The Data Sharing Economy and Open Governance
of Big Data as Public Good

Jung Wan LEE[1]

Abstract

Data-driven markets depend on access to data as a resource for products and services. Since the quality of information that can be drawn from data increases with the available amount and quality of the data, businesses involved in the data economy have a great interest in accessing data from other market players and sharing data with other stakeholders. Despite the growing need for access to data and evidence of the economic and social benefits, data access and sharing remains below its potential. Individuals, businesses, and governments often face barriers to data access, which may be compounded by the reluctance to share, including within and across sectors. To address these challenges, this paper focuses on finding possible solutions for a better data-sharing economy. This paper 1) Discusses opportunities and challenges of open data and the data-sharing economy, limitations of private sector data, and issues with open government data. 2) Introduces open government data initiatives and open governance networks initiatives. 3) Suggests possible solutions, including the governance and management, the legal and policy frameworks, and the technical standards for open data with proposing an open data governance model for the data-sharing economy.

Keywords: Big Data, Open Data, Public Good, Open Government, Open Data Governance, Data Sharing Economy

JEL Classification Code: A13, H41, H44, H49

1. Introduction

From cars to apartments to information, we are sharing just about everything these days. Known as the sharing economy, the concept of sharing access to goods and services has become a major disruptive force in a growing number of industries around the globe. The different facets of the sharing economy offer numerous opportunities for businesses, particularly those that can be distinguished by their creative ideas and their ability to easily connect buyers and senders of goods and services via digital platforms. At the beginning of the growth of this economy, the advanced digital technologies generated billions of bytes of data that constitute what we call "big data".

Data-driven markets depend on access to data as a resource for products and services. Since the quality of information that can be drawn from data increases with the available amount and quality of the data, businesses involved in the data economy have a great interest in accessing data from other market players. However, companies still appear to be reluctant to share their data.

The growth of the digital economy means that sharing data between various stakeholders, whether connected or not, has become crucial for driving value in terms of intra-organizational efficiency, inter-organizational standards and practices, and even solving problems and improving living standards for the wider public. But standing in the way of these goals is a number of roadblocks that prevent the free, unencumbered flow of data between parties. While implementing effective data analytics techniques is frequently discussed as the key challenge when extracting useful information from raw, unstructured data, there now appears a more immediate but arguably less highlighted obstacle in actually attempting to gain access to the data in the first place.

Data-driven innovation forms a key pillar in 21st-century sources of growth, in particular, during the COVID-19

[1]First Author and Corresponding Author. Professor, School of International Economics and Trade, Anhui University of Finance and Economics, China. ORCID ID: 0000-0002-8885-6385. [Postal Address: 962 Caoshan Road, Bengbu City, Anhui Province, 233030, China] Email: jungwan.lee@aufe.edu.cn

pandemic and post COVID-19 era. The confluence of several trends, including the increasing migration of socio-economic activities to the Internet and the decline in the cost of data collection, storage, and processing, is leading to the generation and use of huge volumes of data, - big data. These large data sets are becoming a core asset in the economy, fostering new industries, processes, and products and creating significant competitive advantages (OECD, 2015).

With the growing importance of data for digital transformation and data-driven innovation, access to and sharing of data has become critical. Today, data access and sharing, for instance, are needed to enhance public service delivery and to identify emerging governmental and societal needs. In science, data access and sharing provide a range of benefits for researchers by enabling open science. With the increasing use of artificial intelligence, access to and sharing of data will further increase in importance, even in traditionally less data-intensive fields such as manufacturing, agriculture, and construction. This development has been key for enabling the next production revolution (OECD, 2017).

Despite the growing need for access to data and evidence of the economic and social benefits, data access and sharing remains below its potential. Individuals, businesses, and governments often face barriers to data access, which may be compounded by the reluctance to share, including within organizations and across sectors. Stakeholders face several risks when sharing data, such as the risks of digital security and privacy breaches and the violation of other legitimate private interests such as commercial interests. To address these challenges, some governments and private sector actors have established a wide variety of initiatives, but efforts have been uneven across sectors and countries (OECD, 2021). Furthermore, there remains a need for data governance frameworks to be coherent across economic sectors, society, and countries, including across public sector organizations.

2. Opportunities and Challenges of Open Data and the Data Sharing Economy

In the face of the extreme uncertainty about economic prospects as the COVID-19 pandemic and the shutdown continue, policymakers need as much information as possible. In this midst of providing data and information in a timely manner, private sector data has some limitations compared to official statistics, in particular, not necessarily covering a representative sample, but there is a lot of it and it can show patterns of behavior changing very quickly. For example, the digital giants, Google, Apple, and Microsoft have all made some of their data available for the duration of the crisis. Google has published Community Mobility Reports tracking activity in different places, and the ONS Data Science Campus extracted that data and made it available in a useful format. Users can also use Google Trends to search

for various terms. Apple has made activity data from the use of Apple Maps available in a standard format. Microsoft Azure has created an open access database of COVID-19 articles for use for research based on techniques like natural language processing, while Microsoft Bing has also collated open data resources (Yokoyama, 2020).

