McKinsey&Company

McKinsey Global Institute
McKinsey Center for Government
McKinsey Business Technology Office

October 2013

# Open data: Unlocking innovation and performance with liquid information

### The McKinsey Global Institute

The McKinsey Global Institute (MGI), the business and economics research arm of McKinsey & Company, was established in 1990 to develop a deeper understanding of the evolving global economy. Our goal is to provide leaders in the commercial, public, and social sectors with facts and insights on which to base management and policy decisions.

MGI research combines the disciplines of economics and management, employing the analytical tools of economics with the insights of business leaders. Our "micro-to-macro" methodology examines microeconomic industry trends to better understand the broad macroeconomic forces affecting business strategy and public policy. Current research focuses on four themes: productivity and growth; the evolution of global financial markets; the economic impact of technology and innovation; and urbanization.

For further information about MGI and to download reports, please visit www.mckinsey.com/mgi.

### The McKinsey Center for Government

The McKinsey Center for Government (MCG) is a dedicated center of expertise and innovation within McKinsey's Public Sector Practice to support government leaders around the world in their goal to achieve greater and lasting performance for the citizens they serve. MCG's mission is to help public-sector leaders understand the forces that are driving change, design the best strategies for success, develop the tools and capabilities needed for effective execution, and deploy these tools for impact.

The center focuses on the critical and common challenges that governments face and creates opportunities for government leaders to learn from successful experiences, innovations and different contexts. Drawing on a network of external experts and McKinsey practitioners from across the public, private, and non-profit sectors, MCG provides governments with both new and proven knowledge and tools to confront critical challenges in a context of limited resources and creates opportunities for governments to learn from one other. For more information, please visit www.mckinsey.com/client_service/public_sector/mckinsey_center_for_government.

### The McKinsey Business Technology Office

The McKinsey Business Technology Office (BTO) was created in 1997 to serve executives on their information technology needs in multiple sectors including automotive, energy, finance, health care, high tech, media, mining, pharmaceuticals, public sector, retail, telecommunications and transportation.

Operating in 59 cities in 27 countries, BTO's experts help clients in the following areas: optimizing enterprise architecture, infrastructure, and application management; delivering critical IT projects on time and within budget while capturing maximum business benefit and value; leveraging technology to digitize business processes; developing IT strategies that are tightly aligned to business goals, and designing corresponding organization and governance approaches that deliver the highest value; and understand the long-term technological forces at work that have an impact on their business. For more information, please visit www.mckinsey.com/client_service/business_technology.

McKinsey Global Institute
McKinsey Center for Government
McKinsey Business Technology Office

October 2013

# Open data: Unlocking innovation and performance with liquid information

James Manyika
Michael Chui
Peter Groves
Diana Farrell
Steve Van Kuiken
Elizabeth Almasi Doshi

# *More open data for more users ...*

**40+**
Number of countries with government open data platforms*

**90,000+**
Data sets on data.gov (US site)*

**1.4 million**
Page views for the UK open data site in the summer of 2013

**102**
Cities that participated in 2013 International Open Data Hackathon Day

**1 million+**
Data sets made open by governments worldwide

* As of 2013

# ... *can lead to more value*

## $3 trillion

Approximate potential annual value enabled by open data in seven "domains"

## 3 billion

Metric tons of carbon dioxide equivalent emission reductions from buildings that could be identified through the use of open data

## 35

Hours per year could be saved by commuters from schedule changes based on open data

## 100,000+

Medical, health, and fitness apps for smartphones

## 50%+

Consumer share of potential value of open data

# Executive summary

The global economy is beginning to operate truly in real time, with constant streams of data showing where consumers are shopping, ships are traveling, energy and money are flowing. Big data analytics make it possible to work through massive amounts of real-time and historical information to find unseen patterns and discover anomalies that can indicate opportunities for new products and services and new ways of operating more efficiently. Now, a complementary trend is under way. Open data—the release of information by governments and private institutions and the sharing of private data to enable insights across industries—provides additional depth to big data applications and makes possible entirely new ones, such as the smartphone app that tells commuters when the next bus will arrive. This trend has profound implications for companies, governments, and individuals.

Many open data initiatives, particularly in the public sector, have been motivated by societal goals such as improving the transparency and accountability of institutions, and much has been written about the importance of these efforts. Our research focuses on economic value that can be created by open data. Making data more "liquid" (open, widely available, and in shareable formats) has the potential to unlock large amounts of economic value, by improving the efficiency and effectiveness of existing processes; making possible new products, services, and markets; and creating value for individual consumers and citizens.[1] Realizing this potential will involve creating safeguards for personal privacy and business confidentiality, investments in technology, and changes in mindsets and work processes.

