
Centre for International
Governance Innovation

CIGI Papers No. 273 – March 2023

Could a Global “Wicked Problems Agency” Incentivize Data Sharing?

Susan Ariel Aaronson



CIGI Papers No. 273 – March 2023

Could a Global “Wicked Problems Agency” Incentivize Data Sharing?

Susan Ariel Aaronson

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.

À propos du CIGI

Le Centre pour l'innovation dans la gouvernance internationale (CIGI) est un groupe de réflexion indépendant et non partisan dont les recherches évaluées par des pairs et les analyses fiables incitent les décideurs à innover. Grâce à son réseau mondial de chercheurs pluridisciplinaires et de partenariats stratégiques, le CIGI offre des solutions politiques adaptées à l'ère numérique dans le seul but d'améliorer la vie des gens du monde entier. Le CIGI, dont le siège se trouve à Waterloo, au Canada, bénéficie du soutien du gouvernement du Canada, du gouvernement de l'Ontario et de son fondateur, Jim Balsillie.

Credits

Managing Director of Digital Economy **Robert Fay**
Project Manager **Jenny Thiel**
Senior Publications Editor **Jennifer Goyder**
Publications Editor **Susan Bubak**
Graphic Designer **Brooklynn Schwartz**

Copyright © 2023 by the Centre for International Governance Innovation

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the Centre for International Governance Innovation or its Board of Directors.

For publications enquiries, please contact publications@cigionline.org.



This work is licensed under a Creative Commons Attribution — Non-commercial — No Derivatives License. To view this license, visit (www.creativecommons.org/licenses/by-nc-nd/3.0/). For re-use or distribution, please include this copyright notice.

Printed in Canada on Forest Stewardship Council® certified paper containing 100% post-consumer fibre.

Centre for International Governance Innovation and CIGI are registered trademarks.

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

Table of Contents

| | |
|----|-----------------------------------------------------------------------------------|
| vi | About the Author |
| vi | Acronyms and Abbreviations |
| 1 | Executive Summary |
| 1 | Introduction |
| 2 | Background |
| 4 | Why Are We Talking about Data as a Public Good? |
| 5 | Private Firms Want to Control the Data They Collect and Prefer Not to Share It |
| 8 | What Do We Mean by Data Sovereignty and How Might It Impede Data Sharing? |
| 9 | Voluntary Data Sharing and the Role of Governments |
| 11 | What Could Nations Do to Collaborate in the Global Public Interest? |
| 13 | Works Cited |

About the Author

Susan Ariel Aaronson is a CIGI senior fellow. She is an expert in international trade, digital trade, artificial intelligence (AI) governance and human rights. She is currently writing on public participation in the governance of AI, comparative advantage in extended reality (XR) and the future of data governance.

Susan is also a research professor of international affairs and a cross-disciplinary fellow at George Washington University's Elliott School of International Affairs, where she directs the Digital Trade and Data Governance Hub. The Hub educates policy makers and the public on domestic and international data governance. The Hub also maps the governance of personal, public and proprietary data around the world to illuminate the state of data governance.

Susan is the former Minerva Chair at the National War College. She is the author of six books and more than 70 scholarly articles. Her work has been funded by major international foundations including the MacArthur, Minderoo, Hewlett and Ford Foundations; governments such as the Netherlands, the United States and Canada; international organizations such as the United Nations, International Labour Organization and the World Bank; and US corporations including Google, Ford Motor and Levi Strauss. She loves to do triathlons and study ballet.

Acronyms and Abbreviations

| | |
|---------------|--------------------------------------------------------|
| AI | artificial intelligence |
| DSA | Digital Services Act |
| GDPR | General Data Protection Regulation |
| GovLab | Governance Lab |
| IoT | Internet of Things |
| IP | intellectual property |
| IPRs | intellectual property rights |
| NGOs | non-governmental organizations |
| ODI | Open Data Institute |
| OECD | Organisation for Economic Co-operation and Development |
| UNCTAD | UN Conference on Trade and Development |
| UNHCR | United Nations High Commissioner for Refugees |

Executive Summary

The world is rife with “wicked” problems — problems that no one knows how to solve without creating further problems. Wicked problems such as climate change, terrorism or global poverty transcend nations, generations and traditional solutions. Moreover, there is no one or best way to address wicked problems because they have many different causes and manifest in different contexts (for example, terrorism in the United States is different from that in Ireland or Mali).

Policy makers have a responsibility to attempt to mitigate such problems. The world has the tools, analytical expertise and vast troves of various types of data to assist them in this task. By mixing vast troves of data, policy makers and researchers may find new insights and strategies to address these complex problems. However, national and international government agencies and large corporations generally control the use of such data. These data controllers often hoard data and only occasionally share it. Moreover, some big companies (and likely some governments) do not have a handle on the data they hold, and even where it is stored. These companies have little incentive to review their data and assess how it might be helpful to others.

The author proposes a new international organization, the Wicked Problems Agency, to catalyze both data sharing and data analysis in the interest of mitigating wicked problems. This organization would work to prod societal entities — firms, individuals, civil society groups and governments — to share and analyze various types of data in the hope that such data sharing coupled with sophisticated data analysis could provide new insights.

The Wicked Problems Agency would be different from many other international organizations, as it would be international, cloud-based and focused on analysis, rather than providing public services. It would also serve as a centre for international and cross-disciplinary collaboration and training in the latest forms of data analysis. To entice firms and governments to share data, the Wicked Problems Agency would rent useful data and compensate those entities that control data, in the interest of incentivizing data

sharing. It would also be required to protect that data from theft, manipulation and so on.

Over time, the Wicked Problems Agency is likely to produce additional benefits. It could reduce the data market’s current opacity (most people and firms do not know what data they hold and its current and potential value). Moreover, it could lead entities to hire people who can think creatively about the multidimensional uses of data. Finally, it is increasingly clear that data analysis is the future of innovation. The Wicked Problems Agency could provide a practical example of how data sharing can yield both economic and public good benefits.

Introduction

In 2021, the United Nations Trade and Development Programme (UNCTAD) issued a plea in its 2021 Digital Economy Report: “Global data-sharing can help address major global development challenges such as poverty, health, hunger and climate change...Without global cooperation on data and information, research to develop the vaccine and actions to tackle the impact of the pandemic would have been a much more difficult task. Thus, in the same way as some data can be public goods, there is a case for some data to be considered as global public goods, which need to be addressed and provided through global governance” (UNCTAD 2021, 198).

Global public goods are goods and services with benefits and costs that potentially extend to all countries, people and generations. Global data sharing can also help solve what scholars call “wicked” problems — problems so complex that they require innovative, cost-effective and global mitigating strategies. Wicked problems are problems that no one knows how to solve without creating further problems; they transcend nations, generations and traditional solutions.

Data sharing is not always “good for society.” Both corporations and governments can share data inappropriately, violating human rights and data protection laws. For example, Facebook shared data inappropriately with the British research firm Cambridge Analytica (Confessore 2018). Human Rights Watch reported in December 2022 that the Hungarian government repurposed data it collected

from people applying for government services. The government not only used this information to disseminate the campaign messages of the ruling party Fidesz, but it also allegedly shared the data among agencies, leading to selective enforcement of laws that further benefited Fidesz (Human Rights Watch 2022). In this regard, data sharing among social entities is no different from other activities — all processes and programs have costs and benefits.

But if public, private and non-governmental entities could protect and anonymize personal data (a big if) and share it to achieve public good purposes, the people of the world could reap substantial benefit.

Scholars have tried to quantify the benefits of data sharing, but this author views these studies cautiously since we are in the early stages of measuring and valuing data. The Organisation for Economic Co-operation and Development (OECD) notes that some studies suggest that data access and sharing can increase the value of data to holders (direct impact), but it can also help create 10 to 20 times more value for data users (indirect impact), and 20 to 50 times more value for the wider economy (induced impact). At the same time, data access and sharing may also reduce the producer surplus of data holders (the share of money going to the big companies that have long been collecting personal data). The OECD concludes that these studies suggest that data access and sharing can help generate social and economic benefits worth between 0.1 percent and 1.5 percent of GDP in the case of public sector data, and between 1 percent and 2.5 percent of GDP (in a few studies up to 4 percent of GDP) when also including private sector data (OECD 2019). In addition, data access and sharing can also boost spillover benefits by enabling “super-additive” insights that may be greater than the sum of insights from isolated parts (data silos), leading to increasing returns to scope (OECD 2015).

