PART III

Moving toward an integrated national data system

9. Creating an integrated national data system

Creating an integrated national data system

Main messages

1 By building an integrated national data system, countries can realize the full value of data for development. The system should provide a framework for the trustworthy, equitable production, flow, and use of data.

2 An integrated data system is built on an approach to data governance that is intentional, whole-of-government, and multistakeholder. The steps needed to implement such a system depend on a country's data maturity. What works in one context may not work in another.

3 To be sustainable, an integrated national data system must be continually improved. This will depend on having highly skilled human resources in government, civil society, academia, and the private sector.

4 Robust data protection is critical to building an integrated national data system. As the scope of such a system expands, the economic, social, and development returns increase, as do the data protection requirements.

Toward an integrated national data system

This chapter describes how to create an integrated national data system designed to realize the potential of data for development. Such a system relies on an approach to data governance that is intentional, whole-of-government, multistakeholder, and collaborative. It explicitly builds data production, protection, exchange, and use into planning and decision-making and integrates participants from civil society and the public and private sectors into the data life cycle and into the governance structures of the system. Although such a system is related to a national statistical system, it differs in key areas (box 9.1).

An integrated national data system is all about people. A well-functioning system requires people to produce, process, and manage high-quality data; people to populate the institutions that safeguard and protect the data from misuse; and people to draft, oversee, and implement data strategies, policies, and regulations. A well-functioning system also requires people to use data as a factor of production in both the public sector—for policy design and implementation—and the private sector—for decision-making and innovations in products and services. People are also needed to hold the public and private sectors accountable. All this requires robust data literacy.[1] Meanwhile, at the end of the day, it is people who will benefit from an integrated national data system. They will see better public policies, programs, and service delivery; more business opportunities and jobs; higher market efficiency; and greater accountability.

It is vital that the public trusts that data are being safely stored, exchanged, and used to create value equitably, while protecting against misuse. Thus the social contract for data is built into a well-functioning national data system and should be recognized in national development strategies.

When the foundation of an integrated national data system is in place and a variety of participants are included in the data life cycle, it can yield vast benefits for development. In fact, the more integrated the system and the more participants involved, the higher is the potential return. If two participants safely exchange data with each other, data can flow in two directions. If three participants exchange data, data can flow in six directions, and with four participants, in 12 directions. If data are reused and repurposed, these connections can increase exponentially. When government agencies, civil society, academia, and the private sector securely take part in a national data system, the potential uses of data expand and so does the potential development impact. As the scope of the system expands, so do the data protection requirements and the needs for safeguards against misuse.

An integrated national data system implies that all participants and stakeholders collaborate in a system in which data are safely produced, exchanged, and used. It does not mean that all data are stored in an integrated national database. And while such collaboration requires close coordination and shared governance between the participants, it does not necessarily require a centralized governance structure.

For many countries, a system in which high-quality data flow and are used safely among various participants remains a distant vision. A low-income country suffering from high levels of poverty, fragility, and poor governance may struggle to produce even the most fundamental data, let alone set up a whole-of-government, multistakeholder approach to data governance. Yet keeping sight of this vision matters for all countries, even those struggling the most with data, because it can serve as a guide in making decisions on how to develop their data systems.

Box 9.1 Relationship between an integrated national data system and a national statistical system

A national statistical system is an ensemble of units within a country that jointly collect, protect, process, and disseminate official statistics.[a] As such, a national statistical system is a core part of the more expansive integrated national data system. The scope of the integrated national data system goes beyond official statistics to encompass the data produced, exchanged, and used by participants from civil society and the public and private sectors for a variety of purposes. The blueprint for building an integrated national data system is a national data strategy, which is a country's plan for capturing greater economic and social value from data in line with the principles of a social contract for data. By contrast, the blueprint for building a national statistical system is a national statistical development strategy, which focuses on official statistics.

a. See the definition of a national statistical system proposed by the Organisation for Economic Co-operation and Development (OECD 2002, 220).
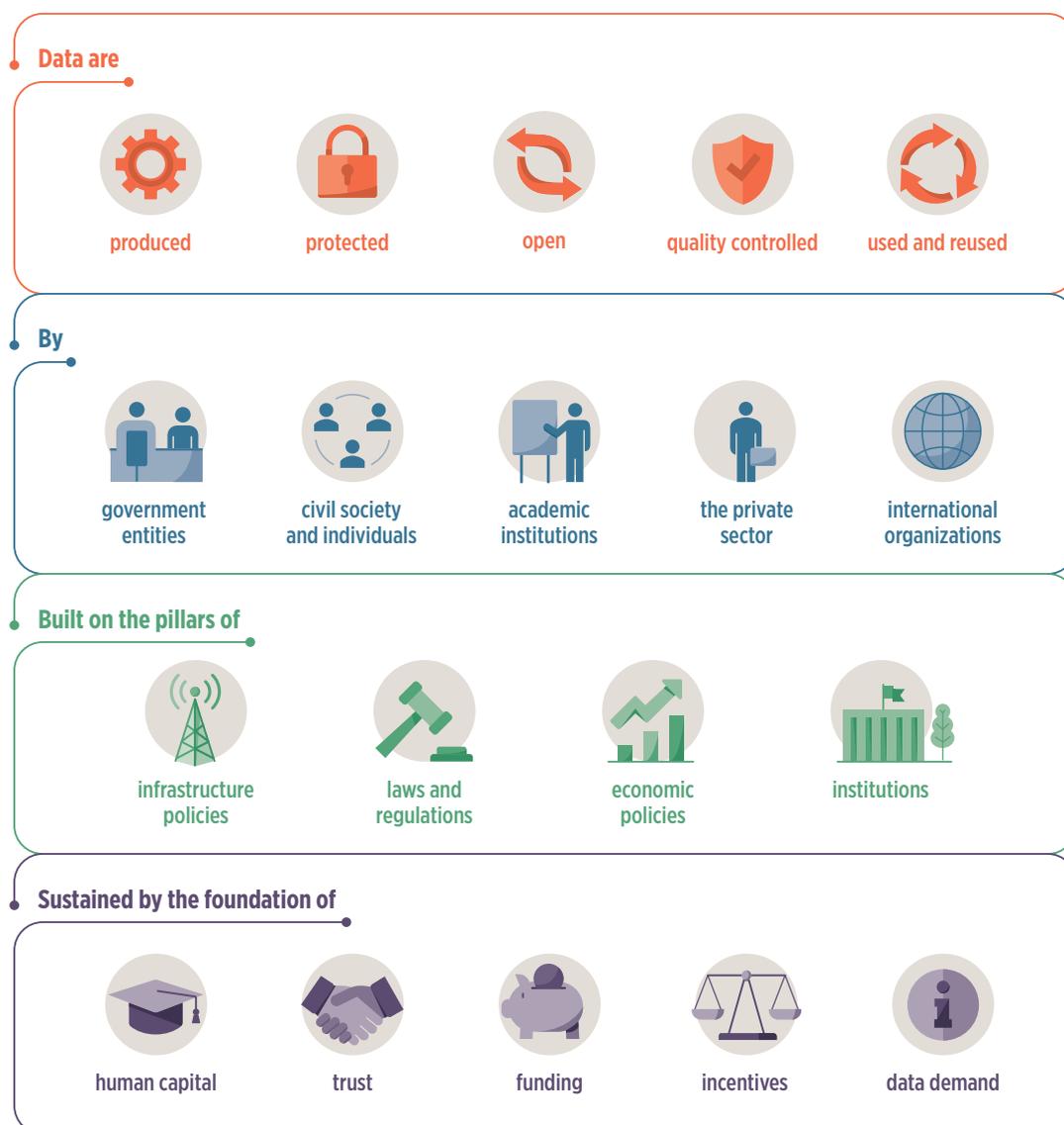
After envisioning what a well-functioning national data system might look like in a frictionless world, this chapter extends the use of the data maturity model of the last four chapters to discuss how countries can move closer to realizing this vision, depending on their context and their level of data maturity. One size will not fit all. Concrete steps to move closer to the vision will critically depend on local factors, many of which are related to political economy issues, such as the strength of institutions and key decision-makers. Another important aspect is the structure of the government: the system will look different in a centralized government structure than in a federal one, for example. But even for countries that remain far from the frontier of good data governance, if their policies address lack of human capital, trust, proper incentives, funding, and a culture of data use, the potential of data for development can be better realized.

The vision of an integrated national data system

An integrated national data system serves a number of important functions; it incorporates various participants from government, civil society, and the private sector; and it is built on the pillars discussed in part II. These pillars rest on a foundation of human capital, trust, funding, incentives, and data demand (figure 9.1).

Figure 9.1 What happens in an integrated national data system?

Data are
produced · protected · open · quality controlled · used and reused

By
government entities · civil society and individuals · academic institutions · the private sector · international organizations

Built on the pillars of
infrastructure policies · laws and regulations · economic policies · institutions

Sustained by the foundation of
human capital · trust · funding · incentives · data demand

Source: WDR 2021 team.

Functions of an integrated national data system

An integrated national data system enables the production of data relevant to development; the equitable and safe flow of data among the participants in the system; and their ability to use and reuse the data while safeguarding against misuse.

Data production. A well-functioning national data system produces data relevant to policy planning, decision-making, and the national discourse. The data meet the needs of the various participants and cover and represent the population of interest. In line with a multistakeholder approach to data production, data produced by private sector entities, civil society organizations, and academia, as well as by citizens, are incorporated where appropriate into the national data system to fill in gaps and enable synergies with government data.

Data protection. To ensure rigorous protection of data and sensitive information, secure storage and transfer of data and safeguards against misuse are in place. This arrangement works as a catalyst for trust and participation in the system. As the types and volume of data expand, the producers and users of data increase, and their interoperability improves, data protection becomes increasingly important for safeguarding the integrity of the system. To the extent possible, robust data protection is achieved by legal and technological solutions before restricting access to data, which is a measure of last resort in an integrated national data system.

Data openness and flow. Open data and interoperability foster the flow of data within government and between the participants in the national data system. Common standards enable the exchange of data across government agencies to improve planning and decision-making as well as to enable cross-border data flows and collaboration.

Data quality control. To safeguard the integrity and quality of the data produced, sound methodological foundations in data production and stringent standards for quality control are adopted. Such foundations also improve the interoperability and comparability of data from different sources. To function

well in this regard, the quality control processes must have a high degree of political independence.

Data use and reuse. The frequent and widespread use and reuse of data propel a successful national data system. A critical aspect is the routine use of data in planning and decision-making across government entities and the use and reuse of data beyond their original purpose, including in business models of data-driven companies, in academic research, in policy making and policy reform, and in informing media content and coverage.

Participants in an integrated national data system

The integrated national data system incorporates participants from the three development pathways—government and international organizations, individuals and civil society, and the private sector. This chapter discusses in more detail five groups of participants: government entities, civil society, academia, the private sector, and international and regional organizations. Although all five groups both produce and use data, each plays a different role in the national data system, which merits separate treatment.

Government entities. Government entities are the primary producers of public intent data for policy and government functions, such as by collecting administrative data through censuses and surveys and through the national statistical system at large. But the role of government entities in the national data system extends beyond producing data for reporting and monitoring to exchanging data across entities and with other participants and using data for policy design and decision-making. Government entities also act as data stewards, setting out the rules that govern data use and ensuring data accessibility and protection. And they act as data managers, laying out and enforcing quality standards and ensuring secure data transactions.

Civil society and individuals. Civil society organizations (CSOs), national and international nongovernmental organizations (NGOs), the media, and individuals play a critical role as the producers and users of data that hold governments and the private sector accountable and highlight issues of public concern. This accountability function also applies to the

production of citizen-generated data that act as a check on official government data if they are in doubt, fill gaps in coverage, or otherwise complement public intent and private intent data. The data produced by civil society are valuable beyond their primary functions, and their value increases through wider use and reuse by other participants in the national data system.

Academia. Academic institutions, think tanks, and research organizations both produce and use data in their research to guide and evaluate policy reforms through impact evaluations and forecasting, for example, and to inform and influence media and the public debate. Academic institutions also provide important education and training for data users and producers in government, the private sector, and civil society, as well as perform data research and development functions in the national data system.

