Centre for International Governance Innovation

Digital Policy Hub

Digital Policy Hub – Working Paper

Trade in Data: Measuring, Commercializing and Regulating

Fabrice Blais-Savoie

Winter term – 2025–2026 cohort

About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

Partners

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

About the Author

Fabrice Blais-Savoie is a Digital Policy Hub master's fellow studying international public policy at the Balsillie School of International Affairs. During his Digital Policy Hub fellowship, he will be examining trade in data and transnational regulatory mechanisms.

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
cigionline.org

Policy Memo

Trade in Data: Measuring, Commercializing and Regulating

Fabrice Blais-Savoie

CIGI | Digital Policy Hub

Bottom Line Up Front

In the twenty-first century, the role of data in the global economy has grown at an unprecedented pace. However, despite data fuelling many recent technological advances, its regulation has mostly mimicked traditional trade rules (which rely heavily on the accurate measurement of the physical flow of goods) while lacking effective tools to implement those rules. Nevertheless, some initiatives have emerged to collect data and identify proxies, informing more rigorous potential measurement avenues, such as collaboration with internet service providers or the expansion of existing trade statistical frameworks for the data era. Domestically, the case of Chinese data markets highlights a potential smaller-scale solution for measuring and regulating trade in data.

Key Points

- The creation of the modern trade environment is the result of more than a century of work. Trade in data only partially benefits from this work, and initiatives must continue to allow for this transition from a goods-based system to an intangible commercial environment.
- Current reporting and analysis of the trade in data relies heavily on proxy variables and antiquated data sources. Reliability and accuracy vary greatly for each country, and statistics are almost non-existent for the Global South.
- Data marketplaces mirror many characteristics of exchanges. As shown by China, government-proximate or closely regulated marketplaces appear to provide a solid avenue for facilitating trade in data while achieving policy goals.

Recommendations

- **Develop an international data measurement scheme:** Establish an international standard for measuring data flows and data industries, akin to the existing Harmonized System (HS) and the extended balance of payments services (EBOPS) classification systems. This standard should disaggregate trade of data into industries and uses, provide metrics for both quantity (bytes) and value (dollars), and mandate reporting of cross-border transactions — even if the data itself does not cross borders.
- **Approach data as a neo-financial product:** Recognize the similarity between the trade of data and the trade of financial instruments, and establish data marketplace authorities to do the following:
  - **Build public regulatory capacity incrementally:** Control and regulate the exchange of data — particularly across international borders — by designing consumer safeguards, developing lists of allowed and prohibited instruments, and building government capacity for audits.
  - **Ensure appropriate burden sharing:** Develop reporting standards, akin to financial regulatory fillings, that require transparent disclosure on the scale of exchanges and source of data.

Introduction

In the twenty-first century, the role of data in the global economy has grown at an unprecedented pace due to the rise of the digital economy (Organisation for Economic Co-operation and Development [OECD] 2022). However, despite fuelling many recent technological advances, regulation has mostly mimicked traditional trade rules (which rely heavily on the accurate measurement of the physical flow of goods) while lacking effective measurement tools to implement those rules (Blais-Savoie 2026). This lacuna restrains the ability of states to implement and evaluate the enforcement of their own policies. Thus, this working paper will draw a parallel between this moment and twentieth-century initiatives to first develop international trade standards, then explore current approaches that utilize existing data sources to measure the data economy, underlining the scale and the growing prominence of this challenge. It will then suggest a path forward inspired by historical successes in traditional trade and contemporary achievements in utilizing centralized data markets, akin to financial exchanges.

Before tackling the policy section of this working paper, it is imperative to specify its scope. The world of data is expansive, spanning from a car's mileage history to military reconnaissance drone video. To present high-level trade policy recommendations, this working paper purposefully narrows this scope by treating all data subjects as generally fungible, and by examining privacy-sensitive data concerns.[1] It focuses on data with a clear commercial purpose, akin to the definition of intangible assets in the International Accounting Standard 38.[2]

History of Trade Measures

Before proposing a new paradigm for trade measurement, it is imperative to understand the trajectory that led to the current trade measurement framework. It allows for the establishment of a common foundation of both understanding and purpose, as well as the contextualization of the tremendous amount of work required to develop a global measurement framework.

