THE RISKS AND REWARDS OF DATA SHARING FOR SMART CITIES

By Massimo Russo and Tian Feng

This is the third article in a multipart series.

LONG BEFORE THE ARRIVAL of COVID-19, cities around the world had started to reinvent themselves through "smart city" initiatives. Beset by growing populations, aging or insufficient infrastructure, and the rising cost and difficulty of meeting the needs of residents, cities were investing in technology to address environmental sustainability, traffic congestion, public transit, and public health, all while facing budgetary pressures.

The pandemic has intensified the need for intelligence and insight as cities struggle to assess the effectiveness of response measures and compliance in real time. Meanwhile, increasing reliance on e-commerce and the shift to remote working are stripping cities of two of their core functions—as places to shop and places to work. Public services are being strained by rapidly growing demand. New challenges, such as the rising volume of last-mile delivery trucks, are adding to the stress.

To develop innovative solutions to problems old and new, many cities are aggregating and sharing more and more data, establishing platforms to facilitate private-sector participation, and holding "hackathons" and other digital events to invite public help. But digital solutions carry their own complications. Technology-led innovation often depends on access to data from a wide variety of sources to derive correlations and insights. Questions regarding data ownership, amalgamation, compensation, and privacy can be flashing red lights.

Smart cities are on the leading edge of the trend toward greater data sharing. They are also complex generators and users of data. Companies, industries, governments, and others are following in their wake, sharing more data in order to foster innovation and address such macro-level challenges as public health and welfare and climate change. Smart cities thus provide a constructive laboratory for studying the challenges and benefits of data sharing.

Why Cities Share Data

BCG examined some 75 smart-city applications that use data from a variety of sources, including connected equipment (that is, the Internet of Things, or IoT). Nearly half the applications require data sourced from multiple industries or platforms. (See Exhibit 1.) For example, a parking reservation app assembles garage occupancy data, historical traffic data, current weather data, and information on upcoming public events to determine real-time parking costs. We also looked at a broader set of potential future applications and found that an additional 40% will likewise require cross-industry data aggregation.

Because today's smart solutions are often sponsored by individual municipal departments, many IoT-enabled applications rely on limited, siloed data. But given the potential value of applications that require aggregation across sources, it's no surprise that many cities are pursuing partnerships with tech providers to develop platforms and other initiatives that integrate data from multiple sources.

For example, Huawei is the chief architect of smart-city solutions in the city of Shenzhen's Longgang District, working with hundreds of partners to access data. Even
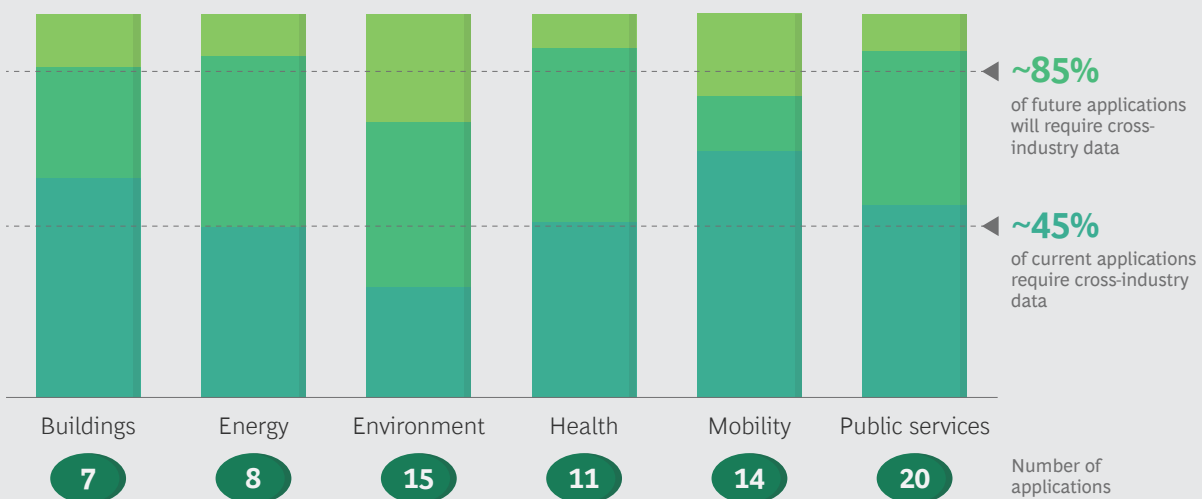
a project with relatively limited scope—the LinkNYC program in New York, which replaced pay phones with Wi-Fi enabled kiosks—required three different companies to provide the necessary data, hardware, and network capabilities. Cities are becoming the managers of data platforms and the orchestrators of burgeoning digital ecosystems. (See the sidebar, "High-Impact Solutions Depend on Sharing Data," page 4.)

