WORLD ECONOMIC FORUM

A Roadmap for Cross-Border Data Flows: Future-Proofing Readiness and Cooperation in the New Data Economy

WHITE PAPER

JUNE 2020

Contents

Preface

Anne Josephine Flanagan,
Project Lead, Data Policy,
World Economic Forum, USA

Nada AlSaeed,
Senior Manager, Bahrain Economic Development Board (and World Economic Forum Fellow), Bahrain

Sheila Warren,
Head, Blockchain, Digital Assets, and Data Policy, and Member of the Executive Committee, World Economic Forum, USA

The World Economic Forum partnered with the Bahrain Economic Development Board and a Steering Committee-led project community of organizations from around the world to co-design the Roadmap for Cross-Border Data Flows, with the aim of identifying best-practice policies that both promote innovation in data-intensive technologies and enable data collaboration at the regional and international levels.

Creating effective policy on cross-border data flows is a priority for any nation that critically depends on its interactions with the rest of the world through the free flow of capital, goods, knowledge and people. Now more than ever, cross-border data flows are key predicates for countries and regions that wish to compete in the Fourth Industrial Revolution and thrive in the post-COVID-19 era.

Despite this reality, we are witnessing a proliferation of policies around the world that restrict the movement of data across borders, which is posing a serious threat to the global digital economy, and to the ability of nations to maximize the economic and social benefits of data-reliant technologies such as artificial intelligence (AI) and blockchain.

We hope that countries wishing to engage in cross-border data sharing can feel confident in using the Roadmap as a guide for designing robust respective domestic policies that retain a fine balance between the benefits and risks of data flows.

Bahrain's interest in this project stems from its recent success in launching national policy frameworks to facilitate the flow of data across its borders, including the Personal Data Protection law, the Cloud Computing Services to Foreign Parties law, the removal of legacy localization requirements, and the expansion of connectivity infrastructure. Bahrain is continuously exploring best-in-class policies on data flows in order to benefit from deep cooperation in the international data economy.

The Project Steering Committee provided a set of global multistakeholder perspectives over the course of the project. This report reflects their extensive input, gathered at workshops around the world, including at the Annual Meeting of the World Economic Forum in 2020, the Sustainable Development Impact Summit 2019 and the Summit on the Middle East and North Africa in 2019.

We believe that the findings from this work are applicable to both emerging market economies and highly developed market economies, and that there are lessons to be learned from each. The World Economic Forum is in discussions with parties around the world that seek to adapt the Roadmap for their own context. We anticipate that the Roadmap will serve as a beneficial, collaborative, inclusive and safe tool to facilitate cross-border data flows given the increased importance of the data economy for global economic recovery and growth, as well as technological and societal development.

A note from the Steering Committee co-chairs

Lothar Determann,
Partner, Baker McKenzie
(Project Steering
Committee Chair)

Leanne Kemp,
Chief Executive Officer,
Everledger (Project Steering
Committee Co-Chair)

Everyone needs data to succeed in today's world economy.

Countries can attract inbound cross-border transfers of data and information technologies only if people, businesses and governments abroad trust them. To earn a reputation as a safe data transfer destination, countries must provide for secure telecommunications infrastructure, respect individual privacy and confidentiality, exercise self-restraint regarding forced data access, and enact laws that also benefit people and organizations outside their borders, including data privacy, security, contracts and trade secret protection laws. Moreover, governments must be transparent, share data, and encourage their people and businesses to share data across borders if they want to participate in cross-border knowledge transfers. Open information societies thrive best in the world economy.

Conversely, people, businesses and governments hesitate to transfer data to countries that maintain weak data security infrastructure, laws and defences; excessively spy and seize data; fail to enforce or comply with laws protecting privacy, confidentiality and contracts; cover up data security breaches and risks; suppress media reporting; or fail to offer foreign businesses and citizens due process and recourse to privacy, contracts and trade secret protection laws. Countries that impose local data storage and retention requirements to secure better access for themselves can expect multinational businesses to stay away and other countries to retaliate. Similarly, countries that regulate data processing too rigidly and with specific restrictions on cross-border data transfers provoke reciprocal restrictions by other countries, resulting in reduced access to global data and technology, pressures for compromises in bilateral trade negotiations, and accumulating complexities. Cross-border data transfers require give and take.

Since the outbreak of COVID-19, governments around the world have started to realize and admit that restrictions on cross-border data flows not only inhibit scientific and economic progress but actually cost lives.

As co-chairs, we thank our fellow Steering Committee members, the larger project community and the staff at the World Economic Forum for their contributions to this white paper on cross-border data transfers. We hope that law- and policy-makers find the data, insights and recommendations helpful and we look forward to receiving feedback and the continued debate on this important topic.

Executive summary

The challenge

The technologies of the Fourth Industrial Revolution, including artificial intelligence (AI), the internet of things (IoT) and blockchain, are exceptionally reliant on accessing and processing data. To realize the potential of such data-intensive technologies, or to fully harness the power and efficiency of cloud computing solutions for start-ups and SMEs, data needs to be able to move seamlessly across country borders. The ability to move, store and process data across borders is foundational to the modern international data economy, and as new global growth relies increasingly on digital growth in the post COVID-19 era, progressive cross-border data flows policy has come into its own as a policy lever for ambitious governments seeking economic recovery.

Despite these benefits, laws and policies that act as barriers to this type of international data sharing are on the rise,[1] threatening to undo this progress, slowing technological innovation and limiting positive societal impact. While some of this friction is based on perception, such as the myth that data is better protected by restricting it to within one country, or a perception that such policies maximize value for local populations, some of it is deliberate and misguidedly protectionist.

The opportunity

Certain regulatory differences across countries cannot be eradicated; they are necessary and appropriate because sovereign nations have different values and strategic priorities. However, in order to create trust between nations when it comes to allowing companies within them to participate fully in the international data economy, there is a clear need for interoperable policy frameworks that can streamline requirements across borders and create mechanisms to reduce regulatory overload. Doing so capitalizes on economies of scale, particularly at regional level, and allows governments to create a friendly policy environment for indigenous and international investment. Investment breeds opportunity, and those countries with a burgeoning technology sector can start to maximize these companies' opportunities on a global scale, enabling them to develop cutting-edge technologies with global impact as well as experiencing potential knock-on economic and societal benefits.

The solution

> Cross-border data flows policy is a foundational prerequisite to a functioning international data economy and thus requires action fro the highest levels of power

Building trust between nations requires both an assurance that countries are like-minded in how they approach supporting their data economy and the implementation of a series of backstops that reduce risk. Our proposed solution is a practical Roadmap for governments of country-level policy building blocks that, when combined, are designed to harness the benefits and minimize the risks of cross-border data sharing.

Cross-border data flows policy is a foundational prerequisite to a functioning international data economy and thus requires action from the highest levels of power. In addition, as we look at all types of data in the economy, not just personal data or proprietary information, it is ultimately governments that are empowered to take action to open the gates and allow data to flow relatively seamlessly across their borders.

In the Roadmap, our project community of globally diverse industry experts proposes what best-in-class data flows policy looks like. For some countries, very little will be needed in the way of upgrading as they have inspired the core principles by their own actions, whereas for others the Roadmap may represent a full suite starting point. In order to cater for varying degrees of ambition, we first crystallize the most essential building blocks, and then offer scope for the most ambitious and advanced economies to future-proof their policy-making in this area.

Introduction

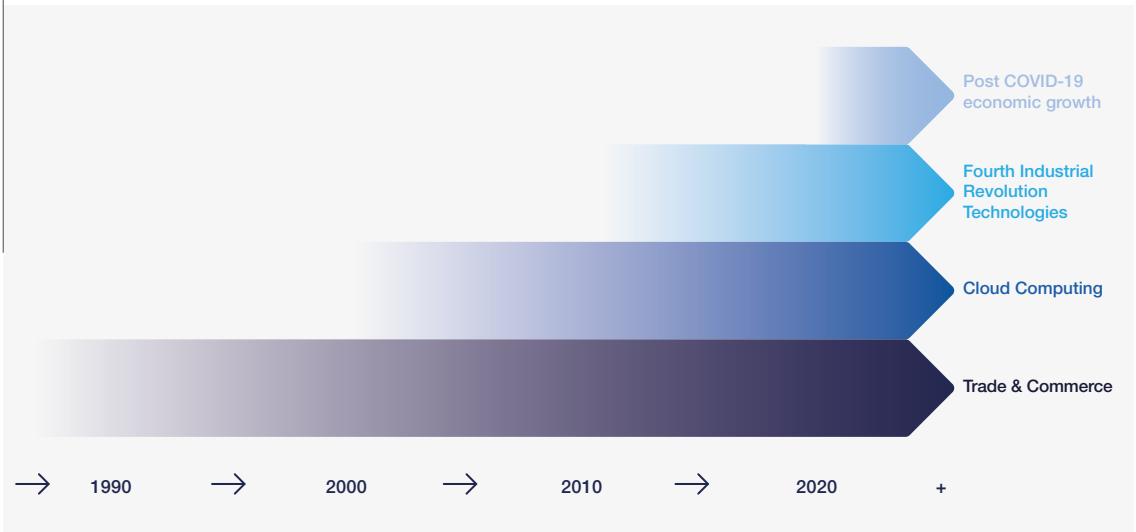
The importance of cross-border data sharing

While cross-border data sharing, i.e. the movement of information across international borders, has long been necessary from the perspective of trade,[2] internet-based services and e-commerce – more recently cloud computing, Fourth Industrial Revolution technologies such as artificial intelligence (AI) and the internet of things (IoT) – rely on access to high-quality data that often resides in more than one territory.

Through the development and deployment of these data-reliant technologies and solutions, nations can expect to derive increased economic and social value. Furthermore, in a world disrupted by the COVID-19 crisis, the data economy has risen in terms of its importance for new economic growth.

Despite such benefits, data localization requirements, e.g. laws, standards or policies which mandate that data be stored within a geographical territory, are on the rise globally, threatening to deter this progress – sometimes intentionally but often unintentionally.

FIGURE 1 *The increasing importance of cross-border data flows over time*



The problems of unjustified data localization requirements

“ For companies, an absence of data localization requirements is akin to having visa-free travel for their data

When unjustified, data localization requirements can prove highly problematic. As well as driving up the cost of cloud computing, upon which SMEs are highly reliant, it becomes difficult to achieve access to high-quality data at scale, upon which technological development relies – and from there, problems quickly amplify. We see associated economic consequences when companies need to create and maintain multiple data centres in different jurisdictions, at great cost in both monetary and environmental terms. Furthermore, companies relying upon such services may find they avoid certain markets altogether due to the increased cost of doing business there. This then has further knock-on effects for the attractiveness of regions when it comes to investment of capital and retention of talent, with data localization restrictions acting as digital walls between countries.

For some economies, a deliberate approach to restricting the international movement of data can be the result of a mistaken belief that localized data reduces risk. From a business perspective, the opposite can hold true: Regulatory certainty breeds business commitment in product and service markets.

For companies, an absence of data localization requirements is akin to having visa-free travel for their data. One still needs a passport (which is represented by the trust mechanisms discussed below), but travel is pre-authorized. Removing barriers to data flows is speedier, cheaper and more efficient than the contrary, and it is hugely beneficial in growing international business regardless of size.

How can we dismantle arbitrary barriers to cross-border data sharing by implementing backstops that provide assurance of appropriate safeguards to governments without undermining global economic growth? What does a practical approach look like? How can countries ensure they have appropriate policy frameworks in place to maximize benefit and minimize risk?

The World Economic Forum has convened a multistakeholder group of businesses, civil society actors, academics and governments globally who were consulted on what makes cross-border data policy fit for purpose and future-proof. The answer was stark: Start by ensuring your own house is in order, otherwise creating trust becomes almost impossible. Without country-level preparedness, international participation proves challenging in this space, and secondly, governments can influence but not fully control the international environment.

Consequently, this white paper does not represent a one-size-fits-all approach to cross-border data sharing, nor does it advocate how countries are to implement the recommendations outlined in the various building blocks that we will discuss, but it does try to identify the policies that countries ought to consider implementing domestically in order to facilitate full participation in the international data economy. As sovereign nations, countries have both their own diplomatic relationships and their own domestic policy considerations to take into account, and the Roadmap is therefore designed to provide a holistic look at what stakeholders believe works when it comes to cross-border data flows.

We assume that all countries and users of this Roadmap wish to have competitive open economies, that they want to exist in a globally competitive region, and they aim to attract both local and foreign investments. We assume also that the countries and users of this Roadmap want to support their technology sector where access to data is a key driver, whether it be in AI, blockchain or IoT applications such as smart cities, etc. We also assume that no country wants to compel its industrial players to share data across borders if they do not wish to do so and that this choice is best left in the hands of the holders of such proprietary data.