The private sector. Firms in the private sector are prolific producers of data for their business processes, needs, and decisions. Some of these data are very valuable to those making public policy and to the public interest (see chapter 3). Thus the private sector is an important contributor to data production in the national data system, whose data are subject to common standards and quality control. In an integrated national data system, businesses gain from a data-driven culture in terms of competitiveness and profitability, and the national data system facilitates the transition to such a data-driven culture. Businesses also routinely rely on public intent data to improve business decisions and processes or to create new products and services.

International and regional organizations. These groups are de facto players in the national data system. Many international organizations require their members to engage in various types of reporting, such as progress in meeting the United Nations' Sustainable Development Goals (SDGs), which affects national data production.[2] International organizations develop methods and tools that cannot be produced efficiently at the national level. They also commonly act as standard setters. Best practice standards and methods are central to the international comparability of the data produced in the national data system (see spotlight 2.2). In low- and middle-income countries, international organizations also frequently act as donors to support data production. Furthermore, given economies of scale, some goals for a national data system may best be tackled at the supranational level (see spotlight 8.1). Regional organizations can be effective mechanisms for data governance and creating economies of scale in data and statistical capacity.

Pillars of an integrated national data system

An integrated national data system builds on the infrastructure policies, laws and regulations, economic policies, and institutions outlined in part II of this Report.

Infrastructure policies. In a well-functioning integrated national data system, hard and soft infrastructure policies are designed to enable the equitable and trustworthy production, processing, flow, and use of data (see chapter 5). People have access to the internet and can use it properly, consuming an adequate volume of data. Gaps in access—in terms of coverage, usage, and consumption—are addressed. These policies enable countries to improve access to international connectivity, favoring competition along the entire infrastructure supply chain. A proper competitive environment facilitates development of the more complex elements of data infrastructure. The establishment of internet exchange points (IXPs), which requires a competitive market for internet providers, helps create a vibrant digital ecosystem. Well-functioning IXPs attract content providers locally and from abroad. The consequent growth of data consumption generates investments from colocation data centers and cloud providers.

Laws and regulations. In an integrated national data system, laws and regulations guiding data openness, usage, and protection are in place (see chapter 6). Open data laws and access to information legislation complement one another by requiring public institutions to disclose data by default while granting individuals the right to compel disclosure. The rights of individuals on the use of their personal data are recognized and reserved, and an independent data protection authority safeguards those rights. Data controllers and processors are held accountable to ensure cybersecurity. Governments play a stewardship role by incorporating soft law around data use

that reflects societal values, including standards, terms and conditions of use, norms, codes of conduct, and other voluntary frameworks. Both state and non-state participants adhere to this body of soft law (see chapter 6).

Economic policies. Executive-level decision-makers in both the public and private sectors view data as foundational for creating value and are committed to implementing policies to maximize the value of data while ensuring that the proper safeguards are in place. In the whole-of-government strategy for data governance, policies set out norms, objectives, and tools. Antitrust tools are adapted to data-driven markets, and antitrust authorities tackle anticompetitive behavior by data-driven firms. Data can flow securely across borders and facilitate cross-border services transactions, such as in the financial services or telecommunications sectors. Tax loopholes for data-driven businesses are addressed (see chapter 7).

Institutions. The institutions required to effectively govern data are in place (see chapter 8). They include those that enact overarching strategic and policy objectives, such as an executive-level cross-functional group of key stakeholders that makes policy decisions, provides strategic direction, and mobilizes the necessary resources. To execute the strategy and manage the national data system, a repurposed existing institution or a newly created data governance office is fully operational. A national statistical office with sufficient financing, independence, and capacity to fulfill its role is in place. Institutions that monitor compliance, such as a data protection authority and an antitrust authority, are operating. Institutions that monitor and evaluate the system as a whole are created—such as oversight agencies that effectively monitor the accountability of data producers and users and nongovernment watchdogs that monitor public and private sector compliance with rules and standards.

Foundations of an integrated national data system

Putting these pillars in place for the national data system is not easy. They need to be anchored in a solid foundation of human capital, trust, funding, incentives, and data demand. Trust in particular plays a critical role in facilitating the integration of participants and their data. It is essential to binding the national data system to the social contract on data (see chapter 8).

Human capital. Human capital underpins a well-functioning national data system. Data producers have the skills needed to produce high-quality data that measure up to best practices and international standards for processing, storing, and ensuring the interoperability of data. The institutions that safeguard data, perform quality checks, and ensure that data flow among participants are staffed with skilled workers, as are the institutions that lay out the policies, laws, and regulations governing data flows. Individuals have the data literacy needed to ensure that data can be used effectively and equitably, to be empowered, and to hold governments accountable. Data literacy should be understood in a broad sense to include understanding basic statistical and numerical concepts; understanding how to analyze, interpret, and communicate data using digital tools; understanding the place of data in decision-making; and understanding data rights and data governance essentials. Finally, the next generation of data users is trained in data literacy through educational curricula, and the next generation of data scientists and statisticians is trained through higher education, ensuring the sustainability of the national data system.

Trust. For data to flow securely within the national data system, participants in the system trust that the data will be protected, that the information inherent in the data will not be misused, and that the value created from the data will be shared equitably. People trust the ability of government, academia, and the private sector to collect, protect, and safely share data gathered from them. Firms trust that their data will be used properly when those data are shared with third parties. And, in general, participants trust that the public sector enforcement systems are robust and that appropriate measures will be taken in the event of data misuse.

Funding. A well-functioning national data system is sufficiently funded. Government agencies have the resources to hire and pay highly skilled data scientists, statisticians, and data collectors at competitive levels, as well as the resources to purchase the technical infrastructure needed to collect, process, and manage data. Likewise, government agencies have the funding needed to achieve the goals set out in the national data strategy and to sufficiently staff the safeguarding and enabling institutions. Academia has funding to create, access, and analyze data. Civil society and individuals have the financial resources

needed to acquire the technology often needed to monitor government data, produce data themselves, and use data from other participants.

Incentives. The right incentives and power balances conducive to the equitable production, exchange, and use of data are in place. To overcome the reluctance of government entities to share data openly because it could expose poor performance, risk data protection breaches with little return, or shrink their power, data exchanges are mandated or encouraged through incentives, where relevant. Incentives are similarly in place for the private sector, encouraging, and where relevant mandating, businesses to exchange data. In the private sector, such incentives deal appropriately with situations in which corporations may have invested capital in systems to accumulate data and wish to earn a return on the investment, keep data out of the hands of competitors, or both.

Data demand. An integrated national data system has a high demand for data and a culture of valuing and prioritizing data use. Data are viewed as foundational for creating public value through improved policy making, particularly by high-level management in both the public and private sectors. In the public sector, central analytical units and technical staff in ministries gather and analyze data tailored to the needs of decision-makers. In the private sector, companies view data as a valuable asset. For individuals and civil society at large, data are viewed as a tool for empowerment and for holding governments accountable. Programs to improve the data literacy of individuals, journalists, and other stakeholders are in place, ensuring future demand and the long-term sustainability of the national data system. Fact-checking is also well established to challenge the misuse of data.

Realizing the vision

For a country suffering from fragility, poverty, and poor governance, this vision of an integrated national data system may seem unattainable. As discussed in chapter 2, for many governments, just producing high-quality data is a challenge. Thus data exchanges and integration among various partners may not seem feasible. Yet any country can take steps toward fulfilling the vision of an integrated national data system. Using the data maturity model, this section describes the concrete steps countries can take to

move closer to the system envisioned here, focusing on how to integrate in the national data system the various participants: government agencies, civil society, academia, the private sector, and international and regional organizations.

Integrating these participants depends not only on a country's data maturity level, but also on other context-specific factors such as the relative strengths and weaknesses of the current institutions and actors. Where relevant, this discussion explores how local contexts might affect progress toward attaining the vision.

The data maturity model is used as an organizing framework to help determine the strengths and weaknesses of the existing data system and identify the sequential steps that can be taken to establish an integrated national data system. The model differentiates three stages. At low levels of data maturity, countries should prioritize establishing the fundamentals of a national data system. Once the fundamentals are in place, countries should seek to initiate data flows. At advanced levels of data maturity, the goal is to optimize the system (figure 9.2).

Figure 9.2 A data maturity model for a hypothetical national data system
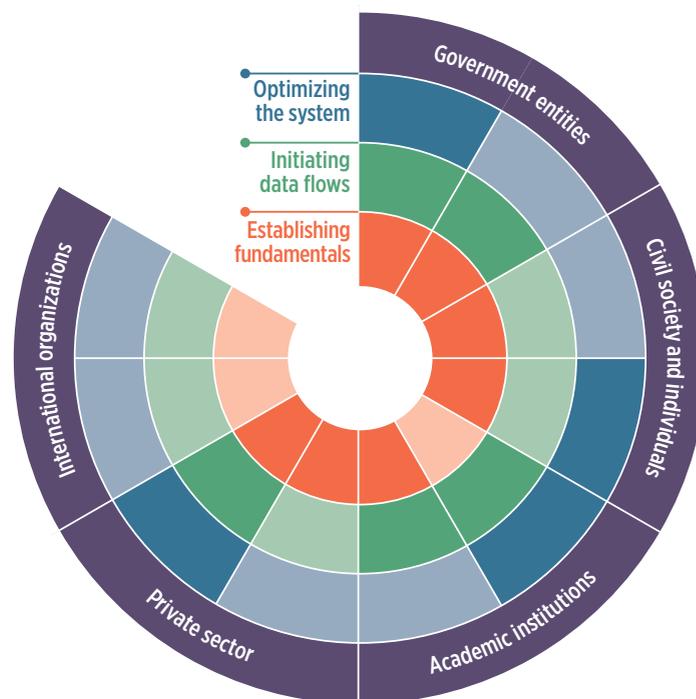
Source: WDR 2021 team.

Note: The figure shows steps in a data maturity model for a hypothetical national data system. The inner circle is the first stage of maturity, the second the middle stage, and so forth. Darker colors indicate steps accomplished; lighter colors indicate steps not accomplished. Thus for each participant, segments may be dark or light. In this way, the figure illustrates that countries may be at different data maturity stages at the same time and that some participants may be more integrated than others.

In practice, deviations from these steps are likely to occur as countries adapt them to their specific circumstances and exigencies. Early steps will likely need to be revisited, refined, and adjusted at later stages. Some countries will be more advanced in some domains but lacking in others—that is, elements identified as fundamentals may not be present in some mostly advanced systems. In a few circumstances, it may be appropriate for a country to change the sequencing of steps in certain domains.
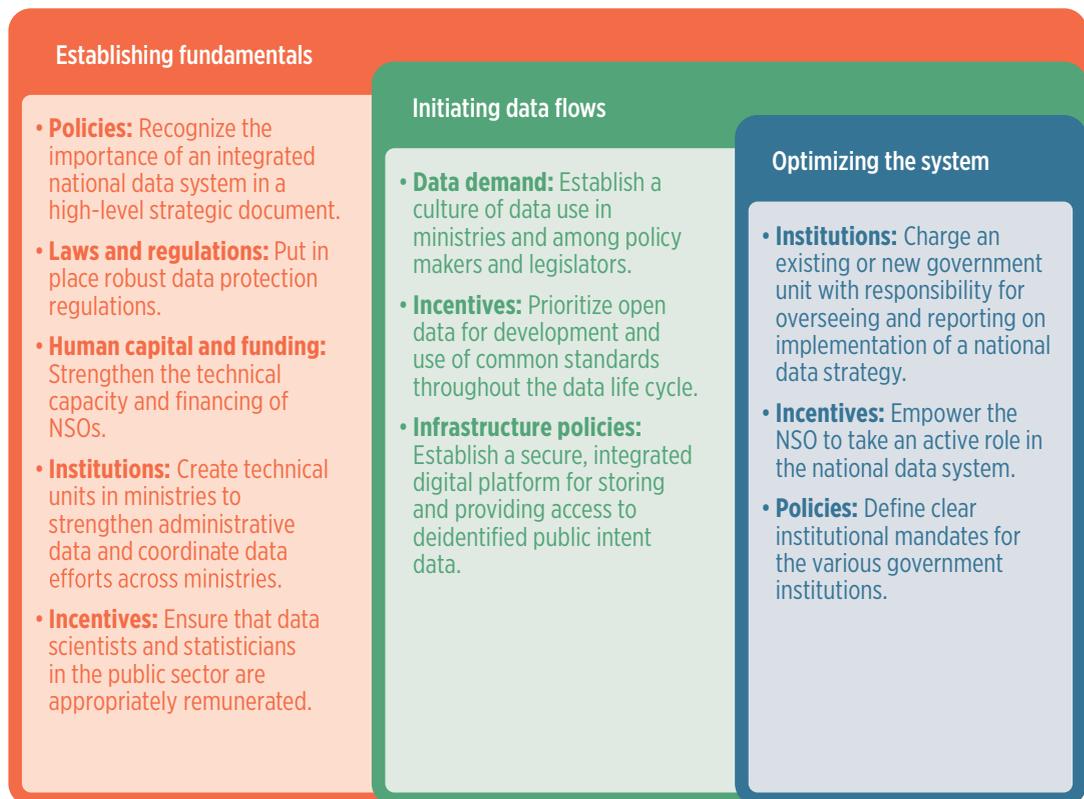
Some countries may not have an intentional whole-of-government approach to data governance but still be advanced in data maturity. Although these countries can have much to gain from taking an intentional whole-of-government approach, an integrated national data system does not call for discarding what has been established, but rather building on its strengths. Regardless of where a country's current data maturity stands, building an integrated national data system will not happen overnight. It is a long-term process of ongoing steps, refinements, and improvements.