The value of making information open and available has long been recognized. Thomas Jefferson, the third president of the United States, understood that the development of the economy and the health of the democracy depended on the free flow of information. Jefferson funded the expedition of Lewis and Clark in 1803 to gather all the information they could on the vast wilderness that had been acquired from France in the Louisiana Purchase—providing data on topology, human settlements, rivers, minerals, soil, flora, fauna, and weather that enabled the settlement of the frontier and the rapid growth of the economy. Governments since have gathered and shared (to varying degrees) vital information: GPS data, weather data, and census information are examples of information sets that are collected by public agencies in the course of their work and then made freely available for use by citizens, businesses, and academics. Open data brings this principle into the 21st century and vastly extends the power of shared information by making available massive, fine-grained, and timely data, which through advanced computing and analytics yield novel insights. Entrepreneurial companies are seizing the opportunity: Climate Corporation, a startup that was recently acquired for about $1 billion, combines 30 years of weather data,

---

1   We use the terms "open data" and "liquid data" interchangeably.

60 years of crop yield data, and 14 terabytes of soil data—all from government agencies—for such uses as research and pricing crop insurance.[2]

We are studying the economic impact of open data now because this is a critical moment, when demand for data-driven insights intersects with more data being made open, and there are rapid advances in analytic capabilities, which accompany the adoption of big data. Although the open data phenomenon is in its early days, we see clear potential to unlock significant value by applying advanced analytics to a combination of open and proprietary data. As data are made more liquid, individuals and organizations can take advantage of the data to create value. The breadth and diversity of information in open sources could make open data a highly cost-effective source of critical insights in many markets.

Put to best use, liquid data can also become a critical element for breaking down information gaps across industries, sharing insights that can raise productivity, enable innovation, and replace traditional and intuitive approaches with data-driven processes. Analytics powered by open data can also help uncover consumer preferences, anomalies in costs, and variations in performance—all of which can inform new products and processes. However, these benefits must be weighed against privacy concerns, as well as loss of control over information about oneself or one's company. And, even if the data are free, there are costs associated with the effort to measure, analyze, and incorporate insights from the data into daily decisions for both consumers and businesses.

This research focuses on quantifying the potential value of using open data in seven "domains" of the global economy: education, transportation, consumer products, electricity, oil and gas, health care, and consumer finance. We identify the "levers" through which open data can create economic value and also explore the barriers to adoption and "enablers" for capturing value by making data more open.

Based on this analysis, we estimate that open data have the potential to enable more than $3 trillion in additional value annually across these domains.[3] This represents our estimate of true economic growth, but it does not account for share shifts that are likely to occur within industries and between consumers and business. Value can arise in a number of ways, including equipping workers with the skills to raise productivity, allowing marketers to micro-segment populations more successfully, and boosting performance across segments by sharing benchmarks, market data, and best practice information. Consumers stand to gain by saving money through greater price transparency and using more information to make decisions.

While use of open data varies by domain, there are common themes in how value is created. Techniques such as segmenting populations and exposing variability, which are commonly used in big data analytics, apply to open data analytics as well. In many instances, consumers will reap the greatest benefits from open

---

2   David Kesmodal, "Monsanto to buy Climate Corp. for $930 million," *Wall Street Journal*, October 2, 2013.

3   Throughout this report we express value in terms of annual economic surplus in 2013 US dollars, not the discounted value of future cash flows; this valuation represents estimates based on initiatives where open data are necessary but not sufficient for realizing value. Often, value is achieved by combining analysis of open and proprietary information to identify ways to improve business or government practices. Given the interdependence of these factors, we did not attempt to estimate open data's relative contribution; rather, our estimates represent the total value created.

data applications, usually by gaining access to information that allows them to make better decisions. Another common theme across the domains is risk. Open data creates new risks for companies and place new demands on government to protect privacy and prevent misuse of information.
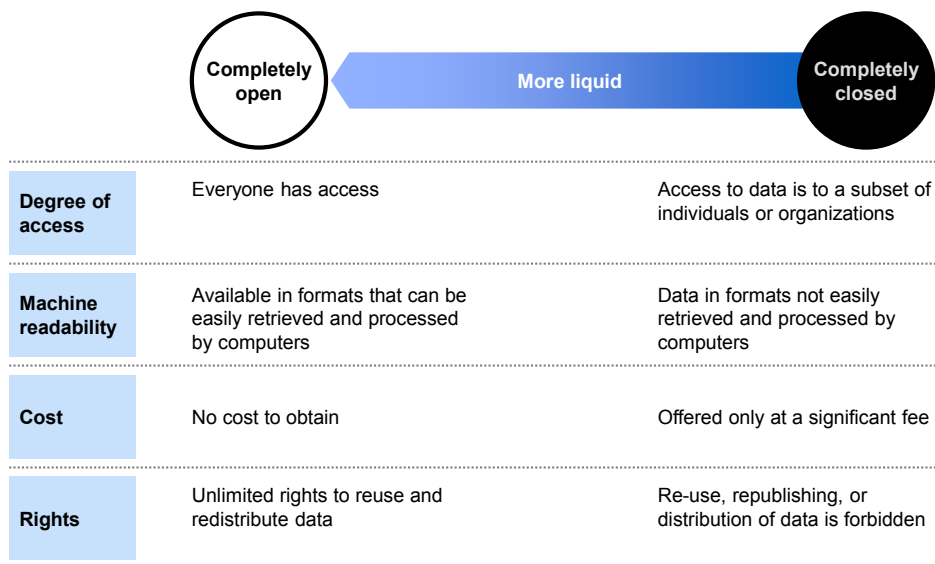
## DEFINITIONS AND EXAMPLES OF OPEN DATA SETS

While it has always been possible to share information openly, until information was in a digital format that could be easily shared and analyzed, the potential for open data was limited. Since the 1970s, however, governments have been able to share more data and more kinds of data electronically. More recently, private players—established companies and emerging data-sharing specialists—have also begun to make their data available to others. These open data sets share the following characteristics[4]:

- **Accessibility:** A wide range of users is permitted to access the data.