Indeed, the crisis seems to have prompted the digital giants to throw their weight behind the importance of open data for the public good, in a very welcome new campaign. Big technology and finance companies as well as many public sector organizations hold a vast amount of data than official statisticians and researchers could use to paint a more accurate picture of today's economy and society, and inform policymakers better and faster. However, the use of data stockpiles by big digital companies to hinder competition for their market has been widely noted. There is a wider point, though, about the potential public benefit left untapped because so much valuable information sits in private silos (Coyle, 2020). The debate about the use of data for the public good was underway before the COVID-19 pandemic. One good thing that could emerge from the current crisis would be greater access to the wealth of valuable information that could help policymakers stimulate and shape the eventual economic recovery.

The Bennett Institute, working with the Open Data Institute, is starting a new project funded by the Nuffield Foundation to address questions on the value of data from a policy perspective (Coyle et al., 2020). Data consists of many different forms of information and meaning so we will be asking questions about how to value data of different types, based on its distinctive economic characteristics. This will affect the policy implications for data governance and regulation as follows:

First, the government has to make policy decisions that rest on the value of data to the economy as a whole. These include decisions to invest in maintaining datasets the government makes available as open data and decisions to regulate concerning data sharing or openness. The HM Treasury (2018) published a discussion paper pointing to the economic potential of data, but also the challenges around unlocking that potential. The "Joint Research Centre of the European Commission" noted the large array of policy questions but ended: "We conclude that there are no easy answers for regulators how to overcome market failures in data and information markets" (Duch-Brown et al., 2017). A greater understanding of the value of data would help identify where the benefits of greater investment in and sharing of data are worth the costs. Given the public good characteristics of data, it seems likely that there is considerable untapped potential value in enabling greater sharing and joining up of data sets. In this paper, "the value of the public good" refers to the economic concept of social welfare, the wellbeing of all society. The value arises from data when businesses create jobs or

become more productive, when governments deliver more effective public services, when our environment is clean and diverse, and when people live happier and healthier (Coyle, 2019).

Second, even where it is recognized that greater sharing of data brings benefits, such as in transport or public health, or for individuals, those making decisions about sharing data hold the need to understand what value they are giving away and what benefits they will receive. The question of how to distribute the benefits arising from data needs to answer the question of what value the provided data has, and how the benefits of data from multiple organizations and individuals can be fairly distributed. These are live policy questions. Implementing data trusts effectively will require an understanding of how to distribute value from users to contributors. The question of market power based on data aggregation is one of the considerations debated by the Furman Review. There is also considerable debate about mechanisms for paying people for personal data.

In this regard, organizations may have questions about how and why they should make data freely available, or open. Here are the most commonly encountered open data misconceptions, and responses to them. Microsoft (2021) tackles some of the most common misconceptions about open data as follows:

First, concerns about open data can run from possible privacy conflicts to a perception that the data itself is too valuable to make more freely available. For example, privacy rightly receives a lot of attention in the health care sector. After all, the details of a person's health are among the most personal of all personal data. But sharing data in this sector could also help accelerate the use of data-driven decisions in tackling disease and thus inform health policy in the future. This is where differential privacy tools can help. A new open source differential privacy platform, developed by "Microsoft and Harvard's Institute for Quantitative Social Science and the School of Engineering and Applied Sciences", was recently released as part of the OpenDP initiative, as a way to give developers globally the opportunity to leverage expert differential privacy implementations and join the community (Fleming, 2020).

Second, the parameters of acceptable use would be established within a data-sharing agreement. Anyone wanting to use other's data would then be bound by the agreement's terms and conditions, which could cover multiple scenarios, such as not using mapping data for oil and gas exploration (Open Knowledge Foundation, 2021). A parallel can also be drawn with the open source software movement, which challenged the notion that a programmer's code was valued and should be protected. Proponents of open source software demonstrated that when code is made freely available to others, new applications can be developed that help solve customers' problems in new ways (Fleming, 2020). Understanding the opportunity costs of being protective of

own data is part of assessing whether the open data route is a good fit for any given organization.

Third, data sharing will put too much power in the hands of a small number of very large tech businesses. If we open our data, we may fear it is going to be hoovered up by one of the very large tech companies currently dominating the data economy. But it could just as easily be argued that this imagined scenario is already too close to becoming reality, which is that a handful of large companies are the only ones who can offer effective data analysis services to the rest of the market (Fleming, 2021). If we are more open with our data and therefore increase the opportunity for people to use it, we will help generate more independent development techniques and technologies. That makes sure that the openness in the data leads to benefits for everyone as opposed to the few.