Coordinated calls for the creation of an international trade measurement system started as early as the second half of the nineteenth century, during a series of commercial and statistical conferences held throughout Europe. However, notably due to the complexity of agreeing on the numerous standards in existence as well as the parallel creation of "minimum lists"[3] and conversion tables (Friedenberg 1968), it took until 1937 for a draft proposal to be distributed by the then League of Nations and until 1959 for the "Brussels Tariff Nomenclature" to come into force. This first nomenclature only included codes for classifying goods, not services or standardized measurements of those goods; these were only formalized as part of the 1988 HS (World Customs Organization 2018). The

1 See e.g. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*, OJ, L 119/1, art 9.

2 See www.ifrs.org/issued-standards/list-of-standards/ias-38-intangible-assets/.

3 The "minimum lists" refer to the minimum level of precision that countries would be asked to report for trade statistics comparability (Committee of Statistical Experts 1938, appendix).

creation of this system was necessary to allow for the collection of statistics, but it also facilitated the negotiation of trade agreements, the control of dangerous products and the implementation of rules of origin (ibid.). Since the creation of these measures, the value of global trade has grown more than 400-fold,[4] more than twice the growth of the previous 200 years (Dedinger and Girard 2017). To note, this period also coincided with the establishment of the modern international liberal order through the creation of major international finance organizations, such as the World Bank and the International Monetary Fund, as well as open trade policies facilitated by the General Agreement on Tariffs and Trade and the WTO. Which of these two is most responsible for this growth matters little in this case — their success was interlinked and neither exist for data. Underlining this interlinking and dichotomy is key as, for comparison, the value of trade in data-driven services is estimated to have grown by more than 6,000-fold in the last 25 years (World Bank 2021, 238). Therefore, the growth of the data economy is the consequence of a completely different context: one with little governance and limited measurements. In this case, the lack of rules — rather than the removal of previously existing rules — allowed for uncontrolled growth. However, as was learned throughout the late twentieth century's rapid growth of global trade, some barriers are necessary to ensure safety and stability (for example, regarding radioactive materials[5] or drugs[6]). Most explicit to the contrasting context, this growth figure comes from estimates drawn from proxy variables, making it practically impossible to discriminate accurately between types of digital trade. The current measurement paradigm relies on existing categories for traditional trade and services as well as catchall "other" classifications. It makes policy enforcement of digital trade impractical — at best — and of trade in data nearly impossible. Governments presently have very limited ability to assess the type of data crossing their borders, its source or its destination. This makes targeted trade controls unimplementable, international negotiations approximate and regulatory design presumptuous.

Emerging Challenges for Measuring Trade in Data

The problem of accurately measuring international trade in data is further complicated by data's ineffable value and geographically agnostic commercial use. On the first point, the actual value of data depends on a myriad of extrinsic and unrelated factors that cannot be determined by the seller and depend on the buyer. For example, a buyer might have access to other data sets that can be linked to amplify any particular data point's value. On the second point, data can be used by a company located in one country without the original data set ever moving. For example, a Canadian company might procure a storage server in Europe, purchase a data set from Asia and then sell a report in South America. The data never entered Canada or South America — it only travelled from Asia to Europe — but its output clearly landed in South America, with the value being captured in Canada.

4 See www.wto.org/english/res_e/statis_e/trade_evolution_e/evolution_trade_wto_e.htm.

5 See www.iaea.org/about/overview/history.

6 See www.unodc.org/unodc/en/about-unodc/index.html.

Valuing Data

Data does not need to cross borders to be used internationally, and the majority of its value is not inherent to itself but rather comes from its relationship to other data. This is because the commercial value of data is generally "inference": the creation of a secondary, transformed product that retains no — or very little — trace of the original resource (traceability is judicially contentious; see Reuters 2025). This challenge also greatly complicates the implementation of modern economic policies such as value-added taxes and, more importantly, supply-chain monitoring. Some economic literature has explored this pricing challenge and acknowledged its unique nature (Mohan, Bharathy and Jalan 2026; Lyu et al. 2025). More specifically, in a literature review, Jiayao Zhang et al. (2024, section 7) identified "arbitrage freeness" — the ability to infer additional data from a data set — as the most commonly used variable for modelling data prices. Alternatively, Jun Hao et al. (2025) identified metadata and other "alternative data" as important variables for accurately pricing data sets. This research supports the hypothesis that much of a data set's commercial value is not intrinsic to the data points. Furthermore, while less quantitative, the literature on privacy — emerging out of the 2001 USA PATRIOT Act (also known as the Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act)[7] and the infamous "Taste, Ties, and Time" data set[8] — has extensively studied the risks associated with inference (King and Forder 2016; Paterson and McDonagh 2018; Dempsey and Flint 2004; du Pin Calmon and Fawaz 2012). Notably, Alicia Solow-Niederman (2022) recently proposed that the advent of machine learning has evolved the data economy into an "inference economy." This is because the "generative" in "generative artificial intelligence" (AI) represents the epitome of inference: a general-purpose model is trained on a data set and then used to create wholly different data. In such cases, the existence of the data itself is of governance interest, not its use, content or meaning. This underlines the policy importance of measuring trade in data beyond mere commercial interests.