Integrating government

Government entities play a central role in the national data system as the main producers of public intent data. To contribute to and sustain a strong national data system, these entities must meet several objectives. They must address shortcomings in the coverage, quality, and usability of public intent data. They need to ensure the effective coordination of public sector data producers and data exchanges and the interoperability of data from various sources. And they must make data available and accessible to stakeholders across the system to promote use, reuse, and repurposing (see figure 9.3).

Data strategy formulation. Recognizing the importance of an integrated national data system in a high-level strategic document, such as a national data strategy, is central to accomplishing these objectives and to garnering the necessary political commitment and resources. The formulation of a national data strategy should be a transparent, collaborative process and include stakeholders from across the government, civil society, academia, and the private

Figure 9.3 Steps to integrating the public sector into the national data system

Establishing fundamentals

- **Policies:** Recognize the importance of an integrated national data system in a high-level strategic document.
- **Laws and regulations:** Put in place robust data protection regulations.
- **Human capital and funding:** Strengthen the technical capacity and financing of NSOs.
- **Institutions:** Create technical units in ministries to strengthen administrative data and coordinate data efforts across ministries.
- **Incentives:** Ensure that data scientists and statisticians in the public sector are appropriately remunerated.

Initiating data flows

- **Data demand:** Establish a culture of data use in ministries and among policy makers and legislators.
- **Incentives:** Prioritize open data for development and use of common standards throughout the data life cycle.
- **Infrastructure policies:** Establish a secure, integrated digital platform for storing and providing access to deidentified public intent data.

Optimizing the system

- **Institutions:** Charge an existing or new government unit with responsibility for overseeing and reporting on implementation of a national data strategy.
- **Incentives:** Empower the NSO to take an active role in the national data system.
- **Policies:** Define clear institutional mandates for the various government institutions.

Source: WDR 2021 team.

Note: Categories overlap and are meant to be illustrative. NSO = national statistical office.

sector to encourage broad-based buy-in. To address shortcomings in public intent data, the national data strategy should reflect the priorities discussed in previous chapters: robust data protection, political commitments to the independence of data producers, adequate and sustained financing, investments in human capital, and efforts to strengthen the data literacy of the general populace, policy makers, legislators, and civil society. This process should also establish a common framework for accountability in and independent oversight of the national data system. To achieve these priorities, the national data strategy should include concrete policy steps, such as the ones that follow, and it should be reflected in the national development plans. For example, in Colombia, National Development Plan 2014–18 was used as a vehicle to formally assign its National Statistical Administrative Department (DANE) the role of coordinator and regulator of the national statistical system.[3]

Data protection regulations. Putting in place robust data and privacy protection regulations is an early priority in establishing an integrated national data system. These regulations should be backed by independent oversight of compliance with them, a function that a data protection authority may serve (see chapter 8). An example of independent oversight is the United Kingdom's Information Commissioner's Office, a nondepartmental body tasked with upholding "information rights in the public interest." Reporting directly to Parliament, it oversees the Data Protection Act, Freedom of Information Act, Privacy and Electronic Communications Regulations (PECR), Environmental Information Regulations, INSPIRE Regulations, and Re-use of Public Sector Information (RPSI) regulations.[4] Although these regulatory steps need to be taken early, they remain relevant for advanced data systems. Legal and institutional arrangements will require adjustment as the data and policy landscapes change.

NSO capacity. Because the NSO fulfills the core function of producing official statistics, it is fundamental that the office be integrated into the national data system. This requires strengthening the technical capacity and financing of NSOs to fill data gaps and produce high-quality official statistics (see chapter 2).

Technical data units. Within government ministries and agencies, the foundations of administrative data should be strengthened. Creating and staffing technical units dedicated to the production and management of administrative data are vital for the participation of ministries in the national data system. Administrative data should be based on common standards promoted across the national data system

for their production, processing, management, and protection. In addition, ensuring the interoperability and accessibility of administrative data must be a priority of administrative data systems. For example, Argentina connects data registers through its data interoperability platform for the public sector (INTEROPER.AR).[5] Statistical units at ministries should centralize cataloging and storage of datasets, including of existing datasets. This will require continually modernizing the technological and data infrastructure for the production, management, safe exchange, and secure storage of data.

Remuneration of data scientists. To ensure a functional integration of government institutions in the national data system, civil servants need the incentives and capabilities to produce, safeguard, and use data. To this end, governments could pursue civil service reforms to ensure that data scientists and statisticians in the public sector are appropriately remunerated. These steps are needed to attract and retain the human capital required to build and sustain a successful national data system. Lack of competitive salaries and career opportunities is a frequently cited barrier to greater institutional performance and the capacity of data producers, for instance in El Salvador, Guatemala, and Peru.[6]

Culture of data use. To initiate data flows, it is vital to establish a culture of data use in ministries and among policy makers and legislators. Institutionalizing data-intensive management practices can jump-start this process (see chapter 2). This effort should be accompanied by ongoing investments in the data literacy of policy makers and legislators. Technical units should be required to periodically deliver knowledge products based on administrative data and disseminated in accordance with a public release calendar. Such products should become an integral part of monitoring, evaluation, and citizen engagement efforts.

Common standards. Open access to public intent data is central to realizing the broad benefits of widespread data use, reuse, and repurposing. On the political front, it is critical that governments prioritize open data for development and use common standards throughout the data life cycle. Government entities should view data stewardship as a strategic function needed for the effective management and use of internal data assets, as well as for seamless data exchanges among entities. To undertake this function, each entity should receive the required financial and human resources and should use common standards for the production, management, quality assurance, and interoperability of public intent data.

Integrated digital platforms. On the technical front, establishing a secure, integrated digital platform for storing and providing access to deidentified public intent data deposited on the platform by producers from across the public sector can initiate data flows and spur further demands for data. The creation of a unified platform should be conditional on putting in place common technological, legal, and institutional standards for safeguarding confidential and sensitive information. An example is Open Data Philippines (ODPH). ODPH was launched in January 2014 as part of the multilateral Open Government Partnership initiative, which also includes Brazil, Indonesia, Mexico, Norway, South Africa, the United Kingdom, and the United States.[7] The ODPH repository works as a core government program ensuring citizens' rights of transparency and access to information. The platform collects more than 1,237 datasets from 99 government agencies and organizations, which allows the disclosure of specific data from different sectors. ODPH acts as an intermediary between the national government and its constituents. It also removes barriers limiting data sharing between agencies.

Data strategy oversight unit. A national data strategy is key to optimizing the national data system. Central to this process is charging an existing or new government unit with overseeing and reporting on implementation of the national data strategy. The unit will play an important coordinating and integrating role in optimizing data flows among participants in the national data system. This role will include acting as liaison with the technical data production teams in ministries and the NSO to support the development and use of common standards for activities across the data life cycle. Of particular importance is ensuring a common and robust approach to the protection of personal data and sensitive information across the system. The institutional home and reporting lines for such a unit will likely differ, depending on the country context. To be effective, the unit should have both the political power to oversee the data agenda of other government institutions and the technical and legal know-how to understand the complexities of data governance.

Many NSOs have the most extensive experience of government agencies in dealing with important data issues. To the extent that a strong and capable NSO exists or that reforms can be readily undertaken to shore up its independence, financial resources, and technical capabilities in line with the recommendations put forth in chapter 2, the NSO may be well placed to co-lead formulation of the national data strategy and possibly to oversee its implementation.[8]

In New Zealand, the NSO is branded as a data agency, and the head of the NSO was appointed Government Chief Data Steward by the State Services Commissioner in July 2017. A cabinet mandate empowers the Chief Data Steward to facilitate and enable "an inclusive, joined-up approach across government to set standards and establish common capabilities, including developing data policy, infrastructure, strategy, and planning."[9]

Estonia has opted for another institutional home for the unit overseeing implementation of the national data agenda. The steering body for implementing the government's digital strategy, which includes overseeing its national data system, is the e-Estonia Council.[10] The council is chaired by the prime minister and organized by the Strategy Unit of Government Office, a public entity charged with assisting the government in designing and implementing policy.[11] Estonia's national data system centers around X-Road, an open-source data exchange layer solution that allows linked public and private databases to automatically exchange information, ensuring confidentiality, integrity, and interoperability among the parties exchanging data.[12] X-Road's cryptography protocols enhance transparency because they log entries into the system and give individuals detailed insights into who is sharing their data and for what purposes.

Similar to Estonia, Argentina has adopted a federated data sharing model that connects data registers and has enabled the development of public services through its data interoperability platform for the public sector. This system includes the Smart Judicial Investigation tool and the National Tax and Social Identification System (SINTyS), which coordinates database exchanges and single data requests at the national, provincial, and municipal levels to support better targeting and monitoring of social programs.[13]

NSO empowerment. Regardless of whether the NSO houses the oversight unit, as the data system matures the NSO should be empowered to take on an active role (see chapter 2). In Estonia, although the NSO does not oversee the national data system, it has been elevated to national data agency status—"a center of excellence" for public sector data administration and research, which would support the making of data-based decisions nationwide, integrating data administration and analytics."[14] In line with this recommendation, NSOs could take on a bigger role in many ways, including conducting data literacy training for stakeholders in government, CSOs, the media, academia, and the private sector, as well as by providing specialized technical assistance to government

departments aimed at improving methods for the production, processing, management, deidentification, and dissemination of public intent datasets. As part of an expanded role in training, NSOs could also develop (or hire) staff to engage in research on new methods for data collection and test the validity of experimental statistics. In addition, NSOs could offer independent quality assurance of the administrative data products and related official statistics produced outside of the NSO.

Institutional mandates. National data strategies should be revisited periodically as the national data system matures in light of evolving data needs and technological improvements. The process of formulating a national data strategy is an opportunity to define clear institutional mandates for the NSO, ministries, and specialized government agencies for the production, quality assurance, exchange, and protection of public intent data. Such mandates can minimize overlapping and duplicated data production, thereby making the whole system more efficient. Defining clear institutional mandates also helps identify the comparative advantages and expected contributions of each institution and simplifies the task of securing financial and human resources commensurate with the mandates of each institution.