This paper discusses why the world has made so little progress in encouraging cross-sectoral and international data sharing. As the UNCTAD report noted, data generated in one country can also provide social value in other countries, which would call for sharing of data at the international level through a set of shared and accountable rules (UNCTAD 2021). Moreover, the world is drowning in data, yet much of that data remains hidden and underutilized

However, individuals and groups of individuals in the private and civil society sector can be motivated by greed and guilt to do things differently. Herein, the author suggests a new international organization, the Wicked Problems Agency. This organization would work to prod societal entities — firms, individuals, civil society groups and governments — to share various types of data in the hope that such data sharing coupled with sophisticated data analysis could provide new insights into the mitigation of wicked problems.

Background

Many factors impede global data sharing for public good purposes; this analysis focuses on two.

First, policy makers generally do not think about data as a global public good; they view data as a commercial asset that they should nurture and control. While they may understand that data can serve the public interest, they are more concerned with using data to serve their country’s economic interest (Aaronson 2022). Second, many leaders of civil society and business see the data they have collected as proprietary data — they collect it, control it and can control the reuse of that data. So far, many leaders of private entities with troves of data are not convinced that their organization will benefit from such sharing, although many private entities voluntarily share some data for social good purposes.

However, data cannot meet its public good purpose if data is not shared among societal entities. Moreover, if data is a sovereign asset, policy makers are unlikely to encourage cross-border data sharing even if it might address problems that transcend borders. As the Bennett Institute for Public Policy and the Open Data Institute (ODI) have argued, “value comes from data being brought together, and that requires organizations to let others use the data they hold” (Bennett Institute and ODI 2020a, 4). But that also means the entities that collected the data may not accrue all of the benefits from that data (ibid.). In short, private entities are not sufficiently incentivized to share data in the global public good.

Researchers have presented several reasons why societal entities should share data for the public good. The Open Data Policy Lab at the Governance

Lab (GovLab) at New York University stressed that gaining access to data sources and other assets held by organizations may facilitate business decisions, find and rectify errors, and gain new insights. These organizations can then test to see if their findings are replicable. Sharing data in the public interest could also enhance reputation and fulfill an organization's social responsibilities. It can help firms recruit and retain talent. Finally, it can help organizations comply with regulations, become more transparent, or otherwise promote responsible data management (Verhulst, Young and Zahuranec 2021). Meanwhile, the ODI found that data sharing in the public interest could improve market reach, help entities benchmark and gain insights into their performance, build relationships with stakeholders, optimize their supply chain, address sector-specific challenges, and build and sustain trust (ODI 2020).

However, policy makers are just beginning to suggest strategies or mechanisms to facilitate data sharing to achieve the *domestic public good nature of data*.¹ These plans describe data as a sovereign asset that should benefit their citizens and local/national organizations. For example, in 2018, the French government asked French politician and mathematician Cédric Villani² to head up a task force on artificial intelligence (AI) for France and Europe's future. After the task force's investigation, Villani and his colleagues stressed that "the benefits of data...are currently enjoyed by a set of a few major stakeholders" (Villani 2018, 6). These stakeholders were not based in Europe. To address this problem of data inequity, he recommended that "the public authorities must introduce new ways of producing, sharing and governing data by making data a common good," where a community would define use and governance (ibid., 8). In addition, Villani stressed that data policies must be designed to safeguard EU sovereignty, protect privacy and foster economic growth among the nations of Europe (ibid.). In another example, Ravi Shankar Prasad, India's minister of law and justice, electronics and information technology, and communications, declared in 2020, "Data is a nation's asset. This great asset of data has to be properly used, processed and value added for healthcare,

agriculture and education. Therefore, not only data ownership but also data sovereignty becomes important" (quoted in Press Trust of India 2020).

While policy makers are thinking of ways to ensure data benefits their constituents, executives at firms are experimenting with a new form of corporate social responsibility called data stewardship, which could facilitate data sharing for both national and international public good purposes. Data stewards are individuals who work to create public value (including official statistics) by reusing data and data expertise, identifying opportunities for productive cross-sectoral collaboration, and proactively requesting or enabling functional access to data and insights from that data (Verhulst 2021a; 2021b).³ Organizations that steward data make important decisions about who has access to it, for what purposes and to whose benefit (Massey 2022). To this end, data stewards have created new entities such as data collaboratives⁴ and data trusts⁵ to facilitate this process (Coyle et al. 2020).

Meanwhile, many private firms that collect data have a "data for good" program (Porway 2021; 2022). However, although these programs can be global, these efforts may not be scalable at the global level. Moreover, these programs are at the behest of the management of the private entity and if economic conditions worsen, executives may be less willing to share. Finally, data for good, data stewardship and data social responsibility strategies cannot alter market conditions that act as a disincentive to data sharing.

Before beginning the analysis, the author makes several caveats. First, the author refers to data as a generic term, but acknowledges that there are many different sources (internet-connected devices and satellites) and types of data (personal and non-personal). Second, the author sidesteps the competitiveness and national security implications

1 Governments are starting to propose strategies for cross-border data sharing, including Japan's free flow of data built on trust and Switzerland's proposal for trustworthy data spaces.

2 See https://en.wikipedia.org/wiki/C%C3%A9dric_Villani.

3 We can see data stewardship in action at data.org, a platform for partnerships such as that between the World Food Programme and Tableau (a data visualization firm), which maps food insecurity after natural disasters or conflict. See www.data.org; www.tableau.com/about/blog/2020/9/inside-look-world-food-programmes-data-driven-response-hunger-during-covid-19.

4 Data collaboratives are a new form of collaboration, beyond the public-private partnership model, in which participants from different sectors – in particular companies – exchange their data to create public value.

5 Data trusts and collaboratives are vehicles for different entities to share data in a trustworthy manner, whereas data stewardship acts as an incentive to share data.

of data, which can affect data controllers' willingness to share data. Third, the paper does not address all of the problems related to multi-sectoral data sharing for the public good — such as making the data findable, useable, shareable (machine-readable format) and useful. Fourth, data portability rules, which enable individuals to move their data from one company to another, are not discussed.

This paper proceeds as follows: First, it examines why/how data can be a public good. Next, the paper discusses why private entities control so much of the world's supply of personal data, as well as other types of data, and why firms are reluctant to share that data. The analysis then focuses on the role of policy makers and their belief that data sovereignty might ensure that their country can utilize data to prosper. Next, it examines the current state of data sharing for global public good purposes and what is limiting in such sharing. The paper concludes with some recommendations.

Why Are We Talking about Data as a Public Good?

In most economies, we rely on market forces — supply and demand — to produce needed goods and services. But markets do not always work efficiently or equitably (Duch-Brown, Martens and Mueller-Langer 2017; Mirando Montoya et al. 2022). The people of the world do not all have equal access, ability and understanding to use data effectively (a data divide). Countries also have different abilities to collect and monetize data to enhance human welfare (UNCTAD 2021). For this reason, some analysts (including this author) describe data as both a commercial asset *and* a public good (Aaronson 2022; Bennett Institute for Public Policy and ODI 2020a).

A public good is a good or service that the “free market” will underproduce because it is non-excludable and non-rival (Duch-Brown, Martens and Mueller-Langer et al. 2017). Data does not quite fit the public good paradigm. Data is inherently non-rival — it does not get used up as most people use it. But users can be excluded through a wide range of strategies,

including intellectual property rights (IPRs), trade rules, data governance rules and price (Bennett Institute for Public Policy and ODI 2020a, 7).

Nobel Prize-winning economist Elinor Ostrom provided a model for governance of such public goods — a commons where various stakeholders would collaborate to ensure that these public goods could be provided and used in an equitable and efficient manner. According to economist Diane Coyle (2020), “Conventional property rights make conflicts over who ‘owns’ this value inevitable, and hence the growing interest in forms of data governance that could deliver trustworthy access to data. A classic commons problem can be tackled by assigning private ownership and access rights; the challenge with non-rivalrous data is to assign common ownership and access rights.”

Scholars and practitioners are developing strategies to govern data based on the commons model.⁶ The commons includes software commons such as GitHub, licensing commons such as Creative Commons, open-access scholarly journals, digital repositories, institutional commons such as digital libraries or Wikipedia, and subject matter commons.

Although shared models for access to data are gaining traction, data sharing and wicked problem solving are hampered by private sector data opacity, models of data ownership and suspected hoarding by firms. Executives at these firms know the data they collect and hold has value, although they might not know how to achieve that value. So they supposedly hoard data. Moreover, these companies are not required to be transparent about the data they collect and how they might use it for commercial purposes (Coyle 2022).

Meanwhile, the public is not demanding that their data, whether held by public or private entities, should be shared to achieve public good purposes. In fact, survey data reveals people are divided — they are open to certain entities sharing data and will support such data sharing if they trust the organization using the data. A 2019 Pew Research Center survey found many Americans do not think they benefit much from the collection of their data, and they find that the potential risks of this practice outweigh the benefits (Auxier and Rainie 2019). However,

⁶ See Bloom (2020) and Bloom et al. (2021).