The State of Trade in Data Monitoring

High-profile scandals in the last decade — for example, Cambridge Analytica or, more recently, Palantir's collaboration with US Immigration and Customs Enforcement (*The Guardian* 2018; Funk 2019) — have mobilized public opinion, but the lack of a comprehensive understanding of the scale of the data economy has greatly restrained government's ability to implement sustainable solutions. These challenges are only amplified by the emergence of generative AI that both requires unprecedented quantities of data (Villalobos et al. 2022; Sevilla et al. 2024) and introduces new avenues for data collection (Gómez Ortega et al. 2025; Alzamil et al. 2025; King et al. 2025). Generative AI has also highlighted the importance of inference for monetizing data:

7 The USA PATRIOT Act, enacted after September 11, was one of the first to expand government's use of data on its own citizens for national security purposes. The 2001 attacks also, more generally, led to a greater use of data in intelligence, as revealed in the 2010s by Edward Snowden (Gellman 2013).

8 A high-profile study developed a supposedly anonymous "Taste, Ties, and Time" data set of students and jobs outcomes that was fully "de-anonymized" by an independent researcher within days of its publication. The data collected was publicly available, but consent for collection was not obtained, raising concerns over privacy (Lewis et al. 2008; Zimmer 2008; Parry 2011).

As with traditional goods, the bulk of the value lies not in the resource (for example, ore, data) but in the transformed product (for example, semiconductors, large language models).

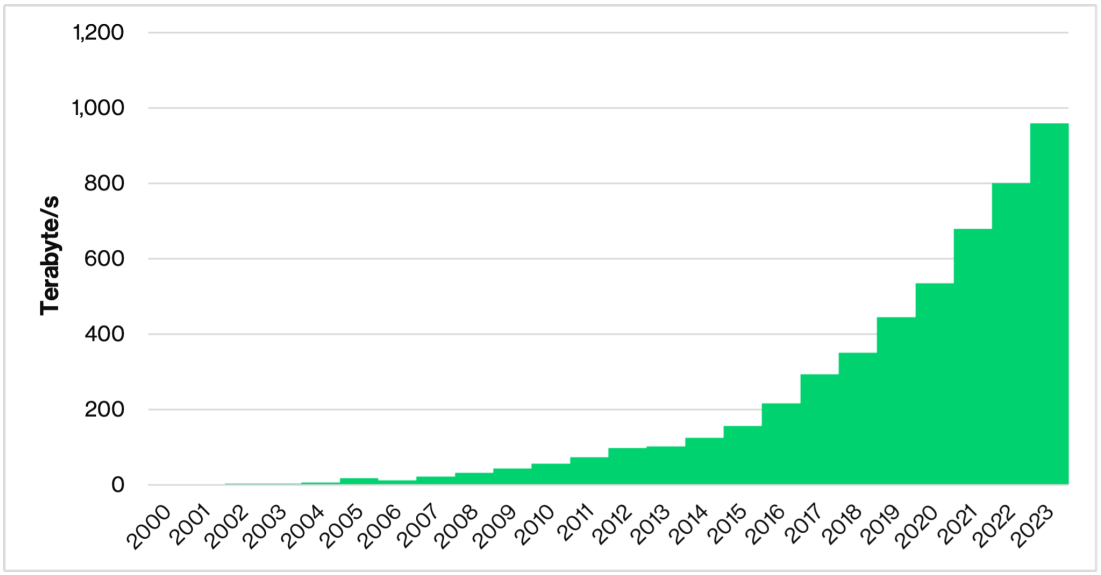
"Measuring" Data

Data can be measured from two angles: value (in dollars) and size (in bytes). At a high level, these two measures provide different policy-making evidence: the first informs economic and tax policy — by measuring economic contributions; the second informs privacy policy and infrastructure needs — by underlining the scale of information collection. There is inevitably some overlap: More valuable data could signal more comprehensive private profiles, while very large data sets could indicate greater levels of market capture. Understanding these distinct measurements is key to understanding the data economy and, in turn, designing informed policy.