Countries wishing to follow the Roadmap are advised to conduct a country-level review of the current legal and regulatory provisions that may complement or obstruct the Roadmap. By definition, progress along the Roadmap indicates progress, with a full suite round-up representing forward-leaning cross-border data flows policy in the Fourth Industrial Revolution.

Finally, while this Roadmap is designed to examine the issue in the context of the international and regional levels, there are learnings here for data-sharing policies at country level.

A Roadmap for Cross-Border Data Flows

A Roadmap for Cross-Border Data Flows

①
Allow data to flow by default

Prohibit data localization requirements except in very specific circumstances in order to create regulatory certainty for businesses.

②
Establish a level of data protection

Establish national legal frameworks that protect the data of private individuals. Complement this with laws that protect proprietary rights.

③
Prioritize cybersecurity

Enact transparent cybersecurity legislation in line with international norms and maintain robust data security infrastructure.

④
Hardwire accountability between nations

Establish cooperation mechanisms between national authorities to hold governments accountable for the security and confidentiality of the data they share, while making allowances for compliance.

⑤
Prioritize connectivity, technical interoperability, data portability and data provenance

Prioritize the development of connectivity infrastructure as a prerequisite to building a local data economy, encourage technical standards to increase interoperability, facilitate data portability at the B2B level to support SMEs, and encourage data publishers to ensure data integrity.

⑥
Future-proof the policy environment

Allow for the possibility of future alternative models (such as federated learning models and data trusts) that can also fulfil the spirit of cross-border data flows.

A Part A: Establishing the building blocks of trust

1 Allow data to flow by default

Policy recommendations

- **Given the overriding disadvantages for respective local consumers, industries, technological development and job markets, both national laws and negotiated cross-border data sharing agreements between governments should prohibit data localization requirements.**
- **Narrow specified exceptions may be allowed in order to achieve a legitimate national security or public policy objective, provided that the measures are: (1) non-discriminatory; (2) not arbitrary; and (3) do not amount to a restriction on trade.**
- **The absence or removal of data localization requirements should not impede national authorities' access to data for law enforcement and regulatory compliance purposes.**

Data localization

"Data localization requirements have the effect of restricting the cross-border flow of data

Data localization requirements are broadly defined as any laws, standards or policies that require an entity to store data on media that is physically present in a specific geographical territory: This can include the infrastructure and services that support the hosting of data, e.g. servers. Data localization requirements have the effect of restricting the cross-border flow of data. They may be either deliberate or unintended, explicit or implicit.

Such requirements can often be found in data and cybersecurity laws, but, in some instances, they may be invisible. By way of example, there may be a clause in a sector-specific law which mandates that a specific type of data be stored in a specific geographic location, such as a clause within a law governing financial services mandating access to records by authorities. In practice, it can, for example, make it illegal to transfer company or employee records across borders or limit an institution's ability to outsource certain functions offshore.

Because data localization requirements can also emerge as a result of indirect policy measures, such as requiring or limiting government procurement to locally established entities only, or tax incentives that favour domestic industries over foreign products and services,[3] they can sometimes be difficult to identify at first. The result is that, even when it is technically legal to do so, the transfer of data across borders becomes impractical and costly, making data localization an inevitable outcome.

The new data localization

We are suddenly witnessing an increase in intentional data localization, spurred on by the increasing importance of data for technological innovation

Some data localization requirements exist as a legacy of older laws written for the pre-internet age and are intended to cover the storage of physical records in a physical territory, but they are interpreted as applying also to electronic records. But not all data localization requirements are legacy-based in nature. We are suddenly witnessing an increase in intentional data localization, spurred on by the increasing importance of data for technological innovation and general economic growth. Some countries that wish to partake in, or in some cases dominate, the new global data economy are adopting increasingly aggressive measures designed to ensure that either they, or their indigenous businesses, have access to vast quantities of data. Aggressive deliberate data localization practices of this nature can have unintended consequences, including raising the price of entry for foreign investment, disadvantaging neighbouring countries (and thus the region as a whole) and, sometimes more insidiously, leading to a potential copycat effect in other nations. The result is an increasingly fragmented deglobalized international data economy that favours larger nations (which naturally have relatively greater amounts of data within their borders than smaller nations) and risks slowing down economic progress for smaller and emerging nations.

In reality, nations of all size can benefit their economies by partaking in cross-border data sharing because it is not a zero-sum game. Unlike commodities, data is not a finite resource. Data begets data,[4] i.e. insights can continuously be derived through combining and analysing quality datasets in greater quantities. This concept is particularly important for data analytics and machine learning, and a key driver as to why data localization policies often backfire.

Debunking myths about data localization

For governments, international data movement raises legitimate concerns, particularly about security and access:

Protecting privacy of citizens' personal data

Data localization practices are often rooted in a desire to protect the personal data of private citizens, but data localization laws cannot effectively address privacy concerns. Doing so relies on robust country-level data protection legislation and controlling access to data, regardless of where it is stored. For example, a company may be compliant with a data localization requirement and store personal data only within a specific territory, but that does not take into account the guarantee that they will otherwise comply with data protection law. Moreover, easier or mandated government access to personal data when stored in a specific territory may ultimately impede privacy interests. In many cases, such measures are simply misguided, but in extreme cases, data localization laws can actually become anti-privacy laws.

Improving cybersecurity

Risk detection, assessment and response to cyberthreats require robust security controls, rather than geographic locality requirements

The cybersecurity world offers lessons on why data localization and residency restrictions can be harmful and costly: Data security issues can arise from storing all data in one geographical territory, which is contrary to the diversification approach most commonly mandated in the cybersecurity industry and often adopted by multinational companies to ensure robust security across a geographically dispersed network. Risk detection, assessment and response to cyberthreats require robust security controls, rather than geographic locality requirements. In addition, distributed and duplicated data on multiple systems leads to variation in local security measures, lower local investment in security and leaves data more vulnerable to breaches. Similar to concerns regarding data protection, data localization requirements in the name of cybersecurity are often misguided policies.

Securing data availability for national security and law enforcement

Cybersecurity concerns are distinct from national security concerns. Countries deploying isolated or outdated technology are less able to protect their national security against foreign military and criminal threats and may instead benefit from the use of cloud services (which are often more affordable when not made to measure and specifically localized).

Most countries consider selective record retention, secrecy and anti-treason laws sufficient to protect national security interests. That is why very few countries have enacted broad data localization laws so far and international treaties such as the Trans-Pacific Partnership Agreement (TPPA) and the EU Free Flow of Non-Personal Data Regulation[5] expressly prohibit member countries from enacting data localization laws or local data centre requirements except where justified. International cooperation between intelligence and police forces – for example, via INTERPOL and regional cooperation arrangements – render even justified cases of data localization less useful than previous instances.

Case study: Cross-border data restrictions and anti-money laundering

Data localization and residency restrictions can severely compromise the ability to detect and monitor fraud, money laundering and terrorism financing activities. By limiting the flow of data across borders, the process of detecting suspicious activities becomes more complex. "A criminal rejected in one country can open a mobile money account and make transactions in another country."[6]

Protecting domestic industries by compelling use of local data centres

Countries can support their local data economy through inter alia, offering education, carefully packaged deregulation, transparent tax codes and strong intellectual property protections.

Introducing data localization requirements in order to create a market for local data centres or the use of locally made technology is not an effective long-term strategy for boosting domestic industries because it limits the growth of the domestic data economy.

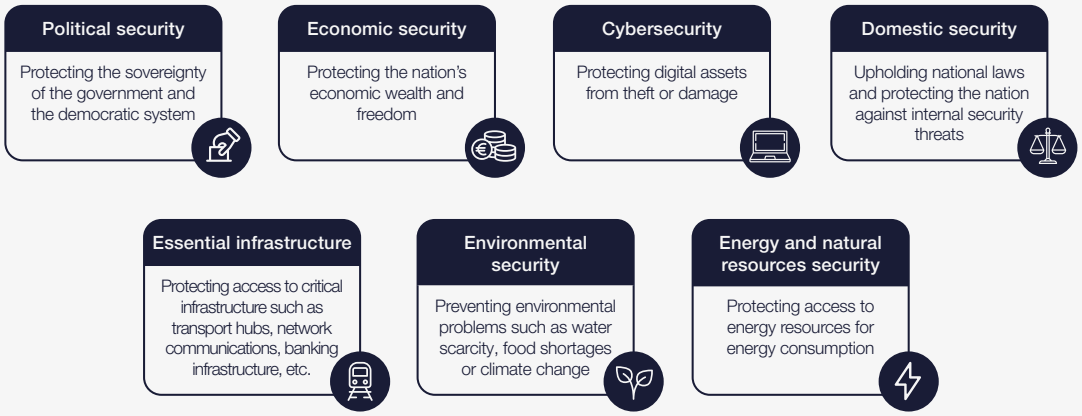
First, local data centre facilities often prove costly and end up not being globally competitive, which slows down local progress in relative terms. Secondly, local industry suffers when other countries retaliate with their own data localization laws or other free-trade restrictions. Thirdly, the lack of access to international markets makes the territory unattractive to foreign investment due to these inherent limitations.

In the same way that countries which allow their economies to trade with partners experience greater economic opportunity, so, too, do countries that allow their economies to participate fully in the international data economy.

Are data localization requirements ever justified?

As mentioned above, every country has the right to secure the infrastructures and assets vital to its national security, governance and public safety. As such, there are legitimate reasons why a country may wish to restrict or scrutinize data entering or leaving its borders.

FIGURE 2 *Government concerns that prompt data localization requirements*



For the purpose of securing government access to data, it is usually sufficient for governments to require companies to guarantee remote access to data (wherever it is stored). There are exceptions to this in the case of hypersensitive data, such as information pertaining to military or defence data. Along similar lines, information that, if accessed by the wrong pair of hands, could take out a national energy grid is data that is beyond the comfort zone of most governments. In line with respecting the sovereignty of nations to protect their national security interests, there are occasions where countries will prefer to insist on localizing data. However, a de minimis approach is recommended here to avoid unintentionally localizing data that is not itself sensitive but which sits alongside sensitive data that does need to be localized. Any requirements should be non-discriminatory, non-arbitrary and should not be disguised restrictions on trade.

Case study: Comprehensive and Progressive Agreement for Trans-Pacific Partnership

The Comprehensive and Progressive Agreement for Trans-Pacific Partnership identifies permissible exceptions for cross-border data flow restrictions in Article 14.11.3 and 14.13.3. Participating members can adopt or maintain data localization measures to achieve a legitimate public policy objective, provided that the measure:

(a) is not applied in a manner that would constitute a means of arbitrary or unjustifiable discrimination or a disguised restriction on trade, and

(b) does not impose restrictions on transfers of information greater than are required to achieve the objective.

How does overreaching data localization really affect supply chains, the workforce, economic growth and global society?

A. Isolation is costly

In fact, most businesses incur extra costs in complying with data localization requirements, both businesses from abroad and local SMEs that would like to use cloud service providers as back-end infrastructure, such as software as a service (SaaS). Those higher fixed costs are ultimately passed on to SMEs, which often have neither the expertise nor the budget to afford their own state-of-the-art mechanisms to store and protect data, making use of the cloud an accessible and effective solution. While foreign businesses look to where their return on investment will be greater, indigenous businesses also become incentivized to locate elsewhere to avoid additional costs, local taxation, local government access to data, and risks of corruption and compliance deficits associated with establishing local presences. Consequently, local companies and consumers lose access to cloud computing capabilities and other advanced foreign information technologies, pay higher prices and become uncompetitive in global markets.

Local labour markets may suffer from withdrawals or reduced job offers from multinationals. Therefore, local consumers and economies lose out – in the form of higher-cost and lower-quality services, as well as lost job opportunities – as a result of the impact of localization requirements on both local businesses and multinationals.

B. It confines e-commerce and supply chains

The overwhelming majority of commercial activities we engage in are virtual, which means they are facilitated by data travelling over fibre-optic networks across the globe every day. Cloud services and particularly SaaS offerings allow businesses of all sizes to access customized enterprise software at relatively low prices. If you prevent data from being hosted outside of your country, most of these services and the technologies that drive them become inaccessible. Data localization can make it impossible for small businesses to get up and running, and it will be impossible for them to scale if they cannot benefit from the economies cloud services provide.

Access to e-commerce, which inherently relies on the flow of data, has helped many developing economies, including countries in Africa, to grow at great pace, as access to mobile methods of payment, access, education and business enables individuals to sidestep infrastructure challenges, and access services directly and affordably. This pace of progress could be quickly unwound were they to introduce arbitrary data localization requirements.

Supply chains are another example. Highly distributed and specialized in nature, these are in many ways the litmus test for market economics: Supply chain nodes that are no longer relevant quickly die, so we can expect supply chains to reasonably reflect real demand for goods and services, at least over time.

