Data-driven Economies: Foundations for Our Common Future

WHITE PAPER

APRIL 2021

Cover: Thulfiqar Ali, Unsplash – **Inside:** Getty Images, Unsplash

Contents

Foreword

Creating a better future together

Jane Allen
Principal, PwC, USA

Kimberly Bella
Project Fellow, Data for Common Purpose Initiative, Data Policy Platform, World Economic Forum LLC; Seconded from Visa

Nadia Hewett
Project Lead, Data for Common Purpose Initiative, Data Policy and Blockchain, World Economic Forum LLC

Cathy Mulligan
Visiting Researcher, Imperial College London, United Kingdom

The COVID-19 pandemic has highlighted the inadequacies of the world's collective approach to data. The inability, and sometimes unwillingness, to share and use data to combat COVID-19 or to protect against predatory uses of data has negatively impacted individuals, private enterprise, citizens, research institutions and governments around the globe. A lack of trust combined with asymmetric economic interests are slowing progress. The importance of finding solutions to improve outcomes in times of crisis is undeniable, but enormous opportunities also exist across a myriad of ordinary use cases and for normal day-to-day outcomes.

Recent advances in technology make our ability to deliver improved outcomes and innovations entirely plausible today. Without proper protocols and governance, however, society risks creating either a world in which access to data is overly restricted and impedes human progress and innovation, or one in which data-sharing solutions are created without properly respecting the rights of the individual parties involved, including businesses. Several solutions, combined with emerging technologies, sometimes called data marketplaces or data exchanges, are starting to emerge. These data marketplaces allow data to be leveraged for broader sets of social outcomes than is currently possible with today's data platforms.

The World Economic Forum launched the Data for Common Purpose Initiative (DCPI) in 2020 to collaboratively solve these challenges. It builds on the work undertaken at the Centre for the Fourth Industrial Revolution over the past two years. This global multistakeholder project community of over 45 organizations in 20 countries is focused on:

- Surfacing opportunities for unlocking data for common purposes
- Creating and promulgating governance models that equitably apportion risks and rewards
- Articulating parameters for responsible, fair and ethical use of data.

The DCPI is exploring government-led data ecosystems as one instance where it could serve as a means to exchange data assets for the common good and to promote the transition to a data-driven economy. This also enables a foundation for evolving discussions on the continued journey to source more trust, plus transparency for economic benefits. Colombia,[1] Japan,[2] India[3] and other countries are actively participating in the DCPI by exploring and/or piloting real-time data marketplaces that will demonstrate the key commercial, technology and policy areas central to the DCPI.

This White Paper is the first of many deliverables of this multi-year initiative to explore policy, technical and commercial enablers for a flexible data governance framework for data exchange. We welcome your participation and contributions.

Executive summary

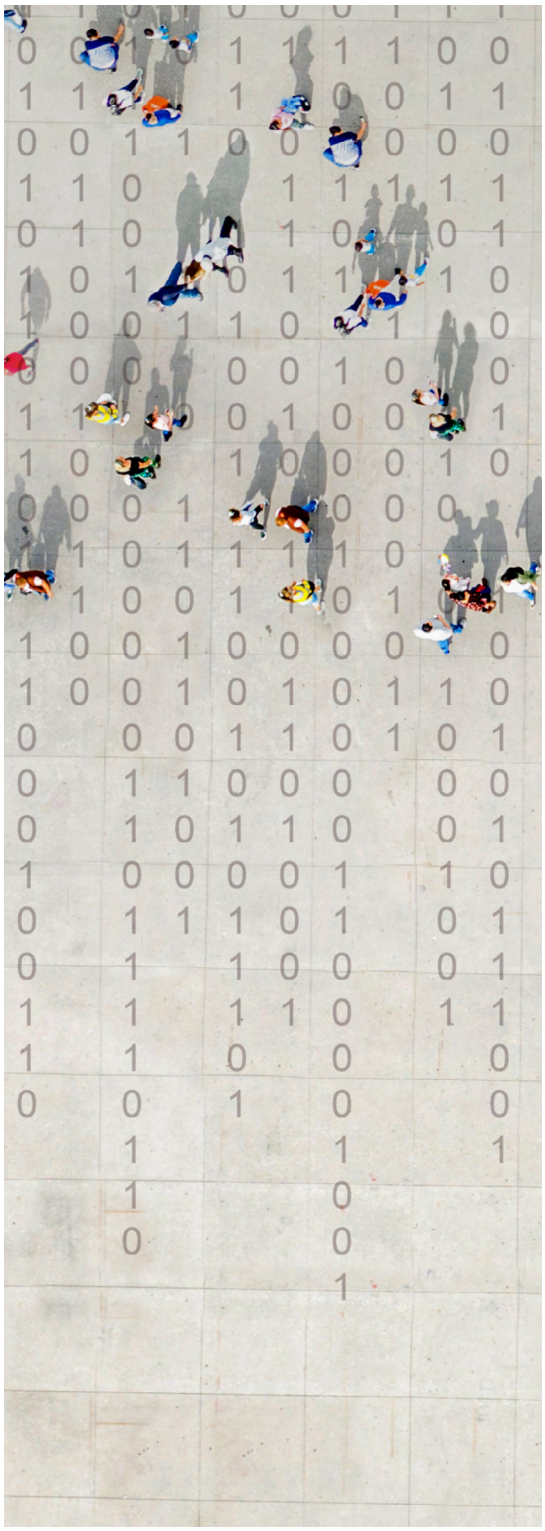
Sharing data with purpose
The world is awash with data and it is growing exponentially. Most of the data in the world (90%) was generated in the last two years alone.[4] Data is mostly siloed, but much of it can be impactful. Organizations and governments should sponsor changes to policies and governance, and invest in new technology to enable data economies to leverage data for better outcomes.

Designing a commercial, technical and policy environment to leverage data for the benefit of society
Accelerating the responsible exchange and use of data can solve critical challenges and fuel innovation for society. Based on the Data for Common Purpose Initiative (DCPI) and multistakeholder community input, this paper introduces five requirements that systems and governance should meet to leverage data for better outcomes. The DCPI explores the concept that data can and should be treated differently depending on its actual and anticipated use.

Paving the way with technological advances
Fourth Industrial Revolution technologies are in the process of enabling data to be treated differently depending on its actual and anticipated use – through "differentiated permissioning". As new uses for an individual's data arise within their *permitted uses*, relevant data can be tagged according to permission choices, automatically encrypted and anonymized.

Enabling the potential value of data exchanges and marketplaces
Currently, platforms (i.e. traditional business-to-business digital platforms) do not offer true data portability, which limits the possibility of combining data across them for multiple purposes. Data marketplaces offer one method to deliver on the five requirements to ultimately unlock data.

This White Paper provides an initial review of the various enablers (commercial, technical and policy) of data exchanges and marketplaces. It examines how they could apply in the context of a broader global data governance framework and offers real-life examples of how they can help to solve some of the world's critical problems.



Below: Akin Bostanci, Getty Images

1

Introduction: Challenge and opportunity

Organizations and governments should sponsor changes to data policy and invest in new technology to enable data economies to develop around a common purpose. These actions will lead to better outcomes of the next pandemic and beyond.

The challenge: Imagine a world...

> **If there is one thing more important than vaccines in this pandemic, it's data: data about transmission dynamics of the novel coronavirus, about symptoms, testing results and hospital admissions.**
>
> Ali Okhowat, Chief Executive Officer, mHealth Global, Canada

...where the COVID-19 pandemic was met with a multilaterally organized global response through the sharing of data. At the time of writing this paper, 2,703,620 lives[5] could have been spared and a significant portion of the 4.3% decline in global output[6] could have been prevented.

The COVID-19 pandemic surfaced many challenges that could have been prevented or alleviated and that are still perpetuated by the lack of data sharing. Health data does not always follow the virus when it crosses borders because many countries make cross-border data sharing cumbersome or, in some cases, illegal. Issues included the trustworthiness of testing kits and personal protective equipment. Contact tracing apps failed to gain the trust of citizens. Successful approaches to treatment took too long to reach other healthcare systems struggling to find answers. Yet the technology to effectively share information exists; what is required are the governance frameworks and data policy to support them.

What would better outcomes have looked like? What policies are required? Can current best practice in other areas provide lessons? While technology and policy enablers will not solve all the challenges and certain territories may be less motivated than others to improve the situation, participating governments have a shared interest in helping the global community. Better outcomes are possible with best efforts among *collaborative* entities.

One type of collaboration that is emerging is data marketplaces (see "Terminology: What is in a word?"). By combining multiple Fourth Industrial Revolution technologies, such as data rights management, machine learning, privacy-enhancing technologies (PETs) and blockchain, these data marketplaces are also being grounded in the responsible, ethical and fair use of data. Technology will not solve everything, but it lays the foundation of support for global collaboration with universally respected norms for personal and sensitive data as well as cross-border data policies and flexible data governance, which could all be combined to provide the better outcomes, such as the following:

- Virus outbreaks are immediately identified and all governments are alerted.
- Data, especially data sets essential to combat global crises, can flow across borders.
- All individuals trust the integrity of contact tracing.
- Strategies and outcomes for treating infected patients are captured and shared.
- All test kits and personal protection equipment meet quality control standards with data integrity.
- Instead of siloed data, systems allow for a combination of data from various origins, including personal, commercial and/or government sources.
- Preventive measures are in place and data is exchanged between countries to better identify emerging outbreaks before they occur.

