In a world where data are becoming an integral part of our lives, data governance has become essential to support the production, use and reuse of data in a safe, ethical, and secure manner. In effect, data governance enforces the social contract around data by applying the principles of trust, value, and equity. This primer provides an overview of the different elements of a data governance framework as discussed in the World Development Report 2021, Data for Better Lives. The implementation of the different elements of the framework will be specific to a country context, informed by local norms, laws, and culture. Using this framework, a series of briefs accompanying this primer, will explore and review the data governance landscape in different regions with a view to facilitating the region’s pathway to a green, resilient, and inclusive recovery. Some of the key overarching messages of this primer are as follows:

- Equitable development of the data economy calls for a foundation of infrastructure to ensure that both poor people and poor countries have affordable access to data services.

- The legal and regulatory framework for data entails a balanced development of enablers that support reuse of data for value creations, and safeguards that create trust in the system.

- As more economic activities shift online, a country’s data governance choices will have important implications for the real economy, in terms of competition, trade, and taxation.

- However sound the legal and regulatory framework, data governance will not get very far without adequate institutions to implement and enforce the rules, and these are often missing.
While data governance is primarily a national concern, there are many areas where international cooperation on data governance is critical and beneficial.

Broad principles on data governance are helpful, but the realities on the ground differ substantially across regions necessitating more regionally grounded analysis and treatment.

The COVID-19 pandemic has brought data governance issues to the forefront of public awareness. The debate about the use of call detail records from mobile phones to support contact tracing to control the virus illustrates many of the challenges posed by data governance. On the one hand, the ability to readily access and repurpose call detail records has added value by helping to arrest the spread of disease in some countries. On the other hand, this practice has raised major concerns about whether people can trust public authorities not to misuse the data in other ways. At the same time, such an approach has little to offer those who may be too poor or disadvantaged to own a mobile phone in the first place, highlighting the issue of equity.
A DEARTH OF DATA GOVERNANCE

Currently, most countries around the world do not have robust data governance frameworks in place. A new Global Data Regulation Survey, undertaken especially for the World Development Report 2021 (WDR 2021), shows that on average countries are only 40 percent of the way toward implementing good data governance practices, ranging from 30 percent in low-income countries to around 50 percent among high-income countries. Countries have made most progress passing data governance legislation, but those laws are not always comprehensive, while the institutions to implement and enforce them are often inadequate.

The WDR 2021 broadly defines data governance to include data infrastructure policy, the legal and regulatory framework for data, the related economic policy implications, and the institutional framework (see Figure 1). These elements constitute the building blocks of an integrated national data system, an aspirational vision that is built on the principles of the social contract for data and seeks to deliver the potential value of data equitably while safeguarding against harmful outcomes. A solid data governance framework is one that creates an environment that allows economic and social value from data to be readily created, provides adequate safeguards to preserve trust in data systems, and levels the playing field so that the benefits of data are equitably shared. The following sections provide a brief overview of each of the data governance pillars that are essential for delivering the potential benefits of data for development while safeguarding against harmful outcomes. Although there is no one-size-fits-all approach, this framework provides a structure for more in-depth regional analysis which will also take into consideration the country context (history, culture, governance, and political economy).

INFRASTRUCTURE CAN ENSURE EQUITABLE ACCESS TO THE BENEFITS OF DATA

Data infrastructure policies, a key building block of the data governance framework, help level the playing field in the modern data economy through infrastructure that enables equitable access to data services for poor people in poor countries. This includes both the policies that make it possible for people to connect to internet data services, and the policies that ensure that countries have adequate infrastructure to exchange, store, and process data efficiently over the internet (see Chapter 5).

worldbank.org/digitaldevelopment
and make services more affordable. In addition, policymakers should increasingly aim to address demand side barriers to service uptake such as expensive devices, poor digital skills, and limited local content.

At the same time, many low- and middle-income countries lack internet exchange points (IXPs), colocation data centers, and access to on-ramps for cloud computing. Countries needing to send their data overseas to access such facilities can face service charges up to ten times higher than those able to perform these functions domestically. Such infrastructure can often be developed by the private sector, but calls for the creation of a supportive enabling environment, including strong governance arrangements as well as availability of renewable electricity supply. Due to economies of scale, small economies may need to club together at the regional level to support the development of such infrastructure.

**FIGURE 1: Data governance layers at the national and international levels**

Source: WDR 2021

[worldbank.org/digitaldevelopment]
Achieving **trust** in data exchanges requires a **robust regulatory environment** for both personal and nonpersonal data. This is the second building block of the data governance framework (see Chapter 6). It involves both rigorous cybersecurity measures to protect the integrity of data systems and strong measures for data protection.