Supply chains are so finely balanced and sophisticated that doing business usually involves interacting with many niche players, many of whom will be distributed globally – and, practically speaking, that requires data to move. Even simply buying goods and services from different places around the globe requires the movement of data.

If you want to buy high-quality industrial ball bearings from Germany for machinery on your factory floor, you must contract with a German supplier. Your German supplier will have sales representatives and engineers who can recommend which ball bearings will work best for you. You will keep their contact information in your vendor management system – a system that is almost certainly electronic and probably located in your vendor's cloud. Your vendor's cloud is powered by infrastructure providers who pass that data back and forth in their data centres and across borders to ensure performance and prevent service interruption or failure. These data centres may be located outside your country and even outside the country where your vendor does business.

One might think that something as small as a micro-sized part of a smartphone may not matter much, but in fact industry is hyperspecialized – and that's a good thing. Performance improves in all of the products and services affected by it, but it also requires commerce to move, and commerce can't move without data moving, too. Every action we take online is tracked and recorded. You can't create an ordering document, make a shipment, record a payment or issue a receipt without data – business, shipping, financial and often personal data moves around the e-commerce circulatory system continuously.

C. It stifles talent

In this hyperconnected, increasingly specialized world we live in, talent matters, is hard to find and is unique. The flight of graduates out of countries that offer few opportunities internally due to localized measures, usually never to return, has a long-term impact on the development of a country's expertise and economy.

At the international level, if you want to develop software in the US or the EU for use in the Middle East, you need experts on your development team who understand the regional languages and cultures. You may want to increase your investment in local talent in the region to do so, and hire them directly or through a third party, but in any case, you will have to onboard them, train them and work with them locally (and often virtually).

This will require you to transfer their personal data and your proprietary data in and out of the country – to their teammates, managers, your customers, vendors and many others. If employee personal data can't be hosted on cloud servers outside their country of residence, how will you gain the local talent you need? Moreover, how will the local talent be available for opportunities if the market is closed to outside vendors? The impact on the local workforce will be profound.

It is not seriously questioned any more whether remote workers are a critical segment of the workforce, and to work remotely, data must travel – financial data, business data, design data, health data, etc. It is all coursing along the information superhighway. Data localization greatly disadvantages the remote worker, foreclosing opportunities for professional and economic growth.

How countries instil data localization requirements

More than 200 countries around the world have enacted data laws of some description,[7] with many similarities and some significant differences between them.

According to data residency and data retention laws, companies must keep data for certain minimum time periods and on national territory to ensure that government authorities can compel access.

Data processing regulations with cross-border data transfer restrictions originated in Europe in the 1970s and have been adopted by more and more countries around the world. These include both broad-brush data residency requirements as well as narrower data retention and residency laws pertaining to communications metadata only.[8]

Countries that require data residency usually also restrict the use of legal instruments that allow for the contractual movement of data across borders between two entities or more (i.e. what is known as cross-border data transfers), but the reverse is not true, e.g. countries may restrict the contractual use of cross-border data transfer mechanisms and yet not have any explicit data residency rules. Data transfer restrictions and data residency requirements are conceptually different. Under data residency laws, companies must process data primarily in a particular territory, but they can also transfer copies of the data abroad. According to cross-border data transfer restrictions, companies must not transfer data to another country except in cases where they can assure adequate safeguards for the transferred data abroad; if companies can meet the requirements for an exception, they are not required to keep a local copy of the data.

Examples of data transfer mechanisms include binding corporate rules (BCRs) or standard contractual clauses that are discussed below in the "Data protection and privacy" section.

Focusing on access

"Ultimately, policy that is in favour of cross-border data access is usually pro-innovation, pro-economic growth and, frankly, pro-people

Ultimately, policy that is in favour of cross-border data access is usually pro-innovation, pro-economic growth and, frankly, pro-people. Governments will righty have concerns about backstops and safeguards to doing so, which are the subject of discussion throughout this white paper. By ensuring that data flow is the default state, governments can concentrate their energy on identifying those very high-risk scenarios where they do consider it appropriate to localize data.

One possible solution to staving off data localization in the data-centre space is the data jurisdiction law that Bahrain has introduced – where foreign governments maintain their jurisdiction over data stored in Bahrain-based data centres. This innovative solution to cloud computing manages to create a level of comfort for governments as the data is not technically stored in Bahrain for legal purposes, even if it physically is. If we consider that cyberspace is everywhere and nowhere, then what ultimately matters is access to the data.

Case study: Bahrain's data jurisdiction law

According to the Legislative Decree No. 56 of 2018 in Respect of Providing Cloud Computing Services to Foreign Parties, data of government and business entities stored in data centres in Bahrain is subject to the exclusive jurisdiction of the foreign state in which the entity is domiciled, constituted or established.

To facilitate cross-border cooperation between authorities, the law allows foreign public authorities to issue binding orders to provide access and disclosure of the data, or requests to preserve or maintain the integrity of the data, as per the laws relevant to the foreign state.

2

Establish a level of data protection

Policy recommendations

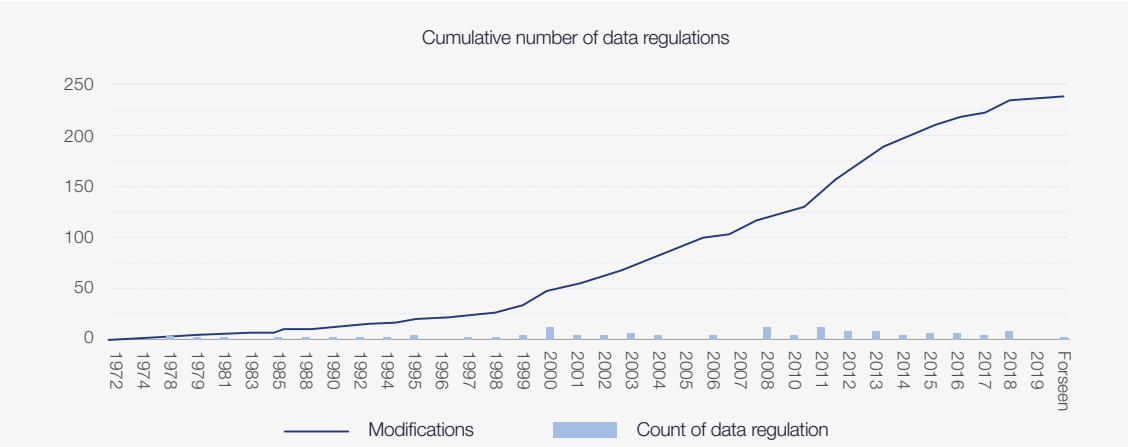
- Participating governments should be required to have national legal frameworks in place that protect the data of individuals, e.g. a data protection law.
- Cross-border transfers of personal data should generally be permitted under national laws.
- A clear cooperation mechanism between national authorities needs to be established to enhance trust and allow for regulatory compliance across borders.
- Compatibility or policy interoperability between data protection and privacy laws is encouraged to ensure certainty and security.
- Governments should investigate the possibility of reaching explicit agreement on the adequacy of other countries' data protection and privacy regimes where the respective legal systems offer substantially similar privacy protections so as to create a common space for the movement of personal data.
- Lawmakers should encourage and enable secure data sharing and focus legislation and law enforcement on abuses such as cybercrime, fraud and harmful discrimination.
- If lawmakers enact broadly applicable privacy laws to define baselines, they should be technologically neutral so as to remain future-proof.

Data protection and privacy: Why it matters for cross-border data sharing

This Roadmap is designed to cover the principles behind the movement of all types of data as it flows across borders. However, personal data or personally identifiable information is a subset of data that is already highly controlled in its cross-border movement. In fact, a significant amount of data qualifies as "personal data" under EU data protection laws and as "personal information" under newer data privacy laws in the US, including the California Consumer Privacy Act. As a result, restrictions on cross-border transfers of personal data affect most data transfers in practice. Below we discuss how the various methods currently in use might be streamlined to extract a fit-for-purpose version of the cross-border flow of personal data.

Mobile phones, fitness trackers, connected cars, medical devices, industrial machines, toys and other IoT devices already generate vast amounts of data and information. The total amount of stored data worldwide is expected to reach 175 zettabytes by 2025.[9] Unsurprisingly, the number of corresponding data laws has exploded in exponential terms in recent years, as seen in the following graph.

FIGURE 3 *A growing number of data regulations*



Source: OECD

28%
of countries do not have data protection and privacy legislation in place

Data protection and privacy laws governing the collection and processing of personal data and personally identifiable information vary from country to country, from the highly sophisticated, such as the EU's General Data Protection Regulation (GDPR),[10] to some emerging market jurisdictions that lack any explicit data protection laws. Almost 72% of countries have full or draft legislation to secure the protection of data and privacy. The 28% of countries that do not have data protection and privacy legislation in place face the risk of missing out on the benefits of cross-border data flows, digital trade and investments in emerging technologies.

FIGURE 4 *Core principles of data protection and privacy*

Under these data protection and privacy laws, organizations usually face restrictions and obligations regarding the collection, use and transfer of data relating to natural persons (personal data). Data protection and privacy laws do not apply to aggregated information, irreversibly de-identified data or data that does not relate to individuals.

The core principles of data protection and privacy as illustrated in Figure 5 above tend to remain fairly consistent from jurisdiction to jurisdiction, though some differences do appear. When these differences are significant, a country's data protection law can end up acting as both a hard and soft barrier to the cross-border flow of data.

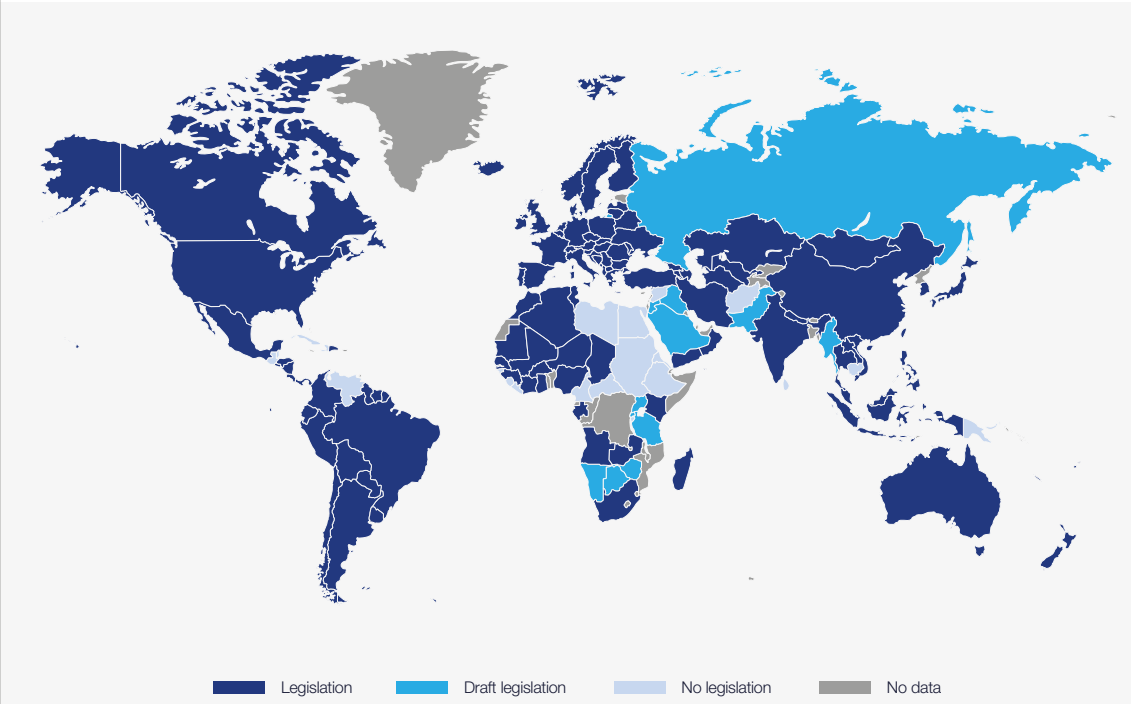
Consider a company resident in Country A that is interested in doing business in Country B. We can assume that doing business will require some sort of cross-border transfer of personal data, such as customer purchasing information.

Asymmetry of approach to data protection between these countries will act as a hard barrier to business in the case where Country A has a robust level of data protection legislation and Country B has a lighter or non-existent level of protection. Because Country A cannot be certain that its citizens' data would be adequately protected in Country B, it may restrict the movement of personal data by the company to Country B.

Asymmetry can act as a soft barrier when both Country A and Country B have robust data protection legislation in place, yet significant differences exist in terms of how those laws are implemented and complied with. Compliance comes with associated cost, and so to minimize costs the company may choose not to do business in Country B, opting instead to stay local or else expand its business into territories whose regimes are more similar to Country A's. The result is that Country B loses out.

