

HOW GOVERNMENT CAN PROMOTE OPEN DATA AND HELP UNLEASH OVER \$3 TRILLION IN ECONOMIC VALUE

Open data has the potential to unleash innovation and transform every sector of the economy. Government can play a critical role in ensuring that stakeholders capture the full value of this information.

MICHAEL CHUI is a partner at the McKinsey Global Institute, and DIANA FARRELL is a director in McKinsey's Washington, DC, office, where KATE JACKSON is a consultant.

¹Additional information about the methodology and approach to size the economic impact of open data can be found in the McKinsey report released in October 2013, *Open data: Unlocking innovation and performance with liquid information*. This report is a joint effort of the McKinsey Global Institute, the McKinsey Center for Government, and McKinsey's Business Technology Office. It provides analysis of the seven domains, including important levers for unlocking open data's value, examples of how the value is being realized today, discussion of how the major stakeholders might become involved in open-data initiatives, and considerations of both the barriers and enablers to unlock economic value.

A recent McKinsey report, *Open data: Unlocking innovation and performance with liquid information*, identified more than \$3 trillion in economic value globally that could be generated each year through enhanced use of open data—increasingly “liquid” information that is machine readable, accessible to a broad audience at little or no cost, and capable of being shared and distributed. The sources of value from open data identified in the report include new or increased revenue, savings, and economic surplus in seven domains: consumer finance, consumer products, education, electricity, health care, oil and gas, and transportation.¹ The domains were chosen to provide a diverse understanding of how

open data, such as census demographics, crop reports, and information on product recalls, can create value in public- and private-sector organizations, in consumer and business-to-business markets, and in products and services.

This report expands on an important topic in the earlier report: the critical role of government in unlocking the economic value of open data, managing risks, and engaging stakeholders (Exhibit 1). As a provider, catalyst, user, and policy maker, government is in a unique position to define an agenda for open data that improves decision making; fosters the growth of innovative businesses, products, and services; and enhances accountability. Government is also well positioned to encourage private companies and other stakeholders to share their data—one of the best strategies for creating value.

In this report, we first explore what open data is and why the trend is taking hold.

We evaluate the primary levers that generate value from open data and government's role in enabling them. The report then turns to the risks of open-data programs; these are top of mind for many citizens, businesses, and policy makers but can be mitigated through approaches we outline. We also describe the needs of stakeholders and opportunities for them to actively create, promote, and participate in open-data strategies and applications. To conclude, we present a discussion of how public-sector leaders can launch, assess, and scale up their open-data programs.

We hope this report helps stakeholders—both inside and outside government—to design, build, and participate in effective open-data programs. For elected officials and civil servants, it can serve as a guide for understanding, evaluating, and promoting open-data initiatives. We include examples of how various governments shape their agen-

Exhibit 1

Government plays a critical role to help enable value creation, manage risks, and engage stakeholders in open data.



Exhibit 2

‘Openness’ or ‘liquidity’ of data can be defined by four characteristics.

	Completely open ← More liquid → Completely closed	
Accessibility	Everyone has access	Access to data is to a subset of individuals or organizations
Machine readability	Data are available in formats that can be easily retrieved and processed by computers	Data are in formats not easily retrieved and processed by computers
Cost	No cost to obtain	Offered only at a significant fee
Rights	Unlimited rights to reuse and redistribute data	Reuse, republishing, or distribution of data is forbidden

das and invest in resources that can be used as models for other agencies or regions. For business, technology, and start-up leaders, our report aims to provide a relevant and perhaps missing perspective on when and how they might engage with government. Last, our report will highlight for citizens, the media, and nongovernmental organizations (NGOs) the potential of open data to improve services that directly benefit consumers.

Open data: What it is, why it has taken hold, and how it varies

What it is. Companies, governments, and NGOs have begun to release and share vast amounts of information. However, the extent to which information is truly open varies in four ways: accessibility, machine readability, cost, and rights for reuse and redistribution. Exhibit 2 shows how data are classified across the spectrum, from completely open to completely closed, using a definition closely related to those of many

international organizations, such as the Open Knowledge Foundation (OKF).²

Open data can be local, regional, and global in scope—from procedure costs at a single hospital to city-level water use to revenues from national tax collection. Governments capture and share large amounts of data on a range of topics, but businesses and other private organizations also have information they could make more liquid, such as data on sales transactions, shipping patterns, and medical records.

The value of data can vary depending on its level of detail, accuracy, and compatibility. For example, a government could annually release 600 data sets that contain aggregated information, such as the number of robberies committed in a month in a particular neighborhood. Although this information is important, it has less potential to generate value than more specific data that are released more frequently, such as a weekly report showing the time of day and

²The definition, from opendefinition.org, is as follows: “A piece of data or content is open if anyone is free to use, reuse, and redistribute it—subject only, at most, to the requirement to attribute and/or share-alike.” For more on the subject, see okfn.org.

location of each robbery, or even a real-time feed of crime reports as they are generated. Similarly, data released in commonly accepted formats that can be easily shared and reviewed by external parties are more valuable than information presented in nonstandard formats. And the release of raw data can create more value than the release of a research analysis, since this allows for more original uses.