- **Machine readability:** The data can be processed automatically.

- **Cost:** Data can be accessed free or at negligible cost.

- **Rights:** Limitations on the use, transformation, and distribution of data are minimal.

Data sets range from completely open to completely closed across these four dimensions. In Exhibit E1 we see how data are open or closed based on the four characteristics that define open data.

**Exhibit E1**

**How data are open or closed, based on four characteristics**

| | Completely open | Completely closed |
|---|---|---|
| **Degree of access** | Everyone has access | Access to data is to a subset of individuals or organizations |
| **Machine readability** | Available in formats that can be easily retrieved and processed by computers | Data 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 | Re-use, republishing, or distribution of data is forbidden |

More liquid

SOURCE: McKinsey Global Institute analysis

---

4   This is similar to the definition put forth in "8 principles of Open Government," which emerged at an early meeting on open data in Sebastopol, California, in December 2007. Four of those principles (data must be complete, data must be primary, data must be timely, and data must be non-proprietary) describe properties of open data that make it more valuable. However, we do not include them in our definition because we believe data can still be classified as open or liquid, even if they do not meet these criteria.

Open data sets also are defined in relation to other types of data, especially big data (Exhibit E2). "Big data" refers to data sets that are voluminous, diverse, and timely. Open data is often big data, but "small" data sets can also be open. We view open and big data as distinct concepts. "Open" describes how liquid and transferable data are, and "big" describes size and complexity of data sets. The degree to which big data is liquid indicates whether or not the data are open.

**Exhibit E2**

**How open data relates to other types of data**                           ILLUSTRATIVE
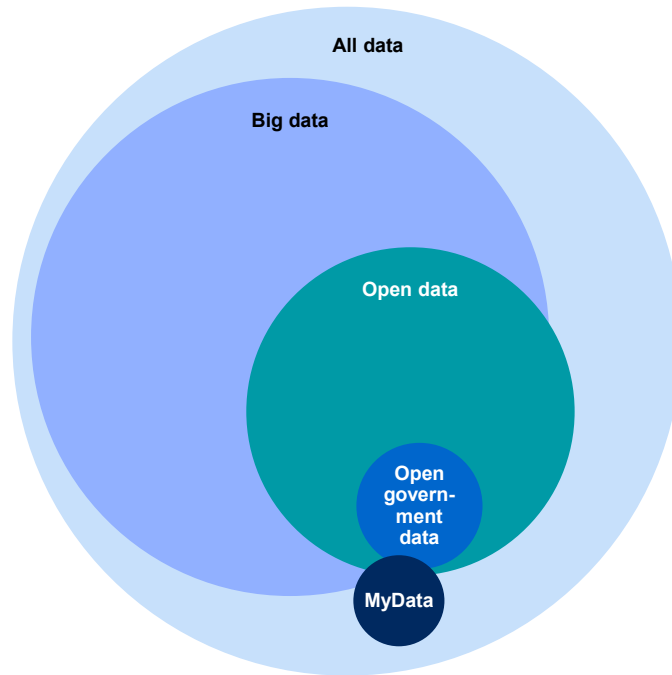


SOURCE: McKinsey Global Institute analysis

Open data sets—whether big or small—can come from the government or other institutions and enterprises, and from individuals. Open data initiatives in the public sector, in which governments release data, are some of the most prominent examples of this trend, but we do not consider open data to be synonymous only with data released by governments.

Finally, the open data concept is associated with "MyData," which involves sharing information collected about an individual (or organization) with that individual. For example, some hospitals now provide individual patients with access to their own medical records data. Providing aggregate statistics (a form of open data) alongside MyData allows useful comparisons; some utilities show consumers how their energy use compares with that of neighbors to encourage conservation. In some cases, individuals are allowed to modify or correct the data provided to them, which improves the quality of the data.

Open data sets vary in scope and source. They can be local, national, or global and can be obtained from both government and commercial sources (see Box 1, "An open data sampler").

## Box 1. An open data sampler

- **New York City building utilities benchmarking.** Detailed information on energy and water consumption for each non-residential building in New York City was released in 2011 and is used by building operators to benchmark the energy efficiency of their buildings and identify opportunities for improvements.

- **Real-time train movements.** Trafikverket, the transportation agency in Sweden, publishes real-time data on train departure and expected arrival times and track numbers for all trains traveling through the country. Third parties have used these data to create applications that allow travelers and shippers to make better-informed decisions on travel modes and routes.

- **Available parking spots.** Real-time open data about available parking locations has been made available in cities such as Singapore, Chicago, and San Francisco. Applications that use this open data help drivers locate parking spaces, reducing parking search time. These data can also be used in infrastructure planning.

- **Census.** Census data are a classic example of open data. In the United States, the federal government must compile and publish census data periodically, as stipulated in the Constitution. The US census provides detailed information on demographic and socioeconomic trends, down to the zip code level, helping government guide delivery of services (for example, locating schools) and enabling stores to customize formats and merchandise. Other countries release similar information.

