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

CHAPTER

8

The Business Implications of Cognitive Computing

We are clearly going through a major transformation in the technologies that are available to change the way we live and work. With the declining prices of software and hardware and the ability to create new innovations with few capital resources, industries across the globe are changing. So, if the difference between success and failure is no longer simply based on how big a company is, how will we differentiate one supplier from the next? Cognitive computing may be the factor that can add a new dimension to the competitive race. Can we make smarter products and services? Can we anticipate what customers and partners will need in the future? In this chapter we will explore the disruptive power of cognitive computing.

Preparing for Change

Businesses have always had more data in their structured databases, document stores, and packaged business applications than they know what to do with. For decades business leaders understood that if they could capture unique insights from that data before their competitors find them, they could have a competitive weapon. Slowly, businesses are beginning to find ways to integrate data across silos so that they can begin taking a holistic approach to gain insights from data. These leaders understand that if they can extract meaningful relationships or patterns from data about customers, partners, suppliers, employees, and overall

market dynamics, they can turn that information into knowledge so that they can anticipate changes and even shape the future. But even with all the progress that has been made, companies are still grappling with how to capture insights that are not obvious. The problem goes beyond speed—it is a problem of how to discover relationships that are meaningful, not simply anomalies.

The risks of inaction have never been higher. Emerging companies with little revenue are disrupting entire industries and markets overnight and causing established companies to scramble to create new strategies on-the-fly. Bookstore owners have found that electronic book stores have often destroyed their business model. Cab companies are being threatened by new ride sharing models. Manufacturing companies have found that innovative new automated processes and new supply chains have caused them to rethink their cost structures overnight. New regulations in healthcare have required healthcare providers to create new cost-effective processes that meet the right quality of care.

These three scenarios are only the beginning of the tremendous market upheavals that are facing companies across the globe. Solving these problems is not easy, but fundamental shifts won't come from looking in the traditional places. Making it faster to find a cab or shared ride isn't a sustainable differentiator. Making it cheaper by lowering driver pay or auto quality or better routing won't do it—every competitor can copy those moves. However, knowing more about the customer so that you can pair the right driver and rider based on knowledge of their experiences and preferences just may make the difference. While every company hopes to be able to better understand customer preferences and behavior, it is not effective to just ask them. To gain deep understanding requires that organizations be able observe or capture data from external sources (such as comments on social media and data from customer transactions). But true competitive advantages come from discovering what customers value, even before they can articulate what they will pay for.

Advantages of New Disruptive Models

On a pure technology level, the advances in deployment models such as cloud and mobile computing are transformational because they enable new disruptive business models. As the market for these types of services has exploded, the cost of computing and storage has dramatically fallen. In addition, emerging deployment models mean that with little capital investment, new companies can create digital assets and gain a foothold in a market in record time. Therefore, companies can no longer assume that a well established installed base or customer loyalty will sustain them in the long run. The answer to this disruption is taking advantage of knowledge in new ways that can support the new realities.

Cognitive computing, as discussed throughout this book, changes the way that data is processed through advanced analytic algorithms and combined

with structured, unstructured, and semi-structured data. With cognitive computing, you can discover insights that were previously beyond your ability to achieve through computation. Without a cognitive system, a human is required to manually discover patterns and insights that were buried inside complex documents, reports, journal articles, and videos images. In legal discovery, lawyers often send the opposition so much information that it cannot possibly process the data in a reasonable amount of time, even when keywords can be scanned automatically. The deeper meaning remains hidden. Even with an abundance of time, researchers may miss key patterns and nuances that are hidden in documents. In contrast, if you process this volume of data with a cognitive system, you can gain insights that would have required an army of researchers that are smart enough to see subtleties across information sources. While one researcher with decades of experience might know precisely what to look for, the average researcher will likely miss important data.

What Does Knowledge Mean to the Business?

Traditionally, companies have relied heavily on past experience in order to predict the future. The fast rate of change in everything from customer preferences to market dynamics to new technology has rendered many of the traditional methods of business forecasting ineffective. Also, traditional approaches are generally inadequate to respond and adapt to external forces, such as extreme weather events that disrupt supply chains or the sudden popularity of an entertainer whose wardrobe will quickly change demand within a specific demographic. All the cues are available in social media data, but traditional systems are not designed to be able to find this data or exploit it so that action can be taken.