Discussions of open data frequently reference “big data,” “big data analytics,” and “MyData,” all of which fall into different dimensions than open data (Exhibit 3):

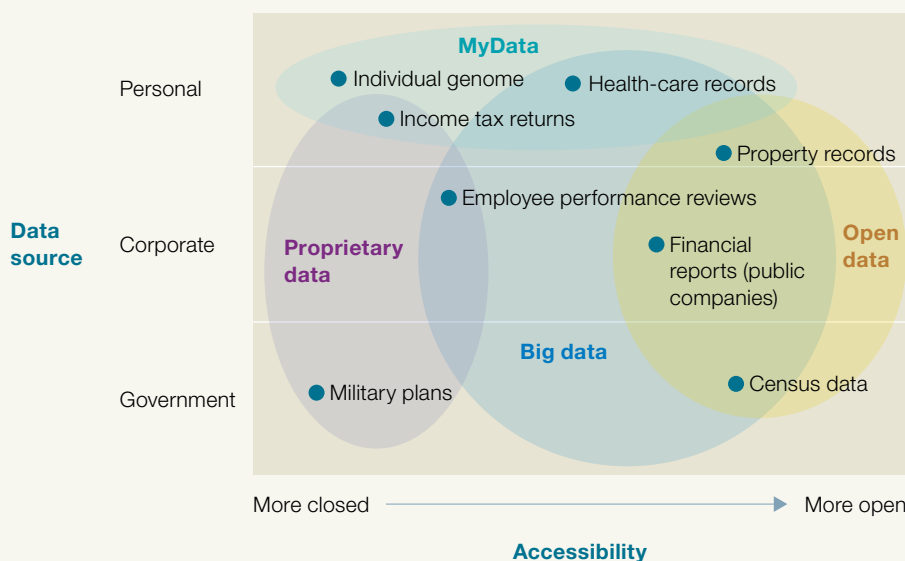
- Big data refers to information that is increasingly voluminous and real-time in

nature, as well as diverse in its sources and types. It represents a source for ongoing discovery and analysis.

- Big data analytics involve examining large amounts of data to uncover insights, such as process efficiencies, as well as connections between events that can generate competitive advantages and improve operations.
- MyData refers to data about an individual consumer to which that person is granted access. Upon receiving access, people can correct inaccuracies, gain real-time and historical information about themselves (such as their physical health and financial standing), and compare themselves with

Exhibit 3

Open data can come from individuals, companies, or governments, and it differs from other data types in its degree of availability.



others. In some cases, individuals can also specify how their personal data can be used.

Why open data is taking hold. Interest in open data has been increasing for many years. We are addressing this topic now because the trend is gaining momentum and is top of mind for businesses, governments, NGOs, citizens/consumers, and the media. Information sources, particularly those from government, are now readily available at lower cost and in machine-readable forms. From electronic health records and student-achievement histories to site-specific energy-consumption readings and mobile-phone location data, a vast and rising quantity of information is collected, cataloged, and shared by companies, governments, and consumers. The expansion of open data, combined with advances in big data analytics, is freeing information that was once trapped inside the dusty pages of overlooked reports, enabling improved decision making, new product and service offerings, and greater accountability.

This change comes at a time of heightened focus on data-driven knowledge and evidence-based decision making. More public- and private-sector leaders are seeking access to information and pressing for standards, guidelines, and regulations that will ensure the sustainability and usability of open-data initiatives. Sitting

back to let open data take its own course, or only letting fragmented efforts take hold, is no longer a viable approach. The question is not whether to have an open-data program but how to build one that works.

How open-data programs vary. Approaches to open data vary across and within countries and change over time. For example, privacy concerns in Germany have limited the government's ability to share individual-level data—but in the United Kingdom, domestic pressures for increased transparency have pushed more government data into public hands each day. The maturity of open-data programs also varies; some countries have much more advanced and comprehensive initiatives than others.

One of the most important differences in programs relates to the degree to which governments encourage citizen involvement. For example, Kenya has launched an open-data portal and encourages public participation, aiming to generate economic value estimated at \$1 billion, mainly through benchmarking appropriate project costs and identifying opportunities for public-sector savings. Several private companies have created innovative apps or tools based on the portal data, including Nairobi-based Upande, which integrates information about water-supply services in different regions that

The expansion of open data, combined with advances in big data analytics, is freeing information that was once trapped inside the dusty pages of overlooked reports, enabling improved decision making, new product and service offerings, and greater accountability.

Exhibit 4

Three overarching value levers of open data can help generate more than \$3 trillion annually for the global economy.

Decision making

Open data provides a fact base to make more informed and objective choices using information that is often available in real time

New offerings

Open data enables organizations to better understand their customers and context and to design new products and services

Accountability

Open data reveals issues in behavior, choices, and spending that citizens and leaders can act on to effect change

would otherwise be difficult for citizens or businesses to obtain on their own.

Such variations, combined with differences in awareness, cultural values, legal systems, and technical capabilities, mean that governments will need to customize their open-data programs. Nevertheless, there are common value levers and roles for government that apply across open-data initiatives, as described in the next two sections.

Value levers of open data

Government plays a critical role in enabling value creation from open data. This value primarily accrues in three ways: decision making, new offerings, and accountability (Exhibit 4). These levers produce benefits for the government itself and for other stakeholders, including private-sector organizations and consumers. In fact, consumers are the main beneficiaries of the potential \$3 trillion in value we identified.

There is some overlap among the three levers—as one example, some new offerings may enable better decision making. We also recognize that a single open-data initiative can simultaneously enable multiple levers, thereby creating value in different

areas. The government of Singapore, for instance, was concerned about the impact of rising residential energy consumption, and so it sponsored a “hackathon”—a community meeting where technologists and startups were invited to explore potential uses of open data and develop new apps. Invitees included energy and sustainability experts, researchers, and developers. The new apps that participants created could help with both decision making and increasing accountability.