- **Social media entries.** Social media are a growing source of wide-ranging information on customer preferences and experiences. Access to the full stream of social media content from a particular platform often requires some kind of commercial arrangement, so it is not fully open along the cost dimension, but it is relatively liquid.

### OPEN DATA INITIATIVES

More than 40 countries—from every region of the world and at every stage of development—have established open data initiatives. These nations are opening up all kinds of data sets to promote economic development, spark innovation, and find ways to make government work better. India has released 3,500 data sets, mostly of agricultural information. Singapore has shared 8,600 data sets from 60 public agencies. The World Wide Web Foundation's Open Data Index rates governments on 14 data-sharing metrics and in 2012 ranked the United States, Mexico, Singapore, the United Kingdom, and New Zealand as the top five most open governments.

In the United States the opening of government data accelerated after President Obama, on his first day in office in 2009, signed an executive order stating that all government information that did not have to be kept secret for security or privacy reasons should be made public. The administration also launched the Open Data Initiative to publish government data and the data.gov website to distribute the data, which has grown from 47 "open" data sets in March 2009 to more than 90,000 sets covering everything from energy use to consumer complaints.[5] The United Kingdom began its open data effort in September 2009 with 2,500 data sets, which grew to more than 10,000 data sets in 2013.[6] Kenya, which in 2011 became the first sub-Saharan African nation to launch an open data initiative, claims that opening up government procurement data and exposing price

5   Vivek Kundra, *Data.gov: Pretty advanced for a one-year-old*, Data.gov.us, October 2013.

6   "Tim Berners-Lee unveils government data project," BBC World News, January 21, 2010.

differences can save the government $1 billion annually.[7] The commitment to making data more open was re-affirmed at the June 2013 G8 summit through the Open Data Charter, which establishes "an expectation" that the default policy should be that all government data be published openly. The charter also notes that there are legitimate reasons why some data cannot be released.

State and local governments are also creating platforms to release their data. In the United States, California and Texas have identified millions of dollars a year in savings by releasing budgetary information and enabling citizens to spot potential opportunities to cut costs. To improve transparency and citizen engagement, cities such as Boston, Buenos Aires, Dublin, Helsinki, and New York have made significant amounts of information public, from restaurant health inspection scores to school performance ratings.

While governments are natural sponsors of open data initiatives—both as sources of data and as regulators of open data use—other institutions and enterprises are also beginning to release their data and invest in programs that rely on the use of open data. In Japan, for example, citizens mobilized to create an open database of radiation readings in different parts of the country following the Fukushima earthquake. Data generated by the automatic identification system (AIS) used by ships, which shows the location of ships at sea for collision avoidance and emergency rescue, have been opened, allowing shippers to see the locations and destinations of all larger commercial ships so they can judge which ports are likely to be congested. A growing industry of third-party brokers is scouring the world's data assets, making data more liquid by aggregating, integrating, and selling access to data (some proprietary, some open), sometimes along with analytics services.

## KEY FINDINGS

We focused our analysis of how use of open data can create economic value globally in seven domains.[8] The domains were selected to give a broad range of examples of how open data can create value in the public and private sectors. These are not necessarily the domains in which open data could have the most impact, but were selected to show a wide range of examples. The following are key findings from our research.

- **An estimated $3 trillion in annual economic potential could be unlocked across seven domains.** These benefits include increased efficiency, development of new products and services, and consumer surplus (cost savings, convenience, better-quality products). We consider societal benefits, but these are not quantified. For example, we estimate the economic impact of improved education (higher wages), but not the benefits that society derives from having well-educated citizens. We estimate that the potential value would be divided roughly between the United States ($1.1 trillion), Europe ($900 billion) and the rest of the world ($1.7 trillion).

---

7   Elana Berkowitz and Renée Paradise, "Innovation in government: Kenya and Georgia," *McKinsey Quarterly*, September 2011.

8   The analysis of health care is based on previous McKinsey research published in *The "big data" revolution in health care: Accelerating value and innovation*, McKinsey Center for US Health System Reform and McKinsey Business Technology Office, January 2013. Our work in the consumer finance domain highlights some key uses of open data but does not size the potential value comprehensively.

- **Open data enhances the value potential of big data analytics and provides additional opportunities.** Often the value from open data is realized by combining open and proprietary data. The additional open information enhances the opportunities from the five big data levers identified by MGI.[9] The levers are: creating transparency to unearth information to make better personal, business, and governing decisions; exposing variability and enabling experimentation to identify areas for improvement; segmenting populations to tailor actions such as creating custom marketing offers; augmenting or automating human decision making; and defining new products, services, and business models. About one-third of the estimated potential value from open data comes from benchmarking, an exercise that exposes variability and also promotes transparency within organizations. In addition to creating value with the big data levers, open data creates value in its own ways. When entrepreneurial companies and established businesses gain access to existing data (often government data) that were previously unavailable, they can significantly refine the process of defining and creating products and services. Open data's transparency also affects markets, fostering competitiveness by making more information available and creating opportunities to better match supply and demand. Liquid data also enables new opportunities for large-scale collaboration among individuals, companies, governments, and other organizations. Finally, open data can enhance the accountability of institutions such as governments and businesses and can raise the quality of decision by giving citizens and consumers more tools to scrutinize business and government.

