has a very strict data localization policy, it is reasonably to conclude that there are a lot of data inflows around the world to China. Additional to the large population incurred data domestically, it is not surprised that China has become an AI superpower today and TikTok has far more superior AI algorithm than other competitors (Coyle and Li, 2021). Second, top four countries in the rank, China, the U.S., U.K., and Taiwan, have greater than 50% of international data flows and top twenty countries have over 80% of international data flows.

Compared to China and the United States, the UK and Taiwan are relatively small countries with smaller population sizes to generate large data flows. Taiwan experienced the world's highest growth in the value of international data flows during 2012-2017, when the average annual growth rate was 154%. In fact, a "quantum leap" of Taiwan's international data flows began in around 2015, aided largely by Taiwan's open and lenient data policy. This policy motivated Google to build three data centers in Taiwan in 2013, 2019, and 2020 to serve their Asia markets, to provide AI training to local engineers for incubating local AI skilled labors, and to establish in Taiwan the company's largest Asian data center and largest overseas R&D center outside the U.S. Taiwan shares its healthcare data with Google in return. Microsoft also announced in 2021 its plan to build a data center in Taiwan for operation starting in 2024.

This analysis at the national level shows that, by adopting an open data policy, smaller countries can increase the growth rates of international data flows, their economic values, and investments. An outstanding question is how large the share of the value of Taiwan's international data flows is created and owned by Big Tech (Coyle and Li, 2021). Nevertheless, incubating a data-sharing environment where data flow freely can encourage firms to invest and cultivate local AI skilled labors, strengthening the country's competitiveness in the AI supply chain. As mentioned earlier, the international data flows only account for less than 10% of the global data traffic, meaning that the majority of the Internet traffic is confined within national borders. Given the fact that data size, data access, market competition, and technological changes can all affect the growth of the value of data (Li and Chi, 2021; Coyle and Li, 2021), it may be beneficial for smaller countries like Vietnam to adopt a more open data policy to attract investments, incubate skilled labors, and enjoy free data flows for the development of local economies. In short, while data flows into or out of a country does not mean that the country can own all the value of the data flows, but any data localization policy that restricts data flows may harm economic growth.

## 5. Conclusion

Data is a key asset for firms to innovate, produce, and compete, and the scale of data flows affects the international division of labor. Understanding the economic values of data and cross-border data flows is important for global data governance, trade, investment, development, and tax policies. At the firm level, the information is important for firms to make investment decisions such as data infrastructure and business models. Moreover, it has been difficult for governments to curtail the negative impacts generated by data-intensive and data-driven businesses. For example, Big Tech's data-driven operation models allow these firms to pay zero or little tax under the current production-based tax system. A key problem is that governments do not know the magnitude of the economic values of data and cross-border data flows to gauge the impact on a firm's profitability (Li, 2022). Therefore, it is important to measure the economic value of data and data flows, especially the value of cross-border data flows, such that policymakers can evaluate how much economic values that may be interrupted, and the ensuing extra transaction costs that may incur for businesses. From the business perspective, the information can also aid in the evaluation of the impacts of data governance policies on firms' investments and transaction costs.

Li et al. (2019), Li and Chi (2021), Li (2022), and this paper have developed a methodology to measure firm-level and industry-level value of data and the economic values of global data flows and cross-border data flows. The methodology can also provide solutions to several other key problems related to the value of data, such as to measure firm-level and industry-level depreciation rates of the value of data, to measure the change in the depreciation rate and the impact on firm's value of data due to an event shock (Li and Chi, 2021), and to further measure the potential size of data markets (Coyle and Li, 2021). To my knowledge, no other existing methodology can estimate the depreciation rate of the value of data and measure the value of global data and cross-border data flows.[4]

In this paper, I find that the value of data flows grows at a slower rate than that of data flows because the value of data depreciates. This effect adds to another controlling factor in the value of data as found by Li (2022) that the global value of data can be affected by the degree of diversification in data access and use. Data gaps exist not only between firms but also between countries. I find that the leading ten countries "control" over two thirds of international data flows, and that the top four countries, namely China (including Hong Kong), the U.S., the U.K., and Taiwan, control over 50% of international data flows. Despite the strict data policy in China, the value of China's international data flows outpaced the US counterpart in 2014, and the gap has been widening. Although ITU data do not indicate the ownership of the data, the value of China's international data flows is the highest in the world. It was worth US $201billion with an annual growth rate of 22.5% in 2017, when the U.S. counterpart was valued at US $82 billion and declining. Because data is the core of the AI supply chain, it is not surprising to see that China's share of journal citations in AI has surpassed that of the U.S. in 2020. Given the fact that data can increase the profitability of firms, productivity, and innovations, controlling data flows can greatly affect the distribution of the benefits of data around the world. As the world increases digitalization, increasing data inequality can limit the potential of data and hinder innovation and growth.

---

[4] The Digitalization Task Team for updating the System of National Accounts (SNA) is in a process of using a sum of costs approach for valuing data in advance of the 2025 update of the SNA, despite the fact that several conceptual and practical issues remain unresolved (Coyle and Manley, 2022). However, it is well understood that the sum-of-cost approach planned is incapable of measuring the value of cross-border data flows or the depreciation rates of the value of data. These major drawbacks will leave policymakers in the dark about the magnitude of the value of cross-border data flows and the related impacts if they are interrupted. Note that the digital trade and services heavily rely on cross-border data flows. Hence, policies on cross-border data flows can affect regional development and investments significantly.