Decision making: Reaching better conclusions more rapidly

Stakeholders across the spectrum, including governments themselves, can use open data to improve decision making. For example, Trafikverket, Sweden’s transportation agency, provides third parties with real-time data on train departure and expected arrival times, allowing passengers to make better travel plans. More generally, companies have developed online programs that use open data on transportation to help travelers and shippers to make better-informed choices about travel modes and routes. The wide release of data also gives businesses

greater insight into the competitive landscape, including current pricing conventions, which helps them make better decisions about contract terms when working with vendors and other parties.

By regularly analyzing open data related to performance, stakeholders can identify gaps and select appropriate improvement strategies. For example, educators can enhance teaching methods through adaptive-learning systems—online tutorials that immediately adjust the presentation of educational material according to students’ learning needs. At Arizona State University, educators used adaptive online-learning programs for students struggling with remedial math. The software tracked metrics, such as keystrokes, to see how students interacted with the material and where they needed extra instruction. These adaptive-learning systems helped improve pass rates from 66 percent to 75 percent.

During the decision-making process, stakeholders can use computer programs to analyze the vast quantities of open data. Such technologies will enable them to make more logical, fact-based decisions faster than manual analysis would allow. For example, city officials in Edmonton, Canada, use application-programming interfaces and real-time updates to analyze the city’s performance on a variety of metrics ranging from public transit on-time performance to utilization of public spaces to 311 call-response time. Researchers can simply and easily visualize more than 400 data sets, empowering them to make better-informed decisions on how to improve performance.

New offerings: Developing products and services that create value for customers

When the government and other stakeholders release data, they help companies, agencies, and individuals to develop innovative apps, products, and services or

improve existing offerings. Consider a few examples of companies that rely on open data for their success:

- The Climate Corporation employs open data to create various tools, including weather-monitoring and yield-forecasting products, which farmers use to make decisions such as where and when to plant crops.
- Propeller Health, a private company, has benefited from access to data from the US Centers for Disease Control and Prevention (CDC). It created a GPS-enabled tracker that monitors inhaler usage by asthmatics. The information is ported to a central database and used to identify trends for individuals, groups, and the overall population. By merging usage data with CDC information about environmental triggers of asthma (for instance, pollen counts in the Northeast and volcanic fog in Hawaii), Propeller Health helps physicians develop personalized treatment plans and spot prevention opportunities.
- The UK-based website findthebest.com used open data from the government to create the UK Car Fuel Economy and Emission App, which helps car buyers compare features such as fuel economy based on their type of commute.
- Emergency-room physicians created iTriage, a mobile app that helps patients understand their symptoms, locate nearby health facilities or physicians, and book appointments. The iTriage app has been downloaded more than ten million times.

Some of the most innovative open-data products are created through cross-sector collaborations involving public and private organizations. For example, after the 2010 earthquake in Haiti, volunteers

across the world combined data from sources such as aerial photography, World Health Organization maps of health facilities, and locations of police facilities from the Pacific Disaster Center. This OpenStreetMap project became a critical source of reliable information to guide workers from government and private-aid agencies in delivering supplies to hospitals, triage centers, and refugee camps. This map also helped responders better match the supply of and demand for various resources in the aftermath of the disaster, significantly improving the services delivered.

Accountability: Increasing transparency and enabling action

By releasing data, businesses and governments can ensure their actions, including their spending practices, are transparent.

This, in turn, can help prevent corruption and waste. Consider a few examples:

- The Brazilian government opened its books via the Brazilian Transparency Portal, which publishes a wide range of information that includes federal-agency expenditures, elected officials' charges on government-issued credit cards, and a list of companies banned from contracting work with the government. Data from the portal are helping journalists and activist groups expose corruption and may reduce unnecessary or suspect spending.
- Beginning in 2011, New York City released detailed information on energy and water consumption for each nonresidential building. Building operators are now using these data to benchmark their own energy



Workers unload an aid shipment from the United Nations in Port-au-Prince, Haiti.

efficiency, prioritize investments to capture savings, and potentially create incentives to promote energy-reducing programs or devices.

- The United States–based Sunlight Foundation, a nonpartisan private organization, works to improve government transparency by publicizing data related to many activities, including food-stamp use, political donations, and Congressional spending.

About one-third of the estimated potential value from open data comes from benchmarking, which exposes variability and opportunities to improve performance. For instance, many schools pay flat, uncompetitive rates for broadband services, without soliciting competitive bids. Estimates in the United Kingdom suggest that schools pay three to four times what private companies with similar capacity requirements pay for broadband service. But benchmarking the price of broadband service offered by local suppliers can significantly reduce costs, and we have seen the benefits in multiple locations worldwide. One school district

in northern California that benchmarked its broadband expenses was able to reduce costs by 30 percent.

When attempting to increase accountability, stakeholders, especially those in government, might consider involving the public. For instance, New York City’s 311 initiative allows residents to report nonemergency complaints—about common problems such as potholes or garbage collection—via a website, a mobile app, text messaging, Skype, or phone. The city now processes 60 percent of service requests online. In addition to lowering transaction and issue-resolution costs, the initiative allows users to track the status of their requests. New York City released the underlying 311 app to other cities around the world, enabling citizen participation in additional locations, including Boston and Buffalo, New York.