49 percent of Americans are comfortable with data sharing to assess potential terrorist threats but only 25 percent think smart speaker makers should share users' audio recordings with law enforcement to help with criminal investigations (ibid.). Deltapoll interviewed 2,002 UK adults online between August 31 and September 4, 2020, on behalf of the British government's Centre for Data Ethics and Innovation. It asked questions relating to data sharing, the impact of data use and sharing on the economy and society. Some 64 percent of those polled stated it could or would bring benefits both to innovation and society as a whole.⁷ A 2021 survey of 2,000 US respondents also revealed ambivalence about data sharing. Some 50–75 percent of survey respondents were willing to share data from their own Internet of Things (IoT) devices for public benefits. Yet respondents stated they trusted government less than sharing data with other organizations with public purposes, such as universities, non-profits and health-care institutions (Mossberger, Cho and Cheong 2022).

In another US-based study, the authors surveyed and interviewed a random sample of Americans. Participants strongly preferred data reuse for public health and research over profit-driven, marketing or crime-detection activities. Participants also strongly preferred data use by universities or non-profit organizations over data use by businesses and governments (Schmit et al. 2021).

Private Firms Want to Control the Data They Collect and Prefer Not to Share It

Markets for data are opaque and, as a result, no one knows the scale, scope or specifics of such markets. It is likely that private entities (firms, civil society groups and so on) control a growing portion of the world's data. Some argue that private sector collection and storage of data today is greater than that of many governments (Bennett Institute

for Public Policy and ODI 2020b, 6). To make this assertion, one must rely on proxies such as who funds and accrues scientific data. For example, the OECD has found that in many fields of science, the private sector has become the main funder.⁸ In the United States, the National Science Foundation found "that federal agencies provided only 44% of the \$86 billion spent on basic research in 2015. The federal share, which topped 70% throughout the 1960s and '70s, stood at 61% as recently as 2004 before falling below 50% in 2013" (Mervis 2017). To the funder go the profits and the data. And that data acts as a force multiplier, increasing the power of the firm. For example, Google and Apple are now creating some of the world's most accurate and widely used maps, outstripping the capabilities of national mapping agencies. In addition, social media platforms such as Facebook, LinkedIn and Twitter know more about the social and work networks of their many customers than government officials (Simon 2019; Schleffer and Miller 2021).

UNCTAD (2021) reports that 70 giant global firms hold the bulk of the world's collections of personal data. These platforms collect, monetize and control the use and reuse of much of the world's data. There is growing evidence that these firms collect every bit of data that they can about their customers (G. 2022; Slynchuk 2022; Freedman 2023; Knowledge at Wharton Staff 2019). The cybersecurity firm Avast notes that through its data collection, employees at Google know what languages you read and speak, your household budget, where you have been (if you use Google maps), and where you live and work (Sidell 2020).⁹ Many of these firms have been collecting data globally, which gives them a huge competitive advantage. Companies can use this data to create new products and services and derive value from data "far beyond initial purposes for which the data has been collected" (Jones Harbour and Koslov, quoted in Taylor et al. 2022).

During the pandemic, the largest digital firms became even larger, more valuable and more profitable (UNCTAD 2021). These firms have many sources of power: gatekeeper power (they are essential distribution channels), leveraging power (use the data they have to compete with the firms that depend on their infrastructure) and information exploitation (to manipulate users to

⁷ See www.gov.uk/government/publications/cdei-publishes-polling-data-on-data-sharing.

⁸ See www.oecd.org/sti/msti.htm.

⁹ Google allows users to delete their data and to turn off ad personalization.

buy more or change their behaviour, and so on). They also use the information they have to create new products or thwart potential rivals. Hoarding may be an essential part of their business model (Khan 2018, 9). According to UNCTAD, control over many data sets gave some countries with such firms both power and competitive advantages in digital technologies such as data analytics, AI, blockchain, the IoT, cloud computing and all internet-based services (UNCTAD 2021, 22, 29, 198). So control of data entails control of downstream sectors. These network effects are leading to a “winner-takes-all” scenario (Acemoglu et al. 2019). Such a winner-takes-all situation has major costs. As they extend control of data online, firms may learn to rely on rents from both their intellectual property (IP) and their monopoly control (Scassa 2021; Ciuriak 2018; Mazzucato, Entsminger and Kattel 2020; Gurumurthy and Chami 2022). Over time, their control over data stores may make it harder for others, whether individuals, firms, civil society groups, international organizations or national governments, to use that data to innovate (Kraemer and Shekhar 2022).

The opacity of data markets may also encourage firms to adopt or continue business practices that are harmful to society. First, these firms continue to use a business model that allows them to collect, use, control and reuse data from individuals but then group the data at scale. They then use complex algorithms to extract value from these troves of collective data. But in most nations, data is protected under rules framed as about privacy, which are premised on a relationship between data controllers and data subjects as individuals, not as members of a group. However, as Martin Tisné (2020, 3) notes, “The era of machine learning effectively renders individual denial of consent meaningless. Even if I refuse to use Facebook or Twitter or Amazon — the fact that everyone around me has joined means there are just as many datapoints about me to target.” As of this writing, most countries have not enacted rules governing collective rights to data, which may increase the reluctance of individuals to allow their data to be shared. Privacy laws may in fact impede data sharing. US and EU data protection/privacy regulations permit using data for profit (Kemp 2020). However, some jurisdictions, such as the United States, make it harder to use the same data for socially beneficial uses (for example, to protect public health). (Schmit et al. 2021).

Second, under international IP rules, large data-rich firms can control the results of the analysis and the reuse of the analyzed data sets (Scassa 2021). Yet even these firms do not really understand the amount or value of the data they collect and hoard. As an example, Facebook (Meta) has no idea where all of its user data goes, or what it is doing with it, according to a leaked internal document obtained by Motherboard, a tech publication. The document stated, “We do not have an adequate level of control and explainability over how our systems use data, and thus we can’t confidently make controlled policy changes or external commitments such as ‘we will not use X data for Y purpose.’ And yet, this is exactly what regulators expect us to do, increasing our risk of mistakes and misrepresentation” (Franceschi-Bicchierai 2022). Meta seems to be arguing that it cannot control its huge trove of data, let alone figure out how to protect it.

Third, corporate interests have extended legal ownership of intangible assets, in particular through copyright. They have restricted access to data using “technology protection measures,” which forbid the right to repair iPhones or John Deere tractors because in so doing you might learn the secret sauce, licensing terms and other “digital rights management” techniques to restrict access to information and control its use. As a result, much online content is treated as private rather than common property (Boyle 2003; Coyle 2022). Moreover, because much of the data is personal data, that data has now become part of corporate property protected under a form of IP rules called a trade secret. A trade secret must have a commercial value and, consequently, firms are allowed to keep such trade secrets “a secret.” Trade secrets are IPRs on confidential information that may be sold or licensed.¹⁰

In the data-driven economy, control over data begets ever more control. Once the company controls some data (for example, data is analyzed under a proprietary algorithm), it can control, mix and reuse the data it collected. Ever so gradually, such firms may control more and more of the world’s information, and can keep that information secret under trade secrets rules (Zech 2016, 6). Given the importance of data and information to human rights, innovation and social progress, firms should not be controlling

¹⁰ See www.wipo.int/tradesecrets/en/.

large swathes of the world's data. Individuals do not generally put their investments in one stock or one bank. Firms that control too much data could become a risk to the system; undermine information sharing and trust; and create social, moral or political concerns (Stucke 2018).

Yet individuals and groups can challenge the practices of the data behemoths. In the United States, some scholars have made the countervailing case for extension of the public domain to proprietary-held data. These scholars often cite a 1966 Supreme Court case that held that Congress may not authorize the issuance of patents whose effects are to remove existent knowledge from the public domain, or to restrict free access to materials already available (Boyle 2003). Scholars who take this position make the case that private companies, including the US-based data giants, benefited from taxpayer investment into the research that underpins the internet and other data-driven technologies. So controlling data under trade secrets is doubly unfair, both to the data subjects and to the taxpayers who funded the original research. It is unclear if these scholars are saying that data should be effectively anonymized, protected and made part of the public domain.

Interestingly, many of the private firms that rely on trade secrets to protect proprietary data and control user data also rely on open-source models.¹¹ Executives understand that although open-source software may be less secure, open-source code, data and algorithms can provide many benefits. For example, open-source software may increase AI adoption by reducing the level of mathematical and technical knowledge necessary to use AI. It advances science by making code available and easier to replicate. Moreover, computer and data scientists can get free feedback on their algorithms if they are open; they can use that feedback to screen potential talent (Engler 2021a).

