Cognitive Computing and Big Data Analytics
By Judith Hurwitz, Marcia Kaufman and Adrian Bowles

CHAPTER

12

Smarter Cities: Cognitive Computing in Government

One of the great challenges of the 21st century is how to leverage technology to solve a variety of problems that accompany the global trend toward urbanization. In cities everywhere, increasing population density strains physical systems and resources. Individual systems have been developed to collect and manage data for each of functional unit. When critical information cannot be shared across critical services, managers often cannot anticipate safety issues or opportunities to optimize services.

The promise of cognitive computing is to enable metropolitan areas to take advantage of data to evolve and become smarter, and deal with expected and unanticipated events effectively. The objective, therefore, is to learn from experience and patterns of data to improve the way cities function over time. This chapter reviews the problems confronting cities and demonstrates how cognitive computing has the potential to transform the way cities operate.

How Cities Have Operated

A city is more than the roads, buildings, bridges, parks, and even people found within its borders. Cities around the world have evolved in a similar way for centuries—agencies are created to provide services to the population in response to changing conditions and technologies. These agencies justify their existence based on their ability to collect the right data and manage that data to support their constituents. For example, population density made the rapid spread of disease a public health issue and created a need for public health data tracking. New modes of transportation such as cars and planes required new transportation management departments that required the collection even more data.

Throughout history, these agencies produced paper records, which dominated the way cities managed processes. The problems with paper records are obvious: they are expensive to store, inefficient to retrieve, and subject to damage or loss from water, fire, or even rodents. Even as paper documents could be scanned via optical character recognition (OCR) to make searching text easier, it did not solve the problem. There was still no way to gain insights into the history, meaning, and context of these documents. The basic issue is that documents contain deep structure, which contains virtual knowledge that is not explicitly captured.

When all data was created manually and managed in the form of paper documents, it was difficult or impossible to recognize the relationships and dependencies between these systems. For example, relationships between education and hygiene, hygiene and disease, crime and poverty, are all obvious to us today. However, without a way to analyze data from different agencies or departments to see these patterns, they depended on insight and hypotheses that drove the quest for supporting data. As cities grew with data in departmental silos, it became increasingly difficult to look at data across departments in order to set budget priorities. There was no systematic way to put the pieces together and learn from experience. The real value in the information resided in the brains of these people.

Improved technology provided more efficient ways to collect, manage, and analyze data, the ability to understand and describe scenarios and predict outcomes improved dramatically. In the last few decades there have been significance advances in data-oriented city management. To support changing needs has required the movement from simply managing data more efficiently in databases to adopting analysis tools to make better decisions based on this data.

As shown in Figure 12-1, data management is on a path from document-centric silos to integrated repositories of standards-based structured and unstructured data. From manual systems to a sensor-based generation of relevant data, the progress has been dramatic and opens up new opportunities to develop systems that learn from their own data and experiences.

For example, a modern transportation department can use sensors or transponders, closed-circuit TV (CCTV) for video images, or perhaps mobile phone tower pings to more accurately determine not only how many vehicles pass a certain point at certain intervals. These systems can accurately track where data originates and where it goes. Knowing "who" is much more valuable for planning and operations than simply knowing "how many." When traffic counts were made by mechanical or manual traffic recorders, traffic workers knew only how many cars had passed. With more details about "who" (even anonymized), "where to," and "where from," they can build models that begin to predict flows much more accurately than simulations, and with machine learning algorithms, even begin to control flow by adjusting traffic signals based on actual usage.

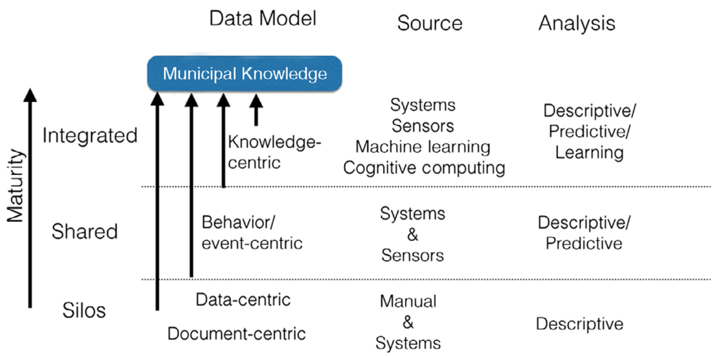

Figure 12-1: Foundations of cognitive computing for smarter cities

As similar advances were made in the data collection opportunities for other agencies, from health to safety to areas such as education, new and improved opportunities for analyses have emerged. As data from all these sources—across departments, structured and unstructured—is captured and made available in standard formats that encourage interagency sharing, cities become an ideal, data-rich environment for developing smarter applications.

The Characteristics of a Smart City

As noted, a city may best be understood as a combination of complex systems that have to work in collaboration with each other—sometimes called a *system of systems*. This is what makes cities and metropolitan areas so difficult to manage. Take a typical large city such as New York or Tokyo. These types of cities include roads and bridges, commercial and residential buildings, public transportation systems, private transportation, water systems, schools, and the public safety infrastructure. Although each of these elements is a universe within itself, they are all interdependent. Operationally, cities rely on smart managers to figure out best practices for managing and improving the way cities work best. But as cities have grown, it is simply impossible for smart managers to approach data driven problems systematically.

A city can become "smarter" if enough data is collected, analyzed, and managed so that critical improvements can be made. What does it mean for a city to be smart? It means that those managing a city collect the right information from a variety of sources and create a unified corpus of data that defines the components that make up a city infrastructure.

Figure 12-2 shows a typical set of government agencies, ranging from the basics of emergency services through utilities, public health, transportation, and human capital management. As individual citizens increasingly have persistent or mobile Internet access and become accustomed to interacting with their government in an ad hoc manner, you can also view community engagement as a function and potential differentiator for cities and other geopolitical units. The following sections break down these functions by task.