- **Consumers stand to gain the most.** Consumers are already beginning to benefit from open data through price transparency (for example, by using online shopping sites that offer price comparisons). Other information about products and services could be made available through open data (e.g., whether trains are running on time or the labor and environmental practices of manufacturers) and could be used by consumers to select the products and services that best match their preferences. Opening MyData gives consumers better visibility into their own consumption, often revealing information that can lead to changes in behavior. Open data also gives individuals (as consumers and citizens) new channels to provide input to improve the quality of goods and services (including public services) and the quality of data. Together, more than 50 percent of the value potential we estimated is in consumer and customer surplus.

- **Open data helps businesses raise productivity, and create new products and services.** Using open data can help companies improve the productivity of current business processes and can lead to new products, services, and entire lines of business for both established companies and entrepreneurs. Open data can also help create more tailored products and services. By sharing data openly (yet anonymously), companies can benchmark performance and share best practices. There is also an open data industry, composed of companies that aggregate and sell data and advise companies on open data use.

---

9   See *Big data: The next frontier for innovation, competition and productivity*, McKinsey Global Institute, May 2011.

- **Open data creates new risks, including threats to reputation and loss of control over confidential information.** Open data can expose individuals and businesses to several risks, especially reputational ones. Opening information such as electricity use or school performance to create aggregated views of population behavior raises serious concerns among consumers who fear that their data will be tied to them and could harm their economic or social standing. Conceivably, credit card companies could raise interest rates on households that waste electricity or the inadvertent release of information about a particular student could lead to bullying. For businesses, open data released by third parties could expose poor environmental or labor practices or show that their products or services compare poorly for price and quality. Companies can also put consumers off by using open data to create online advertisements or marketing offers that show that the company knows too much about the consumer. Another risk arises in sharing benchmarking data among businesses, if the pooled data inadvertently reveal confidential information.

- **Governments have a central role to play as a source of open data and as a regulator.** Government could use policy and other mechanisms to determine the nature of open data within society. To promote open data approaches, government could assuage concerns of both consumers and businesses about the safety of open data and help educate the public about the potential benefits to the economy and society. Governments can use their influence to make data more open, through dialogue as well as regulation, while creating policies that thoughtfully address issues such as privacy, confidentiality, intellectual property protection, and liability. Government agencies also hold vast stores of data, which, if made open, could enable the creation of large amounts of economic value. Making government data available not only enables value creation, but also sets the tone for openness among other institutions.

- **Making data more liquid is necessary but not sufficient for capturing value of open data.** Once data are open, other developments and actions are needed to realize value potential. A vibrant ecosystem of developers will be necessary to transform open data into valuable tools. Thoughtful policies that protect intellectual property and ensure privacy and confidentiality will be needed to give consumers and institutions confidence to move forward with open data. Policies will need to be refined continually to balance the value of anonymous, aggregated information about an individual with the increased risk of identification. Investment in technology is needed to collect data, create suitable platforms for sharing data, and perform analyses to uncover valuable insights. Standards can be developed that will make data from multiple sources comparable. Releasing metadata (data about data) can make open data more usable. An open data marketplace may be needed to provide clear channels for sharing liquid data and to build a community with group norms and rules, which would discourage malicious manipulation of data through economic and social pressure. Third-party services will likely play an essential role in cataloging, cleaning, and parsing information that is not released in machine-readable format. Organizations will also need to acquire and develop the talent, processes, and cultures to complement their technological investments. Key skills include the ability to perform analyses, create useful reports and tools based on open data, and incorporate data into managerial decision-making processes.
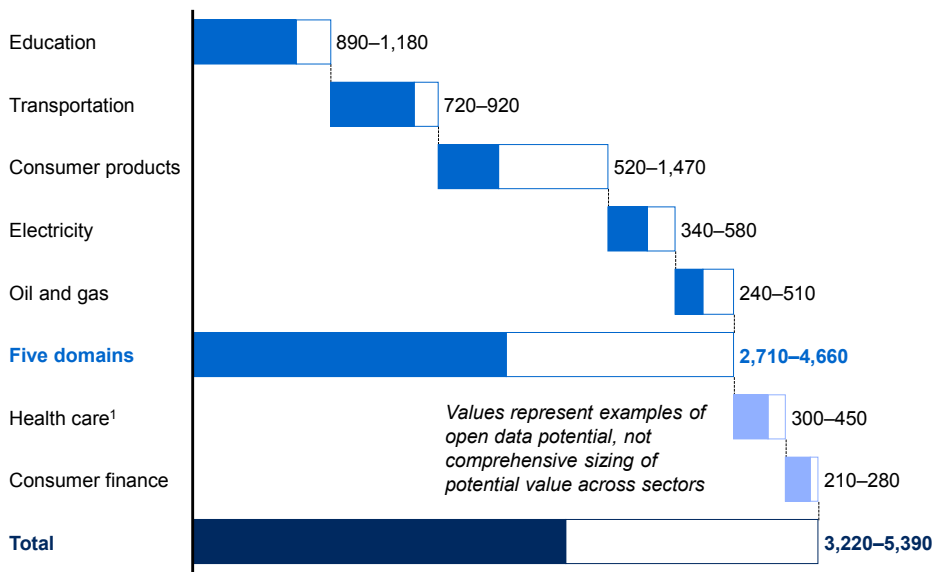
## OPEN DATA IN SEVEN DOMAINS