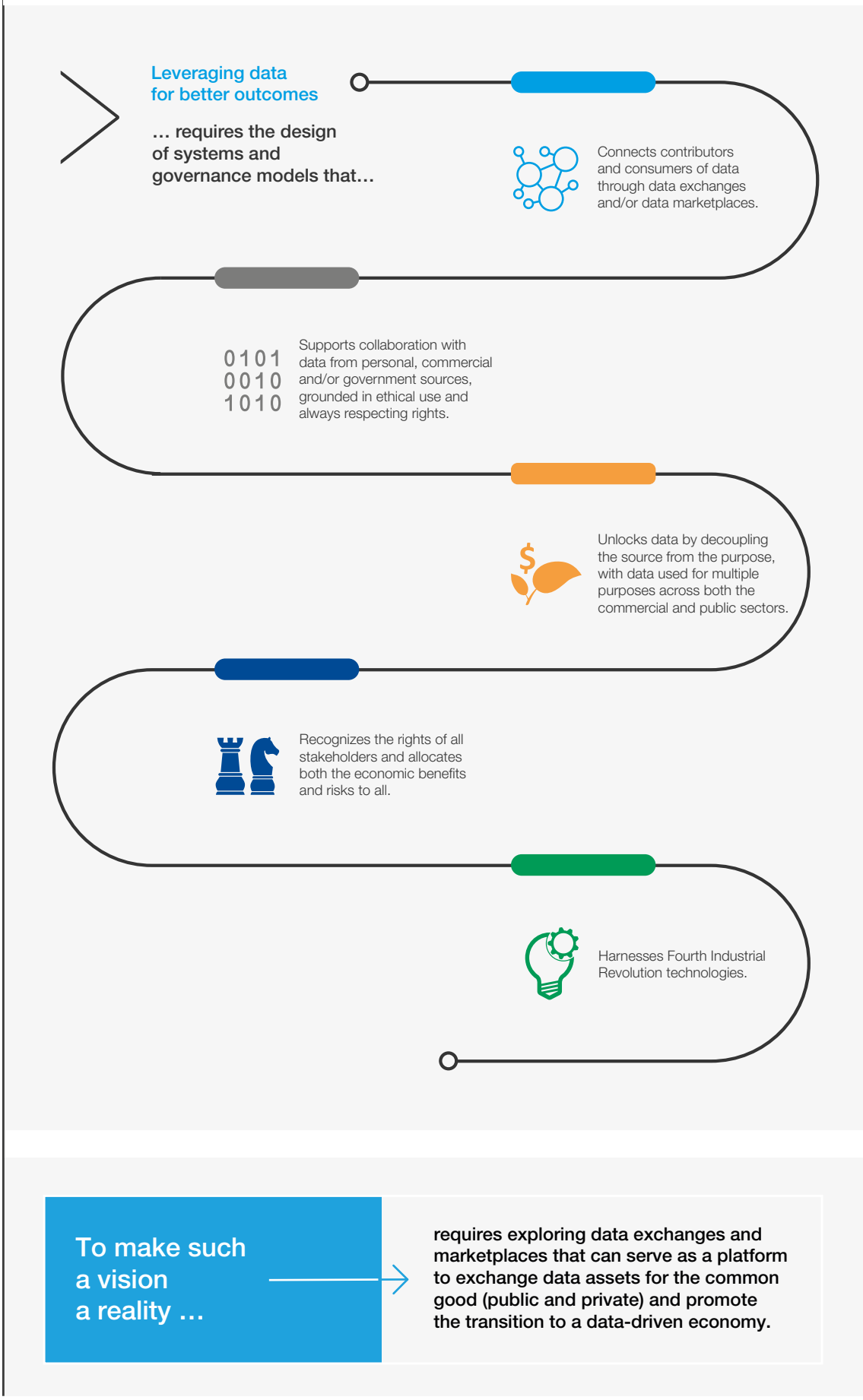
The opportunity: Essentially, accelerating the responsible exchange and use of data can solve critical challenges and fuel innovation for society. As illustrated in Figure 1, five main requirements are involved. One method (but not the only one) to deliver on these improved outcomes for society is through the development of data marketplaces. Correctly constructed, these elements, discussed in more detail in section 1, can provide value by:

- More readily connecting the contributors and consumers of data
- Combining data sets to create new value that is ethical and respects rights
- Evolving policy to focus on intended use
- Enabling the equitable allocation of risks and rewards
- Embracing Fourth Industrial Revolution technologies.

In addition, these systems can create a **commercial, technical and policy** environment to:

- Support the combination of data from multiple origins (e.g. personal, commercial or government sources)
- Allow the same piece of data to be reused many times, for multiple purposes
- Have the flexibility to allow for unforeseen use cases
- Recognize rights
- Provide citizen controls of permissions
- Enable the equitable allocation of risks and rewards.

FIGURE 1 | **Five requirements for systems and governance models to leverage data for better outcomes**



Source: World Economic Forum

The potential value of data exchanges and/or data marketplaces and the terminology related to these concepts are covered in more detail in subsequent sections.

The benefits of well-governed data sharing can include:

- **For individuals**: Increased control of data, potential income generation and greater security that their data is being used appropriately
- **For businesses**: The ability to tailor existing products, develop new products or enrich customer relationships, with knowledge that they are acting in the best interests of their customers and in line with national and cross-national regulation
- **For governments**: Better preparation to develop cohesive globally-facing data policies; creation of a positive data-forward business environment that is attractive to companies and puts people's rights first, which will help to optimize services provided to both individuals and institutions; delivery of high-quality, modern digitally-enabled services to citizens
- **For non-profits**: The possibility to better serve various stakeholders based on insights generated from real-world data.

BOX 1

Resetting communities and redefining well-being

The challenge: The COVID-19 outbreak has been most damaging to vulnerable populations. One of the pandemic's lessons is that government intervention alone can only do so much to address disasters, reduce poverty, support public health and control infectious diseases.

The desired future: Using local government data, personal data, hospital and company data as well as environmental data, it is possible to provide personalized services that go beyond the "one-size-fits-all" approach. Such multistakeholder data sharing can improve many aspects of human well-being, such as maintaining health, saving lives in disasters and preventing poverty.[7]

This paper highlights use cases from multiple industries and jurisdictions around the world, illustrating the possibilities data sharing unlocks for multiple stakeholders and the public good. It discusses the Data for Common Purpose Initiative approach to enabling these data-sharing scenarios on a broader scale, the underlying technology to support them and some foundational considerations for success.

The topics discussed in this paper are described under the following assumptions:

- While misalignment in the terminology still exists, standardization is in flux, and some see "data marketplace" as one type of data exchange. For simplicity, the terms are used interchangeably to cover the broad and nuanced interpretations that exist today.
- This paper does not constitute legal advice. None of what is imagined in an idealistic future is prejudicial to existing or future legal and regulatory practices.
- The paper addresses policy and technology enablers to achieve better outcomes through leveraging data while acknowledging the political, economic and other barriers, including:
  - An imperfect global political landscape, and obstacles that data policy and technology will not address
  - Various entities' economic interests that will not always have the goal of ecosystem equitability.

  Frameworks should be flexible enough to help entities navigate these barriers.
- This paper is designed to be accessible to people with varying degrees of technology and data policy knowledge, including those who are uninformed. The findings in the paper appear in simple terms to bring understanding. For these reasons, the paper does not delve into the multitude of technical layers, complexities and exceptions that exist within the topic areas, though the authors recognize their existence and importance.
- The paper lists key considerations, questions, opportunities, risks and other elements; this list is not exhaustive. These components should not be factored equally but should be weighed appropriately in response to the company-specific use case, stakeholder complexities and other assessments.

BOX 2 | **Terminology: What is in a word?**

The topic of data sharing is subject to multiple philosophical views and assumptions that stem from the relative immaturity of the concept. Quite often the uncertainty arises from the fact that even the terms "data exchange" and/or "data marketplace" mean different things to different people.

(See Box 3 for one example of terminology efforts)

In this paper, the terms "**data exchange**" and "**data marketplace**" are used interchangeably.

Much needs to be considered in any discussion on "**exchanging data**", "**data marketplaces**" and "**multistakeholder data interoperability**", including cultural beliefs, philosophies and nuance, which, it is important to acknowledge, differ from person to person. In this paper, the focus remains on global approaches to realize the potential of these interoperable systems, while respecting the concepts' individualized and localized notions. A broad and global view provides the greatest opportunity to ensure the rights of stakeholders and the equitable allocation of benefits and risks to all.

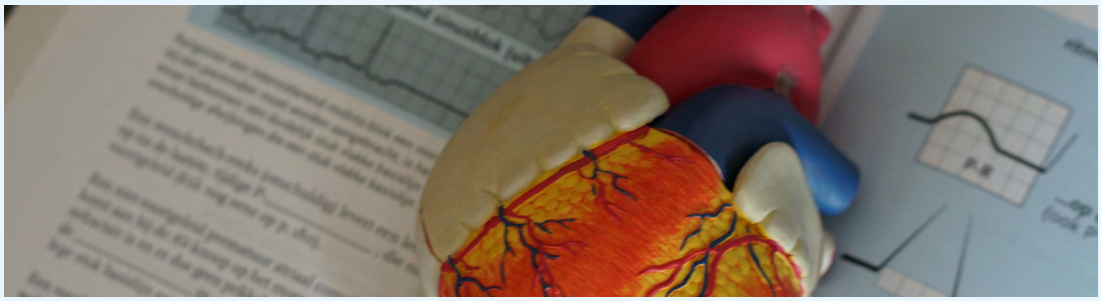
Demonstrating the value of combining data: What outcomes can be achieved?

Shaping a world **in which data governance models allow data from personal, commercial and/or government sources to be combined, while still respecting rights**

Solving rare diseases

The challenge: Approximately 10% of the global population or 475 million people are affected by a rare condition, with an estimated 15.2 million individuals expected to have clinical genomic testing for a rare condition within the next five years.[8]

The desired future: In today's world, researchers and clinicians have access to global rare disease data sets. Genomic data, coupled with phenotypic and clinical data, is a critical resource that can shorten the diagnostic odyssey that patients with a rare disease face. It can also power research and innovation in diagnostics and therapeutics. A federated data system is one method that allows local institutions to protect sensitive personal health data while still providing remote access to data sets for diagnostic capacity.