Government’s critical and unique roles

Sitting at the nexus of citizens/consumers, businesses, and NGOs, government is optimally positioned to extract value from open data—and to help others do the same. Much of this value is consumer surplus,

Exhibit 5

Government can serve as an open-data provider, catalyst, user, and policy maker to create value and mitigate risks.

<p>Provider</p> <ul style="list-style-type: none"> • Capture information electronically • Release data publicly and regularly • Identify ways to improve data quality 	<p>Catalyst</p> <ul style="list-style-type: none"> • Build an open-data culture • Convene stakeholders • Champion the movement 	<p>User</p> <ul style="list-style-type: none"> • Apply sophisticated analytics to improve decision making, offerings, and accountability • Invest in people, tools, and systems 	<p>Policy maker</p> <ul style="list-style-type: none"> • Make rules for internal and external use • Establish standards for data quality and format
---	--	--	--



or economic benefits that are captured by individuals. Government can spur value creation at all levels of society by concurrently fulfilling four important roles: as a provider, catalyst, user, and policy maker. Because technology, capabilities, and cultural norms are always changing, these roles continually evolve (Exhibit 5).

Government as an open-data provider

Across all levels of government, millions of individual data records are collected, stored, and analyzed. From tax returns and unemployment claims to hospital reimbursements and energy use, much of this information is now available electronically and readily shared. In the United States, for instance,

over 85,000 data sets are now available on data.gov, the federal government’s publication arm for open data, along with user guides, event postings, and examples of applications across industries, from agriculture to manufacturing. By making these data available to enterprising companies and individuals, government is spurring private-sector innovation and increasing transparency—two of the most important goals of any open-data initiative.

When prioritizing data for release, government might consider focusing on both the value that can be created as well as the ease with which data can be made available. Often, the data that are easiest to release may contain little that interests stakeholders. To

Attendees work on computers as they participate in a hackathon in Japan.

Many staff members and leaders at government agencies need to significantly shift their mind-set about data from one that has open data as an exception to one where open data is the rule.

avoid wasted effort, government leaders can work with their teams and external stakeholders to evaluate and decide which information to capture, how to standardize and digitize it, and when and where to share it. They can also look for opportunities to migrate data to electronic formats that are suitable for all internal agencies and external audiences. The 2010 US Census effort, which focused on both data quality and accessibility, included dozens of experiments and evaluations of data collection, coding, and analytics to improve the accuracy of the information and to make it available in an easily downloadable format.

Timely data are also essential to accurate research and better decision making. As such, government might consider explicitly setting expectations about how often data will be updated. In addition, government might consider releasing metadata—in other words, data about data. Metadata may describe the accuracy of a data set, the methodology of collection, or other parameters. Government can also assist researchers by publishing data “dictionaries,” which are easily searchable repositories of metadata.

Many staff members and leaders at government agencies need to significantly shift their mind-set about data from one that has open data as an exception to one where open data is the rule.³ Specifically, they can no longer see their primary role as the provision of services; instead, they might consider focusing on helping third parties create innovative products and services.

Government as an open-data catalyst

Government can serve as a catalyst for the use of open data by creating a thriving ecosystem of data users, coders, and application developers. To attract such talent, it can advertise its open-data efforts through press releases or other marketing materials, or even engage in individual outreach efforts. One innovative approach taken by governments worldwide involves organizing “hackathons” and “datapaloozas,” often with prize money and publicity. These events help promote the use of data in the development of innovative products and services. They also allow governments to share newly released or digitized data, draw on the expertise of participants, and provide guidance about privacy and safety. Other marketing platforms to champion open data include interagency meetings, roundtables with NGOs and businesses, and public-awareness campaigns.

There are a range of relevant topics for users to cover at open-data events, including what data could be most valuable if released, the best format for sharing data, user access and permission levels, and potential applications that could be developed from the data. For instance, the US Department of Health and Human Services recently convened 40 leaders from the technology and health-care spheres, presenting them with existing data sets and asking how the data could be used. This resulted in a productive brainstorming session and the development of a number of health-care apps and services.

³The June 2013 G8 summit promulgated an Open Data Charter, which establishes “an expectation” that the all government data be published openly by default.

Governments can also serve as a catalyst by selecting leaders to champion open-data initiatives. In the United States, for example, President Obama has appointed a “White House data evangelist” who oversees and promotes use of information on data.gov. The data evangelist’s team aims to increase public access to open data generated by the executive branch and encourage innovation outside government. The White House also supports innovation fellows—individuals recognized as rising technology leaders—who serve 6- to 12-month rotations in partnership with federal-agency leaders. These fellows help develop pilot programs, prioritize open-data opportunities, and codify best practices.

Government as an open-data user

There are two primary ways that government engages as a user of open data.

Applying analytics. As a data user, government can benefit by applying advanced analytics that improve internal decision making, promote the creation of new offerings, and increase accountability. For example, utilities can identify areas for improvement by sharing data about pilot results, task-time benchmarks, construction plans, performance-management systems, and asset-replacement schedules. In parallel, they can provide MyData to customers, such as information on their electricity-usage patterns to promote more efficient consumption.

Investing in programs. To optimize use of public data within their own agencies, government leaders need to invest in first-class talent, tools, and systems. On the people front, this involves hiring knowledgeable staff at all levels. (Although many cities, counties, states, and agencies today have chief information or chief technology officers, they often lack data expertise

lower in the organization.) Government leaders can also create appropriate incentives to retain employees with strong technology skills, since they may otherwise lose them to the private sector.