In a search for a better way to translate huge amounts of data across business units into a predictable approach to determining outcomes and next best actions, companies are looking at innovative approaches beyond running reports and analyzing data from traditional databases. These businesses are looking for ways to use all types of data to both analyze and continuously learn from that data. Companies are now comfortable using a variety of algorithms to analyze data. The new frontier is being able to take advantage of the variety of data sources—most notably unstructured text data. While it has long been possible to query a text database, leaders want to build a more dynamic and comprehensive knowledge base that can help them anticipate change and take the appropriate action.

The way to find answers to complex issues is changing. The traditional database query is highly structured and takes its direction from the way the database is designed and managed. A SQL query is therefore highly effective when you know that the database includes the data that you need to find. A SQL query is basically a look-up function. When you look for answers in unstructured data,

you typically use a search engine when you have an idea of what you want to find, but you don't know where to locate the source of that information. The search engine relies on tagging and keywords to find possible answers. When someone asks a question, it matches the words in the query with the tags and keywords in the unstructured database. It provides the user with documents that match those words but does not provide insight. Selecting the document that can provide the best answer is left up to the searcher.

The answer is not simply a database that enables a user to ask a question. This approach has been attempted for decades with only moderate success. In fact, queries are the right approach when there is a well-constructed and highly structured data source that is designed to manage a specific set of questions. Knowledge-based or expert systems, supported by a rules engine, were a step in the right direction as an approach to capture and leverage the experience of statisticians. These applications, however, cannot carry on a dialogue with users to refine their recommendations or suggest new questions that would increase the confidence in an answer. For example, to determine the best approach when treating diabetes, the typical physician relies on experience and perhaps consultation with specialists. The experienced specialist who has treated so many patients with the same illness may be able to determine a successful treatment in seconds. In contrast, the physician with only one year of experience will have to spend hours searching literature or calling on specialists to make a diagnosis. But this approach is not scalable because that physician cannot easily transfer that knowledge to a newly minted doctor. Therefore, there is an inherent risk when the student relies only on the expertise of one physician. For example, the doctor with only 2 years of clinical experience will spend hours poring over books, looking at journal articles, and asking colleagues for their opinion about what to do. The new doctor may use a technique that seems right based on his limited experience. In some domains, there are only a handful of specialists who have seen multiple cases of rare diseases. They can have a positive global influence if their cases are shared with less-experienced physicians.

Sharing of expertise via cognitive systems may be useful in any domain in which specialized knowledge can be pooled, from oncology to auto repair, as long as the relevant knowledge is captured in a corpus.

The Difference with a Cognitive Systems Approach

Contrast this with the organization that begins to use a cognitive computing approach. The organization leveraging a cognitive approach will begin by collecting all the data concerning diabetes including treatment options, literature, and clinical trials. That same organization will also leverage the most experienced physicians and begin to codify their knowledge from their years of experience

and case files. Unlike a traditional application that had to be written based on previously known processes, the cognitive system is designed to learn and change based on capturing best practices and knowledge into a corpus or ontology. At the same time, these organizations want to be able to use data from sensors to monitor metrics and correct for changes in patient status.

For decades organizations have tried to find a way to automate the entire business and have a single integrated system that could manage it all. This has never worked. What has worked for organizations is to design systems that are focused on bringing together knowledge related to a specific business outcome (domain knowledge). Therefore, applications focused on accounting, for example, can provide the right information about all aspects and processes needed to manage accounting effectively because processes are concrete and the knowledge is mature and well understood. This type of categorization of knowledge is operationalized in everything from marketing and sales to human resources, finance, operations, and customer service.