> **Solutions to many rare diseases are available, but they are trapped in an isolated clinical record that's often in another country. New solutions to remotely aggregate and access sensitive health data while still adhering to local data privacy and security laws offer an opportunity to finally provide answers to the over 300 million people living without a treatment for their rare disease.**
>
> Lynsey Chediak, Project Lead, Shaping the Future of Health and Healthcare, World Economic Forum LLC

Solving humanitarian needs for food and essential goods

The challenge: World hunger is rising. Approximately 8.9% of the world's population, or 690 million people, go to bed hungry each night. Since 2014, the number of people affected by hunger has been increasing slowly. If it continues at this rate, the number will exceed 840 million by 2030.[9] Oxfam projected in 2020 that up to 12,000 lives per day could be lost due to starvation.[10]

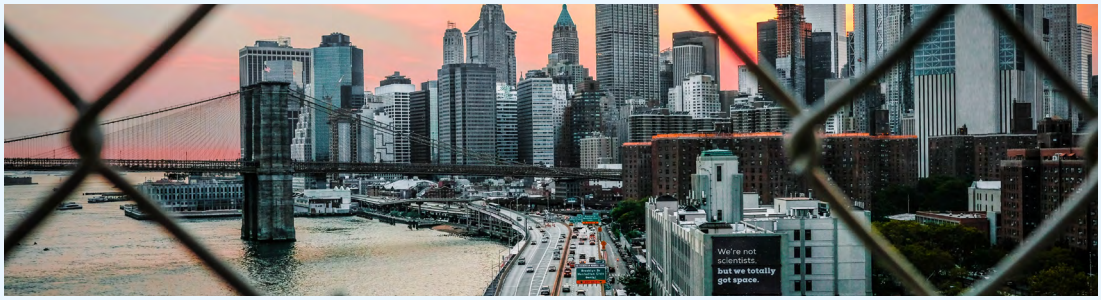
The desired future: Humanitarian, industry and public stakeholders should be provided with the combinatorial power of data to allow near real-time visibility of the flow of food and essential goods at the global and local levels. Identifying and relieving bottlenecks is crucial to getting humanitarian supplies where they are most needed, and aggregated data insights are key to creating future supply system resilience that leads to the protection and preservation of lives at risk, especially during crises.

> **Harnessing data and existing technology platforms to create open-source visibility tools, aggregating 'least granular level', anonymized transport, commodity and disruption data, to deliver global system-wide visibility can enable prompt, even pre-emptive action. Bringing international bodies, industry partners and broader ecosystem stakeholders together to inform the development of such 'meta-layer' visibility solutions provides an opportunity to build systemic resilience that can lead to multiple benefits, not least protecting livelihoods and lives vulnerable to crises.**
>
> Margi Van Gogh, Head of Supply Chain and Transport, Shaping the Future of Mobility, World Economic Forum

Shaping the future of urbanization

The challenge: By 2050, nearly 70% of people will live in cities.[11] The world is urbanizing at unprecedented speed and scale. From public health, infrastructure assets, mobility to transport, the most valuable data for building new services and improving social outcomes resides at the local level, fragmented across multiple agencies and stakeholders.

The desired future: Cities should harness data to grow inclusively and sustainably. Data helps gain citizen participation, better target government support and universal access to services, tailor services for the community and democratize access to city policy-makers.



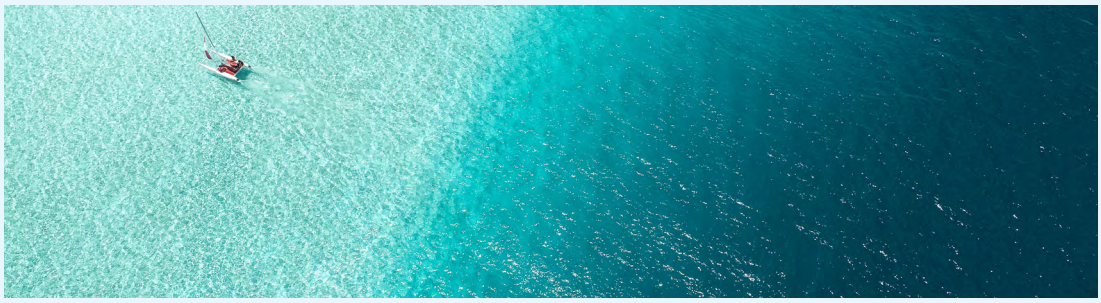
We know that digital and data infrastructure in our cities is crucial to their resilience and sustainability over the long term. The Forum has been helping cities to build that infrastructure through the G20 Global Smart Cities Alliance. To get the most value from smart city technologies, we can go further to unlock data that sits across multiple sectors and stakeholders, in service of the public interest.

Rushi Rama, Lead, Smart Cities, World Economic Forum LLC

Shaping sustainable oceans

The challenge: An abundance of data on the oceans is available (220,000 oceanographic research cruises, 1.95 billion temperature profiles, 1.13 billion salinity profiles, etc.) but data sets are siloed, and many projects are individual and independent efforts.[12]

The desired future: Better access to data sets and technology for a productive and healthy ocean and to improve the environmental footprint of ocean industries is central.



We aim to be the largest and most inclusive hub for ocean data sharing, using new technologies to reduce the time to access data from days and hours to seconds.

Bjørn Tore Markussen, Chief Executive Officer, Centre for the Fourth Industrial Revolution and the Ocean (C4IR Ocean), Centre for the Fourth Industrial Revolution Norway, Norway[13]

Shaping the long-term sustainability of tourism

The challenge: In 2019, the Travel and Tourism industry accounted for 10% of global GDP and employment, but its long-term resilience and role in economic development and connectivity is challenged by issues such as overcrowding, cultural and environmental degradation and overall destination capacity constraints.[14]

The desired future: Public and private stakeholders can integrate their data sets and leverage big data and digital platforms to allow real-time tracking of the supply and demand for destination tourism resources, leading to interventions that enable the sustainable, inclusive and resilient development of the tourism industry.

> **Sustainable tourism is about the efficient use of destinations' limited resources and capacity to maximize benefits for locals and visitors alike. Unlocking the potential of big data and other digital tools is vital for achieving this and making sustainability an opportunity and not a cost for the sector.**
>
> Maksim Soshkin, Lead, Data for Destinations, Shaping the Future of Mobility, World Economic Forum LLC

Below: Ju Sun, Getty Images

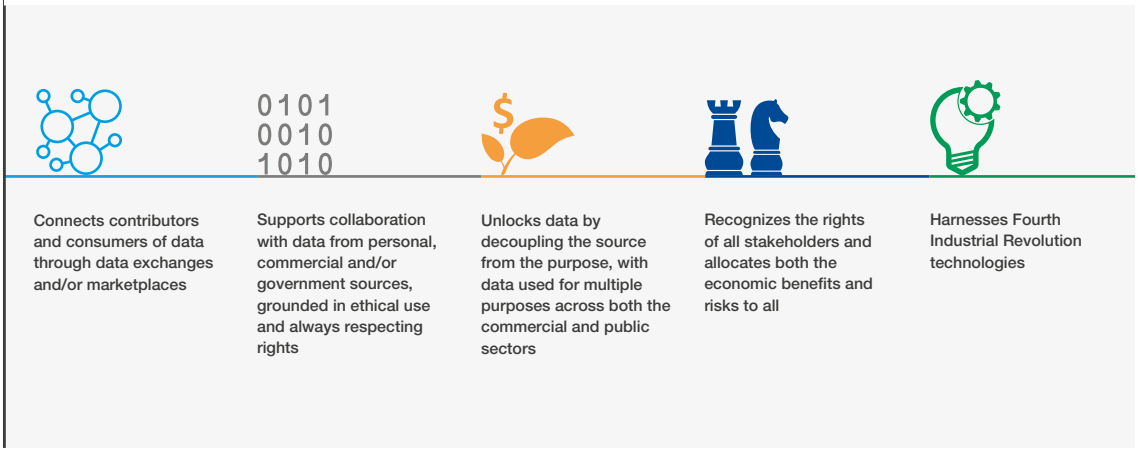
2

Evolving perspectives on data use

Leveraging data for the benefit of society requires designing systems and governance models.

Accelerating the responsible exchange and use of data can solve critical challenges and fuel innovation for society. Leveraging data for the benefit of society involves five requirements, explored in more detail in this section (Figure 2).

FIGURE 2 **Leveraging data for better outcomes**



Source: World Economic Forum

More readily connecting contributors and consumers of data

Connects contributors and consumers of data through data exchanges and/or data marketplaces

To leverage data for the value of full society, the world needs both contributors and consumers of data who are more readily connected. About 2.5 quintillion bytes of data are created each day and this amount is steadily rising with the growth of the internet of things (IoT).[15] Over the last two years alone, 90% of the data in the world was generated.

Much of this data is siloed and inaccessible beyond where the data is being housed, even within a specific jurisdiction or corporate entity. This is due to regulatory and business barriers, and to other factors. Adding to the complexities, the exchange of data must be grounded in responsible and ethical use. As a result, the world has greater potential for data privacy and sovereignty-oriented regulation.

At the same time, much of this siloed data may be helpful to improve society. For example, in agriculture, the data on crop yields is often not synthesized or accessible beyond the immediate jurisdiction of the environment. An ability to roll up data on crop yield at a national level, and even across borders, within responsible, ethical and compliant data use and protection frees up access to data for improved weather pattern prediction, for instance. This can eventually benefit the nature of the agriculture industry. Another example concerns ocean data. A considerable amount of the data resides in territorial waters, and data sharing and access are complicated due to sovereignty rules. To get to clean, productive and sustainable oceans, society needs to produce systems, governance and policy that more readily connect contributors and consumers of data.

0101
0010
1010

Combining data sets to create new value that is ethical and respects rights

Supports collaboration with data from personal, commercial and/or government sources, grounded in ethical use and always respecting rights

It is important to allow for the combination of data from a diverse range of sources, while still respecting rights.

Whether the purpose is providing better outcomes for the oceans, agriculture and public health or leveraging data from smart devices, how can data that is currently or ordinarily siloed be made available for opportunities? Society, governments, civil society and the private sector have the desire to access data for the public good.