Broadband Data

The International Telecommunication Union (ITU) maintains an extensive database of broadband statistics. These have been used to estimate the scale of flows but lack any information on the value or purpose of these flows (OECD 2022). Figure 1 shows international bandwidth usage (the amount of data crossing borders), underlining an increasingly exponential growth in the last few years, likely fuelled by the emergence of general-purpose AI (Villalobos et al. 2022; Sevilla et al. 2024). This approach benefits from the already extensive relationship between governments and "tier 1" internet service providers (Winther 2006). These companies already have some capacity to collect some disaggregated measures of the type of data flowing through their networks, but concerns over privacy have historically restrained these initiatives (Ohm 2009).

Figure 1: International Bandwidth Usage (World, 2000–2023)



Source: Author, with data from ITU DataHub (https://datahub.itu.int/data/?i=242).

Note: Missing data estimated through regression; 14 countries with less than five years of data omitted, accounting for less than 0.1 percent of yearly bandwidth.

Nevertheless, for better or for worse, expectations of privacy have changed since the earlier days of the internet; this avenue might hold potential for trade monitoring, especially if safeguards are put in place to, for example, restrict the collection of data originating from non-commercial IP addresses (for example, households).

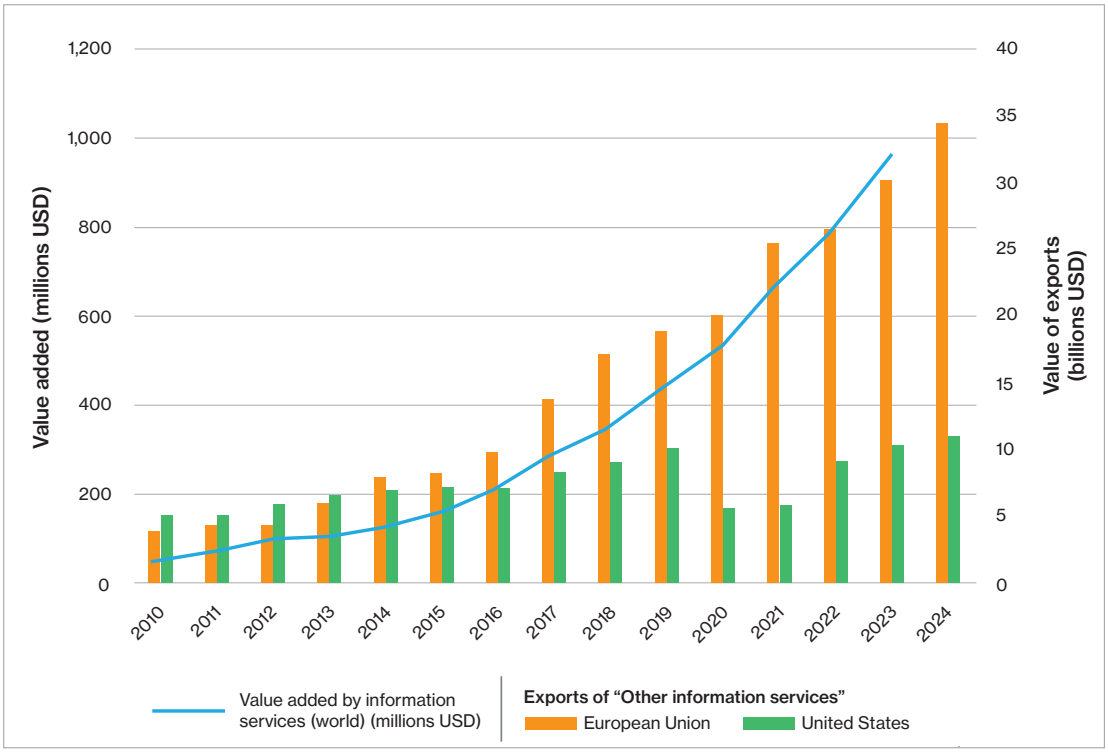
Trade in Services and Value Added

The OECD has been actively working to price international data transfers, identifying "value-added trade" as a useful proxy and EBOPS classification code 9.3.2 (other information services) as mostly representing the value of reported commercial data transfers (OECD 2022). However, due to its reliance on proxies and border declarations, this approach provides only a partial portrait: The HS code likely overestimates the value of reported cross-border data transfers due to its representation of all uncategorized assets, while both indicators fail to account for international transactions where data remains geographically static. Nevertheless, these metrics allow for an overview of the scale of the industry and its rapid growth.