Although the data giants benefit from openness, their executives still decide what IP (and the data underpinning it) should be open and what should not.¹² For example, Tesla agreed to make some of its patents open source and said it would not sue those who sought to build on them. Tesla executives

also asserted that opening the patents gave Tesla better algorithms and employees.¹³ The World Economic Forum described this process as moving from protectionism of IP to democratization of it (Ceulemans et al. 2021). Yet Tesla has been awarded a wide range of patents in AI and energy storage.¹⁴

Taken in sum, the individuals, firms, entities, government agencies and individuals that control data are the puppet masters. They decide how and when to use and reuse data; whether to store or destroy it; whether it can be shared; and whether it can be monetized and sold. Some scholars have decided that mandated data sharing is one way to level the playing field among the data giants and others who seek to use data (Graef and Prüfer 2021).

To change the behaviour of these giant firms, policy makers often focus on data controllers. For example, Europe's General Data Protection Regulation (GDPR) defines a data controller as a legal or natural person, an agency, a public authority or any other body who, alone or when joined with others, determines the purposes of any personal data and the means of processing it.¹⁵ The UK GDPR has a similar perspective. Controllers are "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data."¹⁶ Controllers must comply with the data protection principles listed in article 5 of the UK GDPR.¹⁷

International IP law also targets data controllers. The members of the World Trade Organization have agreed to adhere to the Trade-Related Aspects of Intellectual Property Rights Agreement, article 39(2), which delineates how and when entities may protect information as confidential. It states: "Natural and legal persons shall have the possibility of preventing information lawfully within their control from being disclosed to, acquired by, or used by others without

11 The famous case is the partnership between Apple and Microsoft, where Apple agreed to use Microsoft Internet Explorer browsers on Apple products (Dernbach 2008).

12 See Chen (2020) and Ceulemans et al. (2021).

13 See www.tesla.com/legal/additional-resources#patent-pledge.

14 See Fukuoka and Shiraishi (2021).

15 See www.gdpreu.org/the-regulation/key-concepts/data-controllers-and-processors/.

16 See <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/controllers-and-processors/what-are-controllers-and-processors/>.

17 See <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/controllers-and-processors/what-does-it-mean-if-you-are-a-controller/>.

their consent in a manner contrary to honest commercial practices.”¹⁸ As law professor Teresa Scassa (2018) notes, if the confidential nature of data is lost, the information may lose its value.

The ODI found that although private entities hoard much of their data, they also routinely share or sell data. The ODI examined 270 UK companies and found about 50 percent used data from non-government sources, including other companies (Shadbolt 2015). A 2019 EU consultation of nearly 1,000 small businesses found that one-third acquire data from other companies (European Commission 2019). However, the ODI (2020) found that many firms are reluctant to share or sell data because they see reputational and business risks. Business executives understand that innovation requires new ways of collaborating, including data sharing with ecosystem partners and third-party organizations.¹⁹ But thus far, policy makers have not tried to leverage control of data stores into broad obligations for data sharing for the public good. Instead, they have made a priority of keeping data sovereign.

What Do We Mean by Data Sovereignty and How Might It Impede Data Sharing?

In recent years, policy makers in countries from Australia to Switzerland have talked about the need to keep data sovereign. But data sovereignty is a vague term. Sometimes we think of sovereignty as the state’s regulatory power; at other times, we use sovereignty to describe the state’s ability to act in the digital sphere without being restricted by others (Christakis 2020). For the purposes of this paper, data sovereignty is used to describe policies requiring data to be

stored, processed and handled in its country of origin, in accordance with procedures to be determined by that country (ibid., 65–66).²⁰

The Chinese government was the first nation to adopt policies to promote data sovereignty. In 2010, it declared that although the internet was global, within China’s borders, it was under China’s jurisdiction.²¹ In 2015, Chinese President Xi Jinping explained that respecting cyber sovereignty meant respecting each country’s right to choose its own internet development path and public policies (BBC News 2015). The Chinese Communist Party sees control of the internet and the data underpinning it as essential to both stability and growth. According to *The Wall Street Journal*, President Xi allegedly commented during a private meeting that, “Whoever controls data will have the initiative” (Wei 2021).²²

Data sovereignty is at bottom a strategy for nations to hoard and control data within their borders (Chander and Sun 2021; Christakis 2020). Developing countries, in particular, are concerned that they will not be able to reap the benefits of data for development unless they control data locally. Hence, many developing countries remain unwilling to negotiate the free flow of data until they have figured out how to control the data they create within their borders (Aaronson and Struett 2020; UNCTAD 2021).

Nations espouse different reasons for adopting data sovereignty. Some state that it is a strategy to resist the exploitative and extractive practices of Western (and Chinese) technology giants (Pinto 2018; Couldry and Mejias 2018). Chander and Sun (2021, 16–18) note that in the European Union, it is a means of addressing US/Chinese market dominance, whereas in Russia, it is used to protect against foreign interference in governance and to

18 Agreement Establishing the World Trade Organization, Annex 1C: Agreement on Trade-Related Aspects of Intellectual Property Rights, 15 April 1994, 1869 UNTS 299 art 39(2) [entered into force 1 January 1995], online: <www.wto.org/english/docs_e/legal_e/27-trips_01_e.htm>.

19 See Daugherty, Carrel-Billiard and Blitz (2020) and Ceulemans et al. (2021).

20 Some use digital sovereignty in lieu of data sovereignty, but digital sovereignty generally describes a broader scope of regulations related to the digital economy versus a narrower range of regulations governing data. As an example, Anupam Chander and Haochen Sun (2021, 10) use the term “digital sovereignty” to mean the application of traditional state sovereignty over the online domain, including not only cross-border flow of data through uses of internet filtering technologies and data localization mandates, but also speech activities (for example, combating fake news) and access to technologies.

21 See www.china.org.cn/government/whitepaper/2010-06/08/content_20207978.htm. The document states: “The Internet is under the jurisdiction of Chinese sovereignty. The Internet sovereignty of China should be respected and protected.”

22 See also Bloomberg News (2021).

assert stronger control over local infrastructure. China, Russia and India, for example, cite the need to protect data as a reason to justify data sovereignty. But these same governments have not enacted clear rules governing the use of such data by the public sector. These states may be using data governance to shift power from firms to government. Officials in these nations seem to believe that by controlling large supplies of data, they can achieve economic advantage in the digital economy and will be better positioned to counter the market power of the giant platforms. In a study of national data strategies, six of the 10 countries were focused on achieving competitive advantage in data-driven services such as the cloud or AI, and seven planned to achieve economies of scale and scope in data (Aaronson 2022).

Almost every data strategy and AI strategy is built on this notion that large troves of data will yield prowess in data-driven services. Many countries require that their citizen's data can only be stored and analyzed by local firms (Imbrie, Kania and Laskai 2020; Aaronson 2022; Niklas and Dencik 2020). In addition, a recent research project by George Washington University's Digital Trade and Data Governance Hub found that 43 nations (of a sample of 68 nations plus the European Union) have an AI strategy, and toward that end, policy makers have adopted policies to encourage the collection and hoarding of data locally (Aaronson and Zable, forthcoming 2023). However, the supply of data is one of many indicators that collectively add up to AI prowess (Imbrie, Kania and Laskai 2020). Moreover, there are many AI applications that do not necessitate massive amounts of data, such as transfer learning and reinforcement learning (Chahal, Toner and Rahkovsky 2021). Finally, if nations or firms hoard data, they may reduce data generativity and the public benefits of data analysis at home and abroad.

Voluntary Data Sharing and the Role of Governments

Around the world, private firms, data analysts, non-governmental organizations (NGOs) and government are collaborating to use data to serve those beyond their borders (Alam 2019). For example, Amazon partnered with the Radiant Earth Foundation to help developing countries monitor their sustainable development goals. The two use and share open geospatial data and analytics to measure crops in Africa, and in so doing advance commercial agriculture operations and provide transparency to the agricultural supply chain market in developing countries (AWS *Public Sector Blog* 2019). Swarm, a company that provides satellite data, has partnered with several NGOs to monitor and thwart illegal poaching (O'Callaghan 2022). The GovLab, a New York think tank working on data governance, described how various governmental, NGO and corporate entities have provided and analyzed data that has helped Ukraine and its neighbours during the conflict (Davletov et al. 2022). The UN High Commissioner for Refugees (UNHCR) created a map to visualize migrations happening across Europe as a result of the war, which shows where refugees have crossed borders and where they and the government that accepts them (for example, Poland, Hungary and so on) may need assistance.²³ Satellogic, which collects and analyzes satellite imagery, partnered with the NGO Halifax International Security Forum "to provide the Ukrainian government with a Dedicated Satellite Constellation" (Satellogic 2022). Ukraine is using the satellites to provide actionable information for the Ukrainian government so it can rapidly mobilize resources and evacuate civilians (ibid.). Premise, a marketing platform that pays consumers to provide data on consumption in real time,²⁴ mapped access to food, water and fuel.²⁵ Clearview AI allowed Ukraine to use its data for free to

23 The UNHCR's map to visualize migrations happening across Europe as a result of the war shows where refugees have crossed borders. The data portal aims to monitor the Ukraine refugee situation and reflect recent movements of refugees in order to better aid humanitarian actions. See <https://data.unhcr.org/en/situations/ukraine>.