Despite the challenges of open data, international development and humanitarian organizations are increasingly calling for digital data to be treated as a public good because of its value in supplementing scarce national statistics and informing interventions, including in emergencies. In response to this claim, a responsible data movement has evolved to discuss guidelines and frameworks that will establish ethical principles for data sharing. However, this movement is not gaining traction with those who hold the highest-value data, particularly mobile network operators who are proving reluctant to make data collected in low- and middle-income countries accessible through intermediaries (Taylor, 2016). Taylor (2016) evaluated how the argument for "data as a public good" fits with the corporate reality of big data, exploring existing models for data sharing. Taylor (2016) drew on the idea of corporate data as an ecosystem involving often conflicting rights, duties, and claims, in comparison to the utilitarian claim that data's humanitarian value makes it imperative to share them.

The European Union has made the decision to move away from the traditional economic model, which thrived during the industrial revolution, to the circular economic model in 2015 (The European Data Portal, 2020). The circular economic model is based on the reuse of resources, the regeneration of natural systems, and reducing waste and pollution. In other words, the circular process of the natural living world is adopted in the economy to reduce the negative impact on the environment and the health of all living organisms. The model can be divided into the biological cycle and the technological cycle. Thereby, the circular economy is dependent on cultural change. For example, subscriptions to services or products are better adjusted to the circular economy than the traditional ownership of services or products. In this midst, the European Commission has adopted a New Circular Economy Action Plan, which is one of the main cornerstones of the European Green Deal as of 11th March 2020 (The European Data Portal, 2020). Open data can provide solutions to achieve the goals of

the new strategy by improving decision-making based on data insights. Building on the initial plans of 2015, the new strategy focuses on the design and production for a circular economy, with the aim to ensure that the resources used are kept in the EU economy for as long as possible. Insights provided through open data can improve the decision-making of the efficient use of resources as data can predict certain trends of future supply and demand. For example, According to Jelenic (2019), these open datasets create awareness and government accountability. This open data is used to improve strategic decision-making on the division of resources and efficient use of recycling plants. In addition, during the COVID-19 pandemic, the European Commission has been focusing on keeping the green lanes for waste open so that it can be shipped without delay and can become a resource for another industry (The European Data Portal, 2020).

In summary, the digital revolution is not only reshaping economic development, but also the modes of production and lifestyles at unprecedented speed and scale. Open data has become a new driver to create value. The government, as an important regulator, data resource administrator, service provider, and platform builder, is playing an increasingly crucial role in establishing a regional data ecosystem and promoting the progress of the digital economy.

3. Understanding Open Government Initiatives and Open Data Ecosystems

3.1. Open Government Data Initiatives

Magalhães Santos (2019) presented the results of the analysis of the open government initiative in the city of São Paulo, Brazil. The results indicated the need for the evolution of open government strategies to an ecosystem capable of integrating several parties in the development of public policies; for this, it is necessary to invest in collaborative data between society, private organizations, and the government sector. According to the 2018 report on "Development of Digital Economy in G20 Members" published by the China Academy of Information and Communications Technology, China's aggregate and growth rate in the digital economy stood at USD 4.02 trillion, a total of 16.44%, and ranked 2nd among the G20 members. In spite of a sound momentum in the digital economy, the Government of China sees a wide gap in the leading states regarding digital economy building. Compared with major developed economies, China has a comparatively slow progress on e-Government development, with performance on all the segments ranking lower than the advanced nations (PwC China, 2019). Through a comparative case study of open government data platforms in three Latin American cities, Bonina and Eaton (2020) proposed a theoretical model, which explains how an open government data platform owner is able to govern the demand and the supply side of its platform to facilitate the cultivation of a platform ecosystem.

Open government data initiatives are an emergent platform research topic. However, there is little understanding of how these platforms are governed for the innovation of services using open data, where the cultivation of an installed base of heterogeneous service innovators can lead to an increased usage of open government data. Open-source governance (also known as open politics) is a political philosophy that advocates the application of the philosophies of the open-source and open-content movements to democratic principles to enable any interested citizen to add to the creation of policy, as with a wiki document. Legislation is democratically opened to the general citizenry, employing their collective wisdom to benefit the decision-making process and improve democracy (Rushkoff, 2003). Theories on how to constrain, limit or enable this participation vary (Laffan, 2012; Millard, 2018; Roy, 2014). Accordingly, there is no one dominant theory of how to go about authoring legislation with this approach. There are a wide array of projects and movements which are working on building open source governance systems. Many left libertarians and radical centrist organizations around the globe have begun advocating open source governance and its related political ideas as a reformist alternative to current governance systems. Often, these groups have their origins in decentralized structures such as the Internet and place particular importance on the need for anonymity to protect an individual's right to free speech in democratic systems. Opinions vary, however, not least because the principles behind open source government are still very loosely defined.