But these systems have been designed based on a von Neumann architectural approach in which the logic and processes are designed in a linear fashion. Each business area is self-contained. Accounting does not interact with data from manufacturing except at a transactional level. Although this siloed approach provides the business with high-quality business intelligence, it can be difficult for business leaders to look at the relationships of this data and see patterns across parts of the company. One significant challenge to this approach is that conflicting data definitions across functional units often contribute to a lack of trust in the data and inconsistencies in business knowledge. In addition, this traditional approach to building business knowledge has many other limitations in today's dynamic and fast-paced global markets. For example, knowledge about the business tends to be historically based and internally focused. Today, companies are increasingly recognizing the need to integrate external and more dynamic sources of information about customer preferences and changing expectations into the traditional business knowledge base.

The previous discussion is based on the premise that the business problems that existing systems were designed to address are well understood and well defined. In many situations, traditional systems were designed based on the way the business operated in the past. Therefore, as the business changed, these systems could not easily be changed or be adapted to new innovative business processes.

Meshing Data Together Differently

Traditionally, the systems of record we have been discussing have been designed primarily to support highly structured data. However, the new data environment includes data from unstructured sources that were never considered

to be part of a company's systems of record. These new dynamic and varied sources of information include unstructured data and streaming data, such as call center notes, social media data, news or stock market data feeds, log files, and spatial data from sensors. These newer sources of data add new dimensions, insights, and answers to some challenging questions. There are connections between structured and unstructured data that were often understood but could never be used. For example, executives knew that customer support systems contained extensive notes about the problems and future requirements of customers. However, no one had the time to manually search through these systems to see if there were any correlation between specific customer issues and a drop in the sale of a specific item in a retail store. Likewise, companies typically would save terabytes of log data coming from sensors on machines. Although they might save that data for years, it remained dark. It was simply too big to analyze.

The need—and ability—to gain business value from expanded varieties and larger volumes of data add a lot of complexity to the definition of business knowledge. The structured information that represents the traditional knowledge base about the operations of the business is most frequently stored and managed in a relational database management system (RDBMS). The siloed approach to maintaining business knowledge means that a company may manage its data in many different RDBMSs. For example, transaction data may be stored in one database, while customer information could be stored in another. Because the newer types of data that add to business knowledge are more likely to be unstructured, these data types need to be managed in a range of different types of data stores such as a Hadoop Distributed File System (HDFS), graph databases, or spatial databases. These platforms provide the ability to add structure to data and provide a distributed technique for analyzing that data in context. Companies are including these new data management capabilities so that they can analyze patterns in all different types of structured and unstructured data. These new sources and varieties of data can improve everything from monitoring manufacturing processes to the detection of diseases. Businesses can improve business planning and execution and predict outcomes. In the insurance industry, for example, executives are using Big Data to figure out what product offerings are the best for a certain customer with the least amount of risk.

Companies want to add new types of data to the business knowledge base so that this knowledge can be leveraged and exploited in various situations. Advances in cognitive computing are helping businesses to analyze more business-relevant information and to plan more accurately. Although knowledge about past business performance remains critical, it is no longer an adequate or acceptable way for a business to remain competitive in the future. Companies want to anticipate the actions their customers might

take. They want to predict and thereby prevent infrastructure failures before they happen. Overall, companies want to find ways to derive business outcomes from analyzing large volumes and diverse types of data that are coming from many different sources. Businesses need to learn from data of all types and use that knowledge to optimize both the operational and customer experience.

Cognitive computing can help companies to redefine business knowledge for their organization. With a cognitive system, companies can move beyond analyzing existing information to making inferences and predictions about the future that will improve with experience. For example, a hospital could use a cognitive system to help reduce re-admissions after patients are released. It is much safer and better for patients' recovery to anticipate if they might be at high risk for readmission and take corrective action to keep those patients out of the hospital. There are a number of factors that could place certain patients at higher risk for re-admission than other patients. Factors ranging from smoking, drug abuse, lack of support at home, or hospital or physician error could have an impact on patients' propensity for readmission. Some of these factors might be known to the physician and recorded in traditional patient records, whereas others may not. In many hospital settings, obvious risk factors are overlooked or are simply unknown. A cognitive system could be designed to analyze past cases and search for patterns of risk factors based on reasons for hospitalization and key patient medical and socio-economic factors. The data would include both structured and unstructured information. At the time of discharge, a patient's record can be compared to the database to determine if he is rated high for re-admission risk. If this is the case, the hospital can implement preventive measures that keep the patient from needing re-admission. Because this is a cognitive system, it is designed to learn and get smarter with each patient case. Aggregated case files may create new value for sharable knowledge. Although the model of risk factors was initially based on historical data, it will be automatically updated and reconfigured each time the system is used.