Figure 2 showcases a perfect example of the impact of poor monitoring of international trade in data. The WTO, the OECD and the United Nations all host databases of trade data, none of which provide comprehensive, code-level global data. Even for countries with data, accuracy is mixed. The European Union has been at the forefront of this policy space, heavily funding research and spearheading measurement efforts (notably through the Digital Decade Policy Programme 2030[9]). The United States has been less interested in this area of research and, thus, despite other metrics suggesting it accounts for a much greater segment of the data economy, its contribution is underreported in trade statistics. This lack of information has major implications for the country's ability to regulate its data flows: Despite the American government, particularly the Trump administration, having actively raised concerns over critical data flowing to China, few policy actions have taken place (Harrell 2025).

9 *Decision (EU) 2022/2481 of the European Parliament and of the Council of 14 December 2022 Establishing the Digital Decade Policy Programme 2030*, OJ, L 323, online: <http://data.europa.eu/eli/dec/2022/2481/oj>.

Figure 2: Proxies for Measuring Trade in Data (World, European Union and United States, 2010–2024)



Source: Author, with data from OECD Data Explorer (international trade in services, https://data-explorer.oecd.org/s/473; trade in value added, https://data-explorer.oecd.org/s/474).

Data Markets

Finally, a — so far untapped by governments — source of statistics on the data economy is data marketplaces. The Western market is relatively concentrated: The major platforms hold a very large share of the market, allowing a relatively small sample to provide a solid overview of the overall scale (Azcoitia, Iordanou and Laoutaris 2021, Table 1). However, these data marketplaces remain relatively niche. While reviews on the role of Western exchanges remain limited, in China, where the space is likely much more developed due to the central role given to markets by the government, it has been estimated that they only account for five percent of the country's total data commerce (China Academy of Information and Communications Technology 2024). As for providing measurement estimates, this approach currently has two principal flaws: Data can be present on multiple platforms, inflating the total value measured; and it does not account for direct transactions between businesses or internal transactions within conglomerates or multinational corporations. Compounding the measurement challenge, private data marketplaces make extensive use of cryptocurrencies to facilitate transactions and generally appear to favour decentralized data storage infrastructure, in which the data is either never stored on the marketplaces' servers or is heavily fragmented — possibly across borders (Azcoitia and Laoutaris 2022). However, these challenges are all but solved by the establishment of government-regulated

marketplaces. This approach has been used by China and mirrors the current approach of Western governments in regulating exchanges. Indeed, data markets simultaneously act as commodity exchanges and retailers: They sell the primary input to be transformed, as well as the refined commodity (Arthur 2013; *The Economist* 2017; Medhora 2018).

However, by operating in a new market and navigating these two roles, the data exchange industry has evaded regulation. Traditional exchanges, such as stock and commodity markets, already have regulatory authorities tasked with minimizing fraud and protecting investors[10] (Goshen and Parchomovsky 2006), as well as precise reporting requirements that could easily be extended to data markets[11] (Canadian Securities Administrators 2016; U.S. Securities and Exchange Commission 2021). Similarly, rules for e-commerce,[12] although more novel, are starting to take shape to ensure accurate taxation and protection of consumers (Anggadol 2023). The challenges faced by financial regulators — speed and financial capacity — make them a perfect parallel. Despite some criticism, particularly in the wake of the 2008 global financial crisis (Guynn 2010; Bernanke 2009), these agencies have managed to follow the rapid advancements in the world of finance: from algorithmic trading to ever more innovative securities, as well as crypto and alternative currencies. They have also managed to regulate despite the immense financial capacity gap between themselves and the regulated industry. These strengths come from their independence from politics, certainly, but also from their ability to coalesce domain expertise and their emphasis on self-reporting guidelines, moving much of the implementation burden to the private sector.

Looking at the Chinese approach, in 2022, as part of the "Twenty Data Measures (数据二十条)," the country inaugurated a state-run data marketplace ecosystem (Guo 2026). These markets are managed by arms-length, state-owned enterprises and regulated by a central market authority. In creating this system, China had two clear and explicit goals: increasing governance of exchanges and enhancing the availability of data, particularly government data, for smaller companies and researchers (He and Arcesati 2024; The State Council 2022). Through this system, despite slow uptake, China has managed to grow its data market transactions by almost 50 percent annually (Shanghai Data Exchange 2024) while reforming data privacy and encryption standards. This case exemplifies the route proposed in this working paper for regulating the trade of data while working within the existing private infrastructure. By mirroring an existing regulatory framework, it could minimize pushback and simplify design.

10 See www.cftc.gov/IndustryOversight/MarketSurveillance/CFTCMarketSurveillanceProgram/index.htm; www.sec.gov/about/mission.