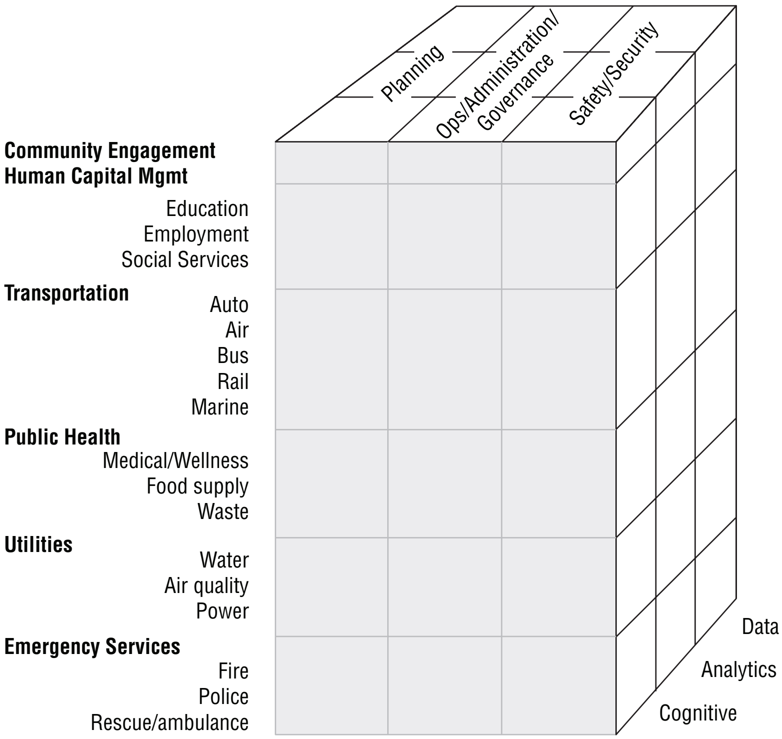

Figure 12-2: Data/knowledge management for cities

Collecting Data for Planning

Each agency in a city collects data for planning, operations, and safety and security, all of which are ongoing activities. The movement toward smarter cities includes better collection and analysis techniques at each stage.

City planning requires that management focus on a broad range of activities in order to promote growth. This growth has to be coordinated with the right scale of operations that allow that growth to be sustainable. At the same time, growth has to be planned so that it can improve and protect the quality of life for citizens. Planners evaluate options and codify them in policies. Planning requires managers to use their intuition and available data to make decisions about the future. The best decision makers are those managers with enough experience to understand what will work and what actions will likely cause problems. However, even the most knowledgeable managers are helped when they get access to better analytical data. What are the trends for extreme weather? What events are happening in the region that might lead to civil unrest, agricultural failures, or drastic population shifts? What are the revenue shifts in industry? What is the outlook for wages? How do all these factors impact the way a metropolitan area can be managed efficiently and effectively? If these professionals managing cities are armed with analytics that can determine patterns and anomalies, they are better prepared for changes.

Cognitive computing applications are well suited to become dynamic planning assistants. Key technologies for cognitive planning systems would include hypotheses generation and evaluation, machine learning, and predictive analytics. The capability of some systems to read vast quantities of unstructured natural language text and analyze it for relevant events and trends will make planners everywhere more effective over time. The cognitive system does not simply analyze data from past events but collects data across anything that impacts the way a city operates. The system looks at the context and relationships between data elements and learns from the data that is ingested and managed over time.

Managing Operations

When policies are in place, various agencies are required to manage daily operations. The use of data and simple analytics has, of course, changed the way most departments operate. Creating repositories of "open data" by public agencies increases visibility into the processes of government and can improve community engagement. At the same time, this open process creates opportunities for commercial ventures to add value through analysis of the data, or even simply improving the way it is presented to the public.

Internally, this new data becomes even more valuable to civil servants because it makes systems based on prescriptive analytics more effective. City managers and planners are constantly trying to anticipate equipment and infrastructure failures and trying to determine how to provide better services without increasing costs. Cognitive computing approach solutions will soon become a mainstay of city operations centers because these systems can learn the rules of the policies while recognizing the realities of the terrain to "think outside the box." Today, manufacturers of complex and expensive machines like airplanes are already using machine learning to predict mechanical failures and ensure that the replacement parts are in the right vicinity. Tomorrow, no modern city will be without systems to do the smart repair and maintenance management from trucks to supplies, such as salt for icy roads.

Managing Security and Threats

Virtually all cities assume some responsibility for public safety and security. Beyond emergency services, protection from natural and man-made threats is an ongoing concern. Identifying potential threats that may occur with little or no warning (from gas line leaks or explosions to hurricane forces), internal and external threats must be monitored, assessed, and mitigated or met with managed responses. Cognitive systems are designed to identify trends and events from a variety of unstructured document sources combined with data from sensors, social media, and community sites. Leveraging all of this data is needed to make safety and security initiatives more efficient and targeted.

Managing Citizen-produced Documentation and Data

For each of the major departments, such as transportation, public health, and emergency services, there is a role for the foundation technologies and cognitive workloads (shown in Chapter 2, "Cognitive Computing Defined," Figure 2-1). Citizens, in these interactions with government, generate a considerable amount of data. All generate or capture data (refer to Figure 12-1), so the foundation structured and unstructured data management workloads are the requirement for advanced systems. At a higher level, senior management in each department has a role in planning and managing operations. By taking advantage of a cognitive assistant that helps to generate and evaluate hypotheses, a manager can gain insights and take corrective actions faster. The wisdom that comes from experience, once held tightly as a job requirement, can be codified in a departmental corpus and used to make workers at all levels more effective.