It is important to create a flexible global data governance model that allows for the combination of data from personal, commercial and/or government sources, while respecting individual rights and complying with any relevant regulation. This model will positively empower a variety of stakeholders and remove unintended barriers. Such innovation could use artificial intelligence and machine learning in many domains (e.g. agriculture, health, the environment, mobility), and it could include new ways to responsibly access the more than 2.5 quintillion bytes of data created each day.

Evolving systems and policy to focus on intended use

Unlocks data by decoupling the source from the purpose, with data used for multiple purposes across both the commercial and public sectors

Policy-makers have focused much attention on data protection/privacy to the detriment of positive regulation that is supportive of specific activity in the data economy. As a consequence, the scale of policy development has not matched the scale of innovation in this area. Furthermore, existing policy and regulatory models usually focus on the source of the data (i.e. the person or the entity that generated the data) and remain relatively silent on intended or eventual use. There is good reason for this: data protection and privacy policies are designed to be both technology neutral and use case neutral. But in a world in which it is no longer realistic to catalogue every possible use of data, is this approach still valid?

The effects of these policies include:

- Failure to harness the full value of data
- Asymmetric policy development, creating fragmented data governance policies that hinder the ability to share data for intended and agreed upon purposes.

To develop a base for the combination of data from a variety of sources, it is necessary to reorient the focus of data policy away from the source of data and towards common purposes (i.e. data can and should be treated differently depending on its actual and anticipated use), notwithstanding the infallibility of data subject rights.

Imagine a world in which devices collect information about a person's vitals and lifestyle, where the individual sets permissions around what their data can be used for, such as research and testing to find a cure for COVID-19, dementia or cancer. As new initiatives are launched for the user's permitted purposes, relevant data will automatically be encrypted, anonymized and transmitted along with digital rights management rules to ensure that the data cannot be used for other purposes (see "Data rights management").

An individual could choose to set permissions for commercial purposes. If someone wanted to use their (anonymized) data for prescribed purposes, such as market research, they could opt in explicitly and get paid. Data valuations could be divided by purpose, allowing the value of the data output to change for specific use cases. Individuals would get paid at the moment of consumption, and financial authorities can simultaneously identify taxable income. With correct governance, this process could reduce bias in data sets and provide streams of income over time.

Enabling the equitable allocation of risks and rewards

Recognizes the rights of all stakeholders and allocates both the economic benefits and risks to all

> **Fundamental questions need to be answered: What can be done with the data? Who has rights over it? Who derives the economic benefit?**
>
> Sheila Warren, Head, Blockchain and Data Policy; Member of the Executive Committee, World Economic Forum LLC

The development and application of data exchange technologies should benefit people and the planet. Harnessing this revolution requires coordination between all stakeholders, including government officials, business leaders, members of civil society, individual citizens and international organizations.

By acting now, government leaders can keep their economies competitive and boost the well-being of their citizens. With multistakeholder coordination, governments can:

- Create a flexible system that equitably apportions risks and reward
- Design granular permissioning mechanisms that support multiple purposes, and inspire citizen trust and participation in the data-driven economy
- Develop market-based mechanisms to attach value to the data and encourage consumption of data that equitably compensates stakeholders
- Find more opportunities to use the same data for multiple purposes, increasing the value of data for all stakeholders.

Understanding the potential of data requires a balance between maximizing its power and delivering transformative change. It is important to:

- Ensure vulnerable population groups are protected from emerging risks associated with unlocking the full potential of data in a data-driven economy
- Build adequate infrastructure and governance frameworks that facilitate the responsible exchange of data based on ethical principles.

In the not-so-distant past, realizing these goals would have been unthinkable, but rapid advances in technology have exposed new opportunities and changed the landscape of what is possible.

Transforming data use visions into technical reality

Harnesses Fourth Industrial Revolution technologies

The world's understanding of the importance and benefits of data sharing has been growing for several years, as have insights into the possible unintended consequences. Many examinations of data marketplaces from both a commercial and a governmental perspective exist.[16] In addition, privacy concerns regarding data use and exchange are not new and have been investigated for a long time.[17] Quite often, the "positive" angles of exchanging data and the overall social value of data have been treated separately from the "negative" views of data and privacy invasion.

Technology innovations required to solve privacy issues have evolved to deliver on the vision of trusted data sharing; it is privacy concerns that have driven the search for PETs. Even simple use cases of exchanging data can have a variety of stakeholders, including citizens, companies or government departments. They can also include data from different sources, including sensors and transactions. This implies various forms of data, various storage and analytics models, as well as varied security and transparency requirements. As technology has progressed to solve these privacy issues, providing an integrated vision and implementing truly usable data marketplaces to deliver social value have become possible.

Summary of requirements: The potential value of data marketplaces

Fulfilling these characteristics might require moving away from closed platform models (i.e. traditional digital platforms) to more open and interoperable systems that compel reimagining permissioning mechanisms beyond the single-use, single-jurisdiction frameworks currently in place.[18] Progressing towards governance frameworks that enable data exchange is one way to materialize this. Government-led data marketplaces are one instance in which they will serve as a platform to exchange data assets for the common good and promote the transition to a data-driven economy.

> **Developing new data frameworks allows us to think differently about information and explore new, exciting uses that could bring needed attention to potentially hidden, previously unexplored value, while doing this with an ethical and managed risk lens.**
>
> Jane Allen, Principal, PwC, USA

Transitioning to a data-driven economy: The Colombia Moonshot[19] project

Currently, vast amounts of data are being created and shared at very high rates.

> **Data gathered from human–human, nature–machine and human–machine interaction and sensing devices have great potential to make well-informed decisions by public and private entities. Colombia is committed to promoting the transition to a data-based economy to improve its competitiveness and become a relevant global player. Understanding data and, more importantly, knowledge generated from data supported by Fourth Industrial Revolution technologies will become a strategic asset to promote social, economic and environmental impact.**
>
> Victor Manuel Muñoz Rodriguez, Director-General of the Presidency of Colombia, Colombia

Transitioning from traditional to data-driven economies is a focus area for leading economies, especially as it can fuel GDP growth in the post-COVID economic recovery. It is imperative to establish governance models, data valuation models, trust entities and ethical, fair, equitable and inclusive frameworks to develop the rules of the game to transition from a traditional economy to an economy based on knowledge data.

At the beginning of 2020, C4IR.CO (an alliance between the Mayor's Office of Medellín and the Presidency of the Republic of Colombia) began to conceptualize a model for a data marketplace for the common good (called the Colombia Moonshot project). In collaboration with the World Economic Forum and the DCPI, the Centre for the Fourth Industrial Revolution Colombia and PWC designed a data exchange platform with multiple stakeholders, aimed at accelerating digital transformation and the shift to a data-based economy, while solving challenges of common interest such as natural disasters, the loss of biodiversity, or even disease outbreaks such as COVID-19.

A subsequent Forum publication in collaboration with the DCPI community and C4IR.CO will be published later this year outlining the Colombia Moonshot project's proposed architecture, legal and ethical framework and needed policy, regulatory, commercial and technical enablers.

Explored in the next section, new technologies provide new opportunities to leverage data for better outcomes while still respecting rights.

Below: iMrSquid, Getty Images

3

Harnessing advances in technology

The dramatic advances in technology over the past few years have enabled truly equitable data exchanges.

Advances in technology (see "Transforming data use visions into technical reality") mean data marketplaces are now feasible – the technology is ready to deliver trusted solutions in this space. To harness these advances, organizations must think through:

- Common digital principles that enable the trusted use of data
- Core technologies that can form the basis of competitive data exchanges.

As technologies develop further, data exchanges must be able to adopt and easily integrate them into the existing governance and technology frameworks, in particular as regulation is also likely to change over time. For these reasons, governance and technology frameworks should be connected but not directly interdependent, enabling policy changes without impacting the technical implementation while supporting technical changes that do not require rewriting governance procedures. This flexibility in the approach to data exchanges is key to allowing stakeholders to select both the technology and governance frameworks that are most appropriate for them.

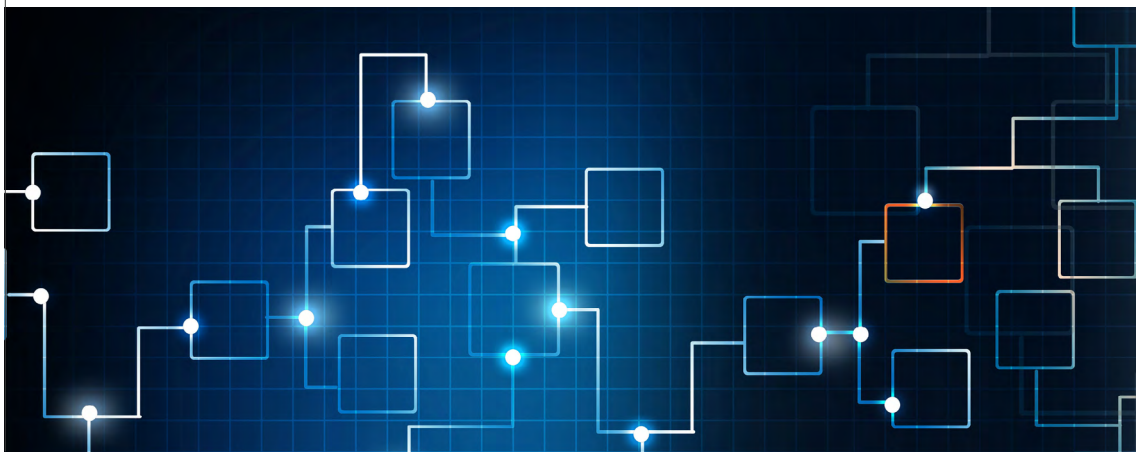
Importantly, the technologies should aim to deliver the required policy and governance outcomes rather than be applied merely for their own purpose. Technology becomes an enabler, rather than a driving force.

Common digital principles that enable the trusted use of data

A large variety of technologies can be chosen from, but any framework will require technology governance that supports compliance with a few principles that enable trust.