3.2. Open Data Ecosystems

Today companies are collecting massive amounts of data to better understand their customers and to make better, more informed business decisions. Frequently, all of this data resides across dozens, sometimes thousands of different sources and in multiple formats, both structured and unstructured. Connecting all of this data and making sense of it is a massive and highly complicated task, but it is essential. To be successful, companies have to be able to connect the dots across varied data sources and data types. Only then they can realize insights and take meaningful action. Over the past decade, a series of technologies have come on the scene promising to solve this problem. Led by the Hadoop movement, it first began in the mid-2000s when products and companies started sprouting up, creating an open data ecosystem. This movement towards composable technologies that integrate with APIs and run on commoditized hardware challenged the status quo of monolithic, interdependent architecture in a big way.

By adopting an open, distributed approach to commoditized hardware, these companies challenged the traditional setup of storing and processing data in proprietary, centralized data warehouses (Wang, 2021). But ultimately, these solutions underachieved their grandiose promise because they became unwieldy, difficult to manage, and economically unscalable.

For decades, companies relied on traditional databases or warehouses, a mostly proprietary, centralized repository where structured data was stored and processed. The traditional data warehouse system required buying pricey on-premises hardware, maintaining structured data in proprietary formats, and relying on a centralized data and IT department to deliver analysis. In the midst, Google wired an ever-expanding number of computers together into a fleet. Eventually, this computing infrastructure grew so big that hardware failures became inevitable, and each programmer had to figure out how to handle them individually. To address these challenges, MapReduce, a programming model and an associated implementation for processing and generating big data sets with a parallel, distributed algorithm on a cluster, was born. It is widely agreed that the first generation of the open data ecosystem is Apache Hadoop. The Hadoop, which evolved into a collection of open source projects, served a significant role in bringing open source from academia to the mainstream (Wang, 2021). The open source framework allowed anyone to process massive datasets distributed across computer clusters, making it a hugely attractive option for enterprises, which were collecting more data by the day.

More recently, we have witnessed the revival of the open data ecosystem. With the rise of the cloud, a proliferation of open source data formats, and the arrival of vendors solving for earlier pain points, we have seen a new breed of open data ecosystem companies emerge and grow in popularity. These new solutions are able to capture the full scope of data that resides within a company, enabling teams to leverage the data to its full advantage. Thanks to services like Amazon S3, Azure Data Lake Storage, and Google Cloud Storage, companies can house structured and unstructured data at scale in cloud-native data lakes. This eliminates the need for expensive, monolithic hardware and enables organizations to scale data volume without the associated management overhead. In addition, the storage costs in the cloud continue to drop. As a result, these storage services have become the default landing zones in the cloud, and are often considered the systems of record. Cloud, with its scale and diversity, inherently encourages disaggregation into best of breed, nimble services. Cloud data warehouses, such as Snowflake, AWS Redshift, and Google BigQuery, while not inherently open, have also tremendously helped bring data to the cloud (Wang, 2021). Going forward, we believe companies will turn to both cloud data warehouses and cloud data lakes to serve different needs and derive value from their data for a long time to come. This time, the open data ecosystem is not going anywhere.

4. Understanding Open Governance Initiatives

4.1. Open Governance Networks

The Linux Foundation has long served as the home for many of the world's most important open source software projects. They act as the vendor-neutral steward of the collaborative processes that developers engage in to create high-quality and trustworthy code. They also work to build the developer and commercial communities around that code to sponsor each project's members. Over the last few years, they have also added a series of projects focused on lightweight open standards efforts - recognizing the critical complementary role that standards play in building the open technology landscape. With Linux's open source software projects, commercial participants' involvement has been critical to driving adoption and sustainability. On the horizon, they envision another category of collaboration, which they are calling "open governance networks" (Behlendorf, 2020).

To address this need, today, the Linux Foundation is adding open governance networks to the types of projects they host. These projects will operate very similarly to the Linux Foundation's open source software projects, but with some additional key functions. These projects will be sustained through membership dues set by the governing board on each project, which will be kept to what's needed for self-sufficiency. Some may also choose to establish transaction fees to compensate operators of peers if usage patterns suggest that would be beneficial. To ensure that these efforts live up to the word "open" and the Linux Foundation's pedigree, the vast majority of technical activity on these projects, and development of all required code and configurations to run the software that is core to the network will be done publicly (Dolan & Nicholas, 2021).

Across these open governance network projects, there will be a shared operational, project management, marketing, and other logistical support provided by Linux Foundation personnel who will be well versed in the platform issues and the unique legal and operational issues that arise. For those organizations acting as governing bodies on such networks today, Linux helps them evolve those projects to reach an even wider audience while taking off their hands the low margin, often politically challenging, grunt work of managing such networks. The Linux Foundation's open governance networks will enable our communities to establish their own open governance network and have an entity to process agreements and collect transaction fees (Dolan & Nicholas, 2021).