Any system that gets input from the public or provides assistance to its residents can benefit from natural language processing (NLP) interfaces. However, the most value is created by systems that actually learn from experience as they build a corpus from use by residents and city employees, and increasingly from sensors, as previously discussed. For example, a transportation system that can accept updates about road conditions and traffic flow in real time from sensors, ad hoc natural language input from citizens, and is connected to the maintenance scheduling system for transportation department machines, will be more effective in dispatching the right equipment at the right time to make repairs, start preventative maintenance, or even schedule system upgrades, than a system that is missing any one of these components. Learning from experience—the defining characteristic for all cognitive computing solutions—is the key technology to improve this performance.

Data Integration Across Government Departments

The importance of interaction between departments cannot be overstated. If water, electrical, or gas lines are to be replaced or maintained, that will often have a major impact on overall infrastructure in other areas. For example, replacing gas lines might mean that a city has to reroute traffic or even resurface major highways. The department responsible for managing utilities will have to share critical information with the transportation planning system so that there is minimal disruption and so that the public is prepared. Predicting major events and planning for alternative routes can help a city manage change. Again, systems that learn from their experiences can better share data among departments to make the overall system more effective. In a major city, no manager can have insights into every planned or unplanned activity. But an integrated cognitive computing system with a common corpus that aggregates data and knowledge from all departments can certainly capture it all and share what needs to be shared between departments.

Chapter 3, "Natural Language Processing in Support of a Cognitive System," (Figure 3-1) discussed the fundamentals of learning systems and the concept of fast thinking versus slow thinking. Fast thinking tasks require intuitive actions that a manager might do without difficult analysis—responding to a citizen complaint or alerting people of an impending weather event. In contrast, slow thinking requires deep thought, analyses, and judgment. In city systems, fast thinking tasks can be automated to make departments more efficient providing deterministic answers to city workers and the public. These systems either require on demand access via the popular 311 information systems in U.S. cities or based on events, such as dispatching alerts to people in a defined perimeter when it's necessary to evacuate for a gas leak, or to remain in place during a police emergency. These answers and notifications can be based on knowledge managed within the cognitive computing system because the system understands the context between elements such as events, people, systems, and the like. The cognitive system is designed to understand relationships and patterns.

For slow-thinking problems requiring consideration of multiple scenarios, or for situations for which there is no single right answer, a cognitive computing system can supply probabilistic responses. Again, knowledge about the user—employee or resident—can make the answers more relevant. For a public health manager trying to determine which course of action to take when facing the possibility of an infectious disease outbreak, confidence-weighted alternatives such as those discussed for healthcare in Chapter 11, "Building a Cognitive Healthcare Application," may help in ensuring that supplies of vaccines or treatments are adequate. Integrating this system with the education system may have a ripple effect on bus planning, which may, in turn, have a ripple effect on transportation logistics. In a city, everything is connected and interdependent. Unified cognitive computing applications, sharing a city corpus and supported by open data, can make those interdependencies a strength rather than a weakness.

The Rise of the Open Data Movement Will Fuel Cognitive Cities

Nations have long shared data valuable to businesses and individuals when it didn't compromise their own interests. From nautical charts in the 19th century to GPS data in the 20th century, the trend toward openness as a default has grown. In the 21st century, this trend has accelerated among cities, aided by better communications systems, standards, and regulations affecting the distribution of publicly held data.

In March 2012, then Mayor Bloomberg of New York signed the New York City (NYC) Open Data Policy (Introductory Number 29-A) and tasked the Department of IT Telecommunications with developing and posting standards for all agencies to make public data available online.

A goal of the initiative was to provide access to all public data from all agencies through a single web portal by 2018. To date, more than 1,000 public data sets are available from NYC agencies, commissions, and other groups, and available for a variety of private and commercial uses.

NYC also has a program called BigApps to "help teams advance new or existing projects that aim to solve civic challenges and improve the lives of New Yorkers." Teams work with civic organizations to develop applications that use this open data, while vying for cash prizes (totaling more than $100,000 in 2014).

Making data available is the first step, but today the problem isn't a lack of data, or even access to data. The problem is the ability to understand what of value is actually *in* all that data. Using the foundational technologies that are part of the fabric of emerging cognitive computing solutions has the potential for revolutionizing how data can improve the way modern metropolitan areas are dynamically managed.

The Internet of Everything and Smarter Cities

Previous chapters explore the technologies that make cognitive computing possible. Here you can find answers to two critical questions for public sector cognitive computing: "Where does the data come from?" and "How is value created?"

Now attack the first question. In a modern smart city, or one with smarter aspirations, data comes from three primary sources: citizens, governments, and businesses, through systems and sensors. For businesses, much of the information comes from smarter buildings, which are managed by systems that coordinate information from all the internal systems and sensors for heating, ventilation and air conditioning (HVAC), water, power, transportation (elevators and escalators), and security. The adoption of standards for enabling all sorts of devices to connect to the Internet has made it easy to envision a future where everything that can be connected will be connected. Assigning an Internet Protocol address (`IP_address`) to a device uniquely identifies it to the rest of the world and allows it to potentially share information with anything else on the Internet. Computers are no longer the only option for Internet communication. Increasingly there are a range of devices that incorporate sensors. Today, devices including refrigerators to smart watches and clothing include the ability to enable machine to machine communication. Summary information is provided to the receiving system or the humans using these sensor-based systems on a need-to-know basis. This so-called "Internet of Things" (IOT) or "Internet of Everything" (soon to be known simply as the Internet) enables businesses and governments to unobtrusively collect or derive all sorts of data about people as they carry on their daily activities. And most of it is given willingly, or without much of a fight. From smart meters that monitor energy usage in the home to transponders that allow us to speed through toll booths without stopping to devices that monitor acceleration and location in our automobiles to closed circuit TV (CCTV) cameras that can monitor group and individual movement, there is no shortage of data being generated.