To understand the potential economic value that might be expected from widespread use of open data, we looked at seven areas of the global economy. For each, we identified the most important levers for unlocking value with open data and we estimated how much annual value to the economy each lever might help enable (Exhibit E3). These levers work by enabling better decision making, such as spurring energy efficiency by revealing relative consumption; by providing the insights for customized products and services, such as personalized learning programs; or by exposing anomalies in performance data that lead to better processes, such as using public budget data to find opportunities to save on procurement. In addition to the potential economic benefits to both the public and private sectors that we describe here and in the following chapters, we recognize that there are separate and potentially powerful societal benefits from making data about government more open (see Box 2, "The societal benefits of open data").

**Exhibit E3**

**Open data can help unlock $3.2 trillion to $5.4 trillion in economic value per year across seven "domains"**

$ billion

| Domain | Value |
|---|---|
| Education | 890–1,180 |
| Transportation | 720–920 |
| Consumer products | 520–1,470 |
| Electricity | 340–580 |
| Oil and gas | 240–510 |
| **Five domains** | **2,710–4,660** |
| Health care[1] | 300–450 |
| Consumer finance | 210–280 |
| **Total** | **3,220–5,390** |

*Values represent examples of open data potential, not comprehensive sizing of potential value across sectors*

1 Includes US values only.
NOTE: Numbers may not sum due to rounding.
SOURCE: McKinsey Global Institute analysis

## Box 2. The societal benefits of open data

The open data phenomenon began with the release of government data and it is still often associated with its potential impact on government. Many of the early successes of open data resulted from government release of information that led to greater accountability within government, such as exposing tax fraud in Canada and monitoring of corruption in Brazil after public records were made broadly accessible.[1]

Opening up public data and arming the public with the information to make concrete policy recommendations has also improved citizen engagement and participation. Peng Shi, a graduate student, created the new methodology for public school assignments in Boston, using information released by the city on the quality and location of schools.[2] His algorithm, which shows parents choices based on school quality and distance, was chosen over five other plans that had been developed through more traditional channels. Open data can also be used to monitor the use of public goods and encourage optimized investments of public goods.

Open data also enables collaborations across sectors in both public and private settings. Following the earthquake that struck Haiti in 2010, volunteers across the world combined data from sources such as satellite maps, World Health Organization maps of health facilities, and locations of police facilities from the Pacific Disaster Center, and data from mobile phone companies. The OpenStreetMap project became a critical source of reliable information to guide both governments and private aid workers and supplies to hospitals, triage centers, and refugee camps. This map helped responders better match the supply and demand for various resources in the aftermath of the disaster, significantly improving the services delivered.

---

1 Marcos Cruz and Alexandre Lazarow, "Innovation in government: Brazil," *McKinsey Quarterly*, September 2012.

2 James Vaznis, "MIT has plan for Boston school assignments," *Boston Globe*, October 28, 2012.

### Education

Using open data in education could enable approximately $890 billion to $1.2 trillion in value annually. The largest potential benefit comes from using open data to improve instruction by identifying the most effective strategies and tools for teaching specific skills and knowledge; students who acquire higher skills can expect higher lifetime earnings. Open performance data can also be used by students and parents to make more informed decisions about choices of schools and academic or vocational concentrations. Barriers include privacy concerns, especially for student and teacher performance data, and concern around the loss of autonomy of teachers in the classroom.

### Transportation

We estimate that $720 billion to $920 billion in value per year could be enabled by the use of open data in transportation. The greatest potential source of value is increased productivity and time saving for individuals from using open data to reduce travel times. Additional value can be gained by using open data to improve the efficiency of public transportation and freight operations, through adjusting train and bus schedules to better match demand and optimizing operations based on industry-wide benchmarks. Open data can also inform infrastructure investments. To capture much of the consumer value of open data in transportation, consumers will need access to easy-to-use applications that provide real-time location and estimated travel times of different transportation options.

## Consumer products

We estimate that $520 billion to nearly $1.5 trillion in value per year could be enabled through the use of open data in consumer products. Consumers could benefit from price transparency, as well as access to additional open data about products (the provenance of packaged food, for example) and suppliers (such as environmental and labor practices). Open data can help manufacturers and retailers generate sales and increase loyalty, by more finely segmenting consumer types and customizing store layouts and assortments for specific neighborhoods. Manufacturers can harness open data in the form of social media chatter or "likes" to inform product design decisions and improve after-sale interactions. While consumers stand to gain the most from open data, privacy concerns represent a very significant potential barrier.

## Electricity

Use of open data in the generation, distribution and consumption of electricity could help bring about value of as much as $340 billion to $580 billion annually. Significant investments in technology and operations will be required to capture the value from conservation and improved operations. By providing consumers with detailed data about their energy consumption and showing how other similarly situated consumers (or businesses) use electricity, customers can discover energy-efficiency opportunities. Utilities can also benefit from sharing benchmarking data to improve project management (e.g., streamlining permitting processes) and operations. Utilities and regulators will need to develop effective programs to use open data, build trust among consumers by establishing clear privacy protections, and encourage organizational cooperation through data-sharing schemes.