24 See www.premise.com/why-premise/.

25 See www.premise.com/blog/introducing-premises-open-build-ukraine-project/.

uncover Russian assailants, combat misinformation and identify refugees and the dead.²⁶

However, such data sharing is voluntary and limited.²⁷ To inspire more data sharing, some scholars have suggested that governments mandate it (Prüfer and Schottmüller 2021; Parker, Petropoulos and Van Alstyne 2021). However, using a game theoretic model to analyze the effect of data-sharing obligations on market outcomes, competition and welfare, one study found that while mandated data sharing increases the level of competition in the secondary market, it lowers the incumbent's incentive to innovate in the primary market (Kraemer and Shekhar 2022, 6, 27–29).

Some governments are experimenting with various approaches to encourage data sharing in the public good, although most focus on data sharing within national borders. The governments taking the most comprehensive approach think of data governance as an ecosystem that has many entities with different objectives, needs and strategies.

For example, the European Union is tackling data sharing through several different laws. First, the Digital Services Act (DSA),²⁸ which came into force in 2022, requires the largest internet platforms to open up their data to independent researchers with European Commission approval: “the DSA would affect companies with at least 10% of EU citizens as active users, which would likely include Facebook, YouTube, Twitter, TikTok, Amazon, and others” (Engler 2021b). Second, the Data Act allows public sector bodies to access and use data held by the private sector in exceptional circumstances, such as floods and wildfires, or to implement a legal mandate if data is not otherwise available (European Commission 2022a). The European Union is also creating shared data spaces for specific sectors in order to facilitate data pooling and sharing. The European Union aims to “overcome existing legal and technical barriers to data sharing and, as such, unleash the enormous potential of data-driven innovation,” and “thereby create the core tissue

of an interconnected and competitive European data economy” (European Commission 2022b).

In the United Kingdom, the Digital Economy Act of 2017 “introduced a framework for sharing personal data for defined purposes across specific parts of the public sector.”²⁹ It enables accredited researchers to gain access to de-identified data for research purposes. The act regulates data-sharing practices for the purposes of research using public data, but it does not govern data sharing in other contexts. Nor does it address the global public good nature of data.

The United Kingdom's Information Commissioner's Office has published and continually updated a Data Sharing Code of Practice since 2011.³⁰ The government plans to establish “data trusts” to facilitate the ethical sharing of data between organizations holding data and organizations developing AI.³¹ A UK advisory agency, the Centre for Data Ethics and Innovation, is responsible for developing these data trusts.³² The Office for Artificial Intelligence also partnered with the ODI between 2018 and 2019 to run three data trust pilots.³³ But here, too, the government is focused on national data sharing.

Canada's Bill C-11 (Digital Charter Implementation Act) proposes establishing public data trusts to allow for reusing de-identified data for “socially beneficial purposes.” But it has not passed Parliament, and it is unclear if socially beneficial purposes extend to non-Canadians.³⁴

Some policy makers are beginning to pay attention to the global public good nature of data. For example, the United Nations and several supporting states created the Digital Public Goods Alliance, a platform where governments can share digital public goods, engage talent and pool data sets.

26 Clearview AI scraped the web for pictures. It claims it has more than two billion images from the Russian social media service VKontakte at its disposal, out of a database of more than 10 billion photos in total (CNBC 2022).

27 See Ragnet et al. (2022).

28 See <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>.

29 See <https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/data-sharing-a-code-of-practice/data-sharing-across-the-public-sector-the-digital-economy-act-codes/>.

30 See <https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/data-sharing-a-code-of-practice/>.

31 See Department for Business, Energy & Industrial Strategy (2019).

32 See Select Committee on Artificial Intelligence (2018).

33 See ODI (2019).

34 See <https://parl.ca/DocumentViewer/en/43-2/bill/C-11/first-reading>. For the purpose of this section, *socially beneficial purpose* means a purpose related to health, the provision or improvement of public amenities or infrastructure, the protection of the environment or any other prescribed purpose.

This platform is open to all, and offers open-source software, open data, open AI models, open standards and open content that adhere to privacy and other applicable laws and best practices, do no harm by design and help attain the Sustainable Development Goals.³⁵ According to the alliance's website, anyone can take, adapt and use these digital public goods. In order for something to be recognized as a digital public good, solutions must demonstrate use of an approved open licence. Once a solution is recognized as a digital public good, it is placed and made discoverable on a public registry, which contains almost 90 digital public goods that any developing country can use.

The website notes that the alliance welcomes participants from private sector technology experts, think tanks, governments, philanthropic donors, international implementing organizations and the United Nations. It is governed by a board that includes representatives from the German Federal Ministry for Economic Cooperation and Development, the Government of Sierra Leone, the Norwegian Agency for Development Cooperation, iSPIRT, the United Nations Development Programme and the United Nations Children's Fund.³⁶ In addition, the United Nations plans a Global Digital Compact to "outline shared principles for an open, free and secure digital future for all," but it is unclear how it will achieve these shared principles.³⁷

Germany sees data as a global public good in its national data strategy. It aims to "ensure that we... can both [add] value...[and] improve the lives of everyone."³⁸ The data strategy examines the role of one kind of IP protection, trade secrets, in preventing data sharing, reducing competition and potentially favouring the creation of monopolies (ibid., 21). Firms can use trade secrets to protect their algorithms and then they obtain control of any data they analyze with such algorithms. Hence, Germany is arguing for greater amounts of data to be viewed as a digital public good that should be shared openly while protecting privacy.

Germany has made the digital public good objective part of its development strategy with

³⁵ See www.un.org/techenvoy/content/digital-public-goods.

³⁶ See <https://digitalpublicgoods.net/who-we-are/>.

³⁷ See www.un.org/techenvoy/global-digital-compact.

³⁸ See www.bundesregierung.de/breg-en/news/data-strategy-adopted-1845882.

its FAIR Forward – Artificial Intelligence for All program.³⁹ Germany is working with six partner countries — Ghana, Rwanda, Kenya, South Africa, Uganda and India — to share open, non-discriminatory and inclusive training data, models and open-source AI applications, as well as digital learning and training for the development and use of AI. It also advocates for value-based AI that is rooted in human rights, international norms such as accountability, transparency of decision making and privacy, and draws on European laws or proposals on AI and data.

What Could Nations Do to Collaborate in the Global Public Interest?

Data is multidimensional — it can simultaneously be utilized as both a commercial asset and a public good. But private entities are collecting, storing and monetizing ever greater amounts of data. These firms are the puppet masters for much of the world's data. They decide when and how to share it. Some of that data could be useful to researchers and policy makers as they attempt to address a wide range of global problems. However, many countries view the data created within their borders as a sovereign asset.

Meanwhile, the world is in the early stages of data governance — no nation knows how to do it effectively or comprehensively in a technically neutral manner (Aaronson, Struett and Zable 2021). Moreover, some governments, such as the United States, may be unable to effectively govern these firms because policy makers recognize that such firms are increasingly essential to the nation's economic stability and national security. The data giant firms are also often the same firms that provide internet infrastructures such as cloud computing to facilitate innovation (Wheeler 2021).

At the same time, global governance has not caught up with the challenge of governing data that is collected and exchanged across borders.

³⁹ See <https://oecd.ai/en/dashboards/policy-initiatives/http:%2F%2Faiopo.oecd.org%2F2021-data-policyInitiatives-26742>.

A growing number of trade agreements have binding provisions encouraging the free flow of data with exceptions for national security, public morals, public health and privacy. However, these agreements generally discuss data as a commercial asset. When they discuss public data, they simply say such public data should be provided in a machine-readable format so those who seek to use and mix that data can do so easily. These provisions essentially reiterate a message that public data is an asset for commercial use. Meanwhile, international institutions focus on voluntary approaches (such as the Digital Public Goods Alliance). As of December 2022, these approaches have not gained wide attention or use.

Some governments are requiring that firms be transparent about the data stores they hold and share some of their data with other firms and/or government. Transparency about data stores will be extremely helpful, but demanding that firms share data could impede firm innovation and competitiveness. Policy makers would need to make changes to corporate governance and IP rules at the national level, which will take time. Meanwhile, firms will acquire ever more data and could restrict access in the interest of competitiveness.