11 Ibid.

12 See www.canada.ca/en/revenue-agency/services/tax/businesses/topics/e-commerce/e-commerce.html.

Conclusion

In sum, despite rapid growth and increasingly frequent scandals, the international sale of data remains a poorly measured sector with limited policy. Some attempts have been made by international organizations to collect data and identify proxies of the sector, providing some valuable avenues for future measurement, such as increased collaboration with internet service providers or the expansion of the existing trade statistical frameworks for the digital era. Finally, the case of Chinese data markets highlights a potential smaller-scale, domestic solution to regulating trade in data.

Recommendations

- **Develop an international data measurement scheme:** Establish an international standard for measuring data flows and data industries, akin to the existing HS and EBOPS systems. This standard should disaggregate trade of data into industries and uses; provide metrics for both quantity (bytes) and value (dollars); and mandate reporting of cross-border transactions, even if the data itself does not cross borders.
- **Approach data as a neo-financial product:** Recognize the similarity between the trade of data and the trade of financial instruments, and establish data marketplace authorities to do the following:
  - **Build public regulatory capacity incrementally:** Control and regulate the exchange of data — particularly across international borders — by designing consumer safeguards, developing lists of allowed and prohibited instruments, and building government capacity for audits.
  - **Ensure appropriate burden sharing:** Develop reporting standards, akin to financial regulatory fillings, that require transparent reporting on the scale of exchanges and the source of data.

Acknowledgements

The author wishes to thank Ken Jackson, Wilfrid Laurier University, for his supervision during this project, as well as Jeni Tennison and Alexander Martin for their comments on earlier versions of this working paper. The author also wishes to acknowledge the feedback provided by Teresa Scassa and Mehrsa Ehsani in their comments on the first working paper produced through this fellowship.

Works Cited

Alzamil, Leena Mohammad, Alawiayyah Mohammed Alhasani and Suhair Alshehri. 2025. “Privacy Concerns in ChatGPT Data Collection and Its Impact on Individuals.” *Future Internet* 17 (11): 511. https://doi.org/10.3390/fi17110511.

Anggadol, Kairos. 2023. “Canada’s e-commerce laws: what you need to know.” *Lexpert*, September 21. www.lexpert.ca/news/legal-faq/canadas-e-commerce-laws-what-you-need-to-know/379856.

Arthur, Charles. 2013. “Tech giants may be huge, but nothing matches big data.” *The Guardian*, August 23. www.theguardian.com/technology/2013/aug/23/tech-giants-data.

Azcoitia, Santiago Andrés, Costas Iordanou and Nikolaos Laoutaris. 2021. “What Is the Price of Data? A Measurement Study of Commercial Data Marketplaces.” Preprint, arXiv, October 25. https://doi.org/10.48550/arXiv.2111.04427.

Azcoitia, Santiago Andrés and Nikolaos Laoutaris. 2022. “A Survey of Data Marketplaces and Their Business Models.” *ACM SIGMOD Record* 51 (3): 18–29. https://doi.org/10.1145/3572751.3572755.

Bernanke, Ben S. 2009. “Financial Regulation and Supervision after the Crisis: The Role of the Federal Reserve.” Speech delivered at the Federal Reserve Bank of Boston 54th Economic Conference, Chatham, Massachusetts, October 23. www.federalreserve.gov/newsevents/speech/bernanke20091023a.htm.

Blais-Savoie, Fabrice. 2026. *Trade in Data: Rapid Flows, Lagging Policy*. Digital Policy Hub Working Paper. www.cigionline.org/publications/trade-in-data-rapid-flows-lagging-policy/.

Canadian Securities Administrators. 2016. “Multilateral Instrument 96-101 Trade Repositories and Derivatives Data Reporting.” www.bcsc.bc.ca/-/media/PWS/New-Resources/Securities-Law/Instruments-and-Policies/Policy-9/96101-MI-July-9-2020.pdf.

China Academy of Information and Communications Technology. 2024. *Research Report on the Development Index of Data Trading Venues (2024)*. August 16. Beijing, China: China Academy of Information and Communications Technology. www.caict.ac.cn/kxyj/qwfb/ztbg/202408/P020240816544947002101.pdf.

Committee of Statistical Experts. 1938. *Minimum List of Commodities for International Trade Statistics*. Studies and Reports on Statistical Methods No. 2. Geneva, Switzerland: League of Nations. https://deriv.nls.uk/dcn23/1903/2276/190322761.23.pdf.

Dedinger, Béatrice and Paul Girard. 2017. “Exploring trade globalization in the long run: The RICardo project.” *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 50 (1): 30–48. https://doi.org/10.1080/01615440.2016.1220269.

