BCG

HOW FAR CAN YOUR DATA GO?

By Massimo Russo and Tian Feng

This is the fifth article in a multipart series.

A FEW YEARS AGO, THE idea that data was the new oil caught on. As consumers revealed more of their behaviors, preferences, and attitudes via their electronic devices, companies realized that data, too, is a plentiful, tradeable, and highly valuable commodity.

But data has distinct characteristics that make it very different from oil—or any other commodity. For one thing, the array of sources and amount of supply are virtually infinite. For another, data can be used (or consumed) more than once, and it can be used in multiple places by multiple parties at the same time. Moreover, while it's hard to hide a tanker full of oil, it's easy to mask a few billion bytes of data—and then put them to uses (beneficial or nefarious) without the awareness of the original data generators. Data is also unlike most commodities in that different types can have different value, and sometimes that value is not immediately clear. Finally, data can be used for lots of different purposes, many of which were neither contemplated nor intended when the data was first generated. We call these "alternative data uses."

The myriad alternative uses of data, the ease with which it can be replicated and shared indefinitely at no cost, and the trillions of bytes of data coming on-stream from the Internet of Things (IoT) pose big and far-reaching questions with respect to ownership, privacy, and value. B2B enterprise data sharing, in particular, is just starting to take these issues into account. As companies prospect for new sources of value, the rules, standards, and conventions governing data ownership rights and the regulatory frameworks for privacy and data sharing have yet to take shape.

In this article, we take a tour of the alternative data use landscape and offer some thoughts for business executives who want to realize value from what will be the predominant resource of the 21st century. How should companies think about use cases that are unknown or do not yet exist? How can they balance the abstract value of future use cases with the actual risk of data misuse?

Collective Value/Private Risk

It's impossible to overestimate the number of potential sources of IoT data and the endless possible combinations of data from

those sources. (By 2025, the volume of data from IoT devices will reach almost 80 billion terabytes, according to IDC estimates.) On a typical city block, sensors on numerous data entities generate data that can be recombined for use in countless areas, such as mobility, public safety, economic development, and health care. This data travels through complex webs of intermediaries leading from the source to the ultimate users and applications.
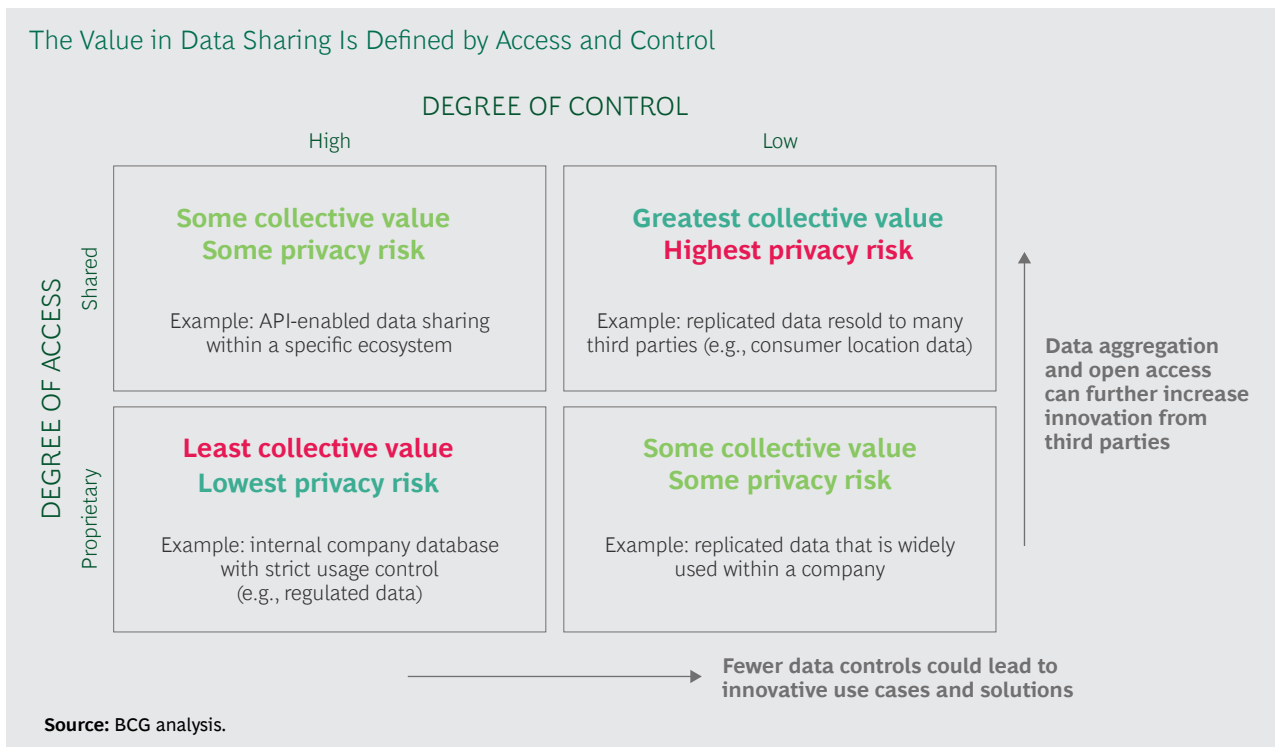
But the use cases that companies have pursued so far are mostly limited to a fruitful but restricted set of opportunities, typically involving players within a single industry. Increasingly, however, companies are starting to move from using their own data to improve internal processes or products to developing applications that require participation in an ecosystem or some form of cross-entity or multi-industry data sharing. Shared data can lead to deeper insights into customers or operations, and data provided to an ecosystem that combines data from multiple sources can enable entirely new applications. One benefit of ecosystems is that the collective contributions of the participants can accomplish more than any individual company could do on its own. They drive innovation and collective value.

