JANUARY 2014
McKinsey Quarterly

What executives should know about open data

Michael Chui, James Manyika, and Steve Van Kuiken

Novel and more accessible forms of information from government and private sources represent a new and rapidly growing piece of the big-data puzzle.

Not all data that's valuable is internal and proprietary. New initiatives by governments as diverse as those of the United States, Mexico, and Singapore are opening the spigots of readily usable public data. Corporate information too is becoming more "liquid," moving across the economy as companies begin sharing data with their business partners and, sometimes, consumers. Also surging is the richness of the information from data aggregators, which are assembling, rendering anonymous, and selling (to interested third parties) a wide range of data flows. Then add huge volumes of data from social-media interactions, available from providers of digital platforms such as Twitter and Facebook.[1]

These new sources of open data represent an expanding trove of largely unexploited value. One everyday illustration of open data at work is a smartphone app that uses real-time data (provided by transit authorities) to tell commuters when the next bus or train will arrive. Using open or pooled data from many sources—all the businesses in a particular sector, for example—often combined with proprietary big data, can help companies develop insights they could not have uncovered with internal data alone.

[1]Data sets range from completely open to completely closed, across four dimensions: *accessibility* (the range of users permitted to access the data), *machine readability* (the ease with which the data can be processed automatically), *cost* (the price to obtain data), and *rights* (limitations on the use, transformation, and distribution of the data).

Demographic data, financial transactions, health-care benchmarks, and real-time location data are among the myriad new information sources a company can exploit to create novel products and services and to make its operations more effective and efficient. New research from the McKinsey Global Institute, the McKinsey Center for Government, and McKinsey's Business Technology Office suggests that $3 trillion or more in annual value could arise from the use of open data in applications across seven domains of the global economy (exhibit). About a third of those potential benefits would involve the use of benchmarks to identify areas for improvement.

Whether or not individual executives at large companies choose to work with open data of various types, the magnitude of the value at stake suggests that some of them will—and that these applications will probably affect a wide range of industries, markets, and customers. Layering open-data mandates into the ongoing development of data and analytics strategies by considering both the use and sharing of more liquid data should therefore become an increasingly important priority for a wide range of companies. Here are a few examples of open data's potential:

- Energy exploration. As new technologies have made it possible to drill in a wider range of geological formations, reservoirs have become more complex. That's raising costs and risks—estimated ratios of prospects to explored targets can be as high as 50 to 1. The sharing of information on drilling permits and on seismic and other data across companies could reduce the number of dry holes and help optimize investments. While the widespread sharing of seismic data is unlikely, sharing among even a few companies could produce significant new value in the oil and gas industry. Governments keen on maximizing resource wealth could take the lead in structuring processes for granting permits so that grants of initial drilling licenses would require greater sharing of seismic data. Sharing data on projected costs and development timetables (through third parties) could establish benchmarks that, we estimate, would reduce per-project costs by 15 to 25 percent.