How Cities Share Data Today

We reviewed the approaches to data management of 30 smart cities around the world and found a number of common practices. (See Exhibit 2.) We also found that these cities need to address more than a few issues as they try to improve their current operations and develop new and better services that depend on data sharing. The issues fall into three categories: how and when municipal departments share data, how cities share data with private-sector vendors, and how cities share data with the public.

Data Sharing By and Within City Governments. The siloed data repositories that municipal departments have built over the years limit cities' ability to develop an end-to-end view of how residents avail
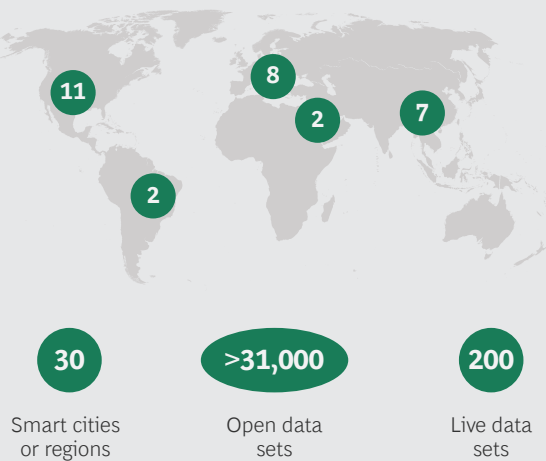

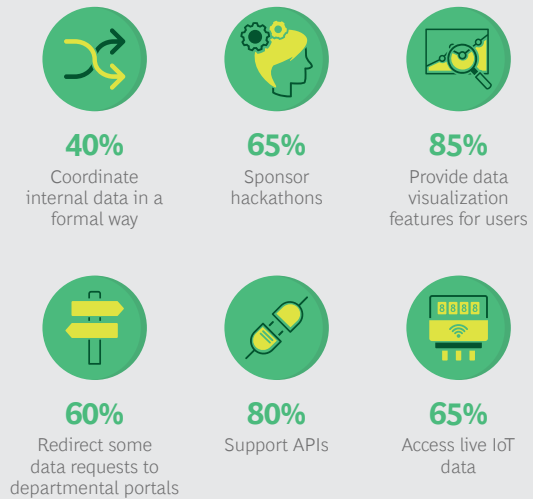EXHIBIT 1 | Smart-City Applications Require Data from Multiple Industries and Platforms

~85% of future applications will require cross-industry data

~45% of current applications require cross-industry data

Buildings	Energy	Environment	Health	Mobility	Public services	
7	8	15	11	14	20	Number of applications

Source: BCG analysis.

EXHIBIT 2 | Smart Cities' Open-Data Platform Practices

We looked at...

11
8
2
7
2

30
Smart cities
or regions

>31,000
Open data
sets

200
Live data
sets

and found that...

40%
Coordinate
internal data in a
formal way

65%
Sponsor
hackathons

85%
Provide data
visualization
features for users

60%
Redirect some
data requests to
departmental portals

80%
Support APIs

65%
Access live IoT
data

Source: BCG analysis.

themselves of services and experience them. Moreover, cities have tended to pursue individual smart-city initiatives rather than citywide transformations. These initiatives typically lack an overarching strategy, reflecting both the fragmented nature of city agencies and legacy technologies. Cities need to overcome both organizational and technological siloes if they are to realize the full potential of smart cities.

Recognizing that citizen-centric solutions require more data sharing within government, cities are starting to invest in internal data exchanges. They are building open-data platforms and data directories that give all departments and agencies greater access to available data. Some 40% of the smart cities we reviewed have open-data platforms with integrated back-end databases that permit access by external applications, a prerequisite for easy data sharing and use. Another 15% have included the building of back-end connectivity as a priority in their digital strategic plans. For example, Portland (Oregon) and Seoul (South Korea) are constructing their own internal data repositories: Portland Urban Data Lake and Smart Seoul Data.