The common tools employed by today's data-sharing ecosystems—such as application programming interfaces and data licenses—control, monitor, and limit data use, and a big part of their job is to mitigate risk. The deeper a company ventures into data sharing and the further afield it ventures from the business it understands best, the more control it must cede to others and the more uncertainty it must endure. While companies can often easily capture the value from data applications linked to their core businesses, it is more difficult to identify and capture the value of distant and novel applications. Risks related to both losing control of data and realizing value from data rise in significance. Collective value and private risk can come into conflict. (See the exhibit.)

Three Features of Alternative Data Uses

Many alternative data uses share three features that distinguish them from more predictable use cases. First, alternative uses are often developed after the data has

---

The Value in Data Sharing Is Defined by Access and Control

DEGREE OF CONTROL

	High	Low
Shared	**Some collective value** **Some privacy risk** Example: API-enabled data sharing within a specific ecosystem	**Greatest collective value** **Highest privacy risk** Example: replicated data resold to many third parties (e.g., consumer location data)
Proprietary	**Least collective value** **Lowest privacy risk** Example: internal company database with strict usage control (e.g., regulated data)	**Some collective value** **Some privacy risk** Example: replicated data that is widely used within a company

DEGREE OF ACCESS

Data aggregation and open access can further increase innovation from third parties

Fewer data controls could lead to innovative use cases and solutions

Source: BCG analysis.

been collected or combined and are therefore difficult to predict. For example, the Automatic Identification System, originally intended to track the location and identity of ships in order to reduce collisions, is now the source of data used in applications from economic analysis and insurance to oceanic research. Similarly, the strategic logic of IBM's acquisition of the Weather Company may not have been obvious on the surface, but weather data can be used in a wide range of applications in such industries as agriculture, aviation, construction, and logistics, as well as in more distant applications in health care, finance, insurance, retail, and utilities. The Weather Company's copious data generation capability augmented IBM's ability to better serve its customers with cloud applications in those and other industries.

The second distinguishing feature of alternative data uses is that they often occur in applications far from the original source of the data, which may pass through a series of intermediaries before it is put to work. Scrubbed, restructured, aggregated, analyzed, and distributed (to potential additional aggregators), the data comes into contact with a complicated web of players. This enables it to get from the point of origin to distant applications, but the process also makes it hard for the owners of the data to track its ultimate uses.

Copenhagen's experience setting up its City Data Exchange highlights the complexities. The city found that not only will a single data set pass through many intermediaries, but each intermediary transforms the data. These transformations take place for several reasons. Some data is fragmented and needs to be aggregated with other data to develop a solution. Some generators of data don't have the capabilities and analytical skills needed to transform data into insights. And different use cases require different forms of processing.

The third feature of alternative data use cases is that they are increasing as more applications shift from technology-driven "push" solutions (with data producers or aggregators building applications for themselves or for others) to use case-driven "pull" solutions (with companies sourcing data for specific use cases that they have identified). We are already seeing this trend in smart cities, where governments are blending data to fashion more citizen-centric services. One example is the variety of new IoT solutions that are helping people stay independent and connected as they get older. Wearables and room sensors can alert emergency services if an accident occurs in an elderly person's home, which makes living alone safer and alleviates the worry of friends and family. Exercise equipment combined with mapping data can help prevent memory loss. Electricity- and water-metering data, originally designed to monitor utility use, can be used to assess activity in a residence and alert social workers in the event of a sudden drop-off.