- **Data permissioning**: Granular mechanisms offer individuals clear control over their data, ensuring it is only used for purposes they have agreed to and allows for attaching permissions for multiple purposes and uses.
- **Data reuse**: These data approaches allow data to be reused across a variety of permissioned purposes and treat data differently depending on its actual and anticipated use, creating the commercial imperative for sharing data.
- **Data protection**: This security focuses on individual rights.
- **Adaptable data processing**: Data can stay at the moment and point of collection or within the original collection system; the analysis of data and its use in processing are performed in a privacy-preserving manner.
- **Data provenance**: Data should be tagged with an identifier of origin to account for rights as it moves throughout the data value chain, allowing the allocation of risk and reward.
- **Economic value and governance**: Data is optimized over time for the different use cases, and governance frameworks are flexible enough to manage the data's use in different cases.
- **Competition**: It is necessary to ensure that data marketplaces are competitive, enabling innovation and new business creation for local economies.

> **Designing the right systems could play a crucial role in rebalancing the data asymmetries currently existing in the digital economy and delivering more equitable data economies, thus promoting competition, enabling the rise of new markets, new market entrants and new services, and ultimately improving welfare for consumers and citizens.**
>
> Gabriele Carovano, Researcher, King's College London, United Kingdom



Fourth Industrial Revolution technologies for data exchanges

A number of technologies enable new ways to exchange data that can manage individual privacy and allocate risk and rewards. Fourth Industrial Revolution technologies allow a truly digital and data-driven economy to emerge – one that integrates multiple systems and solutions. Not all the technologies can be covered here, and significant work already exists on IoT, data cleaning, connectivity and basic security technologies, for instance. The following should be viewed as a suite of possible technologies that can be selected from and applied within any solution and governance framework.

> **The great engineering industries and utilities have been built over many decades on standards and principles. Data exchanges will be no different. To take full advantage of Fourth Industrial Revolution technologies and build trust, value and equitability, the importance of mapping and modelling data flows and their intrinsic relationships and dependencies between people, processes and data exchange technologies cannot be underestimated.**
>
> Fergus Cloughley, Founder and Chief Executive Officer, Obashi, United Kingdom

Data rights management

One of the most complex technical advances is decentralizing data rights management (i.e. allowing the tagging of data and enabling its use across different areas). By "tagging permissioning to data", data is freed to be used for different purposes.

The digital music industry provides one example of digital rights management. From 1999 to 2001, Napster disrupted the music industry with its peer-to-peer file sharing network. Individuals could make digital copies of music and share it with other users without compensating the copyright holders. A federal court in the US shut the service down because of infringement of music industry copyrights.[20] Fast-forward 20 years and technology has been developed to track and trace music across the internet with programmatically-enforced copyrights.

These concepts, tested in the music industry, can be developed for data rights management, enabling the tracking and tracing of data as it moves across data marketplaces, and can provide permissioning solutions for individuals. Combined with PET and/or blockchain, individuals can control granular data permissions, allowing them to share their data according to their wishes.

Data formats and standards

A critical advance over the past few years has been the realization that standards and formats for exchanging data may assist some aspects of data marketplaces. Recent examples include Fast Healthcare Interoperability Resources (FHIR),[21] the Financial Regulatory data standard (FIRE),[22] ISO 20022[23] or World Wide Web Consortium (W3C).[24] Many industries are also developing industry-specific standards made possible through the emergence of industry-, technology- or government-led consortia. These standards enable a base of data exchange, while ensuring innovation and freedom of movement for the companies involved.

> **FIRE is an open-source data standard that has revolutionized data at financial institutions to facilitate the process of sharing financial regulatory data between financial institutions and regulators in the pursuit of financial stability.**
>
> Diana Paredes, Chief Executive Officer and Co-Founder, Suade, United Kingdom

One area in which standard activities related to data marketplaces have already started concerns the terms to describe systems with these characteristics. Multiple standards, guidelines and benchmarks are under way from humanitarian, technology and other areas focused on this task.

Some standards organizations have also created working groups to create definitions, such as the Institute of Electrical and Electronics Engineers (IEEE), to which the DCPI is contributing (see "A terminology standardization effort under way at the IEEE"). All, however, are works in progress.

BOX 3

A terminology standardization effort under way at the IEEE

The IEEE P3800 Data Trading System Working Group is tasked with creating a standard for a data-trading system: overview, terminology and reference model.[25]

The purpose of the standard is to provide the foundation for a data-trading system allowing multilateral exchanges of data. With no known standard for systems that allow trading data across different sectors or industry domains, IEEE saw the need to develop a standard to enable a domain-independent marketplace system.

Analytics and machine learning

As discussed (see "Combining data sets to create new value that is ethical and respects rights"), trusted data sharing across data marketplaces can create new products and digital services, and enable faster and more effective decision-making across both the public and private spheres. A key technology in this area is machine learning, not solely for the automation of decision-making but also for the development of the data marketplaces themselves. As an example, analytics and machine learning enable the assessment of data tags so they can be used at high speed in a variety of uses cases, by diverse service providers and across different economic valuations. Effectively, machine learning breaks data out of being used in a static business model format within a data marketplace.

The increased usage, of course, creates the requirement for both ethical and governance frameworks to be developed; just the mathematical or technical capabilities involved can create systems that are unintentionally biased.[26] Developing frameworks to ensure that any machine learning is used in a socially conscious manner is therefore critical, as is having a broad set of stakeholders in technical discussions.

BOX 4

Data sharing to solve critical global problems

The global food challenge is immense: 2 billion people do not have access to safe, nutritious and adequate food. Certain agricultural start-ups are applying machine learning to improve farms' performance. Tel Aviv-based Prospera, for example, collects 50 million data points from 4,700 fields every day to identify pest and disease outbreaks, cut pollution and eliminate waste.[27]

Digital identity

A digital identity is a set of verifiable digital attributes and credentials, enabling entities (e.g. individuals, organizations, physical goods) to prove who they are in digital systems in the same way as they can in the "real" world. For a long time, the ability to prove one's identity on a digital system was a critical roadblock to the development of heavily automated data exchanges. A variety of new technologies allow robust digital identities to be created and verified. Different digital identity technologies have different governance approaches. For example, some digital identity schemes will be issued by the national government, others are provided by technology vendors and, in some instances, a data exchange may provide the digital identity to facilitate payment for the associated use of data. IDs should ultimately help to ensure the monetization of data is effectively captured on behalf of all stakeholders.

Privacy-enhancing technologies (PETs)

PETs are a broad spectrum of technologies. They are a key enabler of data marketplaces as they help ensure the protection of individuals' data and data rights. Starting with the principles of privacy-by-design is a critical part of any data marketplaces. Within that context, PETs can assist

with four main strategies: 1) hide; 2) separate; 3) control; and 4) enforce.[28] Within these are a variety of technologies for: 1) identification/anonymity; 2) private communications; 3) privacy-preserving computations; 4) privacy within databases; and 5) discrimination prevention.[29] In addition, synthetic data techniques allow inferences to be made where absences occur in data sets, and these inferences can be drawn with a high degree of usefulness. Encryption still plays a role in all data marketplaces, ensuring the protection of data when it is stored and in transit. As discussed below (see "Adjusting security approaches to meet data marketplace requirements"), encryption only works for certain aspects of a data exchange; to ensure that data exchanges are able to work effectively, newer forms of PET will also be required.

Governance frameworks must ensure that advances in technology can be used as they emerge, guaranteeing not just initial data protection when a data marketplace is created, but also *continued* data protection for individuals as exchanges evolve over time (Table 1).

TABLE 1 **Privacy-enhancing technologies**

Strategy	Description	Method
Hide	The confidentiality of data is ensured by encrypting, pseudonymizing or anonymizing data.	– Digital signatures – Zero knowledge proofs – Data stream anonymization
Separate	Personal data is stored and processed using distributed methods.	– Homomorphic encryption – Private communications – Multiparty computation
Control	Respondents can see, modify and delete the information about them.	– Differential privacy – Privacy preserving data mining
Enforce	Privacy policies are in place and enforced.	– Data framework governance

Source: Adapted from Domingo-Ferrer, J., and A. Blanco-Justicia, "Privacy-Preserving Technologies", in M. Christen, B. Gordijn, M. Loi (eds), *The Ethics of Cybersecurity*, International Library of Ethics, Law and Technology book series, vol. 21, 2020

Distributed technologies and blockchain

Distributed ledgers and blockchain technologies are often touted for the development of data ecosystems. Blockchains – used correctly – can ensure the acceleration of responsible data use. This section focuses on what blockchain means for data exchanges.

In combination with PETs described above, blockchain tokenization and smart contract techniques can act as a form of advanced software that ensures what has been agreed is executed upon, enabling the recognition and apportioning of data rights, risks and rewards. It can assist governments and individuals by ensuring that these are apportioned in an automated manner, in turn allowing for a more dynamic exchange of data and exchange models. In the creation of economic value models, blockchain may act as a component/enabler in market-based mechanisms to attach value and permissions to the data.

One example of changing business models enabled by emerging technology is Tip the Farmer projects, which allow an end user who purchases a cup of coffee to economically empower a coffee farmer – to tip the farmer. Through the application of blockchain tokens, it is possible for people who enjoy the coffee to trace the original farmer and provide economic value to them directly.

A primary part of what tamper-proof blockchain can enable is data provenance: proving data is from who it says it is from and that it has not been compromised in transit between originator and receiver. In addition, it is possible to use aspects of blockchain to ensure the automation of payment and indeed the delivery of infrastructure that can fully manage the micropayments required for data exchanges to truly function.

Technology will continue advancing at increasing speeds and has shifted from a barrier to an enabler of data sharing. Forward-thinking governments and organizations will look for opportunities to use new technologies within the confines of governance frameworks to safely and securely expand data marketplaces to unlock new value for their stakeholders.

Below: Selman Keles, Getty Images

4

Laying the groundwork: Key enablers

Organizations must focus on citizen trust, ecosystem growth and data value to support and accelerate the equitable use of data for multiple purposes, especially in a data exchange.