The sharing of some data is inherently problematic. Cities are rightly concerned about exposing personal and other sensitive information externally on open platforms or internally between departments. For example, should medical information be accessible to public agencies? Most people would say no. But what if such data, in aggregated form, could inform public policies designed to curb air pollution or reduce traffic emissions? In Kentucky, the city of Louisville used aggregated data from smart asthma inhalers to monitor the health impact of air pollution in real time. And when are the benefits of sharing data so great that investing in safeguards to anonymize it is justifiable?

Anonymization can be a shallow solution, however—and a potentially reversible one when multiple data sources are involved. Combining anonymized vehicle GPS and address data, for instance, can easily expose people's identity. Moreover, both anonymizing and aggregating data can mask important underlying patterns. As cities build data platforms, they will have to manage the tradeoff between innovative, valuable solutions and privacy risk through encryption, permissions, access restrictions,

HIGH-IMPACT SOLUTIONS DEPEND ON SHARING DATA

As smart cities turn to data and technology for high-impact solutions to high-value use cases, the need for data sharing will grow. Consider two examples: traffic management and elder care.

Current traffic applications use cameras, parking sensors, and telematics data to identify areas of high traffic density. In Austin (Texas), London, and San Francisco, the transportation team runs analytics that, respectively, optimize traffic signals, set dynamic tolling rates, and adjust smart-parking rates. In the future, cities may be able to maintain optimal mobility by adjusting tolls in real time according to actual traffic conditions. Instead of being extrapolated from historic data (as they are today), rates could be based on real-time traffic data using the locations of individual vehicles and their expected destinations (derived from their navigation systems).

Drivers would decide whether to pay the toll to save ten minutes of travel time, for example, or take a detour to avoid it.

Elder care is a growing priority for cities with aging populations. Biometric data from wearables is already being used to monitor the status of high-risk patients and notify them or their physicians of health-related events. Citizen-centric pilots in cities such as Seoul are monitoring activity in the home with motion, electricity, and lighting sensors. Computers recognize changes in the daily patterns of seniors living alone and alert social workers in the event of a potential problem (such as when an individual is suddenly much less active than usual). In the future, wellness recommendations may be personalized according to an individual's condition and goals and combined with activity tracking as he or she ages.

and other methods. And as more and more use cases continue to highlight those tradeoffs, cities will likely need to develop new or adjusted rules and guidelines about what can be shared, when, and with whom.

Data Sharing with Third Parties. Technology and other private-sector companies have long partnered with cities on specific programs. Given the scope of smart-city initiatives—with their ongoing service requirements, live data streams, and often bespoke applications—technology vendors are becoming critical partners in the collection, hosting, and management of data. The responses to Boston's 2017 smart-city request for information document, for example, included numerous proposals requiring comprehensive codesign. Likewise, the Connected Citizens Program of the navigation app Waze works with many cities, including New York, to provide real-time traffic data (such as user inputs on accidents or potholes) in exchange for data on road closures and public-works projects.

Cities also share data with private companies to nurture innovation. Many cities (and other governments) around the world host Startup in Residence programs that give startup companies the space, resources, and access to officials and data that they need to develop public-sector solutions. The City of Boston's Beta Blocks program aims to create a "clearing house, matching platform, or exchange" for civic experiments. The initial request for information generated more than 100 responses from potential partners worldwide.

Many of these partnerships involve data sharing, and the agreements vary widely in scope. Toronto struck a partnership with an Alphabet subsidiary, Sidewalk Labs, for a smart-city pilot in the city's Quayside neighborhood that made Sidewalk Labs responsible for all sensor data collection and analytics. The partnership sparked considerable public debate about data sharing and governance, and these issues contributed to the project's demise earlier this year. In other, narrower initiatives, such as

LinkNYC and Kansas City's Smart and Connected City project, private-sector partners take ownership of the data generated. Cities need to consider whether to concede data rights to third parties and the potential for controversy regarding the sharing of residents' data with commercial entities.

Data Sharing with the Public. Cities initially developed public data-sharing platforms to improve government transparency. Now they are using open-data platforms, open-source methods, and competitions to encourage citizens to develop innovative solutions to urban problems. Cities are providing developer support and application programming interfaces (APIs) that enable citizens to access often live data directly so they can build solutions seamlessly on top of government data.

Almost two-thirds of the cities we reviewed are hosting events such as hackathons to support citizen engagement and innovation, and 80% offer data APIs. Singapore, for example, offers 14 APIs for users to access live data, Dubai has established an API and developer zone along with cloud hosting (for a fee), and Barcelona has developed an open-source data platform, Sentilo, that allows app developers to access a wide variety of sensor data.