Open webOS (formerly HP webOS) was introduced in August 2011. HP announced that they would discontinue operations for webOS devices, specifically the TouchPad tablets and Smartphones. However, HP announced that they will continue to explore options to optimize the value of webOS software going forward. HP stated at the time that they would continue to support Open webOS from a development, engineering, and resources perspective. The Beta release of Open WebOS was announced in August 2012 and comprises 54 webOS components and over 450,000 lines of code (Webinos, 2012). Open webOS includes several projects: Enyo (a JavaScript framework), WebKit/Isis, the Linux Standard Kernel, and the webOS System Manager. HP advised that the open source project will be supported by HP with regard to development effort and resources and that they intend to manage the project openly and transparently.

4.2. The Open Governance of the Internet

Who owns and governs the Internet? In actual terms, no one owns the Internet, and no single person or organization controls the Internet in its entirety. More of a concept than an actual tangible entity, the Internet relies on a physical infrastructure that connects networks to other networks. In theory, the Internet is owned by everyone who uses it. For example, the Internet Corporation for Assigned Names and Numbers (ICANN) that maintains the central repository for IP addresses and helps coordinate the supply of IP addresses. It also manages the domain name system (DNS) and root servers. ICANN currently manages over 180 million domain names and four billion network addresses across 240 countries. With ICANN, DNS became something more than vendor-neutral, international, and accountable to the Internet community. It evolved to develop and manage the root of the domain name system, independent from any company or nation. ICANN's control over the DNS comes primarily through its establishment of an operating agreement among domain name registrars that establishes rules for registrations, guarantees someone's domain names are portable, and a uniform dispute resolution protocol for times when a domain name conflicts with an established trademark or causes other issues (Behlendorf, 2020).

There are similar organizations that interface with open standards and software but perform governance functions. A prime example of this is the 'CA Browser Forum', which manages the root certificates for the SSL/TLS Web security infrastructure. While some cryptocurrency networks claim that there is no need for formal human governance, it is clear that there are governance roles performed by individuals and organizations within those communities. In addition to a good portion of governance being automatable via smart contracts, repairing damage caused by their exploitation, promoting the platform's adoption to new users, onboarding new organizations, and managing hard fork upgrades all still require humans (Behlendorf, 2020).

The 'World Wide Web Foundation' was established in 2009 by Web inventors Sir Tim Berners-Lee and Rosemary Leith to advance the open Web as a public good and a basic right. They are an independent, international organization fighting for digital equality - a world where everyone can access the Web and use it to improve their lives. To deliver digital equality, they aim to change government and business policies for the better. They believe that everyone has the right to access the Internet and use it freely and fully. In recent years, they have influenced policies in over a dozen countries, helping to unlock the benefits of the Web for hundreds of millions of people (Bassi, 2014).

5. An Open Data Governance Model for the Data Sharing Economy

The presence of open data often gets listed as an essential requirement toward "open governance". For instance, an open data strategy is reviewed as a key component of many action plans submitted to the "Open Government Partnership". Yet little time is spent on assessing how open data itself is governed, or how it embraces open governance. For example, not much is known on whether the principles and practices that guide the opening up of government such as transparency, accountability, user-centrism, and demand-driven design thinking, will also guide decision-making on how to release open data.

In the midst, data governance has become more complex and open data decision-makers face heightened concerns with regards to privacy and data protection. The recent implementation of the EU's General Data Protection Regulation has generated an increased awareness worldwide of the need to prevent and mitigate the risks of personal data disclosures, and that has also affected the open data community. Before opening up data, concerns of data breaches, the abuse of personal information, and the potential of malicious inference from publicly available data may have to be taken into account. In turn, questions of how to sustain existing open data programs, user-centrism, and publishing with purpose gain prominence.

To better understand the practices and challenges of open data governance, the audience attending the "5th International Open Data Conference (IODC) in Buenos Aires" deemed governance of open data to be the most important discussion topic. For instance, discussions around the open data charter principles during and prior to the IODC acknowledged the role of an integrated governance approach to data handling, sharing, and publication. Some concluded that the open data movement has brought about better governance, skills,

and technologies of public information management which becomes an enormous long-term value for the government (Eaves & McGuire, 2019).

But what does open data governance look like? The Open Data Research Network (ODRN, 2018) convened a workshop at the 3rd Open Data Research Symposium in Buenos Aires, bringing together open data professionals, civil servants, and researchers. In this symposium, open data governance was defined as the "interplay of rules, standards, tools, principles, processes, and decisions that influence what government data is opened up, how and by whom."

In following up the domain, we suggest possible solutions to start mapping the pillars and layers of open data governance and illustrate the active participants in the decision-making process to build the data-sharing economy as follows (Figure 1).

5.1. The Governance and Management

Governments may decide about the release of data on various levels. An examination of the management side of data governance could provide insight into decision-making methods and devices. For instance, one might analyze how governments gauge public interest in their datasets through data request mechanisms, user research, or participatory workshops. One might also study cost-benefit calculations and similar methodologies to evaluate data, and how they

inform governments what data counts as crucial and is expected to bring returns and societal benefits.