How the distribution of international data flows has evolved in the last decade not only reveals the areas of rising digital economy but also hints the impact of national or regional data policies. With the exception that the US & Canada / Latin America data flow has grown significantly, more new development in the data infrastructure has been taking place in Europe, and Asia and Middle East. My calculations indicate that these countries' international data flows and cross-region data flows can translate to hundreds of billions of U.S. dollars in value. In contrast, the US & Canada / Europe data flow that was once the highest cross-region Internet traffic has been on a declining trend. It is likely that Europe's General Data Protection Regulation (GDPR) contributed to the shift of regional data flows in recent years.

The case of Taiwan provides an interesting implication to the national data policy. Taiwan experienced the world's highest growth in the value of international data flows during 2012-2017, when the average annual growth rate was 154%. The extremely fast growth elevated the amount of Taiwan's international data flows to the fourth place in the world. This example demonstrates that smaller countries can adopt an open data policy and team up with Big Tech to increase the growth rates of international data flows, consequently their economic values, and investments. As indicated by Li and Chi (2021) and Coyle and Li (2021), data size, data access, market competition, and technological change can all affect the growth of the value of data. Because the international data flows only occupy less than 10% of the global data traffic, most of data flows are confined in the borders of largest countries, and smaller countries can possibly access large amounts of data through international data flows. Big Tech may earn a significant share of the increased value of data through the public-private partnership with the smaller country, but, it may still be beneficial for countries with smaller populations and data to adopt a more open data policy and attract investments, incubating local skilled labors and enjoying the access of data flows. Currently, many developing countries including Vietnam and Pakistan are planning to implement stricter data regulations such as data localization policies to restrict data flows across borders. This study can address the core issue relating to how much economic values of data flows are at stake so that policymakers can gauge the potential economic loss when adopting a strict data regulation.

Lastly, it is necessary to note that a more complete set of microdata for international data flows can allow further analysis and provide deeper insight of interest to data policies. The current ITU dataset lack the international data flows for all G7 countries (Canada, France, Germany, Italy, Japan, the U.K., and the U.S.) and most EU countries (such as Austria, Belgium, and the Netherlands) since 2018. In fact, most of these countries do not report international data flow data to ITU, and ITU has stopped estimated the amounts of international data flows for them. Ironically, a number of countries with stricter data flows regulations, like China, Cuba, and Vietnam, continue to report most recent data to ITU. As demonstrated by this study, all detailed information associated with international data flows can help estimate the associated economic value for each country, which can in turn help policymakers make evidence-based data governance policies. In addition to the lack of data for recent years, current publications on international data flows do not provide information on the flow directions between regions and countries, or on the ownership of data flows at the firm-level. In this study, the analysis is based on the published data for countries' international data flows and cross-region data flows without the information of the flow direction. The same approach can be applied to understanding the flow of the value of data and who benefits from the data flow once the microdata for the direction and ownership of Internet traffic become available.

## References


Cory, N., Dascoli, L., & Clay, I. 2022. The Cost of Data Localization Policies in Bangladesh, Hong Kong, Indonesia, Pakistan, and Vietnam, Information Technology & Innovation Foundation Report, December.

Coyle, C. & Li, W.C.Y. 2021. The Data Economy: Market Size and Global Trade, ESCoE Discussion Paper, September.

Coyle, C. & Manley, A. 2022. What is the Value of Data? A Review of Empirical Methods. Policy Brief. Bennett Institute for Public Policy, University of Cambridge, July.

Eisfeldt A., & Papanikolaou, D. 2013. Organizational capital and the Cross-Section of Expected Returns, *Journal of Finance*, 4, August, pp: 1365-1406.

Lev, B., & Radhakrishnan, S. 2005. The valuation of organization capital. In *Measuring Capital in a New Economy*, Corrado, C., Haltiwanger, J., Sichel, D. (eds). Chicago: National Bureau of Economic Research and University of Chicago Press: Chicago, IL; 73–99.

Li, W.C.Y. 2022. Economic Values of Data and Data Flows, and Global Minimum Tax, Moon Economics Institute Discussion Paper, No. 2, January.

Li, W.C.Y. & Chi, P.J. 2021. Online Platforms' Creative "Disruption" in Organizational Capital – The Accumulated Information of the Firm, Moon Economics Institute Discussion Paper, No. 1, August.

Li, W.C.Y. & Hall. B.H. 2020. Depreciation of Business R&D Capital, Review of Income and Wealth, https://doi.org/10.1111/roiw.12380.

Li, W.C.Y., Nirei, M., & Yamana, K. 2019. Value of Data: There's No Such Thing as a Free Lunch in the Digital Economy, RIETI discussion paper 19-E-022.

Prescott, E., & Visscher, M. 1980. Organization Capital, *Journal of Political Economy*, 88, pp:

446-461.

Stanford Human-Centered Artificial Intelligence, 2021. 2021 AI Index Report. https://aiindex.stanford.edu/ai-index-report-2021/

United Nations Conference on Trade and Development. 2021. Digital Economy Report 2021 – Cross-border Data Flows and Development: From Whom The Data Flow, September.

Yi, C.Y. 2021. Firms Change Strategies, Build Their Own Customer Database. *The World Journal*. December 20.