Dempsey, James X. and Lara M. Flint. 2004. “Commercial Data and National Security.” *George Washington Law Review* 72: 1459–1502. https://cdt.org/wp-content/uploads/publications/200408dempseyflint.pdf.

du Pin Calmon, Flavio and Nadia Fawaz. 2012. “Privacy Against Statistical Inference.” Preprint, arXiv, October 8. https://doi.org/10.48550/arXiv.1210.2123.

Friedenberg, Howard L. 1968. *The Development of a Uniform International Tariff Nomenclature: From 1853 to 1967 With Emphasis on the Brussels Tariff Nomenclature*. April. TC Publication 237. United States Tariff Commission. www.usitc.gov/publications/tariff_affairs/pub237.pdf.

Funk, McKenzie. 2019. “How ICE Picks Its Targets in the Surveillance Age.” *The New York Times Magazine*, October 2. www.nytimes.com/2019/10/02/magazine/ice-surveillance-deportation.html.

Gellman, Barton. 2013. “Edward Snowden, after months of NSA revelations, says his mission’s accomplished.” *The Washington Post*, December 23. www.washingtonpost.com/world/national-security/edward-snowden-after-months-of-nsa-revelations-says-his-missions-accomplished/2013/12/23/49fc36de-6c1c-11e3-a523-fe73f0ff6b8d_story.html.

Gómez Ortega, Alejandra, Hosana Morales Ornelas and Uğur Genç. 2025. “Surrendering to Powerlesness: Governing Personal Data Flows in Generative AI.” *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Article No. 232, 1–18. https://doi.org/10.1145/3706598.3713504.

Goshen, Zohar and Gideon Parchomovsky. 2006. “The Essential Role of Securities Regulation.” *Duke Law Journal* 55 (4): 711–82. https://scholarship.law.duke.edu/dlj/vol55/iss4/1/.

Guo, Ran. 2026. “Assetizing, Trading, Franchising: China’s Strategy for Building a National Data Economy.” Asia Society Policy Institute, February 13. https://asiasociety.org/policy-institute/assetizing-trading-franchising-chinas-strategy-building-national-data-economy.

Guynn, Randall D. 2010. “The Financial Panic of 2008 and Financial Regulatory Reform.” Harvard Law School Forum on Corporate Governance, November 20. https://corpgov.law.harvard.edu/2010/11/20/the-financial-panic-of-2008-and-financial-regulatory-reform/.

Hao, Jun, Zeyu Deng, Jin Li and Jianping Li. 2025. “How to price a dataset: a deep learning framework for data monetization with alternative data.” *Humanities & Social Sciences Communications* 12 (1): 1736. https://doi.org/10.1057/s41599-025-06016-y.

Harrell, Peter E. 2025. “Managing the Risks of China’s Access to U.S. Data and Control of Software and Connected Technology.” January. Washington, DC: Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2025/01/managing-the-risks-of-chinas-access-to-us-data-and-control-of-software-and-connected-technology.

He, Alex and Rebecca Arcesati. 2024. “Data Marketplaces and Governance: Lessons from China.” The Role of Governance in Unleashing the Value of Data Essay Series, Centre for International Governance Innovation, November 12. www.cigionline.org/articles/data-marketplaces-and-governance-lessons-from-china/.

King, Jennifer, Kevin Klyman, Emily Capstick, Tiffany Saade and Victoria Hsieh. 2025. “User Privacy and Large Language Models: An Analysis of Frontier Developers’ Privacy Policies.” Preprint, arXiv, September 5. https://doi.org/10.48550/arXiv.2509.05382.

King, Nancy J. and Jay Forder. 2016. “Data analytics and consumer profiling: Finding appropriate privacy principles for discovered data.” *Computer Law & Security Review* 32 (5): 696–714. https://doi.org/10.1016/j.clsr.2016.05.002.

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer and Nicholas Christakis. 2008. “Tastes, ties, and time: A new social network dataset using Facebook.com.” *Social Networks* 30 (4): 330–42. https://doi.org/10.1016/j.socnet.2008.07.002.

Lyu, Yangming, Linyi Qian, Zhixin Yang, Jing Yao and Xiaochen Zuo. 2025. “Optimal pricing approaches for data markets in market-operated data exchanges.” *Statistical Theory and Related Fields* 10 (1): 23–45. https://doi.org/10.1080/24754269.2025.2588857.

Medhora, Rohinton P. 2018. “Data Governance in the Digital Age.” Data Governance in the Digital Age Essay Series, Centre for International Governance Innovation, May 16. www.cigionline.org/publications/data-governance-digital-age/.