This city-centric big data, when used to power predictive analytics algorithms or to develop a corpus for a cognitive computing solution, can provide insights that would never be discovered in time to be useful if the data were kept in departmental silos and analyzed by those with no incentive to share with other departments. It is the integration of this data that enables cognitive computing applications for smarter cities of the next decade.

Understanding the Ownership and Value of Data

In the new cognitive computing era, the traditional questions of who owns such data, who benefits from it, who transforms it into knowledge, and who owns that knowledge takes on new importance as the interconnections create new opportunities to improve the quality of life. At the same time they threaten to take away all semblance of privacy and perhaps security.

Figure 12-3 shows some of the critical relationships between citizens, businesses, and government that enable a virtuous cycle of knowledge creation. With the advent of social media and greater citizen participation in explicit and implicit data reporting, cities get smarter every day. Explicit data reporting includes applications like Street Bump in Boston, a mobile phone app that collects real-time data about city streets by monitoring a user's GPS and accelerometer to identify potential problems such as potholes. (Speed Bump "knows" where the speed bumps are located so it can avoid many false positives.)

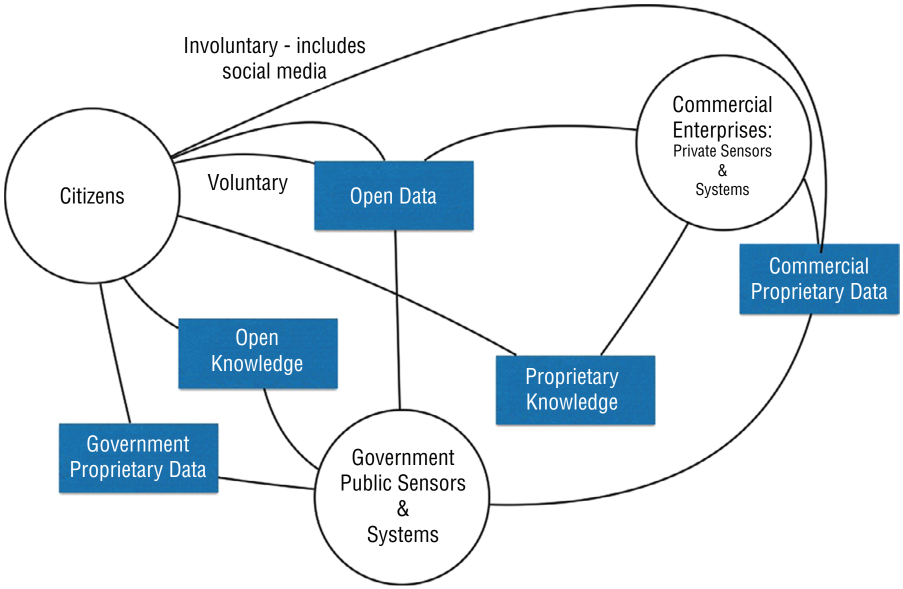

Figure 12-3: Modern city data sources and managers

Implicit data reporting includes activities such as making public comments about government activities on social media that are harvested by government agencies. It would also include systems or sensors that are commonly used with little thought given to the information gathering potential, from loyalty programs at retailers to facial recognition software used in public spaces on CCTV networks.

Cities Are Adopting Smarter Technology Today for Major Functions

This section presents a brief survey of projects in cities that are already working with leading vendors to implement cognitive and precognitive solutions (those based on foundation technologies like predictive analytics that pave the way for future cognitive computing efforts). As the field of cognitive computing is still in the early stages of maturity, most of these projects are ongoing and still being refined as the cities and vendors learn from their experiences. Using the typical urban organizational structure (refer to Figure 12-2), you can focus on a few projects that may form the basis for integrated cognitive computing solutions in the future. In particular, you can consider opportunities to improve the quality of life through analytics and networks that learn from experience.

Managing Law Enforcement Issues Cognitively

It is not surprising that one of the greatest potential opportunities for smarter cities is in the area of law enforcement. This is an area in which there is a huge volume and variety of data that must be managed and analyzed. It is also an area in which patterns and anomalies play a huge role in solving and preventing crime. Leveraging a cognitive approach can benefit not only a single police department but also has the potential for providing repeatable best practices that can have wide use.

The Problem of Correlating Crime Data

The biggest problem facing police forces of metropolitan areas is the difficulty in correlating data from hundreds and even thousands of different data sources. Police departments often have access to large databases that capture historical data from local, regional, and national reports of arrests, crimes, and other recorded incidents. These data sources are often stored in relational databases. Other information sources are unstructured and stored in incident reports, paper files, interviews with witnesses, and the like. Still and video images and audio data must also be analyzed. In addition, there may be situations where a large volume of this data is produced on a daily basis from public and private surveillance cameras. Acoustic event detection monitors makes it impossible to manage manually. Recent advances in facial detection algorithms, however (based on recognition of facial parts and their spatial relationships, along with coloring), have helped to automate this task. The problem of analyzing real-time video as it is captured, however, remains a challenge. Big data provides new opportunities, but big cases still demand a lot of human personnel.

Departments, therefore, need enough time and enough skill to fit the pieces of data together to solve crimes. This is complicated because it is not always possible to correlate data manually. An experienced detective will know how to analyze data and will have mastered the process of connecting the dots. In addition, many crimes generate hundreds or thousands of new reports, pictures, bystander videos, and accounts that have to be considered, but which may be received as unstructured natural language voice recordings or text. The sheer volume and complexity of the data can be overwhelming. Even if the "answers" are in the data, finding them in time to prevent further crime can be impossible if a human has to find patterns between data sources.