Through multistakeholder discussions, dialogues and a series of workshops, the DCPI community identified and framed the most important enablers of data marketplaces to support and accelerate the equitable use of data. Enablers cut across commercial, technical and policy areas, and were framed as opportunities, fears and barriers. Not addressing them can undermine or even stifle the ability of data marketplaces to unlock data for multiple purposes with equitably apportioned risk and rewards.

Framing the most important components to build out multistakeholder data-sharing frameworks, the DCPI community identified the need to focus on the following, among others:

- Access, rights to use, rewards
- Governance and oversight
- A reimagining of permissioning mechanisms
- Inclusion and equitable benefit
- The tying of data value to data use – equitably valuing data
- Data protection and privacy rights, including encryption and anonymization
- Security – building on the ethical use of data.

Underlying all of these components is the need to articulate parameters and controls for the responsible and ethical use of the data.

This section introduces two key themes from the DCPI community's design and scoping phase:

1. Inspiring citizen participation and trust in data sharing

2. Designing systems that scale and support growth.

Subsequent DCPI deliverables will cover these specific enablers and governance considerations in more depth.

4.1 Inspiring citizen participation and trust in data sharing

Citizen engagement can be encouraged by developing trust in the data-driven economy, especially for those whose lack of understanding, fears of data misuse or exposure to security issues have left them disenfranchised from the digital world.

Increasing participation in the emerging data economy through consumer empowerment

Over the past decade, and more recently during the COVID-19 pandemic, the role of the digital economy and data sharing has expanded significantly. Despite the COVID-19-related economic slowdown, digital shopping continues to grow. Digital sales volumes, excluding travel, grew 20% in April, 29% in May and 33% in September of 2020.[30] Consumers worldwide are maintaining habits formed during the shutdown. In parallel, Visa research shows consumers' increased awareness of being asked for personal information when they interact online.[31] For many, it has become routine.

At the same time, about one-third of consumers are hesitant to share personal information and are more likely to deny these requests, even when it is required to use key application features (Figure 3). Consequently, these consumers may be missing the benefits of digital goods and services.

FIGURE 3 **Digitally-hesitant consumers**

Digitally-hesitant segment

33%

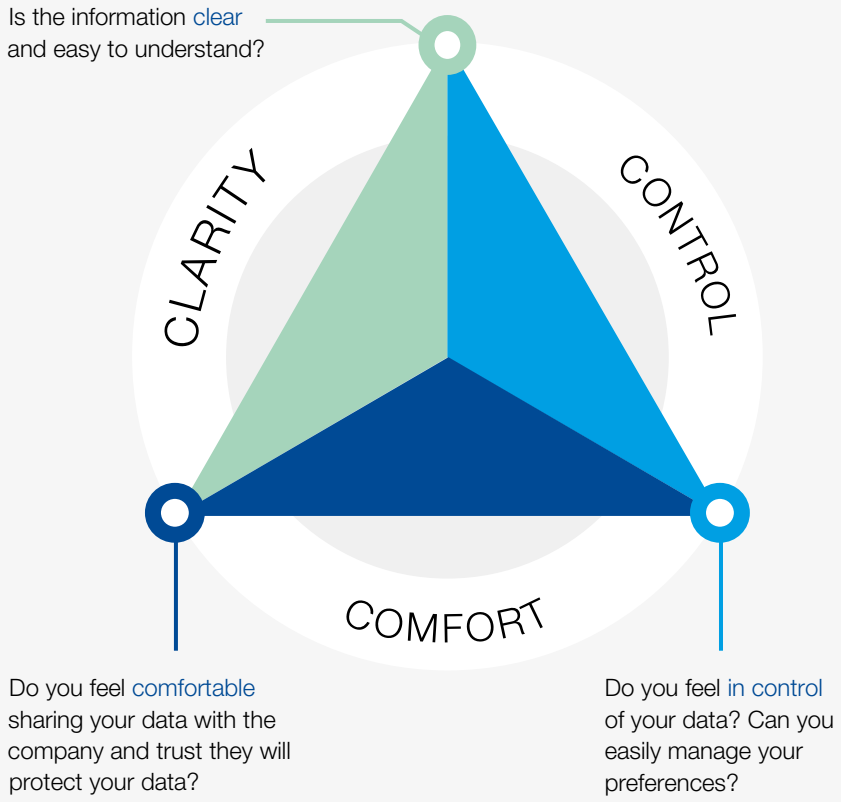
of the population is being "left behind" by the data economy

- 0% feel in control of their data
- Half as likely to read the "terms and conditions" associated with data permission requests
- Half as likely to take proactive measures to protect themselves online
- Only 3% fully understand how data is being used
- Only 2% are comfortable sharing data
- 25% shop online less than once a month
- 86% believe companies benefit more from their personal data than they do
- 82% believe companies use data policies primarily to protect their own legal interest

Source: 2020 Visa Consumer Empowerment Study (unpublished)[32]

To feel more confident sharing personal information and to actively participate in the digital economy, digitally-hesitant consumers need clear descriptions, a sense of comfort and meaningful controls for how personal information may be used (Figure 4).

FIGURE 4 **Consumer data confidence factors**



Source: 2020 Visa Consumer Empowerment Study (unpublished)

The study created and tested alternative permission experiences across a spectrum of proposals (Figure 5).

FIGURE 5 **Spectrum of proposals tested and results**

Spectrum of proposals tested

Traditional terms & conditions (T&Cs)	Granular controls
Legally dense and inaccessible	Clear choices in simple language

Likelihood to grant permission…	Granular controls versus T&C	
	All consumers	Digitally-hesitant
To share the personal information required for the service	16% are more likely	22% are more likely
To share personal information for other purposes (e.g. for marketing, to sell to third parties)	38% are more likely	48% are more likely

Source: 2020 Visa Consumer Empowerment Study (unpublished)

Simply put, consumers – especially digitally-hesitant consumers – feel more comfortable sharing personal information when offered more granular choices to permission it, giving them confidence to expand participation in the digital and data-sharing economy. The best way to address the digital hesitancy that threatens to hold back future growth is to give consumers clarity, control and choices. The benefits of doing so accrue to everyone.

Through the COVID-19 pandemic, the world has seen the countless failures of a disjointed and inefficient approach to educating and empowering citizens to use their own data to protect themselves and those around them. But many of the necessary rules, technologies and forward-thinking organizations to support this new vision already exist. The real work is in creating governance systems to ensure trust is maintained through responsible and ethical data use.

4.1.1 Inclusion and equitable benefit

Designing governance to maintain or to expand inclusion, while also ensuring equitable apportions of risk and reward, can build citizen confidence in the digital economy.

BOX 5 **Choosing use cases to spotlight multistakeholder benefits**

Strong efforts in Latin America aim to reduce socio-economic inequalities, offer higher levels of education to more citizens, supply better internet access, improve inclusion, and support the preservation, conservation and sustainable and responsible use of natural resources. Achieving these goals will require multiple industry sectors to share information and to understand and define the best use cases to pursue. It will also require the right regulations and assurances of data protection for citizen confidence.

When thinking through the various data exchange or marketplace use cases to explore, prioritizing use cases with a broad impact is key. One use case that triggers responses from all sectors is *energy*, a vital and foundational element supporting and enabling all industries, subsectors, data-driven economies and sustainable development initiatives and goals. In addition to generating interest, a broad use case also inspires financial investment from multiple and key stakeholders because it is underpinned by digital transformation maturity, and value creation is driven through collaboration.[33]

Designing inclusion and equitable benefit into data exchanges

Ensuring that data exchanges are inclusive and equitable requires that both governance frameworks and technical architectures are grounded in inclusion and equity. Systems need the ability to price and monetize data in a way that allows anyone to participate, not only large corporations.

To ensure equitable access to data exchanges, data needs to be accessible, usable by large parts of the population, composable so it can be used in a variety of contexts and so it can fit into many different decision-making contexts, traceable, and able to apportion value and risk.

Unlocking data for common purposes could lead to sustainable investment in the core platforms to enable data exchanges. For example, governments could find new revenue sources through government-led data marketplaces that could pay for the regulatory, legislative, technical and operational efforts needed to use and maintain them. This can be one example of addressing the need for entities to share in the investment burden.

BOX 6 | **Eldercare opportunity in Japan[34]**

Famous for having a higher proportion of elderly citizens than any other country, Japan is known as a *super-ageing society*. The country must harness the power of data to solve various social issues associated with ageing while also protecting data privacy through the informed permissioning of each citizen.

The importance of elderly care service has grown, but the pool of skilled caretakers has not. Elderly care service is highly dependent upon the experience and intuition of these skilled caretakers, but nursing homes do not have efficient ways to collect and transfer that knowledge to new cohorts. Many nursing homes are not equipped to use data to enhance the quality of the eldercare because they lack the proper incentive mechanisms and regulatory framework to make the investment.

The technology exists for nursing homes to gather the eldercare data (e.g. treatment data, rehabilitation programme results) through smart devices. Privacy-preserving data analytics support is also available to process the eldercare data to improve treatments, cognitive capacity and, ultimately, quality of life. An approach that combines the skill of caretakers and the power of data will lead to enhanced precision care services for the well-being of all elderly citizens who need care.

4.1.2 Security

Frameworks adopted by data marketplace operators can promote citizen trust by providing strong security while also being agile and adaptable to innovation.

The security of data will be critical to participation in data exchanges and to their continuation. As increasing amounts of value are exchanged in them, they will become increasingly more attractive to dishonest operators, resulting in a rapidly changing threat landscape. Security takes on a new meaning in an "open" data exchange environment, especially when placed in the context of data marketplaces that store or use citizen data or are involved in the delivery of critical national services. Technical, policy, commercial and governance mechanisms will become crucial to effectively coordinate and manage security for the data marketplaces.

Adjusting security approaches to meet data marketplace requirements

Traditionally, data security has focused on two main states of data: data in motion and data at rest. When designing data exchanges, however, two new states of data, data during processing and data post-use, also require security (Table 2).