Mohan, Sai Krishnan, Gnana Bharathy and Akanksha Jalan. 2026. “Enterprise Data Valuation – A Targeted Literature Review.” *Journal of Economic Surveys* 40 (1): 73–92. https://doi.org/10.1111/joes.12705.

OECD. 2022. “Measuring the value of data and data flows.” OECD Digital Economy Papers No. 345. Paris, France: OECD Publishing. https://doi.org/10.1787/923230a6-en.

Ohm, Paul. 2009. “The Rise and Fall of Invasive ISP Surveillance.” *University of Illinois Law Review* 2009: 1417–96. https://illinoislawreview.org/wp-content/ilr-content/articles/2009/5/Ohm.pdf.

Parry, Marc. 2011. “Harvard Researchers Accused of Breaching Students’ Privacy.” *The Chronicle of Higher Education*, July 10. www.chronicle.com/article/harvard-researchers-accused-of-breaching-students-privacy/.

Paterson, Moira and Maeve McDonagh. 2018. “Data Protection in an Era of Big Data: The Challenges Posed by Big Personal Data.” *Monash University Law Review* 44 (1): 1–31. www.monash.edu/__data/assets/pdf_file/0009/1593630/Paterson-and-McDonagh.pdf.

Reuters. 2025. “Key Rulings on GenAI Training and Copyright Fair Use.” *Practical Law The Journal*. July. www.reuters.com/practical-law-the-journal/litigation/key-rulings-genai-training-copyright-fair-use-2025-07-01/.

Sevilla, Jaime, Tamay Besiroglu, Ben Cottier, Josh You, Edu Roldán, Pablo Villalobos and Ege Erdil. 2024. *Can AI scaling continue through 2030*? Epoch AI, August 20. https://epoch.ai/blog/can-ai-scaling-continue-through-2030.

Shanghai Data Exchange. 2024. *2024 China Data Trading Market Research and Analysis Report*. December 12. www.xinhuanet.com/tech/20241212/551d4a39f90e4ff684ff0809fc31985a/c.html.

Solow-Niederman, Alicia. 2022. “Information Privacy and the Inference Economy.” *Northwestern University Law Review* 117 (2): 357–424. https://scholarlycommons.law.northwestern.edu/nulr/vol117/iss2/1/.

The Economist. 2017. “The world’s most valuable resource is no longer oil, but data.” May 6. www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

The Guardian. 2018. “The Cambridge Analytica Files.” News Series. www.theguardian.com/news/series/cambridge-analytica-files.

The State Council. 2022. “China unveils measures to build basic systems for data.” The People’s Republic of China. December 19. https://english.www.gov.cn/policies/latestreleases/202212/19/content_WS63a17f7dc6d0a757729e49bd.html.

U.S. Securities and Exchange Commission. 2021. “Data Management Board Charter.” November 2. www.sec.gov/files/dmb-charter-20211102.pdf.

Villalobos, Pablo, Anson Ho, Jaime Sevilla, Tamay Besiroglu, Lennart Heim and Marius Hobbhahn. 2022. “Will we run out of data? Limits of LLM scaling based on human-generated data.” Preprint, arXiv, October 26. https://arxiv.org/abs/2211.04325v2.

Winther, Mark. 2006. “Tier 1 ISPs: What They Are and Why They Are Important.” May. White Paper. Framingham, MA: IDC. www.gin.ntt.net/wp-content/uploads/2020/01/IDC_Tier1_ISPs.pdf.

World Bank. 2021. *World Development Report 2021: Data for Better Lives*. Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-1600-0.

World Customs Organization. 2018. *The Harmonized System: A universal language for international trade*. December. Brussels, Belgium: World Customs Organization. www.wcoomd.org/-/media/wco/public/global/pdf/topics/nomenclature/activities-and-programmes/30-years-hs/hs-compendium.pdf.

Zhang, Jiayao, Yuran Bi, Mengye Cheng, Jinfei Liu, Kui Ren, Qiheng Sun, Yihang Wu et al. 2024. “A Survey on Data Markets.” Preprint, arXiv, November 9. https://doi.org/10.48550/arXiv.2411.07267.

Zimmer, Michael. 2008. “More On the ‘Anonymity’ of the Facebook Dataset — It’s Harvard College (Updated).” *MichaelZimmer* (Blog), October 4. https://michaelzimmer.org/2008/10/03/more-on-the-anonymity-of-the-facebook-dataset-its-harvard-college/.