Understanding open data governance would also require studying the ways in which open data creation, cleaning, and publication are managed. Governments may choose to organize open data publication and maintenance in-house or seek collaborative approaches. In addition, open data governance may require capacities in government, but could also contribute new capacities. This can apply to staff as well as to resources such as time or infrastructure. How do governments provide and draw capacity from open governance approaches, and what could be learned for other open data governance approaches? However, there are still many challenges to find better solutions. For example, to what extent is management guided by open governance? How are government decision-making processes and devices participatory, transparent, and accountable? How do governments currently make space for more open governance in their management processes? How do governments design routine processes to open up data requests?

5.2. The Legal and Policy Frameworks

Open data policies operate among other legal and policy frameworks, which can complement, enable, or limit the scope of open data. New frameworks such as the "General Data Protection Regulation", and the existing right to information and freedom of expression frameworks prompt
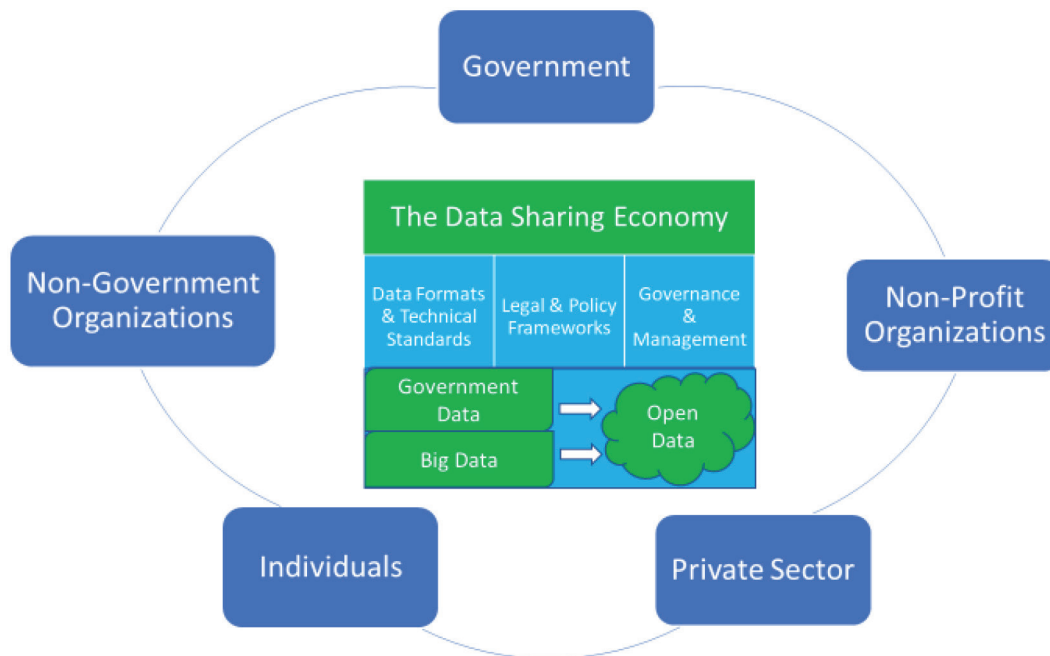

Figure 1: Open Data Governance Model for the Data Sharing Economy

the question of how the legal environment influences the behavior and daily decision-making around open data. To address such questions, one could study the discourse and interplay between open data policies as well as tangential policies like a smart city or digitalization policies (Brandusescu et al., 2019).

Furthermore, we still need to find solutions for: how are open data frameworks designed to guide the implementation of open data? How do they address governmental devolution? Open data governance needs to stretch across all government levels to unlock data from all government levels. What agencies define and move forward open data, and how does this influence adoption and sustainability of open data initiatives? Besides, studying the interaction of privacy protection, right to information, and open data policies, how could open data benefit from policies enabling open governance and civic participation? Do governments develop more integrated strategies for open governance and open data, and if so, how do these laws and policies enable other aspects of open data governance, including more participatory management, more substantive and legally supported citizen participation?

5.3. The Technical Standards

Richter and Slowinski (2019) introduced data sharing platforms, which are emerging as new intermediaries and can play a vital role in the data economy, as they may increase willingness to share data. Sedkaoui and Khelfaoui (2020) emphasized the facilitating role of big data analytics, explaining why and how data analysis algorithms can be integrated operationally, to extract value and to improve the practices of the sharing economy. However, governments may have different technical standards in place for data processing and publication, from producing data to quality assurance processes. Some research has looked into the ways data standards for open data alter the way governments process information (Goëta & Davies, 2016). Others have argued that the development of data standards is a reference to how governments envisage citizens, primarily catering to tech-literate audiences (Brandusescu et al., 2019). Data standards do not only represent, but intervene in the way governments work. Therefore, they could substantially alter the ways government publishes information.