Beyond recruiting talent, government could optimize data use by training current staff—even those not in technical roles—about the overall value of open data, as well as the benefits of providing data to third parties. Training could emphasize how staff can derive insights from open data to improve their own programs or services, as well as strategies for engaging external stakeholders (for instance, by holding the previously mentioned datapalooza events). Agencies that build strong support from the top during these efforts will increase the likelihood that open-data initiatives are viewed as an essential part of their mission rather than side projects.

In an environment where technology is constantly changing, it is important for open-data programs to continue to evolve or value can be lost. Some organizations, such as the Open Data Institute, have established strategies for refining open-data programs that involve investing in computer-storage capabilities, establishing strong technical safeguards to protect privacy, educating staff about confidentiality risks, and building or buying sophisticated software programs. Governments and other stakeholders may improve data usage by studying their strategies.

Government as policy maker

At a fundamental level, government is well positioned to perform several interrelated yet distinct tasks to improve the lives of the public, generate value, and protect individual privacy. These include making rules and helping to establish standards.

Making rules. As a consumer advocate—the watchdog and protector of the public

trust—government can draw on its legislative authority and enforcement powers to ensure safety, security, equity, and justice for all members of society. In the context of open data, its purview includes several elements:

- defining rights to access or use certain sensitive information, including that obtained from medical files, tax returns, and driving records
- deciding what data should be mandatory for individuals and companies to share
- providing guidelines on the collection, management, and dissemination of information
- creating incentives for participation and compliance with data-sharing efforts

In some cases, government may need to limit the use of open data to protect privacy and confidentiality. The United Kingdom, for instance, enacted the Data Protection Act to control how public and private organizations use personal data. Among other provisions, the act stipulates that data should only be used for limited, specifically stated purposes and that they must be stored in a secure location. The act also provides stronger legal protections for particularly sensitive information, such as data related to an individual's health, religious beliefs, or political opinions. Similarly, the US federal government has enacted multiple laws to ensure data privacy, including the E-Government Act, which prohibits the secondary disclosure of information obtained for statistical purposes, among other protections.

There are also situations where the government can serve the public by mandating the release of data. In the United States, the Securities and Exchange Commission (SEC) seeks to protect the public by requiring companies to collect and submit certain financial documents. The SEC then posts documents deemed public infor-

mation on its Electronic Data-Gathering, Analysis, and Retrieval website. The EU has also declared that data about individuals held by private companies must be made available to those individuals upon request (a task that would be made easier through MyData systems, which are central databases that allow individuals to access information about themselves).

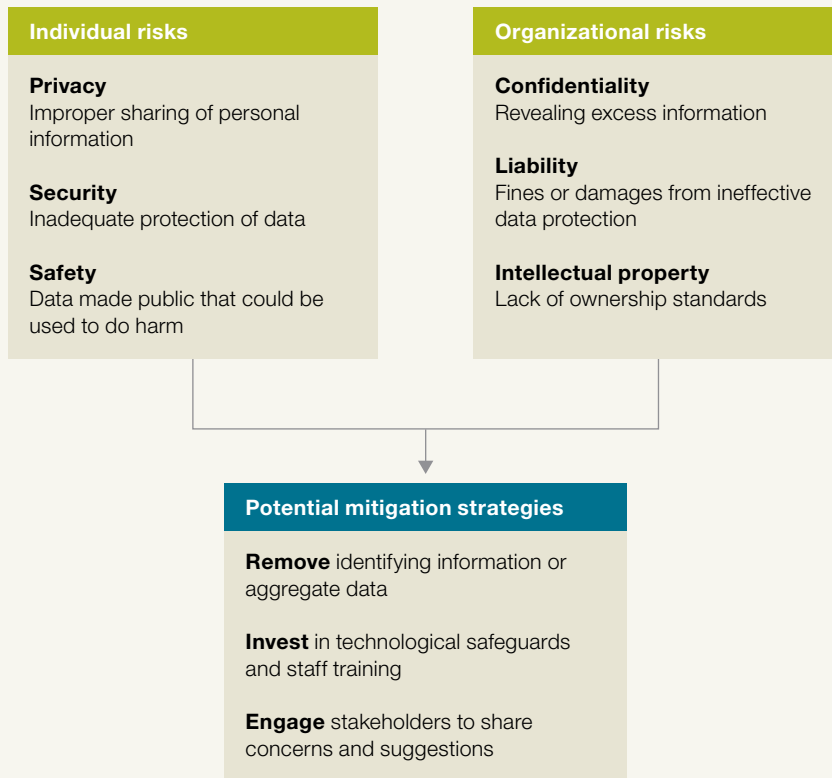
Elected officials may be concerned about public perceptions of privacy issues related to open-data initiatives. In addition to ensuring robust protections are in place around private/identifiable data, government can engage citizens early to address their concerns, seek new ideas, and answer questions about how data can inform policy.

Setting standards. The policy-maker role involves participating in the creation of technical standards about data quality and format, since these can significantly increase and scale the benefits of open data. Public-sector leaders might consider consulting experts to help design standards, which will facilitate use across different computer systems. For instance, they could follow the example of some private organizations by promoting the use of Data Catalog Interoperability Protocol (DCIP), which specifies design criteria for data catalogs published on the web. Among other recommendations, the DCIP promotes the use of JavaScript Object Notation, an easy-to-use data-interchange format.