Finally, but by no means exhaustively, because of the relative importance placed on compliance when it comes to protecting personal data in a dataset (due to regulatory obligations and other penalties, on top of the importance of protecting private individuals), the entire dataset will often be treated as personal data, even if it contains only a small amount of personal data. Larger companies can often design data stacks so that personal data can be stored separately from other kinds of data, but SMEs will usually lack the resources to do so and may effectively end up treating all data equally in line with the highest standard of compliance required, e.g. as if it is personal data. Thus, laws that are intended to apply only to personal data can, in practice, have the effect of applying to all kinds of data.

FIGURE 5 *Data protection and privacy legislation worldwide*



Source: UNCTAD

Cross-border data transfers

Contractual commitments usually require parties to adhere to core principles of various international data protection and privacy laws

Data privacy concerns in respect of the cross-border movement of data can be addressed by mandating contractual commitments by foreign data importers. Contractual commitments usually require parties to adhere to core principles of various international data protection and privacy laws. In this way we can see how existing laws can act as a form of unofficial standardization for cross-border data transfer agreements.

National lawmakers can allow cross-border transfers of personal data, hold data-transferring companies responsible for any consequences caused, and apply and enforce national laws against foreign companies and public-sector entities. The United States has taken this approach and successfully enforced its laws against companies around the world based on their nexus to US markets and jurisdiction. Other countries find it more difficult to enforce their laws across borders if they cannot rely on cooperation from other countries, and the foreign companies involved have less of a nexus to their jurisdiction.

If a country is concerned that it cannot enforce its data protection or privacy laws against companies abroad, and does not trust another country's practices, it can prohibit or restrict cross-border transfers of personal data. A complete prohibition would result in economic isolation because international trade requires communication, collaboration and thus transfers of personal data. Partial restrictions may be helpful to ensure sufficient levels of data protection abroad, using one or more mechanisms:

- More and more countries require companies to provide notice or seek informed consent from data subjects before their personal data may be transferred abroad. Companies administer the notice and consent process, which adds costs to cross-border trade. While the approach is flexible and leaves the decision to individuals, it places a considerable onus on individuals to understand the data value chain and is not optimal in cases where potentially harmful processes are consented to but poorly understood.

- The EU allows data transfers on the basis of consent only in specific circumstances[11] and generally requires that a data importer outside the European Economic Area (EEA) be located in a jurisdiction that the EU has declared "adequate", adopts industry codes of conduct, implements binding corporate rules approved by a data protection authority in the EU or accepts standard contractual clauses (SCCs) promulgated by the EU. Many multinationals view adopting the complex and rigid SCCs as the least burdensome option, even though the SCCs cover only limited data transfer scenarios

and require pass-through to onward transferees, which multiplies costs and burdens. SCCs work less well for cross-border data sharing at scale, and are not ideal for modern-day cross-border data sharing use cases such as machine learning; however, in current use they are an important instrument for SMEs that lack the resources to build tailor-made legal solutions for the cross-border transfer of data.

– The member states of the Asia-Pacific Economic Cooperation (APEC) agreed on a privacy framework in 2004 and cross-border privacy rules (CBPR) in 2011. As of February 2020, eight jurisdictions have implemented these rules (Australia, Canada, Japan, Mexico, South Korea, Singapore, Taiwan and the United States). Only 23 businesses are listed as participants so far, because none of the member states demands cross-border transfers of personal data on participation.

– Other countries have followed the EU approach, with country whitelists and requirements to execute protective contracts. The whitelisting approach provides countries with opportunities for privacy law harmonization and bilateral trade negotiations. Companies can manage contract requirements better if they are generally stated, e.g. not overly prescriptive, but if every country requires its own contract clauses for cross-border transfers at the same level of complexity and word count of the EU SCCs, then multinationals would have to review and execute millions of pages of contract terms, resulting in an undue burden on international trade. A more efficient alternative to country whitelists is mutual recognition of OECD countries or signatories to the Council of Europe's Convention 108 principles. This places the onus on countries to opt in to such an approach.

– Compatibility or policy interoperability between data protection and privacy laws ensures certainty and security in EU-US Privacy Shield programme and Executive Agreements under the US Cloud Act, such as the UK-US agreement, whereby mutual standards are respected regarding the processing of personal data.

Lawmakers should encourage and enable secure data sharing, and focus legislation and law enforcement on abuses such as cybercrime, fraud and harmful discrimination. If lawmakers enact broadly applicable privacy laws to define baselines, they should be technologically neutral so as to remain future-proof.

Each country must find the right balance for its people's privacy and data needs. The legal, cultural and societal differences between nations and regions mean that wholesale adoption of privacy requirements in one country/region may not work for another in the same way. Therefore, countries should reflect on their needs and requirements before using the expertise and experience of others. Every country should aim for a minimum level of data and privacy protection[12] at a domestic level that can then be operationalized in an international relationship for cross-border data transfers.

Fundamental to ensuring that two different nations with two different data protection laws can act harmoniously is ensuring that, at a principle level, there is a degree of commonality. Once this commonality is either recognized or achieved, companies will find it easier to comply with both regimes when constructing private-level cross-border data transfer mechanisms such as standard contractual clauses. However, there are inherent difficulties with this approach, because it requires case-by-case analysis and bilateral arrangements; as the number of countries participating in the process increases, countries should seek to achieve adequacy between them when it comes to a minimum standard of data protection and privacy. In this way, private data transfer mechanisms may still be required at an operational level, but businesses can effectively treat the two or more countries involved as effectively a singular jurisdiction for the purposes of data protection and privacy. Thus, adequacy is a passport that allows personal data to travel across relevant borders. Mutual recognition of agreed international principles as a way of defining minimum standards for data sharing is therefore more practical for the most ambitious nations.[13]

Although there is no uniformly agreed-upon model for data protection, many countries have been adopting European-style concepts in their data processing legislation, including, recently, Brazil and India (a draft bill is pending at the time of writing). Despite the significant differences, many of those models also share communalities in terms of core data protection principles. These can provide a place to start to achieve harmonization and interoperability and reduce friction over cross-border data flows.

Personal data collection, usage and cross-border sharing will increase – in fact, must increase – to better research and cure diseases; treat patients with personalized, precision medicine; develop AI; enable autonomous cars to recognize and protect people; support global communications; create reliable blockchains; ensure the effective fight against financial crime, modern slavery and corruption; enable firms to manage vendor/supplier risk; safeguard against cyberthreats; and protect national and international security. Personal data transfer minimization and prohibitive regulation writ large is counterproductive to pursuing the many opportunities of data-driven innovation, which is why policy-makers should focus on specific privacy harms and craft legislation that balances privacy and other interests proportionally.

Personal data transfer minimization and prohibitive regulation writ large is counterproductive to pursuing the many opportunities of data-driven innovation, which is why policy-makers should focus on specific privacy harms and craft legislation that balances privacy and other interests proportionally

3 Prioritize cybersecurity

Policy recommendations

- Governments should endorse the concept of cybersecurity as a fundamental condition of doing business in an economy.
- Governments should enact robust data security legislation to position themselves as trustworthy data transfer destinations, including data security requirements on public- and private-sector organizations and data security breach notification requirements.
- Governments should create, support and respect robust data security infrastructures and refrain from demanding data access without due process or technology back-door systems.
- Cross-border data sharing agreements between governments should in turn mandate data security measures.
- Cross-border data sharing agreements should contain an anti-snooping clause, i.e. a clause that forbids governments and connectivity providers from viewing the data being transmitted across borders, except in certain prescribed instances.
- A clear cooperation mechanism between authorities needs to be established to enhance trust.

What governments can do

> Having a civilian cybersecurity agency is key to encourage trusted relationships around the world with other cybersecurity agencies, whose role is solely to protect networks, not to attack them

The security of data when it moves across borders is of fundamental concern to both companies and governments, both in terms of risk mitigation and security of proprietary data and intellectual property (IP). The absence of, or the risk of the absence of, security measures further undermines trust and produces friction for cross-border data sharing.

Appropriate data protection technologies exist, e.g. data encryption and data masking. However, the main challenge is when and how to use these measures to create trust within data sharing agreements at a national level. The project community recommended that cross-border data sharing agreements between governments mandate a minimum threshold for cybersecurity, just as already happens for the trade of goods and services. Physical goods are assessed by common product specifications. e.g. origin and weight. Likewise, minimum thresholds for security can be agreed between governments to enable the free flow of data.

The project community further suggested that cross-border data sharing agreements should contain an anti-snooping clause, i.e. a clause that forbids governments and connectivity providers from viewing the data being transmitted across borders, except in certain prescribed instances. It is a well-established principle that connectivity providers should not access content data in transmission (even though they technically have access to it, system controls can be designed to make this access more difficult). Under the EU's ePrivacy Directive,[14] there are stringent penalties placed on electronic communications service providers who snoop on data in transmission on their networks. The analogy to trade of physical goods is also relevant here: Goods are shipped to prevent unauthorized access but inspected at the port of entry to satisfy local law requirements. They are also labelled to generally describe their content. So too could metadata be tagged to provide information about data content without necessarily making it available for review.

If the metadata is tagged with appropriate content notices and securely transmitted using protocols, governments could reliably permit data transfer, while keeping the payload confidential, and preserve the rights to inspection.

To create a trustworthy environment for cross-border data flows by being a trustworthy international player, countries should consider measures at the national level such as enacting national legislation to require data security breach notification for all types of data; ensuring compensation for data subjects or businesses for actual harm caused by data security breaches; and requiring manufacturers of IT products to make secure products by promoting and supporting investment in good security standards and third-party validation. Local laws should protect organizations against cybercriminals and national state espionage, and governments should consider auditing organizations and enforce laws to reduce security breaches and promote trust.

Establishing a sophisticated governmental approach to cybersecurity

Having a civilian cybersecurity agency is key to encourage trusted relationships around the world with other cybersecurity agencies, whose role is solely to protect networks, not to attack them.[15] In establishing its national cybersecurity agency, a government needs to consider the responsibilities of that agency. The role of the cybersecurity agency can vary greatly, and if at first it will be mainly to react to attacks, gradually it should seek to define effective methodologies, inform the private sector about current threats and have a more proactive role.

The second step is to define a cybersecurity strategy and, in that strategy, define how and where the cybersecurity agency should be set up. In the absence of any agency, a white paper by government can build goodwill and lay a foundation for appetite. Such a high-level document, signed off by the government, shows the intention of that government when it comes to cybersecurity. It is a foundation in the cybersecurity framework of a country, showing that the government has identified cybersecurity as a priority, and explicating how it is intending to protect itself and its citizens. It maps the different relationships at national level between the different entities dealing with cybersecurity.

When it comes to increasing a country's cybersecurity posture, it is absolutely vital to define a framework for critical services by mapping the functions that are critical for the country to function. These vary between countries, but usually cover essential services such as the financial sector, energy, water treatment, the military and telecommunications. For each of these sectors, it is then important to define the thresholds at which companies would be considered critical to that sector. Defining these thresholds allows the

Combating cybersecurity risk relies on fast and effective data sharing, globally

government to define which are the companies that will be considered critical infrastructure operators, and based on this, mandate what these companies – those that are considered critical infrastructure operators – need to be doing to achieve a minimum cybersecurity level, e.g. the minimum security requirements that these companies would need to implement, how many audits they'd need to undergo, if they would need access to confidential communication channels to communicate with the government in case of a major crisis, etc. It also specifies, technically speaking, how certain systems need to be hardened: for instance, having them ISO27001 certified. The crux of the debate in such efforts is the proportion of regulation and capacity building. It is easy for a government to pass a law that all companies should be ISO27001 certified, but this can be impossible for some companies. Working with the critical infrastructure operators to define these measures helps ensure they will actually be willing and able to implement them down the line.

Such a framework also paves the way for other solutions – for instance, connecting all of these critical infrastructure companies to a network so that their security can be monitored by the government, or so that the government can push detection rules to their networks and protect them better.

With these elements in place, all stakeholders need to define a variety of plans to face the different cybersecurity issues that may come their way. In its most basic form, a contingency plan ensures that all critical systems have backup systems and that when something happens, a critical company or a critical government service can continue to operate. Second, a crisis management plan is needed. When an incident escalates into an unknown situation, it is no longer an incident but a crisis, and a specific plan is required to manage the type of uncertainty associated with crises, in particular, cyber crises.

Testing the cybersecurity apparatus is the last step towards creating a solid cybersecurity foundation: exercising all of these entities, plans and relationships. Companies need to be empowered to run exercises on their own, as do governments, and both need to be able to run exercises together with international partners.

This is how trust in a national cyber regime is established.