Personal data protection is grounded in a human rights framework, which demands that individual rights are first respected, typically by seeking individual consent, before anything else is done with data. But this type of protection is not without costs both for firms and individuals. A survey of relatively large firms indicates that companies’ cost of complying with the EU’s General Data Protection Regulation (GDPR) range from US$250,000 to almost US$2 million a year (EU 2018). On the consumer side, estimates indicate that it would take the average person 76 days a year to read and study all the consent notices regarding data protection that they routinely click on when visiting different websites (Madrigal 2012). Such concerns have prompted public debates regarding possible alternative approaches to data protection, including placing stronger fiduciary obligations on data processors, upscaling deployment of Privacy Enhancing Technologies (PETs), or introducing Personal Information Management Systems (PIMS) through intermediaries that support people with the management of their data (see Chapter 6).

To make matters even more complex, it is getting increasingly difficult to define personal data. Much nonpersonal data (from mobile phones) is personally identifiable, while even anonymized personal data can increasingly be reidentified by pooling multiple sources of information and applying modern analytical techniques.

Overall, the level of development of regulatory frameworks for personal data protection remains rather uneven around the world. According to the Global Data Regulation Survey, policy makers across the income spectrum have systematically neglected certain important aspects of the data protection framework—such as regulatory limitations on algorithmic decision making, requirements to incorporate privacy by design, and limitations on government exceptions for accessing personal data (see Figure 2).

At present, only a handful of lower-income countries have developed sound data protection legislation. A rare example is Kenya, whose 2019 Data Protection Act includes comprehensive best practice cybersecurity requirements (such as pseudonymization and encryption), as well as a personal data protection framework that incorporates many elements of the EU’s GDPR.

The economic and social **value** of data can be multiplied exponentially as data are reused, repurposed, and recombined with other data sources. Such intensive reuse of data is possible only if the legal framework contains enablers that require data sharing and facilitate data exchange through legal requirements (such as open access) and technical
norms (such as interoperability). Many governments around the world have already adopted open data provisions, often supported by access to information legislation, with valuable results. An impact assessment of the EU’s 2003 Directive on Reuse of Public Sector Information found that it generated €52 billion in value in 2017 (EC 2018).

However, sharing of private sector data (both among firms and with the public) has proven much more difficult due to commercial incentives to retain data and more limited government influence to encourage sharing. In an effort to spur the creation of smart cities, metropolitan areas such as New York, San Francisco, and São Paulo have made legal attempts to require certain private sector platforms to share their commercial data. Among the most comprehensive efforts to date, France’s Law for a Digital Republic (2016) requires that the private sector make certain public interest data sets publicly available based on open standards. The United Kingdom’s Digital Economy Act (2017) takes a somewhat different approach, enabling researchers to gain access to deidentified data for research purposes.

Note: The figure shows the percentage of countries in each country income group that had adopted good-practice legal and regulatory frameworks to safeguard personal data as of 2020.
Data governance also has important implications for the economic fortunes of countries. The expanding role of data in ubiquitous platform business models is reshaping competition, trade, and taxation in the real economy, posing important risks for low- and middle-income countries. This means that the way countries design the safeguards and enablers to govern the data on which these business models depend, will have knock-on effects for the real economy (see Chapter 7).

Competition dynamics in platform-driven businesses tend to lead to market dominance and create an unlevel playing field for entrants in low- and middle-income countries, affecting equity. Half a dozen major US data-driven companies (Google, Facebook, Microsoft, Amazon, Zoom, Verizon) account for 40 percent of global data traffic and two-thirds of the top 25 most visited website across a majority of lower-middle-income countries (see Figure 3). This raises equity concerns as market dominant platforms may be able to extract rent from consumers, and also preclude entry by new players.

**FIGURE 3**: Google, Facebook and Microsoft alone own half of the top websites and are among the top 10 most visited websites in all low-income countries for which data are available.

Source: Alexa (downloaded 2020:Q2)
Antitrust agencies—where they exist—will need to adapt their traditional approach to mergers to address the challenges posed by platform businesses, while balancing the complex interactions between data protection and competition concerns. Proactive adoption of data-sharing regulations may help to restrain the accumulation of market power. For example, in the Uber-Careem merger in Egypt, Uber was obliged to grant future competitors access to Careem’s “points of interest map data” on a one-time basis; provide current competitors with access to trip data (including rider and driver information), subject to data protection laws; and give riders access to their own data.