As tools and technologies become more sophisticated, leaders might consider revisiting existing standards, since better methods for organizing and sharing data may emerge. Government could also reevaluate the level of detail contained within open data, the frequency of its release, and the ease with which data can be integrated across systems.

Exhibit 6

Public-sector leaders can address data risks that could threaten individuals and organizations or undermine open-data initiatives.



Risks in open data—and potential strategies to address them

Public-sector leaders are often called on to protect individuals and organizations from the risks of open data while also advancing open data's potential value. Risks include those that fall largely on individuals, such as privacy, security, and personal safety, and those related primarily to organi-

zations, such as confidentiality, liability, and intellectual property (Exhibit 6).

Individual risks: Privacy, security, and safety

As noted earlier, open data can improve accountability in government, make businesses more competitive, and help public-sector leaders debate issues and identify

sound policies. But total information sharing without safeguards or limits can put individual privacy and security in harm's way. Agencies that release information might therefore need to exercise caution. For instance, regulators in many countries publish data about hospital performance on certain procedures to compare costs and outcomes, but they do not publicly reveal patient-level information. Similarly, 45 US states evaluate student performance according to standards set by the Common Core State Standards Initiative. These standards allow educators to analyze longitudinal student and teacher data from multiple districts, but some citizens have expressed concern that records related to individual performance could damage college prospects if released. In response, officials have developed clear guidelines stipulating that the federal government can only have access to aggregate student and teacher data.

As open data become increasingly important, ensuring cybersecurity will be an ever-more-challenging task. For example, criminals could commit fraud or identity theft if they obtain ATM personal-identification numbers or financial data about a bank's clients. As noted earlier, government could try to address the security challenge in its role as policy maker by establishing limits on data access or usage (such as stipulating that some personal information can only go to law-enforcement officers) or forbidding the release of certain information.

To reduce risk, public-sector leaders could try to anticipate potential uses for newly released information, focusing on the insight it offers and parties that might be interested. For example, government could protect consumers by ensuring that data on individual energy usage is provided to residents as MyData but aggregate the data points when making them publicly available. This would limit the ability of criminals to track when

people are at home (higher energy use) or away (low energy use). Since it is not possible to predict all possible uses of data, government may want to institute policies that regulate their use rather than their collection. This approach may be more pragmatic than expecting to prevent all potential negative outcomes through fine-grained regulation for data collection and access; it also reflects the fact that specific uses of data are often what people find objectionable.

Organizational risks: Confidentiality, liability, and intellectual property

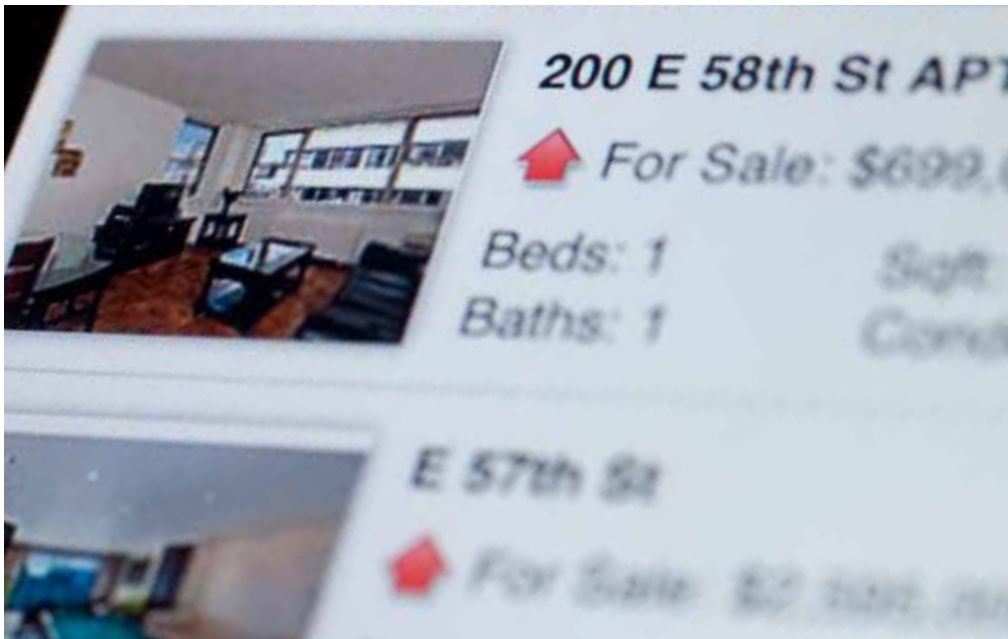
From small start-ups to Fortune 500 companies, organizations stand to generate significant value from open-data initiatives. Government agencies are also poised to reap great benefits. However, public officials can try to actively manage certain risks that occur whenever any organization releases potentially sensitive data. For example, there are concerns that opening performance data to third-party analyses could inadvertently discourage positive change, appropriate risk taking, and innovation. But government could try to balance the public's desire for information with the stakeholders' need to avoid undue criticism—for instance, by instituting regulations that limit liability.