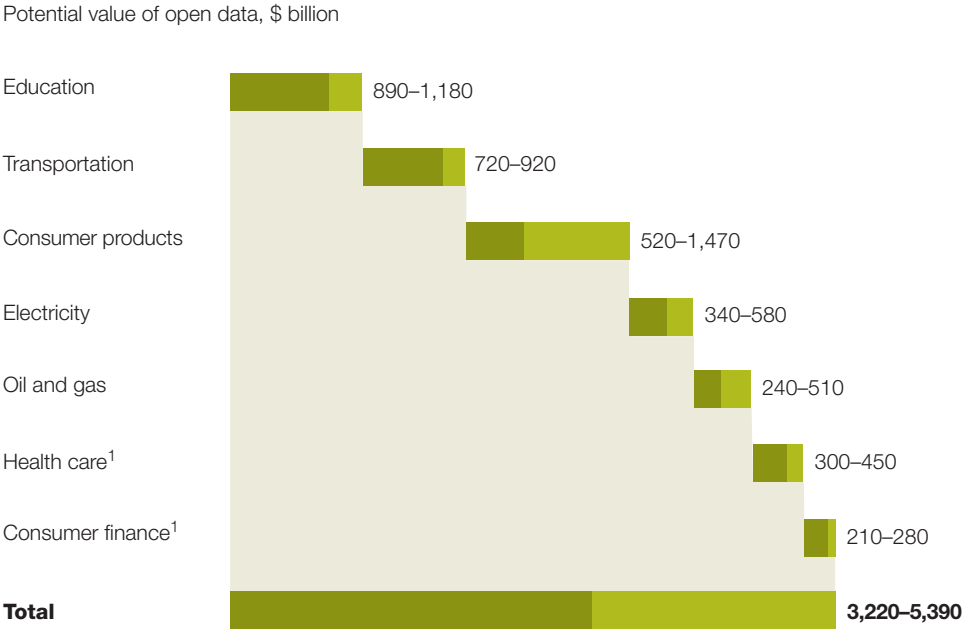
- Consumer insights. In the consumer-products sector, sharing data among retailers and manufacturers in limited circumstances—avoiding exchanges with direct competitors, for example—could lead to marketing approaches not possible with proprietary data

alone. Take Nectar, a UK-based program for loyalty cards, which can be used at Sainsbury's for groceries, BP stations for gasoline, and Hertz for car rentals. Sharing aggregated data allows the three companies to gain a broader, more complete perspective on consumer behavior, while safeguarding their competitive positions.

- Agriculture. San Francisco–based Climate Corporation combined more than 30 years of weather data, 60 years of data on crop yields, and multiple terabytes of information on soil types—all data from public sources. With that reservoir of historical information and real-time data flows, the company offers fee-based advice to farmers and customized crop- and weather-insurance products based on sophisticated algorithms. The company was recently acquired for about $1 billion by Monsanto.

Exhibit

Open data can help unlock $3 trillion to $5 trillion in economic value annually across seven sectors.



[1]Values for health care and consumer finance are drawn from examples of open-data potential and are not a comprehensive sizing of potential value across the 2 sectors; health-care figures are based on US values only.

Source: McKinsey Global Institute analysis

Other possibilities abound. Premium pricing for some goods could be facilitated if companies shared detailed information about products, such as the materials they use or the conditions under which they were manufactured (for example, with renewable energy). On the flip side, open-data applications may also create new areas of consumer value. In a budding trend known as MyData, organizations share information they have collected about individuals with them, in useful forms. Patients could access targeted medical data from a hospital, for instance, to help them manage their health.

Powerful as open data can be, many companies have valid concerns. Consider the sharing of data to establish industry benchmarks. Even if a company uses a third party and gets assurances of anonymity, there's always a risk that its identity might be revealed and that competitors could see how well or poorly it was doing. Shared data also could give away sources of competitive advantage or compromise intellectual property. Similarly, tapping social data could heighten privacy worries among consumers.

Still, it's hard to imagine that the open-data wave will slow down. Third-party open-data aggregators will certainly proceed to sell and publish corporate data, such as customer ratings, safety records, defect complaints and recalls, and comparative price data. Open-data initiatives also continue to proliferate in the public sector. In June 2013, G8 governments adopted an Open Data Charter, which establishes the expectation that the default policy should be the open publication of government data. Traditional competitors and entrepreneurial attackers can take advantage of open-data sources such as social-media comments and crowdsourced ideas to come up with new products and services. Open data, in short, seems to be another of the many relentless shifts in the digital landscape to appear unexpectedly, create new opportunities and strategic complexities, and leave established players with no place to hide.

The authors would like to acknowledge the contributions of Diana Farrell and Peter Groves to this article.

Michael Chui is a principal of the McKinsey Global Institute, where **James Manyika** is a director; both are based in McKinsey's San Francisco office. **Steve Van Kuiken** is a director in the New Jersey office.

For more on this research, see the full McKinsey report, *Open data: Unlocking innovation and performance with liquid information*, on mckinsey.com.