So what can we do? This author aims to promote cross-sector data sharing without stifling innovation. This strategy builds on three pillars: first, the ideas of other scholars who have long worked on data; second, society's failure to address wicked problems in a cooperative manner as discussed above; and third, global society's inability and lack of infrastructure to govern and share data among entities.

First, two scholars, Diane Coyle and Linette Taylor, have greatly influenced the author's thinking about data. Coyle (2022, 356) argues that data is social in nature because any data set is "a limited, encoded representation of reality, embedding biases and assumptions" but "ignoring information that cannot be codified." She means that we cannot rely on markets or command-and-control regulatory mechanisms to coordinate knowledge-intensive activities. She notes that algorithms cannot adjudicate moral or economic trade-offs and conflicts — only humans can do so with legitimacy. Thus, Coyle concludes we need new trust networks to mitigate among data-based systems and human judgment (*ibid.*). Meanwhile, Taylor argues that we are having the wrong debate about data held by

companies. She notes that because our personal data is analyzed in bulk, instead of arguing about individual consent, society should be debating how others can use our personal data in specific contexts for limited specific purposes. Hence, if we want firms to be less opaque about the data they control, we should incentivize data sharing by paying firms for that data — renting it and in so doing incentivizing both data protection and data sharing (Taylor 2016; Taylor et al. 2022).

This author is not the first to say that we need new institutions and policies to address the challenge of data. Ian Bremmer (2019) suggests that the world needs a world data organization to govern data among like-minded nations. But this idea could reduce the generativity of data because it is designed to directly benefit those living in like-minded nations. In another example, the author's CIGI colleagues Bob Fay and Chris Beall argue that "coordination is necessary so that we don't get a digital version of what we have seen with tax havens e.g. a flock to jurisdictions to avoid stronger regulation such as where privacy protection is limited, where governance is limited and so on" (Fay, quoted in Emanuele 2021). The Digital Stability Board would "be a multi-stakeholder forum with a remit to create global governance for big data, AI and the digital platforms, while allowing national variation to reflect different values and cultures" (*ibid.*). It would also coordinate the development of standards, regulations and policies; monitor developments and assess vulnerabilities arising from these technologies; align efforts with other multilaterals; and ensure that civil society and developing countries are represented in the discussions (Beall and Fay 2020).

These entities would be most helpful, but neither proposal aims to promote data sharing to solve wicked problems. However, greed and guilt can motivate people to take a leap of faith and create a new entity fit for this purpose. This author suggests the creation of a cloud-based agency called the Wicked Problems Agency. The agency would not govern data per se, although it could inspire policy makers to adopt more comprehensive data governance rules to facilitate sharing. The Wicked Problems Agency would encourage firms and other entities around the world to provide a list of potential data troves that could be utilized to better understand patterns of human, societal and climatic behaviour that cross borders. The agency would employ a wide variety of

scientists, including social scientists, computer and data scientists, and climate and public health researchers, to seek, collect and analyze data in the interest of mitigating complex wicked problems.

Meanwhile, data stewards at a wide range of entities that may hold relevant data would be incentivized to examine their data and see what troves might be helpful to share. Staff at the agency would examine these data sets and gauge their utility for creating new insights in problems that cross generations and borders. If the data set could be useful, the Wicked Problems Agency would rent that data from these firms, allowing them to maintain the scarcity of the data while also getting good publicity for sharing their data.

The Wicked Problems Agency cannot alter the fundamental economics of data, including market failures such as opacity or hoarding by private entities. Nor can it prod policy makers to stop viewing data only as a sovereign and commercial asset. But it could have important spillovers. By creating an additional market for some of the data, it might slowly reduce data nationalism and encourage global data collaboration. It could prod entities to hire people who can think globally and creatively about data use. Finally, with its work, the Wicked Problems Agency may prod a rethink of how data can be used to better meet the collective needs of the individuals and groups that provided much of that data.

Works Cited

- Aaronson, Susan Ariel. 2022. *A Future Built on Data: Data Strategies, Competitive Advantage and Trust*. CIGI Paper No. 266. Waterloo, ON: CIGI. www.cigionline.org/publications/a-future-built-on-data-data-strategies-competitive-advantage-and-trust/.
- Aaronson, Susan Ariel and Thomas Struett. 2020. *Data is Divisive: A History of Public Communications on E-commerce, 1998–2020*. CIGI Paper No. 247. Waterloo, ON: CIGI. www.cigionline.org/publications/data-divisive-history-public-communications-e-commerce-1998-2020/.
- Aaronson, Susan Ariel, Thomas Struett and Adam Zable. 2021. *DataGovHub Paradigm for a Comprehensive Approach to Data Governance*. Year 1 Report. George Washington University and the Fletcher School at Tufts. <https://sites.tufts.edu/digitalplanet/files/2021/11/DataGovHub-Paradigm-for-a-Comprehensive-Approach-to-Data-Governance-Y1.pdf>.
- Aaronson, Susan Ariel and Adam Zable. Forthcoming 2023. *For the People but Not by the People: Public Engagement in National AI Strategies*. CIGI Paper.
- Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian and Asuman Ozdaglar. 2019. "Too Much Data: Prices and Inefficiencies in Data Markets." NBER Working Paper 26296. September. <https://doi.org/10.3386/w26296>.
- Alam, Mahbulbul. 2019. "Data for public good." *Towards Data Science*, December 11. <https://towardsdatascience.com/data-for-public-good-1414cbc99335>.
- Auxier, Brooke and Lee Rainie. 2019. "Key takeaways on Americans' views about privacy, surveillance and data-sharing." Pew Research Center, November 15. www.pewresearch.org/fact-tank/2019/11/15/key-takeaways-on-americans-views-about-privacy-surveillance-and-data-sharing/.
- AWS Public Sector Blog. 2019. "Geo-Diverse Open Training Data as a Global Public Good." April 22. <https://aws.amazon.com/blogs/publicsector/geo-diverse-open-training-data-as-a-global-public-good/>.
- BBC News. 2015. "China internet: Xi Jinping calls for 'cyber sovereignty.'" BBC News, December 16. www.bbc.com/news/world-asia-china-35109453.
- Beall, Chris and Bob Fay. 2020. "In the Age of Connection, Disconnected Digital Governance Isn't Working." Opinion, Centre for International Governance Innovation, December 28. www.cigionline.org/articles/age-connection-disconnected-digital-governance-isnt-working/.