Another organizational concern relates to questions about intellectual property: who owns rights to the underlying data, who owns rights to the tools used to extract insights from the data, and who owns rights to any products or services created from the data. Currently, permission rights to use open data vary greatly depending on the source. Similarly, some advanced technology, software, and analytics used to derive insights from data are free, as are some apps and other products created from the data; in other cases, organizations levy copyright and licensing fees. Whenever protections or restric-

tions on usage exist, such as contractual agreements that establish whether data can be redistributed, published, or combined with other data, it is important to respect them. If government establishes intellectual-property rules that are transparent, fair, and promote innovation, it could help create a growing ecosystem of “data-preneurs” who turn a profit by using open data to create goods and services.

Some organizations may be reluctant to release proprietary data for fear that common access to their hard-earned knowledge will erode a competitive advantage. Although this is indeed a concern, releasing data may also spur development of solutions that a single company could never devise on its own. For example, the “hive mind” approach, which draws on the talent and knowledge of multiple organizations, is relevant in the information-dependent field of oil and gas exploration. Greater sharing of seismic data

and information about drilling methodologies can increase the success of exploration activities and improve reserve-replacement ratios, reduce costs per barrel, and reduce risk. In its role as an open-data champion, government could encourage private companies to share such proprietary data, including industry benchmarks, in all sectors of the economy. Several protections can help reassure companies that their proprietary data will not be misused. For example, the Yale University Open Data Access (YODA) project provides pharmaceutical and health-care companies a safe route for releasing clinical-trial information that would otherwise go unpublished. People interested in accessing the data must submit a proposal, complete a short course on responsible conduct, and agree to share their research findings. Applicants who seek data purely for commercial or legal purposes are denied access. With such protections in



Public real-estate data is shown on a mobile app.

place, YODA can help ensure that researchers will only use the data to contribute to overall scientific knowledge, thus balancing an individual company’s intellectual-property rights with societal benefits of open data.

External stakeholders in open-data programs

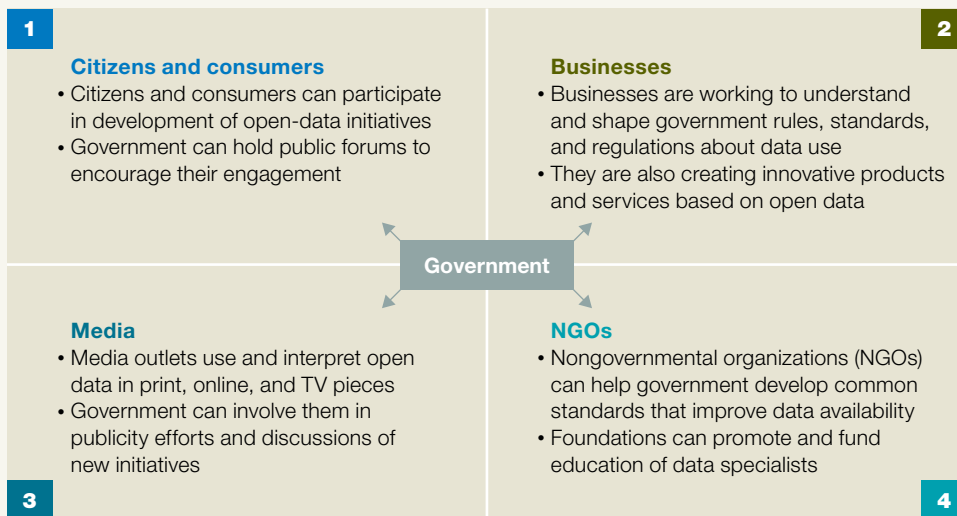
Any government open-data initiative will involve multiple external stakeholders who have different roles, needs, and concerns, as well as varying strategies for learning about data, mitigating risks, and advancing their agendas (Exhibit 7).

Citizens and consumers

Based on our prior analysis of seven domains of the global economy, citizens and consumers stand to gain the most from open data. For instance, they will be able to make more informed purchase decisions since they will have greater price transparency, thereby saving money. They will also have better insights about schools, transportation, and health treatments, which can help them make informed choices. As one example, consumers will have more information about diet and exercise patterns that could reduce their susceptibility to lifestyle-linked conditions and illnesses such as hypertension and diabetes.

Exhibit 7

Among core stakeholders, government is positioned to represent and advance shared interests.



Government could consider engaging consumers by holding public forums that illustrate how open data can improve services—for instance, by helping the public-works department to understand what streets are most in need of repair. In all cases, government could emphasize that private data will be protected, since citizens are often concerned about confidentiality issues.

Consumers themselves can make an important contribution to open-data initiatives by sharing personal information. For instance, consumers who release health information, such as data on daily exercise levels, can help clinicians identify individuals who may be at risk for a particular disease. If individual data are combined with information from other sources, such as demographic records, clinicians can then target specific populations for health-education or prevention efforts, such as disease-screening programs.

Businesses

Across industries, single “data-preneurs,” small start-ups, government contractors, and large multinational corporations are attempting to understand—and in some cases shape—appropriate government rules, standards, and regulations about data use. But even more important, many businesses are now creating innovative products and services based on open data. There are numerous examples we could cite in addition to those mentioned earlier in this report. For instance, Zillow, a US real-estate company, combines data from county land records, satellite imagery, and multiple listing services that advertise homes for sale. Visitors to the Zillow website can thus view images of entire neighborhoods when assessing potential properties. It may be helpful to publicize examples of some of the more successful creations, since other companies may be more inclined to use open data if they see the economic potential.

Business leaders might consider sharing data with other stakeholders or collaborating with them to develop innovative offerings that leverage open data. For example, they could potentially access sources such as Google Transit Feed Specifications, which would allow them to develop services like NextBus, an app that tells riders their projected wait time.