## Oil and gas

Across the oil and gas value chain, open data could help enable $240 billion to $510 billion a year in value, by improving investment decisions about where to explore for new reserves and build downstream facilities. Openly sharing benchmarks can improve investing processes and operations. Sharing consumption data can help consumers make better-informed decisions about energy use (reducing natural gas consumption, for instance). Many benefits depend on companies sharing data that has been regarded as proprietary. Therefore, the business cases and requirements for sharing information will need to be made clear.

## Health care

Previous McKinsey research examined the potential impact of more liquid data in health care and identified $300 billion to $450 billion in annual value that could be unlocked in the United States alone.[10] Potential sources of value include enabling people to take an active role in disease prevention and treatment; helping providers determine what is the most timely, appropriate treatment for each patient; matching patients with the most appropriate providers; ensuring the cost-effectiveness of care; and identifying new therapies and approaches to delivering care. Capturing this value will require changes in how care is delivered and paid for (in the United States), standards for data governance and usability, persuading

---

10 *The "big data" revolution in health care: Accelerating value and innovation*, McKinsey Center for US Health System Reform and McKinsey Business Technology Office, January 2013.

providers to share data, and investing in the capabilities of all players to make the most of open data.

**Consumer finance**

We examined the potential impact of open data in three areas of consumer finance: banking, insurance, and real estate. Though the potential impact in these three areas is not sized comprehensively, we estimate that $210 billion to $280 billion can be captured annually from a few of the largest examples we consider. In banking and insurance, there are significant opportunities to increase value through use of open data to improve product design and underwriting. In particular, open data can be used to assess risks for consumers who do not have a credit history, opening up a large potential source of new business (worldwide, half of adults have no banking relationships). Open data can also be used in fraud prevention and detection: affiliations revealed in social media connections have been used to uncover insurance fraud rings. Consumers can also use open data to choose among a confusing array of complex financial products. In real estate, open data can help match buyers and renters with properties, and inform development and infrastructure investments. Barriers to realizing the value of open data in consumer finance include substantial concerns about privacy by consumers and a lack of legal frameworks regarding the use of new forms of data in banking and insurance.

## IMPLICATIONS FOR STAKEHOLDERS

Governments, companies, and individuals will need to understand how to take advantage of open data. All stakeholders—governments, non-profits, businesses, individuals (as consumers and citizens)—have roles to play in maximizing the benefits of open data. Deriving valuable insights from open data will require new rules and procedures and new attitudes as well as investments in technology and capabilities.

■ **Governments.** Governments can set the tone for open data within a society, both by releasing data and shaping the policy environment. Public-sector agencies can be a key source of open data. An important first step is to set priorities for data release that are based on potential value, rather than ease of "opening" the data for sharing. Agencies can establish clear rules to govern the type of data that should—or should not—be released, with particular focus on safety, national security, privacy, liability, intellectual property rights, and confidentiality. Government leaders can champion the focus on open data across agencies and help make sure that the default decision is to release data whenever possible. As primary sources of open data, governments at all levels can be leaders in developing easy-to-use platforms for accessing open data. Leaders can also seek public-private partnerships or collaborations to support open data activities. For instance, the United States Patent and Trademark Office has worked with Google to post information such as awards and published applications in an easily searchable format.

As the maker of laws and enforcer of regulations, government can shape the legal and economic environment that maximizes the potential societal value from the use of open data, while addressing the legitimate privacy and intellectual property concerns of individuals and organizations. This may include policies on who can access information, the type of information that can be collected or used in certain functions, and protocols for notifying

individuals when information about them is made available. Elected officials and policy makers can convene stakeholders to discuss innovative ideas, set standards, and make data accessible and usable. Regulators can also create policies to encourage companies to collect and release data (with appropriate protections). Additionally, public-sector agencies can help address the need for workers with strong data and analytical skills through education and immigration policies.

- **Entrepreneurs.** Open data creates opportunities for entrepreneurs (or in-house "intrapreneurs" in corporations) to create new revenue streams and increase the productivity of their operations. The availability of open GPS data in the 1980s created whole new businesses in consumer and business GPS and mapping services that today contribute an estimated $90 billion a year in value to the US economy.[11] Other types of open data made available by governments, companies, and third parties (often combined with other data) can be used to create innovative products and services. Particular value may lie in bringing together information from disparate sectors of the economy and combining that information with proprietary data. Companies hoping to capitalize on open data opportunities can seek out and inventory sources of liquid data, develop strategies to influence others to make valuable data more available, invest in the development of tools that can transform open data into products and services, and dedicate staff to developing insights and designing products using open data.

- **Companies.** Open data can bring fresh insights into how companies operate and help management identify unnecessary variations and other barriers to productivity. Companies can choose to share proprietary data to create benchmarks that can improve overall industry performance. Using open data, such as customer discussions on social media, companies can refine product requirements and create new products and services. Companies will also need to have strategies about how, when, and under what circumstances they open their data, taking into account the potential impact the release of their data could have. Companies can choose platforms for data release, participate in the creation of an ecosystem of data users, and consider ways in which to monetize the value of liquid data. In addition, they can participate in standards setting, including for metadata, and provide input into the emerging legal frameworks governing data.