While both the US NIST Cybersecurity Framework and EU NIS Directive provide good policies and procedures for ensuring trust, it is not necessarily the case that organizations implementing them are following them in practice. Therefore, to enable full trust there must be means for trusted third parties to assess whether the required security controls are in place using hard evidence. This audit process is expensive and time consuming and therefore it is recommended that governments follow standards and certifications that are already widely used by the different industries globally.

Suggested examples of globally used well-established cybersecurity standards that are being continuously attested by trusted auditors include the ISO 27001 and AICPA SOC 2 Type 2, the details of which are outlined in Appendix 1.

Developing internationally recognized standards is effective for creating a trusted level playing field and can lead to faster adoption and a higher likelihood of mutual recognition.

Finally, cybersecurity risk applies nationally and internationally. If companies designated as operators of essential services cannot share data about risks internationally, this results in the creation of pockets of vulnerability. Combating cybersecurity risk relies on fast and effective data sharing, globally. This involves both technical and personal data, and the challenges of personal data transfers adversely affect the effectiveness of measures that nations and businesses can take to safeguard their systems and data.

B Part B: Incentivizing cooperation between nations

4 Hardwire accountability between nations

Policy recommendations

- Cross-border data sharing agreements should encourage cooperation between national authorities in order to reduce risk and ensure a level commercial playing field.
- The complementary role of diplomacy must not be overlooked in establishing goodwill.
- Governments should appoint respective authorities to cooperate with each other to support the free flow of data by the private sector across borders. This is especially important in respect of non-personal data where a network of authorities can complement the existing network of data protection authorities.
- Cross-border data sharing agreements should hold parties, including governments, accountable for the security and confidentiality of the data they share, while making allowances for the review and receipt of data as necessary to comply with local laws.
- Transparency of approach should be encouraged to increase trust with other authorities as well as with the private sector.

Hardwiring accountability

Governments are inherently incentivized to implement policies that reduce national risk. Such risks may include companies or persons hiding data offshore or on a foreign cloud service, for the purposes of e.g. tax evasion or evading law enforcement. In order to increase trust between governments, cooperation and accountability mechanisms must be built into any and all data sharing agreements.

Concerns arise on the part of both governments and the private sector when lines of private legal recourse break down across international borders. This is less of an issue when it comes to recourse for theft or the mishandling of proprietary data as the private international legal space is well built out (with some exceptions). Public international law is also well established when it comes to governments sharing information with each other. The complexity arises when these worlds merge, which is the case when dealing with big data or data on the scale that is useful for machine learning, AI development and other advanced use cases such as industrial IoT.

Private stakeholders who share and trade datasets with each other across borders may do so in the

absence of any laws to the contrary. However, in a world where economies that have access to vast quantities of data may have a strategic geopolitical advantage over others, governments take an interest in the relative advantage of their stakeholders in these systems. In this respect, formalizing or recognizing recourse actions for data sharing mishaps could be hardwired at an intergovernmental agreement level so as to assure all stakeholders, including the governments themselves, that the cross-border data ecosystem is on a level playing field and that rights are fairly rewarded.

"Well-established international systems such as currency exchanges teach us the importance of nominated entrusted national authorities when it comes to encouraging good behaviour and securing recourse for bad behaviour

Well-established international systems such as currency exchanges teach us the importance of nominated entrusted national authorities when it comes to encouraging good behaviour and securing recourse for bad behaviour. Currently, such a system is in its infancy when it comes to big data sharing: The closest we come is the network of data protection authorities, which are only informally institutionally harmonized between the countries where they exist, and that have authority only over personal data. In the spirit of realism and assuming a new network of "data authorities" will not spring up overnight, governments could be asked to nominate existing authorities who will act as points of contact on cross-border data sharing for non-personal data. A similar approach has been established in the EU under the Free Flow of Data Regulation. This approach allows for recourse for bad behaviour either between private stakeholders in respective jurisdictions, or as an accountability mechanism for government itself when it institutes justified data localization requirements. This network of authorities can then complement the existing remit of the established national data protection authorities. It can also complement diplomatic efforts, which should not be underestimated in creating goodwill.

At a practical level, technological solutions themselves can lend solutions. Blockchain allows for a hash of data to be stored and moved across borders, rather than the data itself. This offers a high level of security and can, in certain use cases, allow core data to stay on local servers. Blockchain can be used to tokenize data, which can make data ownership, data travel, data velocity and data uses more transparent. Smart contracts can even be used to hard-code agreements between governments. Overall, blockchain can make it easier to audit data transactions and agreements and can ensure the immutability of data (provenance).

Case study: Governance can steer a decentralized future

Innovation that progresses without sufficient consideration for governance and user protection can lead to undesirable outcomes for individuals, companies and societies. For example, blockchain technology, a pillar of the Fourth Industrial Revolution, can not only unlock radical improvements across the public and private sectors, but also enable new business and governance models that enhance security, accountability and transparency for people worldwide. Carefully regulated, blockchain can benefit the widest number of people in the fairest way. However, fractured blockchain systems, with a hyper focus on efficiency rather than transparency, risk losing the trust they can create.

The World Economic Forum recently released a series of principles to help safeguard the promise of this technology, providing a baseline for designing systems that preserve the rights of its users. The underlying aim is to bring greater accountability to the systems that power our societies.

The Presidio Principles[16] establish guidance on the following rights:

Transparency and accessibility – the right to information about the system.

Privacy and security – the right to data protection.

Agency and interoperability – the right for individuals to own and manage their data.

Accountability and governance – the right for system users to understand available recourse.

The importance of accountability for personal, proprietary and sensitive data

When it comes to cross-border transfers of personal information specifically, accountability should operate as the default position. This policy position is warranted both by the high level of community concern attaching to such transfers of personal information and the nature of the risks associated with such transfers. As for the rights and liabilities of the individual stakeholders, it could be argued that data exporters should remain liable for breaches of privacy by data importers under most circumstances. The data exporter is likely to be the entity that is more easily accessible to the data subject. Individuals should not be deterred from exercising their rights due to a perceived or actual impracticality of engaging with a foreign entity acting as the data importer. For example, the APEC Cross Border Privacy Rules (CBPR) approach-based laws recognize that global data flows are facilitated if the laws focus on ensuring that local companies are accountable for data processing activities.[17]

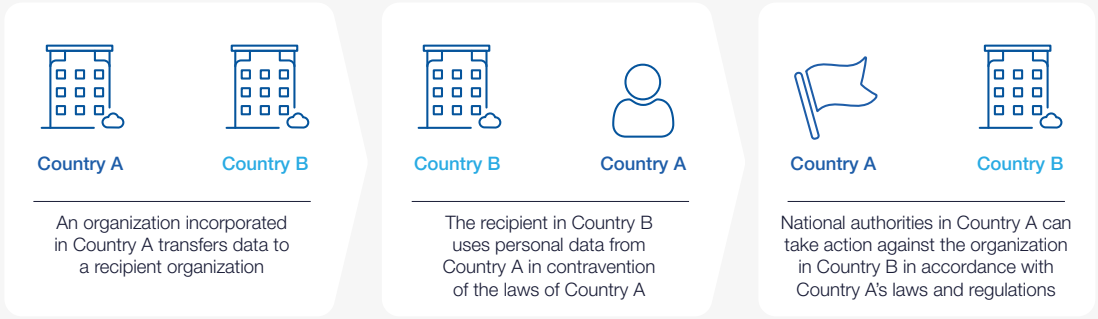
The APEC CBPR also provides a framework, with three circumstances when an agency or organization should not remain accountable. These are when the:

1. Information is subject to a law, binding scheme or contract that effectively upholds privacy protections that are substantially similar to the equivalent in the other country

2. Individual consents to the transfer, after being expressly advised that the consequence of providing consent is that the agency or organization will no longer be accountable for the individual's personal information once transferred, or

3. Agency or organization is required or authorized to transfer the personal information by or under law

FIGURE 6 ***Basic country-level bilateral accountability model***



The general principle of accountability should mean that an agency or organization will be responsible under local equivalent law for the acts and practices of a recipient of personal information that is the subject of a cross-border transfer. That is, where an agency or organization incorporated in Country A transfers information to a recipient in Country B, if the acts or practices of that recipient in Country B in respect of the personal information from Country A would have amounted to an interference with the privacy of an individual if undertaken in Country A, they should constitute an interference with the privacy of that individual for the purposes of Country A's laws. Further, the acts or practices of the recipient in Country B should be taken to be the acts or practices of the relevant agency or organization for the purposes of country A's law.

In theory, when it comes to the movement of proprietary data, the rights and obligations between commercial actors are usually detailed in private commercial contracts, rendering the regulatory lift less necessary than for personal information. However, it goes without saying that countries should adopt policies at the local level which allow for fair and reciprocal access to legal action in the case of commercial disputes.

When it comes to other sensitive data, such as data that could undermine public or national security, as outlined in the data localization chapter of this white paper, sovereign governments may rightly have an interest in restricting the movement of that data beyond a certain jurisdiction. However, as also stated above, it should only do so when, on balance, the risk of allowing that data to move is greater than the risks associated with preventing that data from moving. Moreover, a number of alternative arrangements already exist that help nations mitigate some of the national security risks associated with cross-border data flows, such as INTERPOL, Europol, CEPOL and other regional cooperation arrangements. Many countries choose to enter bilateral and multilateral agreements to solve the operational challenges of cross-border law enforcement. Such agreements are designed to facilitate law enforcement access to data across borders and define mechanisms for coordinated action.

Case study: Cross-border data access for criminal investigations

The UK and the US signed a Bilateral Data Access Agreement to allow law enforcement agencies to gain access to digital evidence held by technology service providers located in these countries. Under the new agreement, law enforcement agencies can bypass the pre-existing Mutual Legal Assistance Treaties (MLAT) process, whereby agencies need to submit requests through the central government of the other country; instead, those agencies can request information directly from the data-holding service provider.

There are, however, limitations to this agreement. Law enforcement agencies will not be able to request data related to residents of the other country (i.e. UK authorities may not request data related to US residents). Moreover, the agreement does not require data-hosting service providers to turn over decrypted data to law enforcement agencies.[18]

Companies could be compelled to seek a derogation for the cross-border transfer of certain non-personal datasets that have been deemed to be highly sensitive by certifying that security and access rights be adhered to, and ensuring that this data is processed offshore or in the cloud for a specific purpose, which is specified at the time of the application. However, assuming other safeguards are in place, such as a minimum level of cybersecurity and tightly adhered to access controls, adhering to the conditions of such a derogation should be relatively straightforward.

Trust mechanisms that encourage cooperation and disincentivize protectionist behaviour

- Governments should appoint respective authorities to cooperate with each other to support the free flow of data by the private sector across borders. This includes cooperation between data protection authorities, competition authorities and law enforcement authorities. The authority taking on this role should ensure that the policy developed through this interaction is consistent with other national policies, and that there is guidance for private-sector players on how cross-border data protection policy interacts with other industry-specific policy. Otherwise, compliance becomes a more burdensome task for businesses.
- Governments should review legislation and regulations that restrict the ability of domestic organizations or individuals to cooperate with foreign authorities in investigations and law enforcement matters.
- National authorities should facilitate a mutual understanding of how national enforcement systems operate in their jurisdictions by sharing information on procedural rules and regulations.
- National authorities should provide each other with the relevant data and information necessary to take the appropriate actions with respect to investigation proceedings. Data sharing should be undertaken on a case-by-case basis between the authorities, and it should cover only information that is relevant to an investigation or proceeding.
- National authorities may choose to impose conditions restricting the further dissemination and use of the data shared with foreign counterparts, or exchange confidential data using confidentiality waivers.
- National authorities should support each other on a voluntary basis by providing investigative assistance as appropriate.

5 Prioritize connectivity, technical interoperability, data portability and data provenance

Policy recommendations

- Governments should prioritize the development of connectivity infrastructure as a prerequisite to building a local data economy.
- Governments should collaborate to develop cross-border data sharing agreements that support similar minimum levels of national and international bandwidth and/or coordinate spectrum usage in order to minimize costs, increase reliability and enhance redundancy (optical fibres, satellite earth stations, IXPs, etc.)
- More ambitious like-minded countries should consider common policies with regard to the deployment of 5G networks, as well as coordinating access to high-performance computing in their data sharing agreements.
- The use of open or standard application programming interfaces (APIs) for data sharing should be encouraged by governments to improve technical interoperability. However, governments should stop short of mandating specific standards that could hinder novel approaches.
- Data portability at the B2B level should be facilitated both domestically and internationally, particularly with a view to supporting start-ups and SMEs.
- Cross-border agreements should contain reference to data provenance and place the onus on data publishers to ensure the integrity of data before it crosses borders in order to avoid bad outcomes for machine learning or contaminated data lakes.