TABLE 2 | **The expanded focus of a secure data exchange**

Data exchange and/or marketplace				
States of data: **Expanded focus**		States of data: **Traditional focus**		
Data during processing	Data post-use	Data in motion	Data at rest	
This is required because data may be processed remotely and potentially used for unintended purposes. Examples include regulatory compliance with General Data Protection Regulation as data is processed.	From a regulatory perspective, this will include the emerging regulatory approaches to privacy frameworks and citizens' "right to be forgotten".	This ensures the security of data while it is transferred from point A to point B either physically or digitally.	This ensures the security of data while it is not being used and is stored; traditionally a database.	**Security needs**

Source: World Economic Forum, adapted from various sources[35]

Optimal security strategies will evolve over time, both for securing the data and for securing the overall ecosystem. Data governance frameworks, therefore, need to allow flexibility for different security mechanisms as they adapt, change and grow in response to the constantly changing threat landscape associated with data.

4.2 Designing systems that scale and support growth

4.2.1 Global interoperability across jurisdictions

More coordinated and harmonized approaches to data governance policy and cross-border data flows[36] could help enable responsible and ethical global interoperability. Also, permissioning mechanisms need interoperability beyond the single-use jurisdictions currently in place. When the next pandemic arrives, governments and industry need to react swiftly and effectively, starting with sharing information without hesitation or friction. Making certain that data localization restrictions, where they occur, are meaningful and flexible will be essential to ensure that global policy interoperability becomes a possibility. Nations are sovereign and free to set their own policies, so the key is multilateral cooperation.

> **In a borderless world, global data flows could be channelled to solve global problems. But the real world is not borderless, which is why achieving this vision depends on interoperable frameworks and trusted data flows that enable global value chains and digital trade.**
>
> Melissa McSherry, Senior Vice-President and Global Head, Data, Security and Identity Products, Visa, USA

4.2.2 Scalability

As with any technology intended for mass adoption, data exchanges must be developed with scalability in mind. There are two aspects of scalability to consider: 1) the scalability of data; and 2) the scalability of the exchange or marketplace itself. The main items to consider during the creation of a scalable data marketplace include:

- **Market depth**: Ensuring sufficient amounts of data for processing, which likely means that a true data exchange or marketplace will need to bring together citizen, corporate and government data to ensure sufficient depth for data processing, intermediation and analysis
- **Speed**: Ensuring transaction speeds meet the needs of different use cases (e.g. high-velocity areas like financial services)
- **Connectivity**: Ensuring systems have the capacity to support simultaneous connections and can expand capacity as the number of participants grows
- **Interoperability**: Ensuring platforms are interoperable to avoid siloed data, to unleash the full value of data economies exploiting cross-sectoral data platforms' synergies, and to keep the data economy competitive by preventing the formation of monopolistic sectoral data platforms.

Scalability can be closely related to the context of the data in question. For example, real-time data streams from sensor networks may need to be processed immediately for the data to be of any use, or the data may become irrelevant and unusable for the question being solved. By contrast, other questions may require a longitudinal data set; such analyses are not time sensitive but will have different scalability requirements.

4.2.3 Competition

Open and interoperable systems could rebalance data asymmetries currently existing between tech giants and markets' new entrants, thus promoting competition and new services, and ultimately improving welfare for consumers and citizens. An important element of data marketplaces will be ensuring proper controls are in place, through Fourth Industrial Revolution technology and data marketplace security measures, to guard against activities that do not comply with regulation.

The role that data marketplaces can play in ensuring equitable access to data is crucial to the new digital era. Ensuring they stay competitive will be a critical factor for their success. For example, if data really is the new oil, why do countries not develop sovereign data funds in a similar vein to sovereign wealth funds?[37] Companies could pay for access to data for a fee that the government can use to improve overall digital infrastructure or otherwise reduce taxes on vulnerable categories.

4.2.4 Governance and oversight

Traditional data governance and oversight defines authority and control over data assets. In a data marketplace, however, data governance takes a broader, more nuanced meaning. In this new context, governance is about managing multiple stakeholders and multiple types of data used in an open environment, while supporting granular permissioning mechanisms.

Several aspects of governance will be required for data marketplaces:

- **Regulatory governance**: The ability to dynamically manage regulatory changes over time and audit systems to ensure that the consumers of the data maintain compliance with the relevant regulatory requirements
- **Self-regulation**: Codes of conduct and standards to support responsible, ethical and competitive data marketplaces
- **Access and control**: System access and technical control over the actual data, requiring technical, policy and business processes that ensure data is used as permissioned by contributors of data and used by consumers of data.

As the DCPI community continues to collect input from around the world and to examine the different enablers of data marketplaces in various contexts, these perspectives will evolve. Nevertheless, all frameworks will be grounded in the ethical and responsible use of data.

Below: Gremlin, Getty Images

5 Looking ahead

The acceleration of Fourth Industrial Revolution technologies and data permeates every aspect of society. Many institutions have focused resources primarily on privacy-based approaches to data. While the ethical use of data and protecting rights are critical, these approaches have resulted not only in a failure to harness the full value of data, but also in asymmetric policy development, creating rapidly fragmented policies that interfere with sharing data for agreed-upon purposes. Simultaneously, the world is navigating through a global pandemic that is causing stress while creating opportunities for change.

This White Paper imagines instances where creating new data governance models that combine data from personal, commercial and/or government sources can positively empower a variety of stakeholders while removing some of the unintended policy barriers. These scenarios are entirely plausible today. However, without proper protocols and governance, society risks creating a world in which access to data is overly restricted, impeding significant human progress and innovation, or in which authorities require data sharing without striking a balance that respects the rights of the individual parties involved, including businesses.

The multistakeholder collaboration, Data for Common Purpose Initiative, which led to the ideas in this paper, represents the continued and tireless commitment of policy-makers, business leaders, academics and others to find ways to exchange data assets for the common good, while protecting rights and the equitable allocation of risks and rewards. This paper is the first step in this project.

In the coming year, through the DCPI community, the World Economic Forum will spur action:

- **Partnerships and public-private cooperation**: Nurturing partnerships and convening the most relevant, influential and dynamic communities of government, corporate, civil society and technical leaders committed to shaping commercial, technical, policy and regulatory governance and the application of data exchanges and marketplaces for better outcomes

- **From local to global**: Leveraging the Forum's global network, laying the foundation for the concepts and data exchange governance frameworks discussed in this paper; demonstrating the interoperability and global applicability that can be achieved

- **From concept to reality**: Piloting and iterating the frameworks to test the hypothesis and demonstrate the "art of the possible" for commercial, technical and policy areas central to the DCPI

- **Subsequent publications**: Providing a synthesis of best practices, emerging trends and thought leadership on the key themes and areas of impact addressed by the DCPI, starting with two publications published in the second quarter of 2021, one by the Centre for the Fourth Industrial Revolution India and the other by the Centre for the Fourth Industrial Revolution Colombia on an enabling framework towards a data economy.

Understanding the potential of data requires focusing on the balance between maximizing its power and delivering transformative change. The World Economic Forum and its partners in this project hope others will join this critical initiative.

Contributors

The Forum's Centre for the Fourth Industrial Revolution Data for Common Purpose Initiative is a global multistakeholder endeavour to co-design, pilot and scale frameworks to create a flexible data governance model that allows for the combination of data from personal, commercial and/or government sources, while respecting individual rights. This White Paper is based on consolidated views of the DCPI community and shared ideas. The opinions expressed herein do not necessarily reflect those of the individual people involved in the project.

Lead authors

Kimberly Bella
Senior Director, Visa, USA; World Economic Forum Fellow

Nadia Hewett
Project Lead, Data for Common Purpose Initiative, Data Policy and Blockchain, World Economic Forum LLC

Cathy Mulligan
Visiting Researcher, Imperial College London, United Kingdom

Acknowledgement

Felipe Bernal Guarín
Manager, Strategy Consulting, PwC, Colombia

Gabriele Carovano
Researcher, King's College London, United Kingdom; World Economic Forum Fellow

Kurt Fields
Director, PwC, USA; World Economic Forum Fellow

Charlie Gerdts
Senior Analyst, Visa, USA

Cindy Hong
Senior Analyst, Visa, USA

Rosetta Jones
Senior Director, Visa, USA

Kibae Kim
Principal Researcher, KAIST, Republic of Korea; World Economic Forum Fellow

Scott Klokis
Director, PwC, USA; World Economic Forum Fellow

Tomoaki Nakanishi
Executive Director, Japan External Trade Organization (JETRO), USA; World Economic Forum Fellow

Yuiko Noda
Director, Government Relations, Visa, Japan; World Economic Forum Fellow

World Economic Forum

Kimmy Bettinger
Project Specialist, Data Policy, World Economic Forum LLC

Anne Flanagan
Project Lead, Data Policy, World Economic Forum LLC

Eduardo Gómez Restrepo
Project Specialist, Internet of Things, Robotics and Smart Cities, Centre for the Fourth Industrial Revolution Colombia

Satyanarayana Jeedigunta
Chief Adviser, Centre for the Fourth Industrial Revolution India

Sheila Warren
Head, Blockchain and Data Policy; Member of the Executive Committee, World Economic Forum LLC

Dimitri Zabelin
Project Specialist, Data Policy, World Economic Forum LLC

Julie Ziskind
Platform Curator, Data Policy, World Economic Forum

The World Economic Forum is grateful for the generous commitment and support of the organizations that dedicate Fellows to the project: KAIST, PwC, Visa, the Ministry of Economy, Trade and Industry of Japan and the Financial Services Agency of Japan.

Special appreciation goes to the global Affiliate Centres for their input and contribution to shaping the outcomes of the frameworks, especially the Centres for the Fourth Industrial Revolution Colombia, India and Japan, which are actively exploring the DCPI's concepts.