Monetizing Alternative Data Uses

Alternative data uses are still relatively uncommon, in part because the path to monetization for the data owners often isn't clear. Our experience and research have identified seven reasons why companies can be reluctant to monetize data:

- **Lack of Imagination.** The companies that generate and own the data are simply not aware of its value generation potential, especially when it lies far from their core business or industry.

- **Unknown or Unidentifiable Value.** The data's value may depend on awareness of its alternative uses or on access to data from other alternative sources, making it difficult for companies to price their own data.

- **Fear of Losing Value.** Companies recognize the potential for future valuable use cases but worry that sharing data with partners may undermine their ability to capture its value themselves.

- **Fear of Losing Control.** Given data's complicated journey from origin to application, companies fear that sharing

their data may cost them control of its ultimate uses—and bring attending financial, brand, or regulatory risks.

- **Data Platform Limitations.** Companies' technological ability to stream, track, control and ultimately charge and get paid for their data is limited.

- **Contractual Limitations.** Companies are constrained by contractual terms from monetizing data originally generated by their customers.

- **Limited Data Inventory and Quality.** Companies do not know what data they own or whether it is of sufficient quality to be useful.

Times are changing quickly, however, and corporate and other data users are overcoming these constraints. For example, BCG has been advising three very different organizations—an automotive OEM, a large commercial trucking company, and a municipal government—on how they can extract value from alterative data uses. All three are following a similar roadmap and process.

They catalogue their data internally to get the best sense of what is available. Like other leading explorers of alternative data uses that emulate digital natives such as Airbnb, Uber, and LinkedIn, these organizations maintain data catalogues that make their data readily available to internal and potentially external users. New data management platforms, such as Talend and Collibra, help sort data to make it more accessible.

They establish processes to think creatively about the value of their data. Conceiving of use cases in adjacent industries is easier, but these organizations pushed themselves to consider the needs of other sectors and the broader economy. They revisit this thinking as use cases enabled by new technologies emerge.

They seek partners that can augment their own capabilities and overcome limitations. Sharing data to find new use cases doesn't just divide the pie, it can enlarge the pie. And while a company may hold a very valuable data set, it may not have the capabilities or market position to make the most valuable use of it. It's important to find the right partners and to understand the different ecosystems that could benefit from the data and the potential solutions that other companies could contribute.

They keep an eye out for changing data uses. IoT and artificial intelligence are continually evolving, so identifying opportunities to monetize data is not a one-time exercise. Companies can establish a formal process, with external advisors and partners, to stay current on how data can be used across a broad range of industries.

They carefully balance opening their data to discovery against exposing it to excessive risk. Holding data too tightly may deprive a company of a revenue stream and society of valuable use cases. But sharing it too broadly may have unintended consequences. A number of methods exist to balance this tradeoff, including data directories, sample data sets, synthetic data, and distributed data models that allow access but prevent the data from being replicated.

ALBERT EINSTEIN USED thought experiments to visualize his theories and understand an untested theory's implications in the physical world. Business leaders may want to use a similar technique to visualize how their data can be used. Like alternative data uses themselves, such experiments may fall outside of normal management practices, but the companies that are first to understand the value of alternative data—and how to realize it—can build new revenue streams and new sources of competitive advantage.

About the Authors

Massimo Russo is a managing director and senior partner in the Boston office of Boston Consulting Group. You may contact him by email at russo.massimo@bcg.com.

Tian Feng is a project leader in BCG's Boston office and a BCG Henderson Institute ambassador. You may contact her by email at feng.tian@bcg.com.

Boston Consulting Group partners with leaders in business and society to tackle their most important challenges and capture their greatest opportunities. BCG was the pioneer in business strategy when it was founded in 1963. Today, we help clients with total transformation—inspiring complex change, enabling organizations to grow, building competitive advantage, and driving bottom-line impact.

To succeed, organizations must blend digital and human capabilities. Our diverse, global teams bring deep industry and functional expertise and a range of perspectives to spark change. BCG delivers solutions through leading-edge management consulting along with technology and design, corporate and digital ventures—and business purpose. We work in a uniquely collaborative model across the firm and throughout all levels of the client organization, generating results that allow our clients to thrive.