  Companies must also evaluate the potential risks posed by open data. Firms should become aware of the types of open data that might lead to reputational harm, hurt their competitive position, or disrupt their industry. Given these risks, companies should participate in the dialogues that set standards, develop legal frameworks and policies, and inform the broader public on such topics as intellectual property, privacy, and confidentiality.

  Data brokers provide data aggregation services and develop other services related to open data (for example, by integrating open data with proprietary data sources, then analyzing those data). Data brokers will have to continue monitoring the open data landscape for more data to collect and aggregate, and continue to develop innovative products and services, as well as participate in policy dialogues.

---

11  Philip Yam, "How to kick-start innovation with free data," *Scientific American*, March 23, 2013.

- **Individuals.** Consumers can capture a large portion of the value potential of using open data. Individuals can seek out applications that use open data and provide feedback to improve these tools. They can also take advantage of MyData to guide their behavior and ensure that data are accurate. Constituencies can also be built to press for release of more open data and creation of more tools from both government and other enterprises.

  While open data has numerous potential benefits, consumers can help to safeguard their interests by monitoring privacy policies and practices to ensure that the data are not used in such ways that lead to social or financial harm, or information that they simply prefer to keep confidential. Citizens can work with government to guide policy around the collection and use of data.

- **Non-profits and NGOs.** Open information on the location of resources—schools, hospitals, roads—and on the quality of health, educational, and economic systems can be used to identify areas of greatest need and calculate the additional resources required. Non-governmental organizations (NGOs) and other non-profit organizations can incorporate this type of information when developing a strategy, and in day-to-day operations. The first step to this will be understanding the data that are available and identifying key gaps. Lobbying governments and private organizations to release information where gaps exist can lead to critical breakthroughs.

  NGOs and non-profits can also organize volunteers with data-related skills who can collaborate to create useful tools. The OpenStreetMap database that was used in Haiti, for example, was produced by volunteers on several continents. NGOs can establish common platforms, set standards, and motivate volunteers. Non-profits can also help fill the talent gap: Code for America, for instance, recruits web developers, designers, and entrepreneurs for a year of service helping cities to use the Web to become more open and efficient. Non-profits also can serve as a neutral organizing force to align international data standards, such as standardizing formats for metadata (data about data) and other data elements.

  While open data can increase the impact of outreach efforts, NGOs and other non-profits must also be aware of the risks and be careful that the data they collect not inadvertently identify or reveal negative information about groups that they are trying to help. Steps should be taken to ensure that privacy is maintained, especially as numerous de-identified data sets are combined, making it more difficult to mask identify.

□   □   □

Our research shows that open data efforts could help to unlock more than $3 trillion in value every year in seven domains of the global economy. The benefits of open data can be self-reinforcing: as individuals perceive benefits from the use of open data, they will help to improve the accuracy and detail of information available, thus increasing the value of the data and the benefits that they can receive. However, this cycle can gather momentum only if private industry and public agencies cultivate a vibrant open data ecosystem and create data policies that provide adequate protection for all stakeholders. Companies will need to put in place the technologies and talent to collect and analyze the data. Individuals—as consumers and citizens—will need to be vigilant and savvy providers and users of open data.

# Related McKinsey Global Institute research

**McKinsey Insights iPad app** (September 2013)

This new app provides mobile access to the latest perspectives from McKinsey, MGI, and the *McKinsey Quarterly*. Our articles and reports address the most challenging issues facing senior leaders around the world, spanning countries, industries, and all business functions to examine leadership and corporate strategy to organization, technology, marketing, and operations. The app is available from the Apple App Store.

**Disruptive technologies: Advances that will transform life, business, and the global economy (May 2013)**

It is critical that business and policy leaders understand which technologies will matter to them and prepare accordingly. This reports cuts through the noise and identifies 12 technologies that could drive truly massive economic transformations and disruptions in the coming years. It also looks at exactly how these technologies could change our world, as well as their benefits and challenges, and offers guidelines to help leaders from businesses and other institutions respond.

**The social economy: Unlocking value and productivity through social technologies (July 2012)**

MGI analyzes the potential impact of social technologies. Based on in-depth analyses of how social technology can be used in five economic sectors (four commercial sectors, plus the social sector), the report identifies a series of value levers that can be applied across the value chain and within and across enterprises.

**Online and upcoming: The Internet's impact on aspiring countries (January 2012)**

This report explains how the Internet today connects about two billion people worldwide. Half of these are in the "aspiring" world—countries as varied as Algeria, South Africa, China, Iran, and Mexico—that are climbing the developmental ladder quickly, with diverse populations and inarguable economic potentialities. It examines the impact of the Internet in populous and fast-growing aspiring countries, where it offers even greater potential than in the developed world.

**Big data: The next frontier for innovation, competition, and productivity (May 2011)**

Big data will become a key basis of competition, underpinning new waves of productivity growth, innovation, and consumer surplus—as long as the right policies and enablers are in place.

www.mckinsey.com/mgi

E-book versions of selected MGI reports are available at MGI's website, Amazon's Kindle bookstore, and Apple's iBookstore.

Download and listen to MGI podcasts on iTunes or at www.mckinsey.com/mgi/publications/multimedia/