Regulations governing the cross-border flow of personal data (see Figure 4) are a critical factor in determining a country’s competitiveness in data-enabled services trade, a major new source of economic value for the developing world. The Bangladeshi firm Augmedix, for instance, offers remote assistance to medical doctors in the United States. Doctors wear smart glasses allowing Bangladeshi assistants to “witness” patient consultations and create associated medical records. This two-way exchange of data, and the high value-added services that it generates, is possible only because both countries allow for such sensitive data to move across borders (open transfers). In other parts of the world, such cross-border data flows may be limited due to government regulations requiring prior approval or mandating local storage of data (limited transfers), or be conditional on a regulatory determination of the adequacy of data protection arrangements in the trading country (conditional transfers).
Taxation of data-driven businesses is posing serious challenges that are preventing an equitable geographic distribution of the tax revenues from data-driven businesses. In the case of the value added tax (VAT), the principle of taxation in jurisdiction where sale occurs is well established, but developing countries struggle to implement it administratively. To illustrate how much revenue is at stake, in Indonesia, the gross VAT revenue potential of the business to consumer (B2C) digital economy is estimated to amount to 0.39 of GDP in 2021 and is projected to rise to 0.65 percent in 2025.

When it comes to corporate taxation, the recent G20 agreement on international harmonization of corporate tax rates is helpful. However, international allocation of taxation rights is unclear and inconclusive talks are underway, facilitated by the OECD. In the meantime, some 30 countries are stepping into the breach with their own ad hoc digital services taxes, creating international tensions. For example, in 2016, India enacted a 6 percent “equalization levy” on digital services linked to online advertising, as a proxy for a corporate income tax on foreign suppliers.
INSTITUTIONS DRIVE IMPLEMENTATION AND ENFORCEMENT

For effective implementation and enforcement of the various building blocks of data governance, a suitable institutional ecosystem that encompasses both state and nonstate institutions and actors must be in place. The roles and responsibilities of institutions and actors are evolving in response to changes in the data economy. However, the main functions (see Figure 5) include developing overarching data strategies and policies; elaborating legal frameworks and guidance on how rules should apply and be enforced if violated; implementing and enforcing compliance with established laws and regulations; undertaking arbitration in case of conflict; and maintaining monitoring, evaluation, and constant feedback loops to promote engagement, learning, and improvements. In some countries, these functions can be performed by existing institutions (such as the national statistical office or relevant sector regulators). In others, there might be a need to create new institutions (such as data protection agencies or data intermediaries).

Source: WDR 2021

FIGURE 5: Functions of data governance
Despite the increasing importance of these institutions, they are not common in lower-income countries. Only 21 percent of low-income countries have established data protection authorities, compared to 81 percent of high-income countries (see Figure 6). Similarly, cybersecurity agencies are relatively widespread in middle- and high-income countries but are present in only 24 percent of low-income countries (see Figure 7). National statistical offices (NSOs) as premier data institutions in countries perform critical data governance functions involved in the production of official statistics. The demands of the changing data landscape require NSOs to modernize their data governance functions to improve data flows both within public sector and with other non-governmental actors. NSOs require sufficient financing, independence, and capacity to effectively modernize countries’ national statistical system (see Chapter 2).
Whatever the roles and functions of individual institutions, the new data economy calls for a whole-of-government approach that enables data flows between different institutions and actors, eliminating siloed decision making, and guiding data management decisions along the data life cycle. For example, Uruguay’s creation of a lead agency — the Agency for Electronic Government and Information and Knowledge Society (Agesic)—close to the Office of the President and acting with a whole-of-government and multistakeholder perspective has been critical in driving the country’s successful e-government reforms and the data agenda since 2007.

No matter what institutional design a country chooses to implement, institutions and actors must have the leadership, technical capacity, resources, and incentives to perform their roles and harness the value of data and shift toward a culture of data use (see Figure 8). Moreover, a multistakeholder approach to data governance is better equipped to govern the complex data ecosystem in a transparent, inclusive, and equitable way that reflects the interests of all key stakeholders (see Chapter 8).

Source: WDR 2021

worldbank.org/digitaldevelopment
DATA GOVERNANCE IS ALSO AN INTERNATIONAL ISSUE

Efficient and effective data governance also calls for closer international cooperation. Such collaboration is needed not only at the global level, but regionally and bilaterally, too.