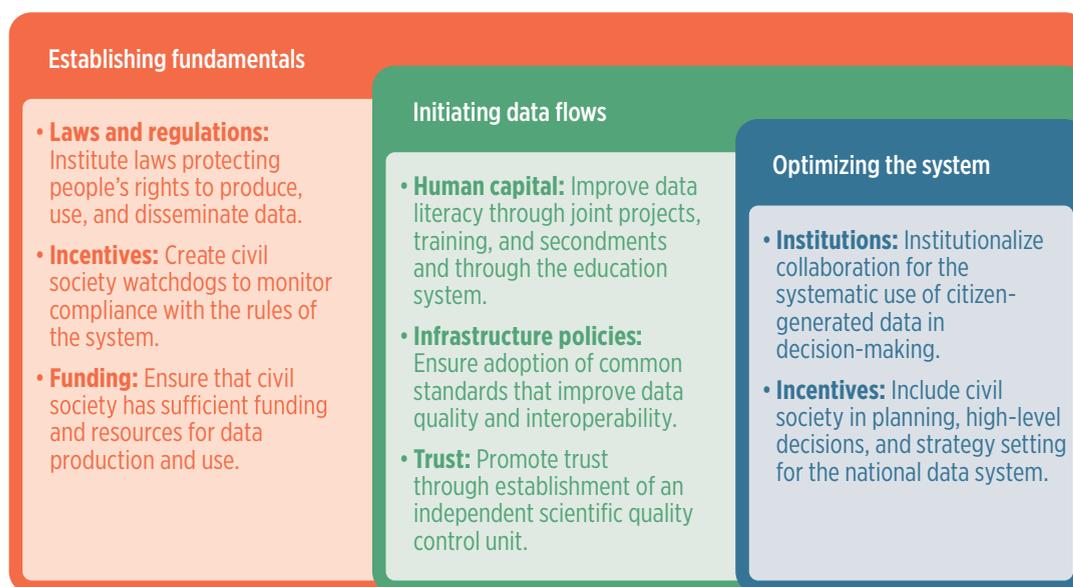### Integrating civil society and individuals

Civil society and individuals should be empowered to participate in the national data system as users of data and as data producers in their own right, whether citizen-generated or collected by CSOs (figure 9.4). A key function of CSOs, national and international NGOs, individual citizens, and journalists and the media is to hold the government and private sector accountable. But the national data system also stands to gain from the systematic incorporation of citizen-generated data for use by other participants. This effort requires collaboration and shared governance arrangements between government and civil society. Because of the importance of civil society's accountability function, these governance arrangements must be set up to reinforce the unconditional independence of civil society data producers and users.

For civil society and individuals to be an integral part of a national data system, several prerequisites must be met.

Legal rights to data production. Laws and regulatory frameworks are needed to protect people's rights to produce, use, and disseminate data. Laws should be amended to support individuals' role in, as well as their accountability function for, the data system. At the same time, laws and regulatory frameworks need to credibly protect data and sensitive information on people so they will trust they can safely participate in the national data system. Instituting civil society watchdogs to monitor public and private sector compliance with the rules and regulations of the data system can act as additional safeguards for independence and data protection.

Figure 9.4 Steps to integrating civil society into the national data system

Establishing fundamentals

- **Laws and regulations:** Institute laws protecting people's rights to produce, use, and disseminate data.
- **Incentives:** Create civil society watchdogs to monitor compliance with the rules of the system.
- **Funding:** Ensure that civil society has sufficient funding and resources for data production and use.

Initiating data flows

- **Human capital:** Improve data literacy through joint projects, training, and secondments and through the education system.
- **Infrastructure policies:** Ensure adoption of common standards that improve data quality and interoperability.
- **Trust:** Promote trust through establishment of an independent scientific quality control unit.

Optimizing the system

- **Institutions:** Institutionalize collaboration for the systematic use of citizen-generated data in decision-making.
- **Incentives:** Include civil society in planning, high-level decisions, and strategy setting for the national data system.

Source: WDR 2021 team.

Note: Categories overlap and are meant to be illustrative.

Funding and resources. Civil society also needs sufficient funding and resources for data production and use. Interviews with representatives of NGOs in Argentina, Kenya, and Nepal revealed that lack of funding can constrain citizen-generated data.[15] Both international donors and national funders could improve funding by including specific budgets for citizen data production alongside funding for institutional data collection.[16] General funds for citizen data collection akin to those for funding of scientific research could also be created. When civil society has limited access to data collection resources, such as smartphones for computer-assisted interviewing or collection of satellite-based global positioning system (GPS) data, such resources could be directly distributed by funders or loaned through organizations dedicated to providing communities with technological tools and training.[17] Open-source software for data collection, such as ODK, as well as free-of-charge software for data collection, such as Survey Solutions and CSPro, could be supported.[18] CSOs also need technical support in adopting and operating such resources and software to ensure they have the capability to produce high-quality data.

Data literacy. Lack of data literacy in civil society is a major barrier to the demand for high-quality, accessible data, and it limits the accountability role that civil society can play. It also leads to low levels of trust in citizen-generated data by other participants, which, in turn, hinders data flows from civil society to the national data system. Improving data literacy through project partnerships, training, and secondments can help address these skill gaps and trust deficits. For example, the Ugandan Bureau of Statistics and Ministry of Education supported the CSO Twaweza in survey and sampling design for a numeracy and literacy survey. Twaweza then independently carried out the data collection and processing, improving the quality of and trust in citizen-generated data. The data were later used by the Ministry of Education.[19] In addition to training, joint projects, fellowships, and secondments of staff from CSOs to data-driven institutions can increase the technical capacities of CSOs.[20] One private sector–led initiative to increase digital literacy in civil society, StoryLab Academy, uses online webinars and face-to-face training to improve digital literacy among African journalists.[21] The academy is a joint initiative of Code for Africa, the World Bank's Global Media Development Program, and Google News Lab.

Data literacy should also be reinforced across society more broadly. One aspect is incorporating data literacy in primary and secondary education curricula (chapter 2). Empirical studies on financial literacy suggest that integrating data literacy into school curricula may ultimately be more effective than targeted adult education.[22] Because of the relatively young populations of lower- and middle-income countries, incorporating data literacy programs in school curricula could reap valuable returns. For example, Rwanda has supported initiatives to build digital skills through its multistakeholder Digital Ambassador Program.[23] In addition to general programs, targeted data literacy efforts to reach traditionally marginalized groups such as women and indigenous communities may be needed to reduce the digital divide.

Common standards. Data flows between civil society and other actors can also be promoted by adopting common standards that improve data quality and interoperability. The efforts to promote adoption of these standards should be augmented by efforts aimed at strengthening analytical capacity to ensure their proper implementation. For example, the collaboration between Twaweza and public institutions helped ensure that data on literacy and numeracy were collected in accordance with official educational standards.[24] Similarly, in Mozambique the standardization of community scorecards used by several NGOs to assess school and health care services at the local level allowed the data to be aggregated to the national level, where they were used in research and advocacy campaigns.[25]

Quality control unit. For citizen-generated data to be reused—such as to inform policy decisions—their quality and representativeness need to be guaranteed. One specific concern is advocacy bias in cases in which the primary purpose of the data produced by civil society is to advocate for certain issues. To this end, the relevant government agencies and CSOs could together establish an independent scientific quality control unit to assess the methodological soundness and representativeness of citizen-generated data for possible use in national data portals and in SDG reporting.[26] Closer collaboration and the adoption of common standards do not mean that civil society ceases to play its critical role of holding other actors accountable by challenging their data, views, or priorities. Methodological rigor and common standards may in fact empower civil society to play an accountability role by increasing the credibility and interoperability of citizen-generated and CSO data. At the same time, not all CSOs may opt for their data to be used for policy purposes or SDG reporting if advocacy is their primary concern. In any case, data should flow to civil society to empower and inform decisions of communities and individuals and ensure that benefits from an integrated data system are broadly shared.

Institutionalized collaboration. To optimize data flows, the relationship between civil society and other participants should evolve from a stage of ad hoc collaboration to institutionalized collaboration for the production and use of data. From the perspective of the government, this means that citizen data production and use are by default integrated into policy-making and administrative processes.[27] An example is the framework developed by South Africa's Department of Planning, Monitoring and Evaluation on how to include citizen-based monitoring in planning, budgeting, and evaluation systems.[28] Institutionalized collaboration allows the government to leverage civil society's unique perspectives, local expertise, and motivation, whereas civil society can influence policies and services, highlighting problems that may otherwise go unnoticed or be ignored. With institutionalized collaboration, civil society could have a designated role of collecting data that would otherwise be too expensive or difficult to collect, such as wildlife counts. Institutionalized collaboration on data would also serve as an incentive for relying on data for policy making more broadly, and thus would bolster demand for data.

Joint planning and strategy setting. Finally, governments should take steps to include civil society in planning, high-level decisions, and strategy setting for the national data system. In Chile, where civil society participation is mandated by the national Law on Associations and Citizen Participation in Public Management, the NSO has put in place a civil society council.[29] Once included in multistakeholder governance forums, civil society can progress from being a standard-taker to a standard-maker. Including citizens in the planning, production, and use of data can also help empower individuals and enhance trust between citizens and their governments.[30]
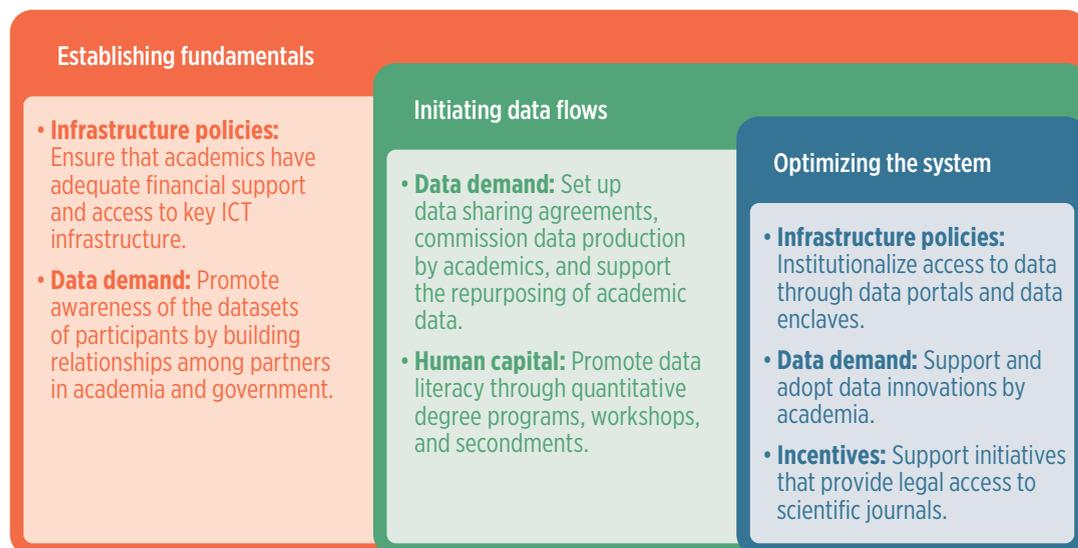
Integrating academia

In a well-functioning national data system, academia (including universities, think tanks, and research organizations) generates data and insights, advances the methodological frontier, and trains other participants in data production and use (figure 9.5).

For academia to realize this potential, several conditions must be in place.

Technological infrastructure. Academics must have adequate financial support and access to key information and communication technology (ICT) infrastructure. Higher education funding should prioritize and include designated budgets for data infrastructure. Where limited, access to ICT infrastructure required for any work involving large datasets should be expanded.

Data awareness. Any form of cooperation and data exchange between academia and other participants in the national data system requires awareness of which databases are maintained by participants. A survey of policy makers in 126 low- and middle-income countries found that they learn about domestic data sources primarily through consultations and informal communications, highlighting the important

Figure 9.5 Steps to integrating academia into the national data system

Establishing fundamentals

- **Infrastructure policies:** Ensure that academics have adequate financial support and access to key ICT infrastructure.
- **Data demand:** Promote awareness of the datasets of participants by building relationships among partners in academia and government.

Initiating data flows

- **Data demand:** Set up data sharing agreements, commission data production by academics, and support the repurposing of academic data.
- **Human capital:** Promote data literacy through quantitative degree programs, workshops, and secondments.

Optimizing the system

- **Infrastructure policies:** Institutionalize access to data through data portals and data enclaves.
- **Data demand:** Support and adopt data innovations by academia.
- **Incentives:** Support initiatives that provide legal access to scientific journals.

Source: WDR 2021 team.