- Bennett Institute for Public Policy and ODI. 2020a. "The Value of Data: Summary Report." www.bennettinstitute.cam.ac.uk/wp-content/uploads/2020/12/Value_of_data_summary_report_26_Feb.pdf.
- . 2020b. "The Value of Data Policy Implications." www.bennettinstitute.cam.ac.uk/publications/value-data-policy-implications/.
- Bloom, Greg. 2020. "The Principles of Governing Open Source Commons." SustainOSS, June 11. <https://sustainoss.pubpub.org/pub/jqngsp5u/release/1>.
- Bloom, Greg, Angie Raymond, Willa Tavernier, Divya Siddarth, Gary Motz, Melanie Dulong de Rosnay and Anouk Ruhaak. 2021. "A Practical Framework for Applying Ostrom's Principles to Data Commons Governance." *Mozilla Blog*, June 11. <https://foundation.mozilla.org/en/blog/a-practical-framework-for-applying-ostroms-principles-to-data-commons-governance/>.
- Bloomberg News. 2021. "Xi's Next Target in Tech Crackdown Is China's Vast Reams of Data." Bloomberg, April 23. www.bloomberg.com/news/articles/2021-04-23/xi-s-next-target-in-tech-crackdown-is-china-s-vast-reams-of-data?leadSource=uverify%20wall.
- Boyle, James. 2003. "The Second Enclosure Movement and the Construction of the Public Domain." Duke Law School Public Law and Legal Theory Research Paper No. 53. December. <http://dx.doi.org/10.2139/ssrn.470983>.
- Bremmer, Ian. 2019. "Why we need a World Data Organization. Now." GZERO, November 29. www.gzeromedia.com/why-we-need-a-world-data-organization-now.
- Ceulemans, Hugo, Mathieu Galtier, Tinne Boeckx, Marion Oberhuber and Victor Dillard. 2021. "From competition to collaboration: How secure data sharing can enable innovation." World Economic Forum, June 27. www.weforum.org/agenda/2021/06/collaboration-data-sharing-enable-innovation/.
- Chahal, Husanjot, Helen Toner and Ilya Rahkovsky. 2021. "Small Data's Big AI Potential." CSET Issue Brief. September. Center for Security and Emerging Technology. <https://cset.georgetown.edu/publication/small-datas-big-ai-potential/>.
- Chander, Anupam and Haochen Sun. 2021. "Sovereignty 2.0." University of Hong Kong Faculty of Law Research Paper No. 2021/041. <http://dx.doi.org/10.2139/ssrn.3904949>.
- Chen, Pearl. 2020. "IBM's open source strategy champions AI trust and transparency." IBM, September 10. www.ibm.com/blogs/journey-to-ai/2020/09/ibms-open-source-strategy-champions-ai-trust-and-transparency/.
- Christakis, Theodore. 2020. "'European Digital Sovereignty': Successfully Navigating Between the 'Brussels Effect' and Europe's Quest for Strategic Autonomy." Multidisciplinary Institute on Artificial Intelligence/ Grenoble Alpes Data Institute. December. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3748098.
- Ciuriak, Dan. 2018. "The Economics of Data: Implications for the Data-driven Economy." In *Data Governance in the Digital Age*, 14–19. www.cigionline.org/publications/data-governance-digital-age/.
- CNBC. 2022. "Ukraine has started using Clearview AI's facial recognition during war." March 13. www.cnn.com/2022/03/13/ukraine-has-started-using-clearview-ais-facial-recognition-during-war.html.
- Confessore, Nicholas. 2018. "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far." *The New York Times*, April 5. www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.
- Couldry, Nick and Ulises A. Mejias. 2018. "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject." *Television and New Media* 20 (4): 336–49. <https://doi.org/10.1177/1527476418796632>.
- Coyle, Diane. 2020. "Common governance of data: appropriate models for collective and individual rights." *Ada Lovelace Institute Blog*, October 30. www.adalovelaceinstitute.org/blog/common-governance-of-data/.
- . 2022. "Socializing Data." *Daedalus* 151 (2): 348–59. https://doi.org/10.1162/daed_a_01921.
- Coyle, Diane, Stephanie Diepeveen, Julia Wdowin, Lawrence Kay and Jeni Tennison. 2020. *The Value of Data – Policy Implications*. Main Report. Bennett Institute for Public Policy and ODI. www.bennettinstitute.cam.ac.uk/publications/value-data-policy-implications.
- Daugherty, Paul, Marc Carrel-Billiard and Michael Blitz. 2020. *Technology Vision 2020: We, the Post-Digital People*. Accenture. www.accenture.com/sg-en/insights/technology/technology-trends-2020.

- Davletov, Behruz, Uma Kalkar, Salwa Mansuri, Marine Ragnet and Stefaan Verhulst. 2022. "Data for Peace and Humanitarian Response? The Case of the Ukraine-Russia War." *Data and Policy Blog*, August 22. <https://medium.com/data-policy/data-for-peace-and-humanitarian-response-the-case-of-the-ukraine-russia-war-a811f690bcc0>.
- Department for Business, Energy & Industrial Strategy. 2019. "AI Sector Deal." Policy Paper. May 21. www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal.
- Dernbach, Christoph. 2008. "MacWorld Boston 1997 – Steve Jobs returns – Bill Gates appears [sic] on-screen." *Mac History*, July 19. www.mac-history.net/apple-history-tv/2008-07-19/macworld-boston-1997-steve-jobs-returns-bill-gateappears-on-screen.
- Duch-Brown, Nestor, Bertin Martens and Frank Mueller-Langer. 2017. "The economics of ownership, access and trade in digital data." JRC Digital Economy Working Paper 2017-01. <https://joint-research-centre.ec.europa.eu/system/files/2017-03/jrc104756.pdf>.
- Emanuele, Marco. 2021. "Towards the Digital Stability Board for a digital Bretton Woods." *The Science of Where*, February 21. www.thescienceofwheremagazine.it/2021/02/01/towards-the-digital-stability-board-for-a-digital-bretton-woods/.
- Engler, Alex. 2021a. *How open-source software shapes AI policy*. Brookings Institution, August 10. www.brookings.edu/research/how-open-source-software-shapes-ai-policy/.
- . 2021b. "A Bold Transatlantic Plan to Open Corporate Databases." *Center for European Policy Analysis*, July 8. <https://cepa.org/a-bold-transatlantic-plan-to-open-corporate-databases/>.
- European Commission. 2019. "SME panel consultation – B2B Data Sharing." October 9. <https://ec.europa.eu/digital-single-market/en/news/sme-panel-consultation-b2b-data-sharing>.
- . 2022a. "Data Act: Commission proposes measures for a fair and innovative data economy." Press release, February 23. https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113.
- . 2022b. "Staff working document on data spaces." February 23. <https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces>.
- Franceschi-Bicchierai, Lorenzo. 2022. "Facebook Doesn't Know What It Does With Your Data, Or Where It Goes: Leaked Document." *Motherboard*, April 26. www.vice.com/en/article/akvmke/facebook-doesnt-know-what-it-does-with-your-data-or-where-it-goes.
- Freedman, Max. 2023. "How Businesses Are Collecting Data (And What They're Doing With It)." *Business News Daily*, February 21. www.businessnewsdaily.com/10625-businesses-collecting-data.html.
- Fukuoka, Kutaro and Takeshi Shiraishi. 2021. "Tesla's strength in patent numbers leaves rivals in dust." *Nikkei Asia*, August 8. <https://asia.nikkei.com/Business/Technology/Tesla-s-strength-in-patent-numbers-leaves-rivals-in-dust>.
- G., Edward. 2022. "Google collects almost 40 data points per user – most out of top tech giants." *Atlas VPN (blog)*, May 25. <https://atlasvpn.com/blog/google-collects-almost-40-data-points-per-user-most-out-of-top-tech-giants>.
- Graef, Inge and Jens Prüfer. 2021. "Governance of data sharing: A law & economics proposal." *Research Policy* 50 (9): 104330. <https://doi.org/10.1016/j.respol.2021.104330>.
- Gurumurthy, Anita and Nandini Chami. 2022. *Governing the Resource of Data: To What End and for Whom? Conceptual Building Blocks of a Semi-Commons Approach*. January. <https://datagovernance.org/report/governing-the-resource-of-data-to-what-end-and-for-whom-conceptual-building-blocks-of-a-semi-commons-approach>.
- Human Rights Watch. 2022. *Trapped in a Web: The Exploitation of Personal Data in Hungary's 2022 Elections*. December 1. www.hrw.org/report/2022/12/01/trapped-web/exploitation-personal-data-hungarys-2022-elections.
- Imbrie, Andrew, Elsa Kania and Lorand Laskai. 2020. "The Question of Comparative Advantage in Artificial Intelligence." CSET Policy Brief. January. *Center for Security and Emerging Technology*. <https://cset.georgetown.edu/publication/the-question-of-comparative-advantage-in-artificial-intelligence-enduring-strengths-and-emerging-challenges-for-the-united-states/>.
- Kemp, Katharine. 2020. "Concealed data practices and competition law: why privacy matters." *European Competition Journal* 16 (2–3): 628–72. <https://doi.org/10.1080/17441056.2020.1839228>.
- Khan, Lina M. 2018. "Sources of Tech Platform Power." *Georgetown Law Technology Review* 2 (2): 325–34. <https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Khan-pp-225-34.pdf>.
- Knowledge at Wharton Staff. 2019. "Your Data Is Shared and Sold... What's Being Done About It?" *Knowledge at Wharton*, October 28. <https://knowledge.wharton.upenn.edu/article/data-shared-sold-whats-done/>.