As well as legal and policy barriers, businesses may experience technical difficulties in moving their data or combining it with collaborators in new territories, and/or using offshore cloud services. A meaningful action governments can take is to incentivize the streamlining and standardization of technical interoperability processes for data movement and sharing, including disincentivizing vendor lock-in. Underlying these policies should be a move towards increased connectivity infrastructure to support and ease real-time cross-border data sharing. Policy-makers and technology providers must work together to optimize the level of data interoperability necessary to meet their objectives.

Connectivity, 5G policy and high-performance computing

The setting of minimum standards for connectivity, including the harmonization of spectrum bands, is a vital step to facilitate data flows. When all players are operating on a minimum (and ideally ambitious) level of connectivity, industry is supported through the removal of connectivity speed bottlenecks both domestically and at cross-border levels. Domestically, high-speed connectivity ensures that SaaS solutions become a more viable option for businesses.

For those countries and regions with advanced ambitions, very high-speed connectivity, 5G and high-performance computing (HPC) offer additional areas of optional harmonization to pave the way for the real-time processing of cross-border data at scale for advanced use cases such as machine learning or AI development. 5G networks, which have perhaps their strongest use case in industrial IoT, can be local in nature, but used cross-border to improve outcomes (due to access to more data), including management of regional resources, or be repurposed for AI development. When such vast amounts of data reach the point where high-performance computing may be required, regional players can collaborate to gain access to, or invest in, these hyper-expensive state-of-the-art machines. In the meantime, governments can facilitate networks of edge computing, both locally and cross-border, to harness the opportunities of computational power at scale.

For 5G builds specifically, encouraging private network investment, including new market entrants, is vital given the relative expense of such a network. Indeed, wealthier nations have responded to the challenge by doing exactly that. Again, economies of scale matter here: A network provider will likely find a region to be a more attractive deployment prospect than a nation on its own. The ability to share data across borders and network infrastructure roll-out at scale go hand in hand.

Technical interoperability

“Fully maximizing the value derived from combining datasets, whether using basic algorithms or AI, usually requires the information to be harmonized, standardized and stored in structured databases

At the network layer, technical interoperability is already an area of international policy harmonization. However, the same cannot yet be said at the application layer.

Technical interoperability for data is defined as the ability to share data between different systems and to enable those systems to make use of the data. Fully maximizing the value derived from combining datasets, whether using basic algorithms or AI, usually requires the information to be harmonized, standardized and stored in structured databases.

Historically, data has been collected and held by a multitude of organizations at the local, national and multinational level across the globe. In many cases, this data is either unstructured or structured in idiosyncratic ways that make it difficult to use cross-functionally with other databases. It lacks a common form or expression, making it hard to use horizontally across varying industries such as healthcare, epidemiology, agriculture and supply chain management as well as the data-intensive use cases of the Fourth Industrial Revolution such as AI and IoT.

Governments can nudge behaviour here by incentivizing companies in their respective jurisdictions to adopt similar approaches to technical standards for data, without necessarily prescribing specific standards, so as to remain future-proof.

Interoperability concept frameworks

Key to any interoperable technical framework is the joining and merging of data from different systems, without losing meaning. The expectations are that data interconnectivity and interoperability should be smooth and seamless when different systems deliver data to those who need it in the form they need it in.

Interoperability can mean different things to different systems. Many National Statistical Offices (NSOs)[19] are adopting open data[20] policies. When they publish statistical data openly, it is vital to identify the needs of different consumers. For example, an analyst may use the datasets in a machine-readable format to test a hypotheses and make predictions; a developer community may need to access the data through an API, and build dashboards, maps and visualization tools; policy-makers may want to access the information through a web search or human-readable reports. Each of these use cases requires different levels of interoperability.

Looking at the challenge from a broader perspective, there are two main types of interoperability:

1. Syntactic interoperability requires multiple systems to communicate and exchange data regardless of the different programming languages.

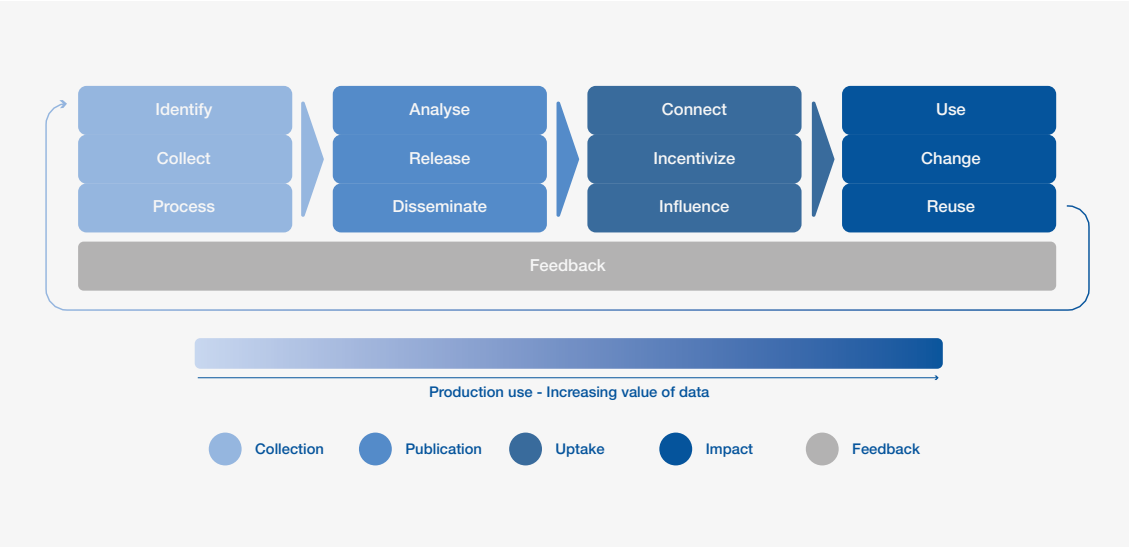
2. Semantic interoperability requires a discrete system to understand and enable the meaningful use of shared data or resources by individuals, organizations and public services.

Depending on the use case for sharing data, the ecosystem could benefit from either syntactic or semantic interoperability. While the real world looks like a syntactic system of dispersed nodes and multiple use cases, in reality most cross-border data sharing takes place in specific use cases and for specific purposes. Semantic models can therefore provide key insights as to what technical interoperability for a specific use case might look like. In addition, AI solutions exist to manage semantic interoperability and governments should be open to facilitating the same.

Open Data Watch's Data Value Chain illustrates the importance of recognizing where in the data life cycle a use case sits so as to optimize interoperability appropriately; the point on the value chain at which the user sits will give clues as to the type of interoperability solution required by them.

As per the model, data increases in value as it moves through the value chain, and can be continuously recycled. In order to maximize such value, both technical interoperability within the system and a level of standardization of the data are required. While this type of model is ideally suited to a data flows arrangement for specific use cases, it demonstrates why interoperability is key to amplifying the value of data in a cross-border data flows scenario.

FIGURE 7 *The Data Value Chain – Open Data Watch*



Data standards

Data standards are a vital component of technical interoperability, ensuring that the data elements and metadata are reusable across different systems in the data value chain. Standards make the data liquid in the sense that value can readily be harvested from standardized and harmonized data. The US Environmental Protection Agency (EPA) data standards include documented agreements on representation, format, definition, structuring, tagging, transmission, manipulation, use and management of data.[21]

Adhering to standards is an essential component of achieving and improving data quality for decision-making. Data standards also ensure consistency in code set use by providing for the maintenance and management of permissible code sets. Appropriate standards must also be flexible enough to accommodate future iterations that occur in response to the changes happening in the data policies, new sources of data and other changes in the community.

Examples of two international standards organizations are the World Wide Web Consortium (W3C)[22] and the International Organization for Standardization (ISO).

For different entities who work with stakeholders to develop common and appropriate technical standards that are regularly updated through a process of expert consultation, governments should not only work towards facilitating the standardizaion of data and increasing the interoperability of data, but also open up access to tools and solutions that enable their internal workforce to work with the data. Open data standards are vital to this approach and should be encouraged.

API interoperability

Application Programming Interfaces (APIs) offer a sophisticated way to make data resources accessible over the web to various applications and users. When APIs behave predictably, they reduce errors and handle users' requests securely and repetitively. The OpenAPI Specification has emerged as the standard format for defining the contract between client applications and services exposed via APIs, making it easier to orchestrate applications as collections of loosely coupled services, each of which supports self-contained business functions).[23]

Integrating legacy system APIs or converting legacy APIs into the new Open API framework standards is costly: Some organizations need to manage multiple API versions or customize services to specific applications. An alternative approach to improve API interoperability is to create an API "middle layer" or "API mashups". These services draw data from multiple APIs (legacy systems or multiple services) and repackage them as a new API endpoint for clients. Such API aggregations improve the integration experience.

Data portability

Data portability, which is the ability to port data from one system to another, is an issue that is top of mind for SaaS customers who may wish to switch services but can be impeded from doing so by back-door data localization restrictions such as vendor lock-in. Lock-in can occur when pricing models penalize or disproportionately price the removal of data from a system, when the physical network infrastructure is not fast enough to allow for a real-time switchover to a new system, or when unfair contractual clauses relating to, for example, bundling, make it difficult to leave the current system. Vendor lock-in can block the movement of data both domestically and internationally and furthermore acts as a barrier to entry for new market entrants. There are further implications for the construction of data lakes that try to centralize data from different sources. Governments can encourage data portability by both disincentivizing vendor lock-in practices and supporting interoperability standards.

Data provenance

Data provenance identifies the origin of the data processor and data owner, and documents a record of the history of the data since collection. Establishing and maintaining provenance protects the authenticity of data.

Blockchain has the capability to document the origin and complete historical record of any type of data in an immutable or tamper-evident record. Every instance of data changing hands or going through any type of operation is traceable.

All modifications to the data could be mandated or required to achieve the consensus of 51% of stakeholders prior to approval. Smart contracts would manage the approval process and document these activities on the blockchain. Changes that were not documented properly or that did not receive consensus approval would be discarded.

Another example of establishing provenance is in the personal health data space, with this type of data often residing in an electronic health record (EHR) system. There is minimal risk in assigning provenance based on an individual initiating a connection to their personal data files, in order to import data from a password-protected health institution. Once provenance is established with EHR data, it is efficient to assume that for any additional data uploaded by that individual, provenance is therefore effectively already established. This approach, using direct or indirect data owner validation, could also be applied to large established institutions sharing proprietary data.

Establishing provenance for de-identified data or other types of data that lacks historic information on its origins is difficult, often impossible. In these cases, it would be beneficial to designate the data as lacking provenance, enabling the users of that data to consider the potential risk to quality when deciding on appropriate data usage.

The majority of data publishers are secondary data producers. As a result, in some cases the journey of a single data point from its origin to its final destination is unclear to the end user, and therefore provenance cannot be easily established. To resolve this issue, the metadata should provide a machine-readable map that makes this information available and traceable across platforms and data producers. When data providers apply semantic web technologies to publish the data, this drastically improves the ability to derive insights from the data, integrate the decentralized data and easily relay the information over the web.

Case study: Blockchain transparency in the diamond industry

For many decades, the diamond industry has operated along opaque supply chains, with data commonly lost, manipulated, suppressed or destroyed. More recently, mining companies, manufacturers and retailers have taken steps to become more transparent about the provenance of their diamonds and jewellery, in response to customer demand for ethically and sustainably sourced stones.

Blockchain has emerged as a secure technology for flowing data in an interoperable way. Major stakeholders such as the Gemological Institute of America (GIA) and the jeweller Chow Tai Fook now register the details of their diamonds on the Everledger platform. By combining blockchain technology with AI, IoT and nanotechnology, Everledger creates a digital twin of every diamond, enabling traceability in a secure, unalterable and private platform.

By sharing provenance data securely, transparency becomes like a two-way street. Information flows securely upstream, carrying insights about the origin and characteristics of the diamond or gemstone. Eventually, the customer at the head of the chain can make their valuable purchase on the basis of increased knowledge and a more thorough understanding of the value of the piece.

Information is also sent back down the chain to help all stakeholders make better decisions. The overall impact is higher clarity in a complex supply chain, which results in closer adherence to the aims of the United Nations' Sustainable Development Goals (SDGs), whether it be gender equality, decent work and economic growth, or responsible consumption and production.

“ While checks and balances are necessary to maintain data integrity, provenance cannot always be easily established

In summary, governments should ensure that they understand the interoperability space and why it matters to ensure that data sharing is nudged towards international interoperability, without specifically mandating the introduction of any one technical interoperability standard. In this manner, the policy environment can remain supportive to technical interoperability without prescribing it with any specification, thus leaving the window open for flexibility among private-sector actors.