The COPLink Project

COPLink is a law enforcement information management system originally developed by the Tucson, AZ police department and researchers at the University of Arizona. It has been commercialized by i2, an IBM company, and deployed in more than 4,500 law enforcement agencies in cities around the United States. This customized project is now being turned into a repeatable cognitive solution under the IBM Watson brand.

COPLink enables individuals, departments, and agencies to gather, share, and analyze historical and current crime information. Users can access data from local and national databases from virtually anywhere, using computers or mobile devices to improve the effectiveness of field personnel. COPLink supports general data standards XML and those commonly used in law enforcement, such as the Logical Entity exchange Specification Search and Retrieve (LEXS-SR), to simplify integration with other applications and to promote data sharing.

The Adaptive Analytic Architecture (A3) enables users to generate leads (hypotheses) for tactical follow-up without extracting data permanently. COPLink also offers agent-based alerts to let officers know when similar searching is made to help coordinate efforts, and can send notifications when new information becomes available based on saved searches.

COPLink uses analytics to actually create suspicious activity reports based on alerts and can share these with the relevant jurisdictions. Simplifying searches and providing this level of database integration makes personnel more efficient. Continuous monitoring of new data and events to improve communications makes them more effective.

There are a number of ways that this type of law enforcement application is being expanded with cognitive extensions, including:

- Providing NLP analyses of unstructured reports as they are created in the field.
- Leveraging hypothesis generation and evaluation technology from Watson to augment the case lead generation process that is currently handled manually by officers.
- Treating the locally generated databases as the foundation for a corpus that can be integrated with cognitive operational systems to plan and manage emergency staff and resources. (COPLink uses analytics to predict crime hotspots, but integrating its corpus—including new knowledge learned from experience with hypothesis testing—with planning tools would enable department managers to benefit from the actual experiences.)
- Integrating sensor data from smart transportation systems in (near) real time to support continuous monitoring of activities that may appear innocuous in isolation, but which may be part of patterns that would otherwise go unnoticed.

Smart Energy Management: From Visualization to Distribution

The importance of visual reports to help people make discoveries from complex data relationships has been discussed. Visual representations of the state of complex systems can also be used to ensure confidence in an operational system and to enable a human to participate in the management process when anomalies arise. For example, allocating energy production resources for a smart grid, or responding to expected changes in demand based on predictive analytics, may create situations that benefit from operator intervention. Visual abstractions make it easier for the operator to detect patterns than simply seeing relevant numbers scroll past on devices. As power systems integrate subsystems with different characteristics or energy sources, visual interfaces are the only practical way to present all the information as a continuous stream. This is much like the approach taken with an automobile dashboard that abstracts some data into simple color maps—red for danger, yellow for caution, and green for normal operations—but provides more detail for other functions, such as a numerical estimate of the distance that may be traveled before refueling. In a single system that balances power generation from water for hydroelectric generation, wind, nuclear generation, and solar and storage in the form of batteries or supercapacitors, visualization is the key to helping the user understand what requires attention as soon as the system can detect it.

The Problem of Integrating Regional Utilities Management

A new smart city is being developed in Kashiwa-no-ha, Japan, through a collaboration between public agencies and private sector enterprises with three stated goals: environmental symbiosis, support for health and longevity, and creation of new industries. Creating a new city by design presents an opportunity to build a utilities-based infrastructure that is based on state of the art technologies with minimal constraints from aging systems.

The Area Energy Management Solutions Project

The Area Energy Management Solutions (AEMS) project in Kashiwa-no-ha, Japan, uses advanced analytics to provide an integrated, comprehensive solution for energy production, provisioning, and optimization. AEMS design and development is led by Hitachi Consulting. Hitachi, the $90BUSD global technology and services provider, has emerged as a leader in the smarter cities movement. Its 2014 Annual Report, titled "Social Innovation: It's Our Future," lays out a vision based on building and leveraging technology to address social challenges, including three that resonate with smart city planners: securing water resources, energy and food; replacement of aging infrastructure systems; and improving transportation systems.

The AEMS project is focused on using analytics to manage energy (including electricity, water, gas, and any other production technology that is ultimately adopted by Kashiwa-no-ha) by forecasting demand and dynamic provisioning based on continuous reporting signals from sensors throughout the city. Creating a solution that incorporates renewable energy sources (solar and wind) with more conventional sources and storing excess capacity through storage batteries allows the system to schedule production when it is convenient and cost effective (daylight for solar and medium-to-high wind for turbines) and puts it into the grid on-demand directly or from batteries that store excess capacity from a period of over-production.

AEMS uses analytics to predict peak loads and evaluate alternative approaches to distribution (from interbuilding sharing of resources, similar to the cloud-based model for sharing physical resources).

By planning for an integrated set of energy systems rather than developing hydroelectric, solar, and other data management systems in isolation, the developers created a design that leverages data from each subsystem and also uses analytics to efficiently load balance the overall system to predict demand, provision resources, and distribute more efficiently. AEMS shares information with building energy management systems in commercial, government, and residential facilities.

The Cognitive Computing Opportunity

From the published plans and reported progress about Kashiwa-no-ha, it is clear that the planners want to take advantage of all the smarter cities' products and practices that can be applied to the operation of their new home. Kashiwa-no-ha is expected to become a model, or test bed, for emerging cognitive technologies. Further integration, to include systems for each of the major constituencies discussed in this chapter, will be a cornerstone of system design and provide opportunities for collaboration. The Hitachi AEMS is new but expected to be integrated with other smart systems in Kashiwa-no-ha over time. For example, integration with its transportation management system, and even weather forecasting and monitoring, would improve the performance of all the systems by sharing new patterns discovered by machine learning algorithms. Consider a day when the forecast called for mild weather, but the actual weather that day is unseasonably wet and cold. An integrated system that has learned from its shared experiences might see a correlation with changing traffic patterns, resulting in changing power demands from public transportation systems. In this scenario AEMS might predict an increase in the ratio of people working from home, and prepare to reallocate power based on the behavior of residents, not a preset model of consumption. Integration of algorithms and data between these systems could use real-time or just-in-time sensor data about the discrepancy between forecast and actual weather to "know" that power demands would change in time to adjust production or distribution patterns.