Endnotes

1. The Centre for the Fourth Industrial Revolution Colombia, in collaboration with a multistakeholder group, is leading the Moonshot Project, aimed at producing a governance framework that allows for a government-led data marketplace in Colombia. The Centre for the Fourth Industrial Revolution Colombia (C4IR.CO) is an alliance that began in 2019 between the Mayor's Office of Medellín and the Presidency of the Republic of Colombia, operated by Ruta N. See www.c4ir.co.
2. The Centre for the Fourth Industrial Revolution Japan has launched the Study Group on Vitalization of Data Marketplace that includes public- and private-sector participants to discuss and explore how to activate a data marketplace in Japan.
3. The Centre for the Fourth Industrial Revolution India is working with the DCPI and a multistakeholder community in the country to make a data economy in India a reality. More details will be available in World Economic Forum publication *Towards a Data Economy through DCPI: An Enabling Framework* (forthcoming).
4. Marr, B., "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read", *Forbes*, 21 May 2018, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#4c8deee860ba? (accessed 18 March 2021).
5. World Health Organization, "Coronavirus disease (COVID-19) pandemic, Numbers at a glance", 21 March 2021, https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed 21 March 2021).
6. Word Bank, "Global Economy to Expand by 4% in 2021; Vaccine Deployment and Investment Key to Sustaining the Recovery", Press Release, 5 January 2021, https://www.worldbank.org/en/news/press-release/2021/01/05/global-economy-to-expand-by-4-percent-in-2021-vaccine-deployment-and-investment-key-to-sustaining-the-recovery (accessed 16 March 2021).
7. The Centre for the Fourth Industrial Revolution Japan is leading a project called "Resetting Communities: Redefining Well-being in the COVID-19 Era". The project asks what the "new normal" for human happiness and well-being should be. More information on this project and contact details are available through the Centre for the Fourth Industrial Revolution Japan website at https://www.weforum.org/centre-for-the-fourth-industrial-revolution-japan.
8. World Economic Forum, "Global Data Access for Solving Rare Disease: A Health Economics Value Framework", White Paper, 2020, http://www3.weforum.org/docs/WEF_Global_Data_Access_for_Solving_Rare_Disease_Report_2020.pdf (accessed 18 March 2021).
9. Reid, K., "5 world hunger facts you need to know", World Vision, 29 October 2020, https://www.worldvision.org/hunger-news-stories/world-hunger-facts#:~:text=World%20hunger%20is%20on%20the%20rise.&text=Globally,%20about%208.9%25%20of%20the,excee? (accessed 19 February 2021).
10. Oxfam International, "12,000 people per day could die from Covid-19 linked hunger by end of year, potentially more than the disease, warns Oxfam", Press Release, 9 July 2020, https://www.oxfam.org/en/press-releases/12000-people-day-could-die-covid-19-linked-hunger-end-year-potentially-more-disease (accessed 18 March 2021).
11. United Nations, Department of Economic and Social Affairs, News, "68% of the world population projected to live in urban areas by 2050, says UN", 16 May 2018, https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html#content (accessed 19 March 2021).
12. See the Centre for the Fourth Industrial Revolution for the Ocean (C4IR Ocean), "ODP and Iconic Global Ocean Dataset", https://www.oceandata.earth/#about (accessed 19 March 2021).
13. For information on the open and collaborative Ocean Data Platform (ODP), see https://www.oceandata.earth (accessed 19 March 2021).
14. World Travel & Tourism Council, "Economic Impact Reports", 2020, https://wttc.org/Research/Economic-Impact (accessed 19 March 2021).
15. Marr, "How Much Data Do We Create Every Day?", op. cit.
16. See, for example, Zuiderwijk, A., et al., "Elements for the Development of an Open Data Marketplace", Conference for e-Democracy and Open Government, 2014; Perera, C., C. H. Liu and S. Jayawardena, "The Emerging Internet of Things Marketplace from an Industrial Perspective: A Survey", IEEE Transactions on Emerging Topics in Computing, 2015; Smith, J., "Data Marketplaces: The Holy Grail of our Information Age", Medium, 7 August 2018.
17. See Xu, H., et al., "Information Privacy Concerns: Linking Individual Perceptions with Institutional Privacy Assurances", *Journal of the Association for Information Systems*, vol. 12, no. 12, 2011; White, C., "Introducing Fairness to the Data Marketplace: Privacy Regulation & Consumer Empowerment", The Henry M. Jackson School of International Studies, University of Washington, 2 July 2019.
18. The World Economic Forum investigates this topic in more detail and reimagines permissioning models in "APPA – Authorized Public Purpose Access: Building Trust into Data Flows for Well-being and Innovation", White Paper, 2019, http://www3.weforum.org/docs/WEF_APPA_Authorized_Public_Purpose_Access.pdf (accessed 23 March 2021).
19. The Centre for the Fourth Industrial Revolution Colombia in collaboration with multiple stakeholders is leading the Colombia Moonshot project, which aims to produce a governance framework that allows for a government-led data marketplace in Colombia.

20. U.S. Copyright Office, “A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001)”, 2001, https://www.copyright.gov/fair-use/summaries/a&mrecords-napster-9thcir2001.pdf (accessed 19 March 2021).

21. Health Level Seven (HL7), Fast Healthcare Interoperability Resources (FHIR), FHIR Release 4, “FHIR Overview”, 2019, https://www.hl7.org/fhir/overview.html (accessed 19 March 2021).

22. GitHub, “Financial Regulatory (FIRE) Data Standard”, 2021, https://github.com/SuadeLabs/fire#:~:text=The%20Financial%20Regulatory%20data%20standard,policy%2C%20monitoring%20and%20supervision%20purposes (accessed 19 March 2021).

23. International Organization for Standardization (ISO), “ISO 20022”, https://www.iso20022.org (accessed 19 March 2021).

24. World Wide Web Consortium (W3C), https://www.w3.org (accessed 19 March 2021).

25. Institute of Electrical and Electronics Engineers (IEEE) Standards Association, “P3800 – Standard for a data-trading system: Overview, terminology and reference model”, 2020, https://standards.ieee.org/project/3800.html (accessed 19 March 2021).

26. Ledford, H., “Millions of black people affected by racial bias in health-care algorithms”, Nature Portfolio, 26 October 2019 update, https://www.nature.com/articles/d41586-019-03228-6 (accessed 19 March 2021).

27. Sonmez, M., “Shaping a data economy”, International Monetary Fund, 2020, https://www.imf.org/external/pubs/ft/fandd/2020/12/shaping-a-data-economy-wef-sonmez.htm (accessed 19 March 2020).

28. Domingo-Ferrer, J., and A. Blanco-Justicia, “Privacy-Preserving Technologies”, in Christen M., Gordijn B., Loi M. (eds), *The Ethics of Cybersecurity*, International Library of Ethics, Law and Technology book series, vol. 21, 2020, pp. 279-297, https://doi.org/10.1007/978-3-030-29053-5_14 (accessed 18 March 2021).

29. Ibid.

30. Visa, “Q4 2020 Earnings Call” Transcript, 28 October 2020, https://s1.q4cdn.com/050606653/files/doc_financials/2020/q4/CORRECTED-TRANSCRIPT_-Visa,-Inc.(V-US),-Q4-2020-Earnings-Call,-28-October-2020-5_00-PM-ET.pdf (accessed 18 March 2021).

31. Visa, 2020 Visa Consumer Empowerment Study, unpublished. In this study, 2,000 adults, balanced to represent populations, were surveyed online in the United States, United Kingdom and Australia from 15-25 September 2020 (margin of sampling error for n2000: +/-2.2%).

32. Consumer survey questions include:
Consumer Survey Q20: Thinking about your data when you conduct digital/online activities like using apps and websites, how much control do you feel you have over what data is collected about you and how that data is used? (N=2,060).
Consumer Survey Q26: When asked by companies on apps/websites for permission to use your data collected when using their product, which of the following best describes how often you read the terms and conditions? (N=2,060).
Consumer Survey Q23: Which of the following best describes how often you grant permission to collect/use your data for something beyond the primary service that you initially signed up for, such as receiving special offers, making service improvements or targeting advertisements to you? (N=2,060).

33. The Centre for the Fourth Industrial Revolution Colombia and PWC are leading the Colombia Moonshot project, aimed at producing a governance framework that allows for a government-led data marketplace in Colombia. PWC provided expertise in developing data-driven and emerging digital markets.

34. The Centre for the Fourth Industrial Revolution Japan has launched the Study Group on Vitalization of Data Marketplace that includes public- and private-sector participants to discuss and explore how to activate a data marketplace in Japan.

35. See Williams, E. A., “Data in Use Is the Point of Least Resistance”, SecurityWeek, 26 April 2019, https://www.securityweek.com/data-use-point-least-resistance; Froehlich, A., “How to secure data at rest, in use and in motion”, SearchSecurity TechTarget, March 2020, https://searchsecurity.techtarget.com/feature/Best-practices-to-secure-data-at-rest-in-use-and-in-motion; Dolci, R., and W. Agahari, “Realizing Control in Data Marketplaces Through Secure Multi-Party Computation (MPC): An Exploratory Study”, Safe-DEED, 7 September 2020, https://safe-deed.eu/realizing-control-in-data-marketplaces-through-secure-multi-party-computation-mpc-an-exploratory-study; Gilad-Bachrach, R., et al., “Secure Data Exchange: A Marketplace in the Cloud”, CCSW’19: Proceedings of the 2019 ACM SIGSAC Conference on Cloud Computing Security Workshop, November 2019, pp. 117-128, https://dl.acm.org/doi/abs/10.1145/3338466.3358924 (all (accessed 19 March 2021).

36. World Economic Forum, “A Roadmap for Cross-Border Data Flows: Future-Proofing Readiness and Cooperation in the New Data Economy”, White Paper, 2020, http://www3.weforum.org/docs/WEF_A_Roadmap_for_Cross_Border_Data_Flows_2020.pdf (accessed 23 March 2021).

37. Mulligan, C., et al., “DataNet – Doing for Data What the Internet Did for Communications”, 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 4-8 May 2020.

WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org