In addition to data management interoperability, backbone infrastructure connectivity (e.g. highspeed broadband or 5G wireless infrastructure) allows for a foundation of data flows activity domestically. Lack of connectivity means that countries cannot be ready to participate fully in the data economy in the first instance, so countries would be wise to prioritize backbone telecommunications infrastructure as well as international connectivity to ensure they can compete fully in the Fourth Industrial Revolution. Given that network investment is expensive, forward-leaning policies such as capitalizing on the economies of scale offered by regions rather than individual companies can strongly encourage private capital investment in connectivity infrastructure such as 5G.[24]

Furthermore, governments should be cognisant of the risk of inaccurate data crossing borders. While checks and balances are necessary to maintain data integrity, provenance cannot always be easily established. Policy-makers should take a balanced approach to minimize the risk of data integrity concerns, without restricting the movement of data across borders. Meanwhile, companies can use reliability and/or validation studies when data is exchanged between entities. These types of studies are highly common in scientific research and will likely become more common in the data sharing space.

C

Part C:
Future-proofing international data sharing policies

6 Future-proof the policy environment

Policy recommendations

- **Cross-border data sharing agreements should explicitly recognize the possibility of alternative models (such as federated learning models and data trusts) that can also fulfil the spirit of cross-border data flows.**
- **Leaving the door open to future models ensures that policy in this space remains future-proof and robust.**
- **Newer, more sophisticated models can be helpful in establishing trust across borders and their use should be evaluated on a case-by-case basis.**

“Sometimes the algorithm itself can travel, rather than the data

Technological solutions exist that allow data insights to flow between entities/countries without the data leaving the local server. Local data is tagged or referenced and the reference can then be shared. The reference may be completely random or may contain minimal information about the original data and may therefore be considered to be less sensitive or be classified as a lower-risk piece of data. The data reference can be used by the country of origin and, if necessary, it can be used to make a request to the data repository in the receiving country. Each of those requests can then be authorized by the local country. Innovations such as this are especially useful for machine learning on sensitive data. Indeed, sometimes the algorithm itself can travel, rather than the data, and in doing so the algorithm learns by moving from dataset to dataset, while the data itself remains on site wherever it is stored. The result is that it is the insights from the data, the metadata or the IP that move.

Techniques such as the above open the door to a myriad of new options and can increase comfort levels when it comes to data sharing. But what would happen if and when a government decides to restrict the movement of the hash, the algorithm or the insight (IP) from the data rather than the data itself? Such a scenario would undermine data sharing policies at the cross-border level, potentially acting as a back door to backward-leaning protectionist policies and more importantly would reduce the value of the insights that may be gleaned from a critical mass of data.

The solution would seem obvious: Intergovernmental agreements for data should not only cover the movement and protection of data, they should also recognize and protect the free movement and protection of proprietary algorithms, their distributed nature and value.

Below, we explore how these models work and why it is in our view necessary for governments to offer affirmative cohesive policy frameworks that consider such models.

Federated learning models

The solution to enabling broad data sharing policies lies in part in understanding who the true data rights holders are – whether they be governments, businesses, institutions or individuals – and their needs for data to move across any and all borders in a way that protects those rights and interests. This becomes ever more complex when it comes to trawling datasets for machine learning purposes, effectively creating a distributed data lake. It is usually the case that more data leads to better learning outcomes, provided that data is accurate and somewhat useful to begin with. Even if a country finds that it has "enough" data within its own borders to achieve meaningful learnings by algorithms and is thus not highly incentivized to develop proactive cross-border data sharing policies, it is still often the case that the relevant insights from analysing the data lake, or the algorithm that is developing and refining on the basis of those learnings, would be improved through access to data in another jurisdiction. An obvious situation where this might be crucial is in the analysis and investigation of genetic markers for rare disease or pools of biodiversity genomic data across entire biomes such as the Amazon Basin to mine for novel protein and biological molecules with high value to society and the economy. Such an approach increases the intelligence of end-user applications both at the edge and system-wide.

Considering the real scaling potential of federated data learning, a systems-view approach is needed to consider how to best manage federated learning at an international level, including attributing rights appropriately. This approach enables the application of machine learning to a data pool that can exist in many locations simultaneously, including across borders, while also protecting against privacy breaches.

Data trusts

A data trust is a system or entity that manages the rules of the game in a created data ecosystem, particularly the contributed data on behalf of the data suppliers. Trust ownership rights can be exchanged for data contributions by the data supplier. Compensation to data suppliers can accrue because interested parties such as governments and businesses want the insights derived from the data inputs. Depending on the scenario, they may pay for access and use.[25] Data trusts are often enabled by federated learning techniques, as above, as well as by private computing architectures. Data trusts may exist somewhere or in many places, usually determined by where the data suppliers (inputs) are located.

Data trusts do not need to be international in nature. However, to enable countries that wish to capitalize upon innovations and research learnings from data trust models that require access to data in different jurisdictions, it is essential that such activity does not inadvertently become illegal. Currently, international legal norms on cross-border data trusts are simply not explicitly accounted for in many instances, putting their legality in question.

Should governments wish to explicitly authorize or encourage the use of an international data trust framework, governance rules and a fit-for-purpose system-wide architecture would address privacy protection, transparency and end-to-end traceability as well as the data supplier's ability to exercise rights. Of course, underlying all of this is the fact that such a framework must be allowed to happen legally, whether explicitly mandated or not.

An alternative model to an international data trust is a series of national data trusts that reside in different jurisdictions. Under this model, multiple trusts may be hosted within sovereign borders, when required by local regulations or other key requirements (which, as argued earlier, should be discouraged in general and applied only in the narrowest justifiable circumstances, such as for national security concerns) and federated as part of a single system. This type of solution can enable "learning sovereignty" with fair and equitable value attribution for data suppliers while at the same time enabling important societal benefits and commercial innovation opportunities that otherwise would not be possible in current settings without data sharing. Federating data between independent data trusts can provide access to data that would otherwise be isolated. In many cases, this data might not be available at all, as inter-entity trust is required to readily access many types of private or sensitive data.

Case study: LunaPBC

An example of a data trust is LunaDNA, a genomic and health database community. LunaDNA, a community-owned platform company, is managed by LunaPBC, a public-benefit corporation. In this example, the US Securities and Exchange Commission (SEC) qualified LunaDNA to provide ownership shares in the company using shared data as the purchasing currency. Thus, the data contributors own LunaDNA. Ownership percentages are based on various types of health, genomic and self-reported data and the amount that is shared. In terms of protecting the rights of individuals contributing data to LunaDNA, the SEC agreement calls out protections that are contractually guaranteed to LunaDNA members under the LunaDNA consent, privacy and subscription agreements. Shared data is consented for use in health and quality-of-life studies. All studies are approved by an outside administrative body established to protect the rights and welfare of individuals participating in the research (e.g. an institutional review board) and data can be used only at an aggregate or population level. The raw data remains within LunaDNA, and never leaves the platform. Approved researchers and discovery partners bring queries or analysis tools to the data, using secure sandboxes and only answers are returned. These answers could come in the form of statistics, metadata or data models.

LunaPBC is contractually licensed to manage data in LunaDNA on behalf of the community. The management agreement explicitly details the operational and legal responsibilities of LunaPBC. It also includes requirements such as providing transparency, account control and financial attribution to LunaDNA members, as well as the conditions outlining acceptable data usage. The LunaDNA model also provides a privacy-preserving pseudonymous method of re-contacting members, enabling a researcher or digital community sponsor to share important information, send communications, request additional data or to solicit consent for direct participation in studies. These requests bridge the gap between protecting member privacy and enabling members to consent to activities that could potentially compromise their privacy, such as pharmaceutical clinical trials.

Hashing, data trusts and federated learning models all offer innovative solutions for unlocking siloed data, which is ultimately a key aim of cross-border data flows policy

Hashing, data trust models and federated learning all offer innovative options for unlocking siloed data, which is ultimately a key aim of cross-border data flows policy. By creating space for such innovative, new or alternative models (including other and future models for unlocking data or data learnings or even algorithms across borders), governments can ensure that their digital borders remain open for business and that both their international and national policies remain future-proof in respect of the data economy.

While the specificities of recognizing any one model are at a sovereign government's discretion, it is recommended that the existence of such models be considered as a reality when constructing data flows policy so as to avoid any unintended consequences and potentially choke sectors of the economy. To give the example of AI, machine learning could be severely curtailed by the prevention of algorithms or findings travelling across borders should that ever be a policy position.

Conclusion

Operationalizing the Roadmap

> There is no one-size-fits-all approach; after all, every inter-country relationship is unique

Empowering governments to adopt robust but safeguarded cross-border data sharing policies is of critical importance to ensure that economies do not get left behind in the Fourth Industrial Revolution.

Building trust between governments is ultimately a matter that rests with sovereign nations. Nevertheless, businesses can help governments set fit-for-purpose policy that allows for interoperability from country to country or region to region. There is no one-size-fits-all approach; after all, every inter-country relationship is unique.

However, by considering a common set of policy levers – as represented in the Roadmap – nations can feel confident that they are engaging in the relevant analysis needed to both build trust with their counterparts in the digital economy space and facilitate their own domestic data economy.

The Roadmap deliberately does not prescribe the implementation of these measures at a country level, as such measures are highly context-specific and countries need to conduct their own analysis of their readiness in the various policy areas examined. However, the gauntlet is firmly thrown to countries to stress-test the Roadmap and determine how to implement it. This exercise will be critically important as the world moves towards post COVID-19 economic recovery and the need for cross-border data flows becomes ever more acute.

This paper is part of a series by the Centre for the Fourth Industrial Revolution focusing on data policy in a post COVID-19 world.

Appendix

Glossary of terms

AICPA SOC 2 Type 2: SOC stands for "system and organization controls", and are a series of standards designed to measure how well a given service organization conducts and regulates its information. The purpose of SOC standards is to provide confidence and peace of mind for organizations when they engage third-party vendors. A SOC-certified organization has been audited by an independent certified public accountant who determined that the firm has the appropriate SOC safeguards and procedures in place.

More specifically, SOC 2 is designed for service providers storing customer data in the cloud. It requires companies to establish and follow strict information security policies and procedures encompassing the security, availability, processing, integrity and confidentiality of customer data.[26]

Anti-snooping clause: A clause that forbids governments and connectivity providers from viewing the data being transmitted across borders, except in certain prescribed instances.

APEC CBPR: Asia-Pacific Economic Cooperation Cross-Border Privacy Rules.

Artificial intelligence (AI): The capacity of a machine to imitate intelligent human behaviour.

Binding corporate rules (BCRs): A cross-border transfer mechanism of the GDPR whereby multinational corporations can seek explicit approval for their actions.

Cloud computing: On-demand computer and storage systems that are managed by a third party and often exist across multiple data centres in multiple locations.

Confidentiality: Property such that information is not made available or disclosed to unauthorized individuals, entities or processes.

Cross-border data flows: The regular unimpeded movement of data across international borders.

Cross-border data transfer restrictions: Restrictions on companies that mean they must not transfer data to another country unless they can assure adequate safeguards for the transferred data abroad; if companies can meet the requirements for an exception, they are not required to keep a local copy of the data.

Data jurisdiction law (Bahrain): According to the Legislative Decree No. 56 of 2018 in Respect of Providing Cloud Computing Services to Foreign Parties, data from government and business entities stored in data centres in Bahrain is subject to the exclusive jurisdiction of the foreign state in which the entity is domiciled, constituted or established.

Data localization requirements: Any obligation, prohibition, condition, limit or other requirement provided for in the laws, regulations or administrative provisions of a jurisdiction or resulting from general and consistent administrative practices in that jurisdiction and in bodies governed by public law, including in the field of public procurement that imposes the processing of data in a specific territory or hinders the processing of data in any other territory.

Data portability: The ability to port data from one system to another; this is an issue that is top of mind for B2B customers of data hosting services.

Data porting: Moving data from one backbone system to another in order to use it on that different system, which may or may not be compatible.

Data processing: As defined in the GDPR (Article 4), any operation or set of operations that is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

Data processor: As defined in the GDPR (Article 4), a natural or legal person, public authority, agency or other body that processes personal data on behalf of the controller.

Data protection and privacy laws: Laws that govern the collection and processing of personal data and personally identifiable information and which vary from territory to territory. These differences can act as both a hard and soft barrier to the movement of data across borders and can cover personal and/or non-personal data.

Data provenance: Identifies the origin of the data processor and data owner and documents a record of the history of the data since collection.

Data residency laws: Under these laws, companies must process data primarily in a particular territory, but they can also transfer copies of the data abroad.

Data stack: An abstract concept that determines the order in which data is stored within a system. Stacking enables compartmentalization of the dataset and is used in the concept or privacy engineering whereby the order and location in which the data is stored on a system enables the data to be treated in a certain way, e.g. to meet data protection obligations.

Data subject: As defined in the GDPR (Article 4), an identified or identifiable natural person; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Data trust: An entity or group of entities that manages the rules of the game in a created data ecosystem, and particularly the contributed data on behalf of the data suppliers.

ePrivacy Directive: Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).