Protecting the Power Grid with Machine Learning

In the United States, power grids are included in the energy sector of critical infrastructure that must be secured against cyber attacks under provisions of the National Cybersecurity and Critical Infrastructure Protection Act of 2014. That requires the Secretary of Homeland Security to conduct activities to "protect from, prevent, mitigate, respond to, and recover from cyber incidents." Energy companies are turning to machine learning algorithms to keep up with the volume of data that must be continuously monitored to mitigate risk and maintain compliance with increasingly stringent regulations. One emerging software company called Spark Cognition in Austin, Texas has developed a cognitive system intended to help secure Internet of Things environments such as power grids.

The Problem of Identifying Threats from New Patterns

Electrical grids are attractive targets for a variety of physical threats, from vandalism to terrorist attacks. More and more, however, the threats are from cyber attacks that attempt to disrupt service. The scale of the physical infrastructure and complexity of connections and dependencies make automation of threat and breach detection critical and potential threat detection a requirement. As new patterns for threats and vulnerabilities emerge, utilities must prepare responses before they materialize.

The Grid Cybersecurity Analytics Project

C3 Energy (C3) is an energy information management firm founded by the founders of Siebel Systems with a mission to make energy systems more efficient and secure through analytics. C3, working with researchers at the University of California and a National Science Foundation Cybersecurity Center called TRUST, developed Grid Cybersecurity Analytics (GCA)—a smart grid analytics application that uses machine learning algorithms to identify and detect potential threats. GCA was designed to understand the characteristics of normal operations (communications traffic levels, asset activity, and such) and identify potentially threatening activities or anomalies. The machine learning algorithms enable GCA to adapt and identify emerging threats as they evolve or mutate over time by reading up to 6.5 billion records/hour and providing petabyte-scale analysis. That level of performance is required for large grids serving cities.

The Cognitive Computing Opportunity

The Grid Cybersecurity Analytics system already leverages machine learning algorithms and the latest research in threat assessment and response. Natural extensions to such a system beyond its application in grid security would center on leveraging data and lessons learned as the system gains experience with threats that have commonalities with potential attacks on other critical infrastructure assets, such as bridges, tunnels, and information networks. A scenario is also envisioned in which systems such as GCA share data with systems such as COPLink and sensor-based smart asset-monitoring systems to help prevent or mitigate attacks by sharing data on individual or groups whose behavior raises a threat alert. On the mundane side, data from the underlying analytics platform could also be shared with a corpus for a resource management system that could be tied to incentives for better energy management and to planning systems to recommend incentives for electric vehicles.

Improving Public Health with Cognitive Community Services

Public health in cities is concerned with wellness and medical care. Wellness includes access to information and preventative care, feedback on behavior that impacts health, and the availability of a full range of care when prevention is not enough. In some jurisdictions it may also include monitoring or managing risks in the food supply and waste management that can improve or diminish overall health. Some of the earliest adopters of cognitive computing technology have been commercial health management firms that have offered personalized recommendations on activities and nutrition, and sometimes incentives for good behavior. In cities, however, many efforts—such as restricting food sodium content or limiting soft drink sizes—have taken a one-size-fits-all-citizens approach. Cognitive computing technology, which can tailor findings and recommendations for diagnoses and individual wellness plans, is the next logical step.

Smarter Approaches to Preventative Healthcare

In many established cities, access to preventative care is seen as a social service for economically disadvantaged people or offered as a perk through private or public employers' insurance policies. As the cost of personal health monitoring drops, through the availability of free or modestly priced online educational resources and access to sensor-based feedback devices, local governments can begin to justify community-oriented health services that offer personalized care driven by analytics.

The Town Health Station Project

Kashiwa-no-ha, Japan, has worked with academia (The University of Tokyo) and businesses such as Mitsui Fudosan to develop the latest in community healthcare supported by cognitive computing solutions to aid in preventative care. Its Town Health Station is a model for public health that is expected to be emulated around the world. The project was designed from the outset to embrace the Internet of things and take advantage of new, continuous sources of health information such as the output of personal devices, from mobile phones to exercise bracelets to workplace and residential sensors. The health station is the center of a partnership between local government, the University of Tokyo Institute of Gerontology, the Center for Preventive Medical Science at Chiba University, and residents, to promote long-term health.

The Cognitive Computing Opportunity

The Town Health Station project (and ancillary programs such as community exercise activities) are already creating a corpus of individual and community data from sensors and professionals that can be shared with other communities to improve the quality of care locally and remotely. As a planned smart city, Kashiwa-no-ha uses analytics to manage traffic and even plan for shared transportation resources in the spirit of the emerging global sharing economy. Integrating these systems in a common corpus that learns from the experiences and behavior of its residents is a natural evolution for the town, and one that would keep it at the vanguard of smarter cities. Similar to the integration envisioned for the Hitachi AEMS solution, the learning systems that share new knowledge between the town, academia, and the medical community is also expected to share the anonymized unstructured health and wellness data with every system that can use it to improve performance.

Building a Smarter Transportation Infrastructure

Intracity transportation management is, in many ways, more difficult than regional and national transportation management due to population and structure density. As cities build up, it is increasingly important to manage transportation and traffic flow through better use of information. When it becomes prohibitively expensive and disruptive to add infrastructure such as additional lanes or levels to roads and rails, getting smarter about who can go where, and when, will drive cities to cognitive computing solutions.