On the other hand, most government data systems are not designed for open data. Too often, governments are struggling to transform huge volumes of government data into open data using manual methods. Legacy IT systems that have not been built to support open data create additional challenges to developing technical infrastructure, but there is no single global solution to data infrastructure. How could then governments transform their technical infrastructure to allow them to publish open data efficiently? If standards

can be understood as bridge-building devices, or tools for cooperation, how could open governance inform the creation of technical standards?

6. Conclusion

Open data is information that is available to the public to use, no matter the intended purpose. Big data, on the other hand, big data is data that contains greater variety, arriving in increasing volumes and with more velocity. Put simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just cannot manage them. Big data that is open can be useful to businesses, regardless of size or industry. Apart from in-house data such as sales information and website traffic, there are a number of sources that can be used for a variety of business purposes. With this open data platform, we can look at things like profit and loss accounts, balance sheets and cash flow, employee details, technology insights, group structure, and more. Perhaps one of the most obvious is social media; trending topics on Facebook and Twitter can offer great insight into potential opportunities to market a timely product or service; Google Trends can also work well for this. Public data sources such as weather info offered by the Met Office can be used for marketing purposes as well as sales predictions for weather-related products like sun cream, etc.

In the midst of building the data-sharing economy, perhaps the biggest obstacle that entities face in this regard is the inaccessibility to data in the form of data "silos"—that is, the closed walls that exist around data that ultimately makes it a costly, resource-intensive process to obtain data. Such insular systems are typically unable to operate with other systems and thus prevent important data from being shared. This creates gross inefficiencies across organizations, as management is prevented from having access to the data of all the business divisions that it needs for comparative analysis. Silos are nothing more than the barriers that exist between departments within an organization, causing people who are supposed to be on the same team to work against one another. As such solutions demonstrate, data sharing problems cannot be solved by a single stakeholder alone. It will require community-wide and industry-wide solutions and collaboration among governments to encourage widespread sharing across organizational and geographical boundaries.

The suggested model in this paper shows the relationship between the two forms of data and various stakeholders associated with the data-sharing economy. With the model of open data governance aspects as a starting point, a growing body of ethnographic research suggests that tech innovations such as algorithmic decision-making, open data, or smart city initiatives are 'multiples' — meaning

that they can be practiced in many ways by different people, arising in various contexts. With that understood, this model may elicit how open data governance becomes salient for governments, under what circumstances, and why. Ultimately, working together will make the biggest difference in addressing individual stakeholder needs and preferences and ensure that ethics, privacy, and regulatory concerns are suitably addressed.

In conclusion, there is no doubt that data is the future. While the amount of big data continues to grow considerably as the Internet of Things expands, it is open data that offers real power to businesses; even more so when the two are combined. When we are able to firmly define "big data as a public good", and also "government data as a public good", and eventually "data as the public good", by having free access to large datasets, companies can gain insights that can have a great impact on their marketing, sales, and customer service strategies, as well as overall business development. Governments could provide better services for the wellbeing of society, including creating jobs, becoming more productive, delivering more effective public services, and providing an environment where people live happier and healthier.

References

Bassi, B. (2014). *Who governs the Internet?* https://www.commonplaces.com/blog/who-governs-the-internet/

Behlendorf, B. (2020). *Introducing the open governance network model*. The Linux Foundation. https://www.linuxfoundation.org/blog/introducing-the-open-governance-network-model/

Bonina, C., & Eaton, B. (2020). Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City, and Montevideo. *Government Information Quarterly, 37*(3), 101479. https://doi.org/10.1016/j.giq.2020.101479

Brandusescu, A., Iglesias, C., Lämmerhirt, D., & Verhulst, S. (2019). *Open data governance and open governance: Interplay or disconnect? The World Wide Web Foundation*. The Web Foundation News. https://webfoundation.org/2019/02/open-data-governance-and-open-governance-interplay-or-disconnect/

Coyle, D. (2019). *The value of data*. https://www.bennettinstitute.cam.ac.uk/blog/value-data/

Coyle, D. (2020). *Big data for the public good*. https://www.bennettinstitute.cam.ac.uk/blog/big-data-public-good/

Coyle, D., Diepeveen, S., Wdowin, J., Kay, L., & Tennison, J. (2020). *The value of data: Policy implications*. Cambridge, UK: The Bennett Institute for Public Policy. Department of Politics and International Studies, The University of Cambridge.