Fourth Industrial Revolution: A way of describing the blurring of boundaries between physical, digital and biological worlds created from advancements in AI, IoT and other technologies.

General Data Protection Regulation 2018 (GDPR): Regulation number 2016/679 entitled Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

Internet of things (IoT): A network of items – each embedded with sensors – that are connected to the internet.

ISO 27001:[27] ISO/IEC 27001:2013 specifies the requirements for establishing, implementing, maintaining and continually improving an information security management system within the context of the organization. It also includes requirements for the assessment and treatment of information security risks tailored to the needs of the organization. The requirements set out in ISO/IEC 27001:2013 are generic and are intended to be applicable to all organizations, regardless of type, size or nature.

The information security management system preserves the confidentiality, integrity and availability of information by applying a risk management process and gives confidence to interested parties that risks are adequately managed.

This international standard can be used by internal and external parties to assess the organization's ability to meet the organization's own information security requirements.

Open Data Watch's Data Value Chain: The Data Value Chain (Open Data Watch 2018) framework[28] helps technical practitioners to understand how interoperability adds value to data on the data value chain. The data value chain describes the four major stages: collection, publication, uptake and impact. It is essential to reference interoperability at each stage, when the handshake happens between systems to either consume or deliver data in the value chain. For example, it will define classifications and standards to be followed, while collecting and storing the data. Importantly, it describes how downstream systems should use the data. The interoperability checklist must also reflect the organizational practices and data management plans that cover the entire data value chain.

Personal data: As defined in the GDPR (Article 4), any information relating to a data subject. It is important to note that information that relates to a data subject, even without a name, can qualify as personal data under the GDPR.

SaaS: Software as a service; software solutions that reside in the cloud but, due to high-speed connectivity, can be used in real time as if they resided locally.

Service provider: An entity that delivers application functionality and associated services across an IT network to multiple service consumers.

SMEs: Small and medium enterprises.

Statistical Information System Collaboration Community's National Data Backbone Framework: The OECD's Statistical Information System Collaboration Community (SIS-CC) 2018 model of the National Data Backbone addresses issues of fragmented data ecosystems and data/operational silos in order to enable interaction within an international reporting framework composed of different requirements.[29] This model emphasizes the semantic interoperability of the various components, particularly organizations and institutions.
The approach uses the existing open-source community for official statistics and data solutions at the national level, with a globally hosted platform, and increases the volume of contributions to open-source projects. The component architecture enables countries and organizations to use open-source assets while also tailoring and even creating their resources, provided the necessary expertise is available within the various organizations to do so.

Technical interoperability: The ability to share data between different systems and to enable those systems to make use of the data.

Threat: Potential cause of an unwanted incident, which may result in harm to a system or organization.

Contributors

Lead authors

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

Nada AlSaeed
Senior Manager, Bahrain Economic Development Board; and World Economic Forum Fellow

Lothar Determann
Partner, Baker McKenzie and Adjunct Professor, Free University Berlin; University of California, Hastings College of the Law; Lecturer, Berkeley School of Law

Leanne Kemp
Chief Executive Officer, Everledger

Steering Committee

Lothar Determann
Partner, Baker McKenzie and Adjunct Professor, Free University Berlin; University of California, Hastings College of the Law; Lecturer, Berkeley School of Law (Chair)

Leanne Kemp
Chief Executive Officer, Everledger (Co-Chair)

Juan Carlos Castilla-Rubio
Chairman, Space Time Ventures

Elizabeth Davies
Senior Director, Data Protection, Splunk

Jan Huizeling
Vice-President Digital Farming, Yara International

Robert Charles Kain
Chief Executive Officer and Co-Founder, LunaPBC

Kimmo Kasslin
Vice-President of Labs, CUJO AI

Contributors

Vivienne Artz
Chief Privacy Officer, Refinitiv

Julien Chaisse
Professor, School of Law, City University of Hong Kong

Evîn Cheikosman
Project Coordinator, Data Policy, World Economic Forum

Indre Deksnyte
Senior Vice-President of Marketing, CUJO AI

Michael Holsey
Project Manager, EMD Digital, Merck

Francis Jee
Fellow, Blockchain and Distributed Ledger Technology, World Economic Forum

Rosetta Jones
Senior Director, Global Strategic Initiatives, Visa

Maisie Mahoney
Head of the Digital Project Management Office

Allan Millington
Director, Data Management, EY

Adrien Ogée
Chief Operations Officer, The CyberPeace Institute

Diana Paredes
Chief Executive Officer and Co-Founder, Suade

Sheila Warren
Head, Blockchain, Digital Assets, and Data Policy, and Member of the Executive Committee, World Economic Forum, USA

Christian Wickert
Head of Corporate Policy USA, Merck

Acknowledgements

Rima Al Kilani
Director, Bahrain Economic Development Board

Omar Sultan Al Olama
Minister of State for Artificial Intelligence of the United Arab Emirates

Mohamed Al Qaed
Chief Executive, Information & eGovernment Authority, Bahrain

Khalid Al Rumaihi
Chief Executive Officer, Bahrain Mumtalakat Holding Company

Alain Bejjani
Chief Executive Officer, Majid Al Futtaim

Andrew Berkley
Project Lead, Data Science and Analysis, World Economic Forum

Rohit Chopra
Commissioner, Federal Trade Commission

Frans Cronje,
Co-Founder and Chief Executive Officer, DataProphet

Patricia Ellen da Silva
Secretary of Economic Development, Government of the State of São Paulo

Catalina De la Rocha Gonzalez
Research Assistant, C Minds

Gonzalo De Romana
Chief Executive Officer, TASA

Simonetta Di Pippo
Director, United Nations Office for Outer Space Affairs

Tala Fakhro
Chief Project Officer, Bahrain Economic Development Board

Ziyang Fan
Head of Digital Trade, World Economic Forum

Laura Gallagher
Global Head of Corporate Citizenship, AIG

Leticia Gasca Serrano
Chief Executive Officer, Skills Agility Lab

Khalid Humaidan
Chief Executive Officer, Bahrain Economic Development Board

Austin Hunter
Project Specialist, Data Policy, World Economic Forum

Jesse Lin
Project Specialist, Digital Trade, World Economic Foru

Eric Loeb
Executive Vice-President, Government Affairs, Salesforce

Alistair Millen
Designer, Studio Miko

Ali Moore
Editor, Astra Content

Sara Pantuliano
Chief Executive, Overseas Development Institute

Avinash Patwardhan
Managing Director, Smart Cities, Jacobs

Linda Pawczuk
Global Co-Lead, Blockchain, Deloitte

Rhiannan Price
Director, Sustainable Development Practice, Maxar Technologies

Claudia Sadoff
Director-General, International Water Management Institute, Sri Lanka

Rishi Saha
Director, Public Policy, Middle East and Africa, Amazon Web Services

Taro Shimada
Corporate Vice-President and Chief Digital Officer, Toshiba

Magdalena Skipper
Editor-in-Chief, Nature

Jan-Gunnar Winther
Director, Centre for the Ocean and the Arctic

James Workman
Founder, AquaShares

Endnotes

1. Ferracane, M.F. (2017), Restrictions to Cross-Border Data Flows: A Taxonomy,European Centre for International Political Economy (ECIPE): https://ecipe.org/publications/restrictions-to-cross-border-data-flows-a-taxonomy/ (link as of 15/5/20).
2. World Economic Forum (2019), Exploring International Data Flow Governance: https://www.weforum.org/whitepapers/exploring-international-data-flow-governance (link as of 19/5/20).
3. Ursic H., Nurullaev R., Olmedo Cuevas M. and Szulewski P. (2018), Data Localisation Measures and Their Impacts on Data Science. In: Mak V., Tjong Tjin Tai E. and Berle A. (eds.) (2018), Research Handbook in Data Science and Law. Research Handbooks in Information Law. Cheltenham: Edward Elgar: pp. 322–353: https://ssrn.com/abstract=3102890 (link as of 15/5/20).
4. Press, G. (2020), 6 Predictions About Data in 2020 and the Coming Decade, Forbes: https://www.forbes.com/sites/gilpress/2020/01/06/6-predictions-about-data-in-2020-and-the-coming-decade/#3d3c658c4fc3 (link as of 15/5/20).
5. Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union.
6. Scharwatt C. (2019), The Impact of Data Localisation Requirements on the Growth of Mobile Money-Enabled Remittance, GSMA: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/03/GSMA_Understanding-the-impact-of-data-localisation.pdf (link as of 15/5/20).
7. Casalini F. and Lopez Gonzalez J. (2019), Trade and Cross-Border Data Flows, OECD Trade Policy Papers, No. 220, OECD Publishing, Paris: p. 15.
8. Determann, L. (2020), Determann's Field Guide to Data Privacy Law, 4th Ed., #0.14 et sequ.: www.elgaronline.com/view/9781789906189/9781789906189.xml (link as of 15/5/20).
9. Reinsel, D., Gantz, J. and Rydning, J. (2018), Data Age 2025: The Evolution of Data to Life-Critical, IDC: https://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf (link as of 15/5/20).
10. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
11. Article 49 (1) (a) EU General Data Protection Regulation.
12. Determann, L. (2019), Healthy Data Protection, SSRN: http://ssrn.com/abstract=3357990 (link as of 15/5/20).
13. This approach is similar to the EU-US Privacy Shield, whereby mutual standards are respected regarding the processing of personal data.
14. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).
15. ENISA (November 2016), NCSS Good Practice Guide: https://www.enisa.europa.eu/publications/ncss-good-practice-guide; whitehouse.gov (September 2018), National Cyber Strategy of the USA: https://www.whitehouse.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf; ITU (2018), Guide to Developing a National Cybersecurity Strategy: https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-CYB_GUIDE.01-2018-PDF-E.pdf (links as of 18/5/20).
16. World Economic Forum (2020), Presidio Principles: Foundational Values for a Decentralized Future: https://www.weforum.org/communities/presidio-principles (link as of 28/5/20).
17. APEC Policy Support Unit (2018), Study on Single Window Systems' International Interoperability: Key Issues for Its Implementation: https://www.apec.org/Publications/2018/08/Study-on-Single-Window-Systems-International-Interoperability (link as of 18/5/20).
18. Thomson C., Ludlam J., Peddie J. and Grimmer T.J. (23 October 2019), UK and US Sign Data Access Agreement to Expedite Digital Evidence – Sharing in Criminal Investigations, Baker McKenzie: https://www.bakermckenzie.com/en/insight/publications/2019/10/uk-us-data-access-agreement (link as of 18/5/20).
19. UN, National Statistical Offices: https://unstats.un.org/home/nso_sites/ (link as of 18/5/20).

20. The Open Definition: Opendefinition.org (link as of 18/5/20).

21. United States Environmental Protection Agency (10 January 2020), Data Standards: https://www.epa.gov/data-standards (link as of 18/5/20).

22. Berners-Lee, T. (27 July 2006), Linked Data, World Wide Web Consortium: https://www.w3.org/DesignIssues/LinkedData.html (link as of 18/5/20).

23. Vasudevan, K. (18 April 2017), Microservices, APIs, and Swagger: How They Fit Together, Swagger Blog: https://swagger.io/blog/api-strategy/microservices-apis-and-swagger/ (link as of 18/5/20).

24. GSMA (April 2019), The 5G Guide: A Reference for Operators: https://www.gsma.com/wp-content/uploads/2019/04/The-5G-Guide_GSMA_2019_04_29_compressed.pdf (link as of 18/5/20).

25. World Economic Forum (23 January 2018), New Partnership Aims to Sequence Genomes of All Life on Earth, Unlock Nature's Value, Tackle Bio-Piracy and Habitat Loss: https://www.weforum.org/press/2018/01/new-partnership-aims-to-sequence-genomes-of-all-life-on-earth-unlock-nature-s-value-tackle-bio-piracy-and-habitat-loss/ (link as of 18/5/20).

26. evariant.com, What is SOC 2 Type 2 Certification?: https://www.evariant.com/faq/what-is-soc-2-type-II-certification (link as of 18/5/20).

27. ISO, ISO/IEC 27001:2013: https://www.iso.org/standard/54534.html (link as of 18/5/20).

28. Open Data Watch (2018), The Data Value Chain: Moving from Production to Impact: https://opendatawatch.com/publications/the-data-value-chain-moving-from-production-to-impact/ (link as of 18/5/20).

29. Statistical Information System Collaboration Community (SIS-CC) (2018), Building National Data Backbones: Empowering Countries Through Capacity Development and Technology. Workshop Highlights Report from the 2018 SIS-CC Workshop. Paris, France: https://siscc.org/wp-content/uploads/2018/10/sis-cc-workshop-2018-highlights-report.pdf (link as of 18/5/20).

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