Bilateral cooperation between counterpart agencies in different countries can be an important means to address many data governance challenges. Law enforcement agencies need to collaborate to combat cybercrime, which moves effortlessly across borders, by means of Mutual Legal Assistance Treaties. Tax administrations need to share financial information of firms to support mobilization of VAT revenues from sales of businesses registered in third countries but undertaken across digital platforms. Antitrust agencies may benefit from coordinating regulatory approaches when determinations made in one country affect firms active in other jurisdictions.

For smaller economies, regional cooperation can play a valuable role by reducing costs of data infrastructure through development of shared regional facilities, achieving scale economies through physical and regulatory integration of markets, and enhancing voice and negotiating strength in global processes.

Global cooperation is needed for key areas where free rider behavior (such as protectionism or tax evasion) is a risk. At the global level, key agreements on taxation and trade for data-driven businesses are pending international agreement. Multilateral efforts can also do much to reduce transactions costs through harmonization of technical standards, often led by industry bodies.
A DATA GOVERNANCE VISION

While data governance challenges may seem daunting, the experience of leading-edge countries offers a concrete vision of what can be achieved when these issues are systematically addressed. In Estonia, for example, the government has set up a national data system to safely manage citizens’ personal data for use by government agencies and participating businesses. X-Road is an open-source data exchange layer solution that allows linked public and private databases to automatically exchange data, ensuring confidentiality, integrity, and interoperability. It combines a technical solution (software architecture and protocols) with a governance solution (the once-only principle legally obliging public sector agencies to refrain from duplicating data requests). X-Road’s cryptography protocols also enhance transparency because they log entries into the system and give individuals detailed insights into who is sharing their data and for what purposes. X-Road embodies the three main principles of any social contract for data governance. Its transparency engenders trust. Its national scope and universal availability promote equity. Its comprehensiveness and ease of use generate value.
A well-designed data governance framework brings the social contract to life. It allows countries to capture the full economic and social value of data, while creating trust in the integrity of the data system as well as ensuring that the benefits of data are equitably shared. This involves implementing policies that will close the digital divide, providing universal broadband access, and ensuring that all countries can access modern infrastructure to exchange, process, and store growing volumes of data. It also entails developing regulatory frameworks and institutions to safeguard data through cybersecurity and data protection measures, while enabling the safe sharing and reuse of data. Moreover, data governance has important economic implications, with the design of safeguards and enablers to govern data in platform business models having a real effect on economies. In addition, internationally coordinated action—on antitrust enforcement, regulation of platform firms, data standards, trade agreements, and tax policy—is critical to ensuring efficient, equitable policies for the data economy that respond to countries’ needs and interests. Ultimately, effective data governance is critical for reaching the aspirational goal of an integrated national data system.

Each region or country will design data governance policies, considering local norms, values, and culture. Understanding the nature of the data governance challenges which differ significantly across regions will be critical for designing these policies. In Sub-Saharan Africa, universal access challenges remain paramount. In the Middle East and North Africa, open and transparent access to public sector data to support value creation has been a longstanding concern. In Latin America, a central concern is how to leverage the economic opportunities posed by digital platforms and trade in data-enabled services. To do justice to these wide-ranging circumstances, this primer will be accompanied by a series of regional briefs which will provide a more in-depth analysis of the different elements of the data governance framework, and conduct benchmarking of countries within specific regions. These will bring the overarching messages of the WDR 2021 to a more concrete local level and support regional and national policy dialogue on the way forward for data governance.
ACKNOWLEDGMENTS

This primer is based on findings from World Development Report 2021: Data for Better Lives. The note was authored by Vivien Foster (Chief Economist, Infrastructure Vice-Presidency) and Malarvizhi Veerappan (Program Manager, Development Economics Vice-Presidency), with editorial support and guidance from Rami Amin (ET Consultant, Digital Development).

The team acknowledges the core WDR2021 author team for their contribution to the overall report, as well as to those who contributed to the report through consultations, the work of which underpinned the basis for this document. The WDR Core Author Team included Robert Cull, Vivien Foster, Dean Jolliffe, Malarvizhi Veerappan, Jaffar Al-Rikabi, Adele Barzelay, Miriam Bruhn, Rong Chen, Niccolo Comini, Samuel Fraiberger, Craig Hammer, Talip Kilic, Jan Loeprick, Daniel Gerszon Mahler, Michael Minges, Martin Molinuevo, Nancy Morrison, David Newhouse, Sara Nyman, Selome Paulos, Vinny Ricciardi, David Satola, Dorothe Singer, Philip Wollburg, and Bilal Zia.

REFERENCES


worldbank.org/digitaldevelopment