Dolan, M., & Nicholas, S. (2021). *Understanding open governance networks*. The Linux Foundation. https://www.linuxfoundation.org/en/blog/understanding-open-governance-networks/

Duch-Brown, N., Martens, B., & Mueller-Langer, F. (2017). *The economics of ownership, access, and trade in digital data* (JRC Digital Economy Working Paper 2017-01). Seville, Spain: European Commission, Joint Research Centre. https://ec.europa.eu/jrc/sites/default/files/jrc104756.pdf

Eaves, D., & McGuire, B. (2019). *The first decade of open data has been a win - but not for the reasons you think*. Digital Government Newsfeed. https://apolitical.co/solution-articles/en/the-first-decade-of-open-data-has-been-a-win-but-not-for-the-reasons-you-think

Fleming, S. (2020). *Open data: Your questions answered.* Microsoft News. https://news.microsoft.com/on-the-issues/2020/07/06/open-data-campaign-privacy-misconceptions/

Fleming, S. (2021). *How open data is driving change.* Microsoft News. https://news.microsoft.com/on-the-issues/2021/04/23/open-data-melting-glaciers-covid-19-ai/

Goëta, S., & Davies, T. (2016). The daily shaping of state transparency: Standards, machine-readability and the configuration of open government data policies. *Science & Technology Studies, 29*(4), 10–30. https://doi.org/10.23987/sts.60221

HM Treasury. (2018). *The economic value of data* (Discussion Paper No, 39107). London, UK: The Government of the United Kingdom. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/731349/20180730_HMT_Discussion_Paper_-_The_Economic_Value_of_Data.pdf

Jelenic, M. C. (2019). *From theory to practice: Open government data, accountability, and service delivery* (Policy Research Working Paper 8873). Washington, DC: The World Bank. https://documents1.worldbank.org/curated/en/410191559670657041/pdf/From-Theory-to-Practice-Open-Government-Data-Accountability-and-Service-Delivery.pdf

Laffan, L. (2012). *A new way of measuring openness: The open governance index.* Technology Innovation Management Review. https://www.timreview.ca/sites/de fault/files/article_PDF/Laffan_TIMReview_January2012.pdf

Magalhães Santos, L. G. (2019). Toward the open government ecosystem: Connecting e-participation models and open government to analyze public policies. In R. M. Bolívar, K. Bwalya, & C. Reddick (Eds.), *Governance models for creating public value in open data initiatives. Public Administration and Information Technology* (vol. 31, pp. 144–196). Cham: Springer. https://doi.org/10.1007/978-3-030-14446-3_4

Microsoft. (2021). *Open data: Progress and unrealized promise.* https://news.microsoft.com/opendata/

Millard, J. (2018). Open governance systems: Doing more with more. *Government Information Quarterly, 35*(4), S77–S87. https://doi.org/10.1016/j.giq.2015.08.003

Open Data Research Network (ODRN). (2018). *The social dynamics of open data*. http://www.opendataresearch.org

Open Knowledge Foundation. (2021). *Open data commons public domain dedication and license (PDDL) v1.0.* https://opendatacommons.org/licenses/pddl/1-0/

Organization for Economic Co-operation and Development (OECD). (2021). Data governance: Enhancing access to and sharing of data. https://www.oecd.org/sti/ieconomy/enhanced-data-access.htm

Organization for Economic Co-operation and Development (OECD). (2017). The next production revolution: Implications for governments and business. https://www.oecd.org/sti/the-next-production-revolution-9789264271036-en.htm

Organization for Economic Co-operation and Development (OECD). (2015). Data-driven innovation: Big data for growth and well-being. https://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm

PwC China. (2019). *Building a transparent digital government, creating a diversified data ecosystem, and developing a robust digital economy*. https://www.pwccn.com/en/research-and-insights/digital-government-diversified-data-ecosystem-digital-economy.pdf

Richter, H., & Slowinski, P. R. (2019). The data-sharing economy: On the emergence of new intermediaries. *IIC - International Review of Intellectual Property and Competition Law, 50*, 4–29. https://doi.org/10.1007/s40319-018-00777-7

Roy, J. (2014). Open data and open governance in Canada: A critical examination of new opportunities and old tensions. *Future Internet, 6,* 414-432. https://doi.org/10.3390/fi6030414

Rushkoff, D. (2003). *Open-source democracy: How online communication is changing offline politics.* London, UK: Demos.

Sedkaoui, S., & Khelfaoui, M. (2020). Sharing economy and big data analytics. Hoboken, NJ: Wiley.

Taylor, L. (2016). The ethics of big data as a public good: Which public? Whose good? *Philosophical Transactions of the Royal Society A, 374*, 126. http:/doi.org/10.1098/rsta.2016.0126

The European Data Portal. (2020). *Open data and the circular economy*. https://data.europa.eu/en/highlights/open-data-and-circular-economy

The International Open Data Conference. (2018). *The future is open*. https://opendatacon.org/

Wang, C. (2021). *What is the open data ecosystem and why it's here to stay*? Sapphire Ventures. https://sapphireventures.com/blog/what-is-the-open-data-ecosystem-and-why-its-here-to-stay/

Webinos. (2012). *Open governance* (Webinos White Paper FP7-ICT-2009-5 257103). EU: Webinos. https://www.w3.org/WoT/IG/wiki/images/a/a5/Webinos-whitepaper-Open-Governance.pdf

Yokoyama, J. (2020). *Closing the data divide: The need for open data*. Microsoft. https://blogs.microsoft.com/on-the-issues/2020/04/21/open-data-campaign-divide/