Note: Categories overlap and are meant to be illustrative. ICT = information and communication technology.

role of personal interactions and social capital.[31] At low data maturity levels, promoting awareness of the datasets of participants by building relationships between partners in government and academia, such as through workshops, is therefore a priority. As in civil society, laws and regulations need to protect academics' rights to collect and share data.

Data exchange agreements. To initiate data flows, researchers can leverage relationships with other participants in the national data system and set up project-based data exchange agreements—for example, to access administrative data collected by the government. Other participants may take advantage of academia's expertise and commission the generation of data for their needs. Data generated by academics as part of their research could be repurposed by other participants in the national data system, from government policy makers to students pursuing higher education. An example of how researchers can make their data available for repurposing is the Datahub for Field Experiments in Economics and Public Policy, a public searchable database that researchers affiliated with Innovations for Poverty Action (IPA) and the Abdul Latif Jameel Poverty Action Lab (J-PAL) use to publish their data from impact evaluations.[32] It is critical to ensure that the data made available for downstream use are sufficiently documented, not only to confirm replicability of past findings but also to properly inform future use.

Data literacy. Academia also plays an important role in the flow of human capital and the promotion of data literacy. Tertiary education programs should be geared toward training professionals skilled in using data who are prepared to join institutions in the data system. At lower data maturity levels, any programs providing skills in quantitative fields, such as statistics, economics, or computer science, will be useful. Beyond formal tertiary education, academia can train participants from government agencies, the private sector, or civil society by means of training courses, workshops, and seminars. For example, the Data Literacy pillar of the Data-Pop Alliance—an alliance created by the Harvard Humanitarian Initiative, MIT Connection Science, and Overseas Development Institute—has developed a framework and tools to establish core competencies toward becoming data literate.[33] The Nepal Data Literacy Program, established in 2019 through a partnership between the Nepalese government, the World Bank, and the United Kingdom's Department for International Development (since incorporated into the Foreign, Commonwealth and Development Office), comprises a 100-hour modular, customizable pedagogy to support both building technical skills and efforts to enhance a culture

of data use among Nepalis.[34] The program is now partnering with the Kathmandu University School of Management to incorporate data literacy toolkits into university programs and develop a data-driven course that will be free to other institutions and thousands of students as a result.

Embedding a team of researchers in public institutions is another effective way of transferring skills. In Peru, IPA and J-PAL partnered with the Ministry of Education to embed a team of researchers in the ministry. They then worked with public officials to conduct several impact evaluations using administrative data. The ministry subsequently scaled up three programs based on the evaluation results, and the unit is now government-run.[35]

Data portals and enclaves. To optimize data flows to academia, academia's access to public intent data could be institutionalized by establishing data portals and data enclaves. The latter enable researchers to use confidential microdata from surveys and government censuses behind secure firewalls. For example, Mexico's National Institute of Statistics and Geography (INEGI) offers researchers data access through its Microdata Laboratory, which is located in secure enclaves on its premises, provided they undergo an application and training process.[36] Similar institutions have been set up in other countries, such as DataFirst in South Africa[37] and the Scientific Research Center at the Palestinian Central Bureau of Statistics in the West Bank and Gaza.[38] Implementing these models requires a secure infrastructure, skills in the deidentification of data, trust, and sanctions for misuse and attempts to reidentify individuals. When public intent data are made accessible to researchers, data producers should require that insights gained from the data flow back to them such as in the form of technical briefs or open-access journal articles. This requirement helps ensure that data exchanges serve broader development objectives. Microdata access should also extend to metadata and syntax files to increase ease of use and transparency.

Data innovations. Finally, innovations emanating from academia should be supported and, where relevant, adopted. Academia can play a role in transferring and applying global knowledge to local contexts. For example, randomized experiments in international development research were originally pioneered by academics at elite universities, but since then they have proliferated and been adopted as a decision-making tool by many governments, including those in low- and middle-income countries.[39] A prominent example is Mexico's National Council for the Evaluation of Social Development Policy (CONEVAL), which was set up in 2004 with the mandate to coordinate

evaluation exercises of the National Social Development Policy, as well as to provide guidelines to define, identify, and measure poverty.[40] The agency, endowed by the government of Mexico with budgetary, technical, and management autonomy, implements or commissions evaluations of the social policies developed by the Mexican government.

Access to scientific journals. Local research institutions, partnering with governments, have acted as important knowledge brokers in this process. Domestic researchers can also push the methodological frontier and develop methods tailored to the specific country contexts. For academia to play this role, it needs not only sufficient funding and academic freedom, but also access to international journals, databases, and exchange opportunities. For example, in light of high access fees, especially for scholars and institutions in low- and middle-income countries, initiatives that provide legal access to scientific journals should be supported and scaled. One example of such initiatives is Research4Life.[41] Similarly, researchers from high-income countries should be more responsive to requests for full-text publications from other researchers on platforms such as ResearchGate or LinkedIn.

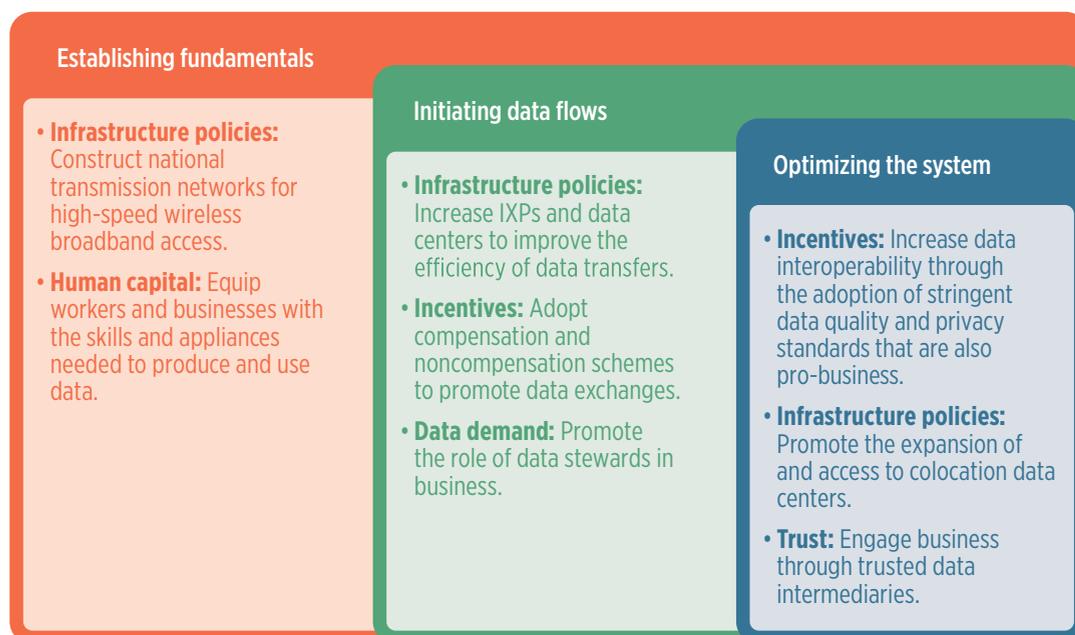### Integrating the private sector

Targeted policies, initiatives, and incentives are needed to support businesses through the transition to a data-driven culture and enable them to become active participants in the national data system (figure 9.6).

High-speed wireless broadband. Establishing reliable, efficient physical infrastructure is foundational to the production and use of data and is an obvious prerequisite to integrating the private sector into a national data system. As data traffic expands globally in volume and velocity, businesses can participate in national and international data systems only if they have access to reliable, affordable high-speed wireless broadband.[42] Notwithstanding the high investment costs involved in the construction and operation of the national transmission networks (backbones) connected via fiber-optic cable and satellite to international links, governments need to find ways to support this vital infrastructure (see chapter 5).[43] Incentives, including subsidies, can be used to encourage existing or new operators to invest in less lucrative geographical areas. And there may be opportunities to leverage fiber-optic infrastructure in other sectors such as utilities and railways. Fiber can also be installed cost effectively in conjunction with new road construction.[44]

Data literacy. Another priority in the early stages of the integrated national data system is to equip workers and businesses with the skills and appliances needed to produce and use data, including having access to mobile devices and computing and data management infrastructure. Data systems

Figure 9.6 Steps to integrating the private sector into the national data system

Establishing fundamentals

- **Infrastructure policies:** Construct national transmission networks for high-speed wireless broadband access.
- **Human capital:** Equip workers and businesses with the skills and appliances needed to produce and use data.

Initiating data flows

- **Infrastructure policies:** Increase IXPs and data centers to improve the efficiency of data transfers.
- **Incentives:** Adopt compensation and noncompensation schemes to promote data exchanges.
- **Data demand:** Promote the role of data stewards in business.

Optimizing the system

- **Incentives:** Increase data interoperability through the adoption of stringent data quality and privacy standards that are also pro-business.
- **Infrastructure policies:** Promote the expansion of and access to colocation data centers.
- **Trust:** Engage business through trusted data intermediaries.

Source: WDR 2021 team.

Note: Categories overlap and are meant to be illustrative. IXPs = internet exchange points.

require workers with specialized skills in statistics, economics, computer science, geographic information systems, and data science to allow businesses to collect, store, and process data in the first place. Government-led or -funded training initiatives on data literacy for the current labor force are essential to equip workers with the skills required in today's labor markets and create the necessary data literacy and demand for data. In particular, much more needs to be done to scale up the availability—and gradually the sophistication—of public data literacy programs for the workforce outside of formal secondary or tertiary education systems. Government-funded examples include the Nepal Data Literacy Program and the Sudan Evidence-Base and Data Literacy Capacity Development Program, which has developed an introductory data literacy course and an intermediate-level "data storytelling" curriculum.[45]

Installation of IXPs. Increasing the efficiency of data transfers is central to initiating data flows within the private sector and across the integrated national data system. Governments can take steps to encourage private-led investment in the installation of IXPs, which reduce internet access costs and improve performance for users. For example, it is estimated that the absence of an IXP regional interconnection infrastructure forced Latin America to pay nearly US$2 billion in international traffic costs in 2014; increasing IXPs in the region could reduce overall traffic costs by one-third.[46] At the technical level, building an IXP is relatively simple and inexpensive, but establishing and maintaining the necessary level of trust and collaboration between stakeholders can be a challenge.[47] To establish an IXP, internet service providers (ISPs) and other actors (many of whom are competitors) must agree on IXP location, mode of operation, and management structure—all of which should be neutral to ensure buy-in.[48] For example, when the first IXP, KIXP, was established in Kenya in 2000, a legal challenge filed by the incumbent telecom operator led to its immediate closure. Only after a year of appeals and persistent lobbying was KIXP allowed to reopen.[49] In Mexico, the first IXP was not established until 2014 because of problems around trust and a lack of collaboration between ISPs. The second was installed in 2018.

Compensation and noncompensation schemes. To initiate and maintain flows, proactive compensation and noncompensation schemes to encourage private sector data exchanges are critical. Business-to-government (B2G) data exchanges are generally based on voluntary contractual agreements. The government should typically take a somewhat restrained approach to measures that force the private sector to exchange data (see chapters 6 and 7). Although data are a nonrival good and the reasons in favor of data sharing are compelling, the private sector does not necessarily have to provide the government with access to its data for free. The European Commission has suggested that compensation schemes could be in the public interest. Options include free data; free of charge plus tax incentives; marginal cost pricing; marginal cost pricing plus a return on investment markup; or market price.[50] The Contracts for Data Collaboration (C4DC) initiative has created an analytical framework and an online library of key elements of data sharing agreements to reduce the transaction costs of negotiating data sharing between the private sector and policy makers.[51]

Other types of incentives to increase B2G data exchanges without compensation include public recognition programs that showcase engagement. Such a program could enhance the reputations of companies and reduce the amount of compensation expected (if any). By increasing the transparency of B2G data exchange arrangements, societal expectations about the utility of private intent data for public policy, as well as expectations about B2G engagement, could gradually shift, thereby encouraging other businesses to join. Governments could also consider marketing mechanisms such as labeling schemes that could be used to highlight B2G data exchanges undertaken to pursue public policy goals.[52]

Companies could be incentivized to share their data through corporate social responsibility programs. For example, Facebook's Data for Good initiative is offering innovative datasets intended to aid public policy decisions. Other private companies may be encouraged to share their data at a reduced cost for public initiatives, with special grants for researchers or tax breaks for the data provider. If the private sector considers the risks of data exchanges to be too great, data intermediaries can facilitate arrangements in which the relevant algorithms are sent to companies directly for local data analysis. For example, OPAL (Open Algorithms), which describes itself as a "non-profit socio-technological innovation," provides a platform that allows researchers to send companies certified open-source algorithms that are then run on big data behind companies' firewalls.[53] This arrangement allows governments and others to analyze and gain insight from granular datasets collected by private companies that otherwise would be unavailable for legal, commercial, or ethical reasons.