Managing Traffic in Growing Cities

In cities everywhere, traffic congestion leads to frustrated drivers; excess energy consumption, inefficient commerce as people spend more time driving and parking when they could be working and shopping; delayed emergency responses; and higher pollution. It also requires assets to manage peak loads that may be underutilized in general. Almost any change to flow involves trade-offs that inconvenience people and interrupt commerce, while potentially slowing emergency responses. In 2014, lane closures on a heavily traveled bridge between two metropolitan areas in the United States nearly became a political scandal when paramedics took twice as long as usual to respond to an accident with multiple injuries.

The Adaptive Traffic Signals Controller Project

The city of Toronto, Canada, adopted the Multi-Agent Reinforcement Learning Integrated Network of Adaptive Traffic Signal Controllers (MARLIN-ATSC), a system for smarter traffic management, with impressive results. Toronto reports that downtown delays have already been reduced by 40 percent on an average day by simply managing traffic signals more effectively. Xerox, the $23 billion information and document management company, worked with the University of Toronto's Intelligent Transportation Centre to develop and deploy the system, which incorporates camera images and machine learning chips to enable real-time communication between traffic lights to detect patterns and dynamically adjust their timing. Helsinki has a similar traffic project underway with Xerox, with comparable results and similar opportunities to expand into a cognitive computing environment. As Xerox continues its focus on business transformation and migrates away from a dependence on manufacturing devices, it is building up its credentials in smarter city professional services and integration with projects such as MARLIN-ATSC.

The Cognitive Computing Opportunity

MARLIN-ATSC is already an example of machine learning and adaptive, integrated devices. In the next several years, many opportunities may be available to integrate systems such as this with other cognitive computing municipal management systems, and to extend its capabilities by collecting data from external systems such as the cars themselves.

For example, the U.S. Transportation Department is working on a "connected vehicle" initiative that would promote the use of car-mounted wireless devices that communicate with each other and with traffic signals to increase the effectiveness of systems such as MARLIN-ATSC. In addition to the obvious issues of individual privacy and security, such a program faces years of testing and changes in legislation before the government would require the use of these devices in new cars. It is, however, technically feasible now.

Transportation management is one of the most developed domains for analytics and cognitive solutions. In part, this is due to the ready availability of sensors and systems to collect and share data. Virtually every component of a transportation system is amenable to measurement, from cars and buses to planes, trains, and boats, plus the highways and ports that support them. Combining data from this type of signal management with data from CCTV or transponders—even anonymously—could further improve the capability of cities to reduce congestion while providing input for security and planning. Integration with security systems such as COPLink to identify patterns after a criminal event could help law enforcement's capability to predict future events. Integration with systems such as the Grid Cybersecurity Analytics may have limited threat improvement opportunities, but offer real potential to improve the use of electric vehicles by better understanding their usage patterns to optimally locate charging stations and introduce variable pricing to encourage usage patterns that minimize infrastructure requirements.

This type of integration may also help speed the way for the adoption of autonomous vehicles (self-driving cars, buses, and so on) by promoting communication between driverless and human-piloted vehicles to make a smoother transition.

Using Analytics to Close the Workforce Skills Gap

Human capital management in cities requires management of significant interdependencies among employment, education, and social services. Cities also need to find ways to try to lower unemployment, because it can increase crime and lower the tax revenue at a time when assistance is needed the most. Neighborhoods with high concentrations of unemployment face additional problems as lower income disproportionately impacts small businesses and real estate values, too. As businesses tighten their skills and experience requirements for hiring, it is important that new workforce entrants and the unemployed have the right skills when openings appear.

Identifying Emerging Skills Requirements and Just-in-Time Training

A representative from Boeing, one of the largest aerospace firms noted, "We don't have necessarily a labor challenge, we have a skills challenge." Even with thousands of job openings and thousands of applicants, many jobs at this firm still go unfilled. Preparing residents for opportunities by identifying and training them on appropriate skills before the need arises requires an understanding of emerging skills requirements supported by analytics. Matching skills training to aptitude in an applicant pool is a perfect opportunity for cognitive computing applications.

The Digital On-Ramps (DOR) Project

Leaders in the city of Philadelphia, PA, with a population of approximately 1.5 million, recognized that their residents were falling behind the digital literacy and technical skills demands of industry and recognized that significant action had to be taken. In 2011, it was estimated that by 2030, it could have 600,000 citizens who were unprepared to compete effectively for new jobs. At that point, only 41 percent of homes in Philadelphia were connected to the Internet, but increasing mobile access was not being leveraged effectively to improve training or education. Working with IBM under a grant from its Smarter Cities Challenge program and support from the Clinton Global Initiative, Philadelphia developed Digital On-Ramps (DOR), a learning delivery system that includes in-person and digital education, tailored to anticipated industry needs and individual aptitudes and learning styles. Similar to the electronic health records (EHR) discussed in Chapter 12, DOR creates a universal ID and "digital record depository" for each learner that tracks school, training, and work experiences and accomplishments. Captured as structured and unstructured data in a personal portfolio, it is used to guide the learner toward a job goal with the advice of a counselor. Today, the counseling is done by human practitioners, but the system is capturing data that is useful for evaluating progress using descriptive analytics. This data should soon be useful as input to predictive analytics when the last two stages in the DORS process—building individual skills and matching a skill set to jobs and networks—have generated sufficient data to build a corpus for a cognitive computing solution.

Early results from DOR indicate that creating individualized programs for future industry needs while leveraging a variety of disparate resources, from the free library to local schools and corporate philanthropic programs, can be a successful endeavor. Human will reinforced by analytics has already enabled Philadelphia to secure a MacArthur Foundation grant to develop a certification process to help guide employers when hiring graduates of the DOR programs.