City open-data platforms include features to manage data access and permissions, integrating functionality that provides varying degrees of access, control, and security. Amsterdam's API can be used by anyone, but access to certain data sets requires that users provide authentication and authorization. In Texas, the city of Austin requires a login to download certain data sets, with varying levels of permission. New York's transit API requires a free key registration so the city can track data use.

A few cities (and other government entities) are beginning to crowdsource citizen and enterprise data. Columbus (Ohio), San Francisco, Toronto, and all of Estonia are among the few that allow the public to upload their own data sets, which are then vetted by the government. In 2015, Copenhagen partnered with Hitachi to pilot a

data exchange where the government and private-sector players can buy and sell data, but the project has yet to reach significant scale.

Real-time IoT data sets are just beginning to be integrated into open-data platforms. While 65% of the open-data platforms we surveyed featured some sort of rapidly updated sensor data, only a small number of live IoT data sets are available to the public, mainly in areas of high consumer interest such as public transportation, traffic, and the environment.

The Challenges of Broader Data Sharing

While cities are at different stages of maturity when it comes to data sharing, there is a broad trend toward more aggregation and sharing of data within city governments and with external parties. We see four big challenges that cities need to address.

How to Break Down Data Silos. In all the smart-city open-data platforms we observed, the city is the central collector, curator, processor, and host of a variety of data from different departments. But this approach, with each city creating its own bespoke data system, raises issues of expertise and scale. As data volumes increase, does every city really need to become a technology and data expert? Commercial technology partners can provide ways to handle common needs and help more cities gain access to innovative solutions.

How to Ensure Data Quality and Control. Managing data quality and controlling access without unduly limiting innovation raise issues that are just coming to the fore. They apply to all manner of data sharing—within cities, with third parties, and with the public. Addressing these issues will likely require the development of national rules, regulations, and standards to avoid a complex tangle of individual municipal solutions. Technology providers are developing more sophisticated answers to questions of data access and control in order to encourage innovation while protecting privacy. The proliferation of COVID-19 contact-tracing

solutions, and the challenges of adoption due to concerns about how the data could be used, provide a strikingly relevant illustration of the problem.

How to Balance Innovation, Data Ownership Rights, and Value.
With increased data sharing, commercial relationships between cities and private-sector companies must address ownership rights. Who owns and controls what data? Who should profit from any insights generated by the data? What liabilities are associated with the use or misuse of data?

How to Manage the Tradeoffs Between Data-Sharing Benefits and Privacy Concerns.
Cities are making tradeoffs between the value of innovation and the value of privacy. Singapore published data-sharing principles, for example, that emphasize the accessibility, co-creation, timeliness, and machine readability of data; the city also wants data to be "as raw as possible." However, more open and detailed data comes with privacy worries, and IoT sensor data can introduce additional concerns because of its sheer volume, real-time nature, and ubiquity. How can cities leverage governance and technology to balance the tradeoff between value and privacy? Various initiatives and organizations, such as New York University's GovLab and the UK's Open Data Institute, are pursuing data governance models and

data collaboratives (in which participants exchange data for public value).

CITIES ARE AT the forefront in aggregating data to address large-scale challenges. As more cities seek to become "smart," they will face a growing conundrum. On the one hand, truly smart cities will develop more citizen-centric IoT solutions, and even solutions that address a single problem will involve more aggregation and sharing of data. Cities will become better equipped to address higher-value use cases such as public health and welfare. On the other hand, as more data is aggregated, the risk of misuse, privacy violations, and misappropriation of value will rise. Enterprises orchestrating ecosystems of partners face similar challenges in balancing innovation and risk.

Municipal leaders, corporate executives, government policymakers at all levels and their private-sector partners need to start considering how cities can create innovative solutions to complex and changing needs while limiting risks and protecting privacy. More than ever in the midst of a global pandemic, smart cities provide a real-life, real-time lab for all of us.

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

About BCG Henderson Institute

The BCG Henderson Institute is Boston Consulting Group's strategy think tank, dedicated to exploring and developing valuable new insights from business, technology, and science by embracing the powerful technology of ideas. The Institute engages leaders in provocative discussion and experimentation to expand the boundaries of business theory and practice and to translate innovative ideas from within and beyond business. For more ideas and inspiration from the Institute, please visit https://www.bcg.com/featured-insights/thought-leadership-ideas.aspx.