Data stewards. Data exchanges can be advanced by businesses designating data stewards within their

organizations to oversee internal data governance and engage with others. Data stewards play a vital role in establishing good data management in the private sector and in pursuing and facilitating sustainable data exchange arrangements.[54] Alongside their business-led functions, data stewards could be tasked with identifying data that could be shared to promote the public interest and identifying and nurturing potential collaboration with the government or others, such as data collaboratives. Data stewards can also lead efforts to ensure that any insights gained from exchanges are acted on. Although primary oversight of data protection issues should be assigned to a chief privacy officer, in the context of data exchange and reuse data stewards should be responsible for protecting potentially sensitive information and ensuring the protection of data when reused.[55]

Common standards. Transitioning to high data maturity levels requires facilitating the interoperability of public and private data through the adoption of stringent data quality and privacy standards that are also pro-business. Like all participants in the national data system, businesses must be incentivized to adopt stringent standards for data quality and interoperability to facilitate integration with public systems. Businesses' adoption of such standards could be promoted through some of the incentive mechanisms discussed earlier in this chapter and rolled out through advanced training programs. Uptake by businesses can also be increased as countries adopt international standards that improve the cross-border interoperability of data in the commercial sphere—such as the International Organization for Standardization (ISO) standard for electronic data interchange between financial institutions (ISO 20022).[56] For private sector data to be safely integrated in the national data system, businesses further need to comply with the data protection and privacy regulations put in place as a fundamental step in moving toward an integrated national data system.

Colocation data centers. As firms become more reliant on data systems, they will need access to colocation data centers to help manage their data processing and storage needs and to reduce the costs associated with running and certifying internal data centers.[57] Access to data centers in lower-income countries remains poor, in part because of these countries' fragile business environments and low demand for data (see chapter 5).[58] Appropriate measures to promote the expansion of data centers and increase access will depend on the context. In high-capacity, high-demand contexts, data centers may need to be located relatively close to users to maximize cost savings and speed, and changes to the local business climate

may be sufficient to encourage the necessary investments. Where local capacity and demand are low, however, regional efforts to promote investment in regional data centers and other digital infrastructure may be more appropriate. Governments might also explore opportunities to work with large businesses in-country that already have in-house data storage systems and that could, if aggregated, create the necessary scale for colocation data centers to meet local needs. So long as the business of running carrier-neutral colocation data centers proves profitable in a particular context, securing private investment should be straightforward.[59] In Africa, at least 20 new private sector data centers are expected to come online by 2021, which will bring the total to more than 100 across the continent.[60]

Data intermediaries. Trusted data intermediaries can be used to optimize B2G data flows in more mature data systems (see chapter 8). Data intermediaries can provide sophisticated, data-driven businesses with the assurances they need about the security of their data, combined with strong accountability and transparency mechanisms that grant them more control over and visibility of data use.[61] For example, the Reserve Bank of India (RBI), in response to national guidelines introduced by the Ministry of Electronics and Information Technology (MEITy) for standardizing consent for data sharing, is institutionally separating the collection of customer consent from data processing to enhance trust in their data management processes and use.[62]
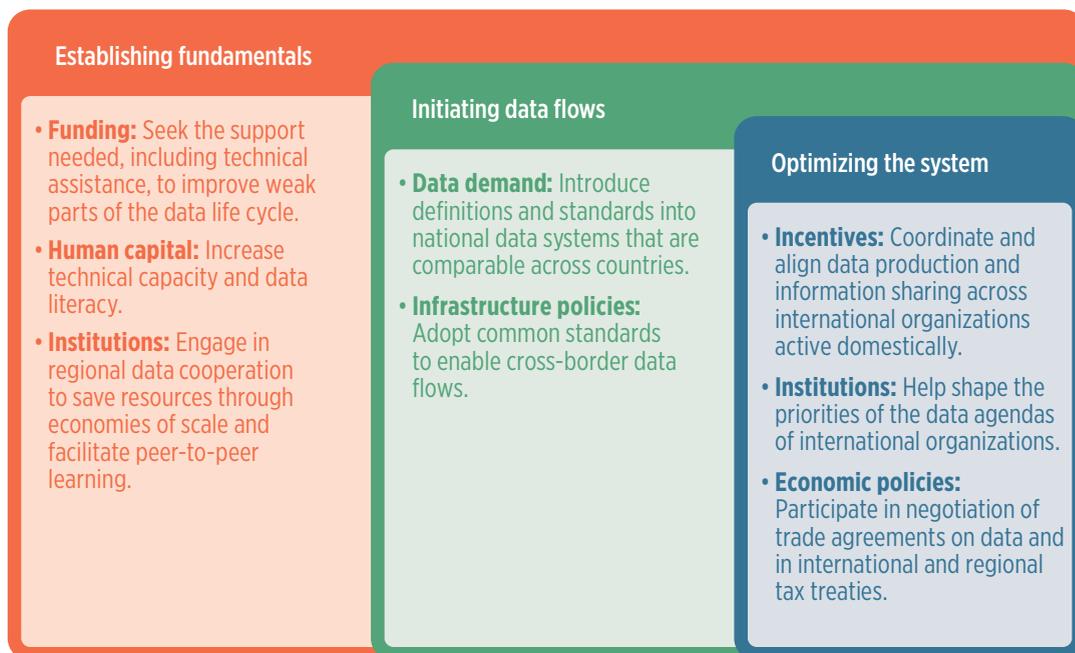
Integrating international and regional organizations and collaborating across borders

International and regional organizations, donors, as well as international NGOs are important participants in the national data system by collecting their own data, funding country-level data collection, setting international standards, and using country data for monitoring and analysis. International and regional organizations are also forums for cross-border collaboration on data production and exchange as well as data governance.

Although national governments have rather limited control over international organizations and their agendas, they can take steps to integrate these institutions into the national data system in a beneficial fashion (figure 9.7).

Technical assistance. At low levels of data maturity, countries often struggle with limited resources for setting up a national data system, but they could utilize funding, technical assistance, and global public

Figure 9.7 Steps to integrating international and regional organizations into the national data system and collaborating across borders

Establishing fundamentals

- **Funding:** Seek the support needed, including technical assistance, to improve weak parts of the data life cycle.
- **Human capital:** Increase technical capacity and data literacy.
- **Institutions:** Engage in regional data cooperation to save resources through economies of scale and facilitate peer-to-peer learning.

Initiating data flows

- **Data demand:** Introduce definitions and standards into national data systems that are comparable across countries.
- **Infrastructure policies:** Adopt common standards to enable cross-border data flows.

Optimizing the system

- **Incentives:** Coordinate and align data production and information sharing across international organizations active domestically.
- **Institutions:** Help shape the priorities of the data agendas of international organizations.
- **Economic policies:** Participate in negotiation of trade agreements on data and in international and regional tax treaties.

Source: WDR 2021 team.

Note: Categories overlap and are meant to be illustrative.

goods from international organizations to address weak spots in the data life cycle. International and regional organizations are well placed to level the playing field by putting in place the conditions that would enable countries with the least data maturity to begin catching up to their more data mature peers. For some countries, this may mean obtaining funding for core data production. For others, it may be seeking assistance with data storage and management. For still others, international organizations can help deidentify datasets, assist in the adoption of improved methods and tools for data production, and suggest modernization of statistical laws and regulations to ensure they are conducive to safe data exchanges. For these steps to be effective, governments need to assess where in the data life cycle they might need support.

Data literacy. At the early stages of data maturity, governments may also rely on international organizations and development partners for programs aimed at improving technical capacity and data literacy. The former could include short-term training, such as the World Bank's C4D2 Training Initiative, which provides statisticians in low- and middle-income countries with specialized training in the collection, analysis, and use of microdata. It could also

include twinning arrangements between statistical agencies that could create opportunities for on-the-job or postgraduate training and staff exchanges or secondments, such as those carried by Statistics Norway.[63] One example of a regional institution created to address data literacy, among other things, is the African Union Institute for Statistics.[64] It is important to keep long-term sustainability in mind for such programs. Training a handful of staff in an agency is of little value if the retention rate of these staff is low. Through the Data for Policy Initiative, the World Bank has committed to sustainable technical support of national statistical systems in at least 30 low-income countries.[65] Long-term institutional relationships with such agencies increase the ease with which they can adopt demanding international best practices.

Regional cooperation. Starting in the early stages of data maturity, countries can use cross-border collaboration to save resources through economies of scale. Countries can cooperate in setting up certain functions of the national data system at a supranational level when trying to perform these functions in each country individually would be inefficient and could precipitate balkanization. For example, through the Statistics for Development Division of the Secretariat

of the Pacific Community, nations are working together on data collection, analysis, dissemination, and methodology, thereby reducing costs across the data life cycle.[66] Regional collaboration in data-related regulatory matters can also be beneficial. The African Tax Administration Forum (ATAF) is one example (see spotlight 7.2).[67]

Most countries are members of a regional network where peer-to-peer learning can facilitate the adoption of best practices. In these networks, countries can learn from peers that are one step further down the road. This learning can spur innovation and help countries move up the data maturity model. For example, experienced member countries of the Programme for International Student Assessment (PISA), a benchmarking initiative of the Organisation for Economic Co-operation and Development (OECD), have shared their experiences with new program members, facilitating comparable measurement of educational outcomes internationally.[68] In the area of competition, international exchanges of knowledge would be particularly useful in improving understanding of antitrust issues in data-driven markets.

Common standards. After the resource limitations are addressed, data should flow from international organizations to domestic participants and vice versa, as well as across borders. To successfully and securely initiate data flows to and from international organizations, data must be internationally comparable and anchored in common standards. National decision-makers could insist on introducing cross-country comparable definitions and measures into project monitoring, evaluation, and high-level strategic documents. Working closely with international standard setting organizations is instrumental to this end. For example, the System of National Accounts (SNA), the international standard for measuring economic activity, includes a set of internationally agreed-on concepts, definitions, classifications, and accounting rules. The SNA has facilitated the comparability of macroeconomic statistics internationally, with 90 percent of countries using at least the 1993 SNA standard.[69] Common standards also facilitate comparisons across countries, allowing international organizations to better prioritize resources.

Similarly, international organizations can play an important role in coordinating and supporting the development of national statistical systems that are comparable and compatible across countries. For example, the Cape Town Global Action Plan for Sustainable Development Data, adopted by the UN Statistical Commission, provides a roadmap for the funding and modernization of national statistical systems

needed around the world to monitor the Sustainable Development Goals.

Coordination. To integrate international organizations into the national data system and to avoid overlapping and conflicting initiatives, domestic actors need to ensure that the data roles and responsibilities of international agencies within a country are coordinated. In India, this challenge was solved by creating sectoral committees in which the country offices of various United Nations (UN) organizations, ministries, and research institutions participated. Through these committees, the SDG-related activities and technical support of the various international agencies were divided across regions and domains in a nonoverlapping manner, anchored in the UN Resident Coordinator's office.[70] A similar model could be replicated or refined to ensure the efforts of international agencies are coordinated. Although this step is needed to optimize flows, such coordination is crucial in countries with less developed national data systems where many donors are active.

Data agendas of international organizations. As countries build their capacity and obtain more resources, their scope for influencing the data agenda of international organizations increases. Countries can work to ensure that the agendas of international organizations are guided by country needs and priorities. Such an effort can minimize competing agendas and better align data needs and data gaps between national and international agencies, maximizing, in turn, the relevance of data and thus data exchanges among participants.