The Cognitive Computing Opportunity

A strong workforce attracts businesses, which strengthens the workforce, creating a virtuous cycle. DOR was launched in Philadelphia, PA, but from the outset it was considered to be a model for other cities struggling with the mismatch between skills and jobs in industry. Integrating cognitive computing functionality into DOR will make it more powerful for those in the program, but also for ongoing city planning and management. For example, adding NLP capabilities for interactions between individuals and the system will lower barriers to participation and provide more tailored responses.

As participants become comfortable with sharing information about personal aspirations and experiences with the system (in natural language). It could learn from experience what works and what doesn't (self-training), and begin to make better recommendations and suggestions for midcourse corrections if it detects changing employment conditions. From reading and interpreting labor statistics, job ads, and editorials in unstructured text from around the country and around the world, such a system could provide ongoing personalized guidance to a population that would otherwise be left to its own devices.

Finally, an integrated cognitive computing environment that shares data across all major city systems could use DOR data to improve transportation planning and operations, utilities demand forecasting, and public health as the impact of new jobs propagates throughout the city and region.

Creating a Cognitive Community Infrastructure

As you have seen, cognitive computing solutions benefit from retaining knowledge created during operation as hypotheses are tested and refined over time. When natural communities—groups of people with similar interests, experiences, or simply geographic proximity—communicate, they raise the collective intelligence of the community. Professional communities that communicate with or through a cognitive computing solution can amplify the learning, as you have seen through medical diagnostics. Physical communities, such as city residents, can likewise benefit from collaboration via cognitive computing.

The Smart + Connected Communities Initiative

In New Songdo City, South Korea, a partnership between the Korean government and private industry is building a new, green, smart city designed from the outset to make extensive use of sensors and analytics. The $35 billion (U.S. dollars) project is one of the largest new city ventures in the world. Cisco, the $48 billion U.S. dollars global networking and communications firm, has wired every home and business in New Songdo City with video screens and systems to facilitate collaboration and a sense of digital community within a physical community expected to have a population of 65,000 by 2016. This Smart + Connected Communities initiative is the first in what Cisco and the development organization hope is dozens of similar projects throughout Asia. Cisco is a leader in developing devices for the IOT and was an early pioneer of the connected world concept. Cisco is promoting New Songdo City as a showcase for the benefits of networking and remote control of residential and commercial energy management and security. These systems will be integrated from the outset, and machine learning tools such Cisco's Cognitive Threat Analytics—which learns and adapts as new threats are identified—will be deployed as New Songdo City nears completion.

The city, set to be operational in 2015 and home of the largest skyscraper in South Korea by 2016, includes an international school for children of visiting business executives, which is connected at all levels with a sister-school in California. Extensive use of video for collaboration and communication is planned to develop a community of learners within the general population. As video analysis techniques mature, that technology will be integrated to continue the automation of learning support.

The extensive use of telepresence and sensor data to deliver more personalized services to every residential and commercial location should foster communication while discouraging energy-intensive travel. This approach fits with New Songdo City's goal of world-class sustainability.

The Cognitive Computing Opportunity

New Songdo City is starting with many advantages over existing cities—before the people arrived, the initial wiring was in place to encourage communication and capture sensor-based, video-based, and system-based information. Beyond community sharing, these systems integrate with transportation, energy, and water management systems, providing an unprecedented opportunity for a city whose population is supported by and bound together by cognitive computing technologies.

The Next Phase of Cognitive Cities

As more of the global population moves into urban areas, the importance of city-oriented cognitive computing will increase. Making better use of the big data generated by the daily activities of individuals, public sector agencies, and businesses will support differentiation for cities as they compete to lure valuable talent and businesses and increase the effectiveness of collaborative efforts for alliances.

For each area, however, you can expect to see an ongoing conflict over private versus public ownership of the data and knowledge, as competition to market intelligent solutions to the densely populated urban residents increases. This is not an issue for cognitive computing per se; rather it reflects the great value that will be created as a result of smarter processing of dumb data at scale.

Most of the examples in this chapter have come from the introduction of analytic and cognitive computing technologies to address straining physical and information infrastructures in established cities. That will continue to be an issue for the foreseeable future, but the desire for urbanization is also driving the creation of new cities to meet the demand. Although planned communities of the last century, from Reston, Virginia, to Celebration, Florida, were built around architectural and open space patterns harvested from the study of centuries of organic city growth, the next generation of planning will be based on communication and collaboration needs, much better suited to cognitive computing analyses than to nostalgic looks at successful town squares.

Kashiwa-No-Ha, Japan, and New Songdo City, Korea, are essentially new cities unencumbered by aging infrastructure and preconceptions. As these two cities mature, and others like them spring up in the next two decades, they will demonstrate the value of integrated analytics and cognitive computing to improve municipal life, and even the most conservative cities will have to adapt or see their relevance erode in the age of cognitive computing.

Summary

Smarter cities have moved beyond collecting data such as vehicle and animal registrations and tax-based data on homes and income. Now, city managers are collecting data through collaboration between departments with optimization in mind. Some leaders have appointed chief data officers to look for opportunities to share data whenever new systems are created. A new public health record database, for example, may be designed to capture elements that can be shared with an education database, or a security database. Most cities are actively making as much data as possible open for free use by the public and by innovative entrepreneurs who can build new services on this data. Every shared use or new commercial use can potentially produce value for residents.

As smart city managers, including elected officials and civil servants, begin to understand the power of data to transform city life, they are embracing advanced analytics and cognitive computing as a source of added value. Smarter cities make better use of all their resources, and having good data is at the heart of these better decisions. Today, the best systems for creating this type of data are the class of emerging cognitive computing applications that learn from their experiences and can guide their users to better decisions and outcomes. In the future we may take for granted the cognitive computing applications that make cities safer and more efficient, while anticipating our needs as residents. Today, the journey is just beginning but the benefits are already clear.