Media

In all formats—including print, radio, TV, online video, tweets, and blogs—media outlets use and interpret open data. Some also closely follow and report on developments in this sphere. OKF’s recently released 2013 rankings of the most and least “open” countries received extensive global media coverage. In addition, media outlets have intensely monitored national-security breaches in which data are compromised, as the WikiLeaks and Edward Snowden incidents made evident. This coverage spurs public dialogue about the standards needed to protect individual or national-security interests.

Many journalists use open-data sets to identify trends, patterns, and behaviors that deserve attention, a trend known as data journalism. For example, they could use open data to highlight the time of day when most traffic accidents occur or to illustrate crime patterns within a city. Articles on such topics may serve as catalysts that spur officials or private citizens to take action. Acknowledging the importance of open data, newspapers and other media sources are dedicating increased resources to the creation of infographics based on open-data sets. Some of those most interesting ones are interactive and allow readers to find specific data. In London, the *Guardian* recently published an infographic that shows temperature changes over the past 100 years in most countries. Readers can type in

a city and country to view location-specific data.

Media can also serve a vital function by publicizing success stories about open data and reporting on trends related to data use. Government leaders could thus engage journalists, commentators, and technology and government reporters in early discussions of new programs or policy changes, as this may help raise public awareness about and engagement with the initiatives. Publicity about open data could also stimulate crowdsourcing, since more people will become aware of available information and suggest new ways to use it.

The media can even help with job-recruitment efforts. For instance, news articles may raise awareness of government open-data initiatives and related job opportunities for coders or app developers.

NGOs

With access to information from multiple governments and constituencies across the globe, NGOs (including universities, advocacy groups, and think tanks) represent some of the most active collectors, disseminators, and users of open data today. The International Monetary Fund, the World Bank Group, and the World Health Organization all have vast stores of information that they share with the public, allowing others to analyze data and create tools and products similar to those based on data from businesses and governments. Given their presence across the globe and in multiple jurisdictions, NGOs can play a crucial role by encouraging policy makers to adopt common standards that improve the availability and use of open data.

Foundations that invest in education can also support open-data initiatives by promoting training efforts for data scientists. Such education is more crucial than ever,

since most countries lack skilled workers who have training in the science, technology, engineering, and mathematics fields. Other NGOs, such as “Code for America,” sponsor programs that place people with data skills in positions where they can help create applications enabled by open data.

Developing a customized open-data approach

Although interest in open data is gaining momentum worldwide, countries and agencies are not at the same starting point. Some, such as Denmark, the Netherlands (which topped the 2013 OKF rankings), Norway, the United Kingdom, and the United States, already have well-established initiatives that touch almost every sector of government, and they are well positioned to remain on the forefront of value creation. Other countries have various open-data programs across agencies and are still attempting to integrate them. Still others are just beginning to investigate the use of open data or have a few nascent programs. Regardless of their current capabilities, however, all governments will need to develop a clear understanding of their strategy and goals.

To begin, we suggest examining the three critical components of any successful initiative—people, tools, and systems. The diagnostic can include five elements:

- inventorying available data to assess their value
- analyzing initiatives that are in progress
- identifying gaps in knowledge, capabilities, technical infrastructure, management, engagement, and other areas
- prioritizing actions and investments
- evaluating how third parties are using an organization’s data and whether they are fully exploiting the information’s potential

For each topic, a list of questions can help government devise an improvement plan and meet the needs of all stakeholders. In the people category, for instance, government leaders can assess the external talent pipeline, describe the skills needed for internal employees, evaluate the degree of leadership support, and rate the strength of their open-data culture.

As noted earlier, government can also advance its open-data agenda by convening experts at working sessions at the start of—and throughout—open-data initiatives. Given the ever-changing nature of technology and the increasing availability of data in electronic format, they might consider treating open-data efforts as an iterative process. Ongoing engagement will also help government understand what consumers, NGOs, and the private sector truly value, allowing strategies to be refined. In the United States, for example, government officials within the Health Data Initiative convene an annual conference for companies that are investigating innovative strategies for using health data in tools and applications. More than 2,000 data experts, technology developers, entrepreneurs, policy makers, health-care-system leaders, and community advocates attended one recent forum. In addition to speeches, breakout sessions, and presentations, the forum allowed companies to demonstrate their products and work on them in “code-a-thons” that brought innovators together for live collaboration.

While some countries may not be able to convene such large conferences, they can still benefit from gathering smaller groups

of experts. In Rwanda, the government is working with the World Bank Group and the local start-up community to define the open-data agenda. In Uruguay, the government invited representatives from NGOs and regional technology leaders to help outline a road map for its open-data program and select high-value data sets to release.



The use of open data is a relatively recent phenomenon, but as with many technological advances, it is growing in relevance and prevalence—in other words, it is becoming the “new normal.” The benefits of open data are significant and include the potential for more than \$3 trillion in economic value annually, much of which will likely go to consumers. Yet the success of open-data programs is not guaranteed. It takes real effort to engage an external community to use open data. Risks, such as threats to privacy and intellectual property, can be actively and continually addressed.

Encouragingly, governments around the world have taken steps toward developing responsible and robust open-data programs. We hope this report will inspire and enable more government leaders and stakeholders across the spectrum of business, citizens and consumers, the media, and NGOs to engage in the conversation, invest in the work, and promote open data in a way that helps unlock economic and societal benefits. ■