Trade agreements. At this stage, countries can also seek to leverage international and regional organizations to participate in the negotiation of trade agreements on data aimed at facilitating cross-border trade in data. This may be tackled bilaterally—such as the Digital Trade Agreement between Japan and the United States[71]—or attempted through the World Trade Organization (see chapter 7 and spotlight 7.2). To capture tax revenues from the multinational digital economy, an international tax treaty will be necessary. Similarly, countries should seek to coordinate their antitrust authorities' regulatory actions on data-driven businesses across borders (see spotlight 7.2).

Integrating the national data system

A successful national data system creates an environment in which the value of data for development can be maximized. The impact of data on development increases with the number of participants safely producing, exchanging, using, reusing, and repurposing the data. Incorporating the various participants in

the national data system is a central task in building an integrated system. This chapter has laid out steps for how to approach this task. Some cross-cutting themes that have emerged from this discussion have the potential to tie together and strengthen the entire system:

- *Data literacy and data education* are prerequisites for people's participation in the national data system. Better data literacy improves policy-making and business decisions and strengthens efforts to hold governments and the private sector accountable. Data literacy also boosts trust in data.
- *Stringent and shared approaches to data protection* are necessary for participants to trust the integrity of data production and use and initiate data exchanges.
- *Data interoperability, comparability, and reliability through common standards and quality control* allow data to be integrated from different sources and boost their usefulness. These standards and quality controls may need to be developed. Data stewards can play an important role in ensuring quality and interoperability.
- *Data openness and accessibility* through means such as digital platforms make widespread data use and reuse possible.

Investing in these cross-cutting steps can have wide-ranging benefits for all participants in the national data system and for development, but they will require commensurate financial and political commitments.

Meanwhile, countries are already constructing national data systems, whether intentionally or not. Some of this is happening through day-to-day government activities, such as service delivery and monitoring of programs. Some of this is happening because events such as the COVID-19 pandemic are accelerating change. And some of this is happening because technological advances are ushering in sweeping transformations on an unprecedented scale. As the country examples in this Report show, pushes and pulls across the economy and society are shaping the construction of national data systems implicitly or explicitly.

This *World Development Report* on data for development advocates an intentional, comprehensive, multistakeholder, collaborative approach to constructing an integrated national data system that aims to maximize the development benefits of data while minimizing the risks. This approach takes into account those now left out of or marginalized in the data economy.

It prescribes a data-driven culture that can creatively and constructively use, reuse, and repurpose data. It calls on countries to shape their system based on their own circumstances, including their own capabilities, values, and political economy. It recognizes the complexity of this endeavor but recommends a phased approach to make it happen. It moves away from reactive steps to proactive ones. It calls on the international community to help countries take these steps and to provide the standardization, harmonization, and tools necessary to make it work. And, not least, it is the considered aim of this Report to foster a global discussion that can truly help data improve lives.

Notes

1. A similar argument is made in MacFeely (2020).
2. International and regional organizations are also key producers of data. One important role they have is producing transnational data—flows, interactions, and links between countries or phenomena that are difficult or impossible for a country itself to record. For example, the United Nations Office on Drugs and Crime collects data on illicit activities between countries, such as trafficking in drugs, that one country alone cannot collect because the object of interest leaves the country of origin without being detected.
3. Dargent et al. (2020).
4. British Academy and Royal Society (2017).
5. OECD (2019a).
6. Dargent et al. (2020).
7. See Department of Information and Communications Technology, ODPH (Open Data Philippines) (dashboard), https://data.gov.ph/. See also Aceron (2018), Warwick (2017), and an example of open government data in Australia at Digital Transformation Agency, Search for Data (dashboard), https://data.gov.au/.
8. United Nations Statistical Commission (2021) contains case studies of the role of national statistical offices in national data systems.
9. Stats NZ (2018).
10. MKM (2018).
11. See Government Office, Government of Estonia, "E-Estonia Council," https://www.riigikantselei.ee/en /supporting-government/e-estonia-council.
12. Kivimäki (2018); World Bank (2021).
13. OECD (2019a).
14. MKM (2018, 22).
15. Piovesan (2015).
16. Gray, Lämmerhirt, and Bounegru (2016); Lämmerhirt et al. (2018).
17. Lämmerhirt et al. (2018).
18. See ODK (dashboard), https://getodk.org/. See also US Census Bureau, Census and Survey Processing System (CSPro) version 7.5.0 (dashboard), https://www.census .gov/data/software/cspro.html; World Bank, Survey Solutions version 21.01 (dashboard), https://mysurvey .solutions/en/.

19. Carranza (2018); Gray, Lämmerhirt, and Bounegru (2016).
20. Wilson and Rahman (2015).
21. For more information, see Code for Africa, Academy Africa: Courses (dashboard), https://academy.africa/courses.
22. Bruhn, Lara Ibarra, and McKenzie (2014); Bruhn et al. (2016); Frisancho (2018); Lührmann, Serra-Garcia, and Winter (2018). However, the long-term effects are not yet known (Entorf and Hou 2018).
23. Bizimungu (2017).
24. Gray, Lämmerhirt, and Bounegru (2016).
25. Lämmerhirt et al. (2018).
26. Cázarez-Grageda et al. (2020) and MacFeely and Nastav (2019) present more elaborate proposals for how data from civil society can be used for tracking SDGs and in official reporting, and they establish quality frameworks enabling NSOs to engage with CSOs.
27. Lämmerhirt et al. (2018).
28. Lämmerhirt et al. (2018).
29. Carranza (2018).
30. Misra and Schmidt (2020).
31. Masaki et al. (2017).
32. Harvard Dataverse, Datahub for Field Experiments in Economics and Public Policy (data repository), https://dataverse.harvard.edu/dataverse/DFEEP?q=&types=dataverses%3Adatasets%3Afiles&sort=dateSort&order=desc&page=1.
33. Data-Pop Alliance brings together researchers, experts, practitioners, and activists to change the world with data through three pillars of work: diagnosing local realities and human problems with data and artificial intelligence (AI); mobilizing capacities, communities, and ideas toward more data-literate societies; and transforming the systems and processes that underpin societies and countries. In 2016 Flowminder Foundation joined as the fourth Core Member. For more information, see ThoughtWorks, Data-Pop Alliance (dashboard), https://datapopalliance.org/.
34. See Nepal Data Literacy Program, Data Literacy for Prosperous Nepal (data literacy portal), https://dataliteracy.github.io/.
35. Ministry of Education, MineduLAB (dashboard), http://www.minedu.gob.pe/minedulab/.
36. Volkow (2019).
37. University of Cape Town, DataFirst (data repository), https://www.datafirst.uct.ac.za/.
38. Palestinian Central Bureau of Statistics, "Research Center," West Bank and Gaza, http://www.pcbs.gov.ps/site/lang__en/598/default.aspx.
39. Many governments have established dedicated monitoring and evaluation agencies or even ministries. For Spain, see "Building a Monitoring and Evaluation Framework for Open Government," chapter 4, pages 117–40, in OECD (2019b).
40. See Consejo Nacional de Evaluación de la Política de Desarrollo Social (National Council for the Evaluation of Social Development Policy), "About Us? Features" (¿Quiénes Somos? Funciones), Mexico City, https://www.coneval.org.mx/quienessomos/Conocenos/Paginas/Funciones.aspx.
41. Bohannon (2016). For information on Research4Life, see https://www.research4life.org/about/.
42. Katz and Callorda (2018).
43. Vertically integrated operators have constructed most national transmission networks. As long as there is strong competition among several players, final consumer prices can be affordable and networks can be resilient. Where private investment is lacking, a government could construct a state-owned transmission network, although the implications for public debt could be severe. Alternatively, a government could take on a coordinating role among operators to create a heterogenous backbone, requiring open access to and cost-based pricing for operator fiber routes.
44. An example is landlocked Mongolia's north-south fiber-optic backbone connecting it to China and the Russian Federation runs along the railway (Tsolmondelger 2019).
45. For a description of the Sudanese program, see Moscoso (2016). For elements of the training course, see "Welcome to the Sudan Evidence Base Programme–Data Literacy Training," https://sudanebp.tuvalabs.com/.
46. Agudelo et al. (2014).
47. Rosa (2018).
48. Kenya's IXP, launched in 2000, was Africa's first IXP. It was established and is run by the local ISP industry association, the Telecommunication Service Providers of Kenya. See Jensen (2012); Technology Service Providers of Kenya, "KIXP Background," https://www.tespok.co.ke/?page_id=11651.
49. Jensen (2012).
50. High-Level Expert Group on Business-to-Government Data Sharing (2020).
51. Dahmm (2020). Contracts for Data Collaboration (C4DC) is a joint initiative of SDSN TReNDS, New York University's GovLab, the World Economic Forum, and the University of Washington.
52. High-Level Expert Group on Business-to-Government Data Sharing (2020).
53. OPAL was created by groups at MIT Media Lab, Imperial College London, Orange, the World Economic Forum, and Data-Pop Alliance. For more information, see the OPAL website, https://www.opalproject.org/home-en.
54. High-Level Expert Group on Business-to-Government Data Sharing (2020).
55. GovLab (2020).
56. See International Organization for Standardization, About ISO 20022: Governance (dashboard), https://www.iso20022.org/about-iso-20022/governance.
57. A colocation data center is a facility equipped with networked computers providing remote storage, processing, and distribution of data where multiple data service providers may colocate. They are mainly operated by global information technology (IT) companies, governments, and enterprises that host other companies' data (known as colocation). For the relevant standard of the International Organization for Standardization (ISO), see ISO, "ISO/IEC 27001:2013(en)," at OBP (Online Browsing Platform) (database), https://www.iso.org/obp/ui/#iso:std:iso-iec:27001:ed-2:v1:en.

58. This is often attributed to a lack of demand, as well as aversion to a country's perceived high risk of natural disasters, unpredictable political environment, barriers to doing business, and unreliable energy and internet infrastructure (C&W 2016).
59. Munshi (2020).
60. ADCA (2020).
61. World Bank (2021).
62. RBI (2019).
63. SSB (2020).
64. More broadly, the Global Network of Institutions for Statistical Training works to build sustainable statistical capacities through efficient and harmonized training programs.
65. Dabalen, Himelein, and Rodríguez-Castelán (2020).
66. Statistics for Development Division, Pacific Community, "Pacific Statistics Methods Board (PSMB)," Nouméa, New Caledonia, https://sdd.spc.int/pacific-statistics-methods-board-psmb.
67. ATAF (2020).
68. OECD (2018).
69. Fifty percent of countries are using the latest—the 2008 SNA standard. See Statistics Division, Department of Economic and Social Affairs, United Nations, National Accounts (database), https://unstats.un.org/unsd/nationalaccount/; World Bank, Statistical Performance Indicators (database), http://www.worldbank.org/spi.
70. Recently, a coordination forum was instituted in which key stakeholders collaborate on issues related to support for statistical monitoring of SDG goals and targets, including the use of new technologies and capacity development to track SDG-related outcomes.
71. Office of the United States Trade Representative, U.S.–Japan Digital Trade Agreement Text (dashboard), https://ustr.gov/countries-regions/japan-korea-apec/japan/us-japan-trade-agreement-negotiations/us-japan-digital-trade-agreement-text.

References

Aceron, Joy. 2018. "Independent Reporting Mechanism (IRM): Philippines End of Term Report 2015–2017." Open Government Partnership, Washington, DC. https://www.opengovpartnership.org/wp-content/uploads/2018/10/Philippines_End-of-Term_Report_2015-2017.pdf.

ADCA (Africa Data Centres Association). 2020. "State of the African Data Centre Market 2020." ADCA, Abidjan, Côte d'Ivoire. http://africadca.org/en/state-of-the-african-data-centre-market-2020.

Agudelo, Mauricio, Raúl Katz, Ernesto Flores-Roux, María Cristina Duarte Botero, Fernando Callorda, and Taylor Berry. 2014. *Expansión de Infraestructura Regional para la Interconexión de Tráfico de Internet en América Latina.* Infraestructura Series. Caracas, República Bolivariana de Venezuela: Development Bank of Latin America. http://scioteca.caf.com/handle/123456789/522.