- Kraemer, Jan and Shiva Shekhar. 2022. "Regulating Algorithmic Learning in Digital Platform Ecosystems through Data Sharing and Data Siloing: Consequences for Innovation and Welfare." July 14. <http://dx.doi.org/10.2139/ssrn.4163171>.
- Massey, Joe. 2022. "Designing a framework for responsible data stewardship." ODI, August 9. <https://theodi.org/article/framework-for-responsible-data-stewardship/>.
- Mazzucato, Mariana, Josh Entsminger and Rainer Kattel. 2020. "Public Value and Platform Governance." UCL Institute for Innovation and Public Purpose Working Paper WP 2020-11. www.ucl.ac.uk/bartlett/public-purpose/wp2020-11.
- Mervis, Jeffrey. 2017. "Data check: U.S. government share of basic research funding falls below 50%." *Science*, March 9. www.science.org/content/article/data-check-us-government-share-basic-research-funding-falls-below-50.
- Miranda Montoya, María Consuelo, Jackeline Bravo Chamorro, Luz Marina Leegstra, Deyanira Duque Ortiz and Lauren Maxwell. 2022. "A blank check or a global public good? A qualitative study of how ethics review committee members in Colombia weigh the risks and benefits of broad consent for data and sample sharing during a pandemic." *PLOS Global Public Health* 2 (6): e0000364. <https://doi.org/10.1371/journal.pgph.0000364>.
- Mossberger, Karen, Seongkyung Cho and Pauline Hope Cheong. 2022. "The Public Good and Public Attitudes toward Data Sharing Through IoT." Paper prepared for presentation at the Telecommunications Policy Research Conference, American University, Washington, DC, September 16. <http://dx.doi.org/10.2139/ssrn.4183676>.
- Niklas, Jędrzej and Lina Dencik. 2020. "European Artificial Intelligence Policy: Mapping the Institutional Landscape." Working Paper. https://datajusticeproject.net/wp-content/uploads/sites/30/2020/07/WP_AI-Policy-in-Europe.pdf.
- O'Callaghan, Jonathan. 2022. "Swarms of Satellites Are Tracking Illegal Fishing and Logging." *Wired*, August 30. www.wired.com/story/satellite-constellations/.
- ODI. 2019. "Data trusts: lessons from three pilots." <https://docs.google.com/document/d/118RqyUAWP3WlyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#heading=h.3fngvdcfo2cs>.
- . 2020. "Sharing Data to Create Value in the Private Sector." https://docs.google.com/document/d/1UbUA1fNOFF66xWYOE03fvUQi2UwQJ6_iQFHorgElalw/edit#.
- OECD. 2015. *Data-Driven Innovation: Big Data for Growth and Well-Being*. Paris, France: OECD Publishing. <http://dx.doi.org/10.1787/9789264229358-en>.
- . 2019. *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*. November 26. www.oecd-ilibrary.org/science-and-technology/enhancing-access-to-and-sharing-of-data_276aaca8-en.
- Parker, Geoffrey, Georgios Petropoulos and Marshall Van Alstyne. 2021. "Platform mergers and antitrust." *Industrial and Corporate Change* 30 (5): 1307–36. <https://doi.org/10.1093/icc/dtab048>.
- Pinto, Renata Avila. 2018. "Digital Sovereignty or Digital Colonialism." *Sur: International Journal on Human Rights* 15 (27): 15–27. <https://sur.conectas.org/wp-content/uploads/2018/07/sur-27-ingles-renata-avila-pinto.pdf>.
- Porway, Jake. 2021. "Charting the 'Data for Good' Landscape." *Data.org*, July 14. <https://data.org/news/charting-the-data-for-good-landscape/>.
- . 2022. "A Taxonomy for AI/Data for Good." *Data.org*, February 7. <https://data.org/news/a-taxonomy-for-ai-data-for-good/>.
- Press Trust of India. 2020. "Data Nation's Asset, Must Be Secured: Ravi Shankar Prasad." *NDTV*, July 19. www.ndtv.com/india-news/data-nations-asset-must-be-secured-ravi-shankar-prasad-2265241.
- Prüfer, Jens and Christoph Schottmüller. 2021. "Competing with big data." *The Journal of Industrial Economics* 69 (4): 967–1008. <https://doi.org/10.1111/joie.12259>.
- Ragnet, Marine, Yuki Mitsuda, Andrew Zahuranec, Behruz Davletov and Stefaan Verhulst, with contributions from Susan Ariel Aaronson, Peng Hwa Ang, Silvana Fumega, Fabrizio Scrollini, Krishna Sood, Ivy Tang and Ales Versic. 2022. "State of Open Data Policy: Repository of Recent Developments." May. https://docs.google.com/document/d/1J_YJb_mPwFLXVSuJBK0pudIGKT9ARK0Gbw4ojW6qQo/edit#pli=1.
- Satello. 2022. "Satello and HFX Collaborate to Equip Ukraine with Satellite Tasking Capabilities." Press release, March 21. <https://satello.com/news/press-releases/satello-and-hfx-collaborate-to-equip-ukraine-with-satellite-tasking-capabilities/>.
- Scassa, Teresa. 2018. *Data Ownership*. CIGI Paper No. 187. Waterloo, ON: CIGI. www.cigionline.org/publications/data-ownership/.

- . 2021. "Rights in data, the public interest, and international trade law." In *Addressing impediments to Digital Trade*, edited by Ingo Borchert and L. Alan Winters, 195–216. <https://cepr.org/publications/books-and-reports/addressing-impediments-digital-trade>.
- Schleffer, Guy and Benjamin Miller. 2021. "The Political Effects of Social Media Platforms on Different Regime Types." *Texas National Security Review* (summer): 77–103. <http://dx.doi.org/10.26153/tsw/13987>.
- Schmit, Cason, Theodoros Giannouchos, Mahin Ramezani, Qi Zheng, Michael A. Morrissey and Hye-Chung Kum. 2021. "US Privacy Laws Go Against Public Preferences and Impede Public Health and Research: Survey Study." *Journal of Medical Internet Research* 23 (7). www.jmir.org/2021/7/e25266.
- Select Committee on Artificial Intelligence. 2018. *AI in the UK: ready, willing and able?* Report of Session 2017-19. House of Lords. <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>.
- Shadbolt, Nigel. 2015. "Open data means business." ODI, May 31. <https://theodi.org/article/open-data-means-business/>.
- Sidell, Elle Poole. 2020. "What Does Google Do With Your Data?" Avast, December 18. www.avast.com/c-how-google-uses-your-data.
- Simon, Matt. 2019. "What This CIA Veteran Learned Helping Facebook With Elections." *Wired*, July 24. www.wired.com/story/facebook-knows-more-about-you-than-cia/.
- Slynychuk, Andriy. 2022. "Big Brother Brands Report: Which Companies Access Our Personal Data the Most?" *Carlio Blog*, November 29. <https://clario.co/blog/which-company-uses-most-data>.
- Stucke, Maurice. 2018. "Here Are All the Reasons It's a Bad Idea to Let a Few Tech Companies Monopolize Our Data." *Harvard Business Review*, March 27. <https://hbr.org/2018/03/here-are-all-the-reasons-its-a-bad-idea-to-let-a-few-tech-companies-monopolize-our-data>.
- Taylor, Linnet. 2016. "The ethics of big data as a public good: which public? Whose good?" *Philosophical Transactions of the Royal Society A* 374 (2083). <https://doi.org/10.1098/rsta.2016.0126>.
- Taylor, Linnet, Hellen Mukiri-Smith, Tjaša Petročnik, Laura Savolainen and Aaron Martin. 2022. "(Re)making data markets: an exploration of the regulatory challenges." *Law, Innovation and Technology* 14 (2): 355–94. <https://doi.org/10.1080/17579961.2022.2113671>.
- Tisné, Martin. 2020. "The Data Delusion." Luminata, July 14. <https://luminategroup.com/storage/1023/The-Data-Delusion---July-2020.pdf>.
- UNCTAD. 2021. *Digital Economy Report 2021 — Cross-Border Data Flows and Development: For whom the data flow*. <https://unctad.org/webflyer/digital-economy-report-2021>.
- Verhulst, Stefaan. 2021 a. "Data Stewardship Re-Imagined — Capacities and Competencies." Medium, October 8. <https://medium.com/data-stewards-network/data-stewardship-re-imagined-capacities-and-competencies-d37a0ebaf0ee>.
- . 2021 b. "Re-imagining Data Stewardship: Capacities and Competencies." Presentation at UN World Data Forum, October 6. www.slideshare.net/StefaanVerhulst/reimaginig-data-stewardship-capacities-and-competencies.
- Verhulst, Stefaan G., Andrew Young and Andrew J. Zahuranec. 2021. "The '9Rs Framework': Establishing the Business Case for Data Collaboration and Re-Using Data in the Public Interest." Medium, October 26. <https://medium.com/data-stewards-network/the-9rs-framework-establishing-the-business-case-for-data-collaboration-26585455ccc0>.
- Villani, Cédric. 2018. *For a Meaningful Artificial Intelligence: Towards a French and European Strategy*. Mission assigned by the Prime Minister Édouard Philippe. March. www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf.
- Wei, Lingling. 2021. "China's New Power Play: More Control of Tech Companies' Troves of Data." *The Wall Street Journal*, June 12. www.wsj.com/articles/chinas-new-power-play-more-control-of-tech-companies-troves-of-data-11623470478.
- Wheeler, Tom. 2021. "A focused federal agency is necessary to oversee Big Tech." Brookings Blueprints for American Renewal & Prosperity. February 10. www.brookings.edu/research/a-focused-federal-agency-is-necessary-to-oversee-big-tech/.
- Zech, Herbert. 2016. "A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data." *Journal of Intellectual Property Law & Practice* 11 (16): 460–70. <https://doi.org/10.1093/jiplp/jpw049>.

**Centre for International
Governance Innovation**

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

 @cigionline