ATAF (African Tax Administration Forum). 2020. "Domestic Resource Mobilisation: Digital Services Taxation in Africa." Policy Brief 01, ATAF, Pretoria, South Africa.

https://events.ataftax.org/index.php?page=documents&func=view&document_id=61.

Bizimungu, Julius. 2017. "New Initiatives to Drive Digital Literacy and Use of Tech Services." *New Times*, September 30, 2017. https://www.newtimes.co.rw/section/read/220887.

Bohannon, John. 2016. "Who's Downloading Pirated Papers? Everyone." *Science* (blog), April 28, 2016. https://www.sciencemag.org/news/2016/04/whos-downloading-pirated-papers-everyone.

British Academy and Royal Society. 2017. "Data Management and Use: Governance in the 21st Century: A Joint Report by the British Academy and the Royal Society." British Academy and Royal Society, London. https://royalsociety.org/topics-policy/projects/data-governance/.

Bruhn, Miriam, Gabriel Lara Ibarra, and David McKenzie. 2014. "The Minimal Impact of a Large-Scale Financial Education Program in Mexico City." *Journal of Development Economics* 108 (May): 184–89. https://doi.org/10.1016/j.jdeveco.2014.02.009.

Bruhn, Miriam, Luciana de Souza Leão, Arianna Legovini, Rogelio Marchetti, and Bilal Zia. 2016. "The Impact of High School Financial Education: Evidence from a Large-Scale Evaluation in Brazil." *American Economic Journal: Applied Economics* 8 (4): 256–95. https://doi.org/10.1257/app.20150149.

C&W (Cushman and Wakefield). 2016. "Data Centre Risk Index." C&W, London. https://verne-global-lackey.s3.amazonaws.com/uploads%2F2017%2F1%2Fb5e0a0da-5ad2-01b3-1eb8-8f782f22a534%2FC%26W_Data_Centre+Risk_Index_Report_2016.pdf.

Carranza, Javier. 2018. "Citizen to Government Data Partnerships: What Can We Learn from and Recommend to Civil Society Groups Working in the Official Statistics Domain?" Eurostat Statistical Working Paper KS-TC-18-007-EN-N, European Union, Luxembourg. https://doi.org/10.2785/728477.

Cázarez-Grageda, Karina, Julia Schmidt, and Rajiv Ranjan. 2020. "Reusing Citizen-Generated Data for Official Reporting: A Quality Framework for National Statistical Office–Civil Society Organization Engagement." PARIS21 working paper, PARIS21 Consortium, Paris.

Dabalen, Andrew L., Kristen Himelein, and Carlos Rodríguez-Castelán. 2020. "Data for Policy (D4P) Initiative." Poverty and Equity Note 23, World Bank, Washington, DC. https://openknowledge.worldbank.org/handle/10986/33857.

Dahmm, Hayden. 2020. "Laying the Foundation for Effective Partnerships: An Examination of Data Sharing Agreements." UN Sustainable Development Solutions Network's Thematic Research Network on Data and Statistics (SDSN TReNDS) on behalf of Contracts for Data Collaboration (C4DC).

Dargent, Eduardo, Gabriela Lotta, José Antonio Mejía-Guerra, and Gilberto Moncada. 2020. "Who Wants to Know? The Political Economy of Statistical Capacity in Latin America." Inter-American Development Bank, Washington, DC.

Entorf, Horst, and Jia Hou. 2018. "Financial Education for the Disadvantaged? A Review." IZA Discussion Paper 11515, Institute of Labor Economics, Bonn, Germany.

Frisancho, Verónica. 2018. "The Impact of School-Based Financial Education on High School Students and Their Teachers: Experimental Evidence from Peru." IDB Working Paper IDB-WP-871, Inter-American Development Bank, Washington, DC. https://doi.org/10.18235/0001056.

GovLab (Governance Lab). 2020. "Wanted: Data Stewards; (Re-)Defining the Roles and Responsibilities of Data Stewards for an Age of Data Collaboration." GovLab, Tandon School of Engineering, New York University, Brooklyn, NY. http://www.thegovlab.org/static/files /publications/wanted-data-stewards.pdf.

Gray, Jonathan, Danny Lämmerhirt, and Liliana Bounegru. 2016. "Changing What Counts: How Can Citizen-Generated and Civil Society Data Be Used as an Advocacy Tool to Change Official Data Collection?" DataShift, Civicus, Johannesburg. http://civicus.org /thedatashift/wp-content/uploads/2016/03/changing -what-counts-2.pdf.

High-Level Expert Group on Business-to-Government Data Sharing. 2020. *Towards a European Strategy on Business-to-Government Data Sharing for the Public Interest.* Luxembourg: European Union. https://www.euractiv.com /wp-content/uploads/sites/2/2020/02/B2GDataSharing ExpertGroupReport-1.pdf.

Jensen, Mike. 2012. "Promoting the Use of Internet Exchange Points: A Guide to Policy, Management, and Technical Issues." Internet Society Report, Internet Society, Reston, VA. https://www.internetsociety.org/wp-content /uploads/2012/12/promote-ixp-guide.pdf.

Katz, Raul, and Fernando Callorda. 2018. "The Economic Contribution of Broadband, Digitization, and ICT Regulation." Thematic Report: Regulatory and Market Environment Series, International Telecommunication Union, Geneva. https://www.itu.int/pub/D-PREF-EF .BDR-2018.

Kivimäki, Petteri. 2018. "X-Road as a Platform to Exchange MyData." Nordic Institute for Interoperability Solutions, Tallinn, Estonia. https://www.niis.org/blog/2019/10/30/x -road-as-a-platform-to-exchange-mydata.

Lämmerhirt, Danny, Jonathan Gray, Tommaso Venturini, and Axel Meunier. 2018. "Advancing Sustainability Together? Citizen-Generated Data and the Sustainable Development Goals." Global Partnership for Sustainable Development Data, United Nations, New York. http://www.data4sdgs.org/resources/advancing -sustainability-together-citizen-generated-data-and -sustainable-development.

Lührmann, Melanie, Marta Serra-Garcia, and Joachim Winter. 2018. "The Impact of Financial Education on Adolescents' Intertemporal Choices." *American Economic Journal: Economic Policy* 10 (3): 309–32. https://doi.org/10.1257 /pol.20170012.

MacFeely, Steve. 2020. "In Search of the Data Revolution: Has the Official Statistics Paradigm Shifted?" *Statistical Journal of the IAOS* 36 (4): 1075–94.

MacFeely, Steve, and Bojan Nastav. 2019. "You Say You Want a [Data] Revolution: A Proposal to Use Unofficial Statistics for the SDG Global Indicator Framework." *Statistical Journal of the IAOS* 35 (3): 309–27.

Masaki, Takaaki, Samantha Custer, Agustina Eskenazi, Alena Stern, and Rebecca Latourell. 2017. "Decoding Data Use: How Do Leaders Use Data and Use It to Accelerate Development?" AidData, Global Research Institute, College of William and Mary, Williamsburg, VA.

Misra, Archita, and Julia Schmidt. 2020. "Enhancing Trust in Data—Participatory Data Ecosystems for the Post-COVID Society." In *Shaping the COVID-19 Recovery: Ideas from OECD's Generation Y and Z.* Paris: Organisation for Economic Co-operation and Development.

MKM (Ministry of Economic Affairs and Communications, Estonia). 2018. "Digital Agenda 2020 for Estonia." MKM, Tallinn, Estonia. https://www.mkm.ee/sites/default/files /digital_agenda_2020_web_eng_04.06.19.pdf.

Moscoso, Sandra. 2016. "Increasing Data Literacy to Improve Policy-Making in Sudan." *Governance for Development* (blog), March 15, 2016. https://blogs.worldbank.org /governance/increasing-data-literacy-improve-policy -making-sudan#.

Munshi, Neil. 2020. "Africa's Cloud Computing Boom Creates Data Centre Gold Rush." *Technology Sector* (blog), March 2, 2020. https://www.ft.com/content/402a18c8 -5a32-11ea-abe5-8e03987b7b20.

OECD (Organisation for Economic Co-operation and Development). 2002. *Measuring the Non-Observed Economy: A Handbook.* Paris: OECD. http://www.oecd.org/sdd/na /1963116.pdf.

OECD (Organisation for Economic Co-operation and Development). 2018. "Peer-to-Peer Learning to Strengthen Dissemination of PISA for Development Results." PISA for Development Brief 27, OECD, Paris. https://www .oecd.org/pisa/pisa-for-development/27_Peer_learning _results_dissemination.pdf.

OECD (Organisation for Economic Co-operation and Development). 2019a. *Digital Government Review of Argentina: Accelerating the Digitalisation of the Public Sector.* OECD Digital Government Studies Series. Paris: OECD. https://www.oecd-ilibrary.org/governance/digital -government-review-of-argentina_354732cc-en.

OECD (Organisation for Economic Co-operation and Development). 2019b. *Open Government in Biscay.* OECD Public Governance Reviews Series. Paris: OECD. https://doi .org/10.1787/a70e8be3-en.

Piovesan, Federico. 2015. "Statistical Perspectives on Citizen-Generated Data." DataShift, Civicus, Johannesburg, South Africa. http://civicus.org/thedatashift/wp-content /uploads/2015/07/statistical-perspectives-on-cgd_web _single-page.pdf.

RBI (Reserve Bank of India). 2019. "Master Direction–Non-Banking Financial Company–Account Aggregator (Reserve Bank) Directions." Document RBI/DNBR/2016-17/46, Master Direction DNBR.PD.009/03.10.119/2016-17, RBI, Mumbai. https://www.rbi.org.in/Scripts/BS_View MasDirections.aspx?id=10598.

Rosa, Fernanda R. 2018. "Internet Node as a Network of Relationships: Sociotechnical Aspects of an Internet Exchange Point." Paper presented at TPRC46: Research Conference on Communications, Information, and Internet Policy, Washington College of Law, American University, Washington, DC, September 21–22, 2018.

SSB (Statistisk sentralbyrå, Statistics Norway). 2020. "Annual Report 2019: International Development Cooperation in Statistics Norway." Plans and Reports 2020/1,

SSB, Oslo. https://www.ssb.no/en/omssb/om-oss/vaar
-virksomhet/planer-og-meldinger/_attachment/416480?
_ts=1711b58ce28.

Stats NZ (Statistics New Zealand). 2018. "Data Strategy and
Roadmap." Fact Sheet, Wellington, New Zealand. https://
www.data.govt.nz/assets/Uploads/fact-sheet-data-road
map-12422-oct-18.pdf.

Tsolmondelger, Odkhuu. 2019. "ICT Infrastructure along
Transport Network." Presentation, Information Commu-
nications Network LLC, Ulaanbaatar, Mongolia, Novem-
ber 20, 2019. https://www.unescap.org/sites/default
/files/ICT%20Infrastructure%20Along%20Transport
%20Network%2C%20Mongolia%20NetCom.pdf.

United Nations Statistical Commission. 2021. "Approaches
to Data Stewardship." Paper prepared for the High-Level
Group for Partnership, Coordination and Capacity-
Building for Statistics for the 2030 Agenda for Sustain-
able Development. https://unstats.un.org/unsd/statcom
/52nd-session/documents/BG-3a-DataStewardship-E.pdf.

Volkow, Natalia. 2019. "Harnessing the Potentiality of
Microdata Access Risk Management Model." Paper
presented at Joint UNECE/Eurostat Work Session on
Statistical Data Confidentiality, Session 1.1, Conference
of European Statisticians, The Hague, Netherlands,
October 29–31, 2019.

Warwick, Mara. 2017. "Philippines: Open Data Launch."
Speeches and Transcripts, World Bank, Washington, DC.
https://www.worldbank.org/en/news/speech/2017
/03/02/open-data-launch.

Wilson, Christopher, and Zara Rahman. 2015. "Citizen-
Generated Data and Governments: Towards a Collabo-
rative Model." DataShift, Civicus, Johannesburg, South
Africa. http://civicus.org/images/citizen-generated%20
data%20and%20governments.pdf.

World Bank. 2021. *Unraveling Data's Gordian Knot: Enablers
and Safeguards for Trusted Data Sharing in the New Economy.*
Washington, DC: World Bank.