Using Business Knowledge to Plan for the Future

As technology advances the relationships between sources of knowledge, we can begin to understand the insights that had previously been inaccessible. Shaping outcomes is the heart of the change offered by a cognitive approach to computing. It is useful to evaluate the four stages of maturity in analytics capabilities to understand where we have come from and where we are going.

THE FOUR STAGES OF MATURITY IN ANALYTICS

While businesses are discovering the need for advanced analytics to be able to manage change, organizations do not necessarily all have the ability to leverage data in the same ways. Over time, as analysts gain more expertise they will be able to gain new levels of sophistication. Therefore, we have offered four stages of maturity that companies go through in gaining insights from their data.

Stage 1: Collecting, cleansing, integrating, and reporting on data. During this stage you can use this data to query and analyze current and past business performance. Before you can make sense of data, it is critical to understand what problem your company is trying to solve. What do you want to do with this data and why? When the organization understands the business goals, it is the right time to put a strategy in place. This stage is important in creating a baseline of consistent and trusted knowledge about the business. You can't begin to accurately predict where your company is headed if you don't have a clear understanding of where you have been. Data cleansing and data integration ensure that senior executives understand that reports from business units such as sales, operations, and finance are accurate. Any predictions of future performance based on current and historical information make the implicit assumption that business operations are stable. This approach assumes a Bayesian approach that begins with a hypothesis and then uses a statistical analysis to reach conclusions. Therefore, the rate of change is considered only for the rate of growth or decline in the overall market.

Stage 2: Trend analysis for forecasting. This stage uses basic modeling capabilities to make business forecasts based on an analysis of historical trends. For example, a clothing buyer for a retail chain looks at past sales across the company's stores and forecasts next year's sales by store prior to placing a new order. The model is likely to account for various factors that differ across each of the stores such as climate, store location, and demographic characteristics of shoppers. The buyer may apply what-if analysis to adjust the sales forecast based on changes in selected variables. For example, what if next season has 5 or 10 additional heavy snowfall days? The forecast can be adjusted downward to account for less traffic in the store due to snowstorms. Although creating a forecast for the future based on past performance is a good place to start, these models were not designed to capture and account for change as it is happening. For example, the buyer in this example may end up with a lot of unsold merchandise after overlooking rapid changes in fashion trends among a certain demographic. The outcome from these systems tended to be based on the ability to codify current knowledge and report from those findings. In essence, the results of leveraging these systems were not predictive in nature. Rather, the results are based on a structured and well-defined set of problems.

Stage 3: Predictive analytics. This stage is defined by the use of statistical or data-mining solutions that consist of algorithms and techniques that can be applied to both structured and unstructured data. Multiple sources of both structured and unstructured data types can be used individually or together to build comprehensive models. Some of the statistical techniques used in this phase include decision tree

analysis, linear and logistic regression analysis, data mining, social network analysis, and time series analysis. A key factor in predictive analytics capabilities is having the ability to incorporate predictive models with business rules into the operational decision-making process. This makes the modeling process more actionable and helps businesses to improve outcomes.

The focus of predictive analytics is on anticipating trends before they happen so that you can act to minimize risk for the business. Although predictive analytics has been used for many years by statisticians in certain industries, advances in software tools combined with increasing compute power has made this technology more accessible and more widely used by business users. Predictive analytics models are designed to analyze the relationships among different variables to make predictions about the likelihood that events will take place. For example, an insurance company may build a model that analyzes the components of fraudulent claims and use this model to flag new claims that have a high probability of being fraudulent. Another common use case for predictive modeling is to help call center agents understand the next best action to take when they are interacting with customers. Based on the individual customer's profile, specific product recommendations can be made at the point of interaction between the agents and customer.

Stage 4: Prescriptive and cognitive. Prescriptive and cognitive approaches take predictive analytics to the next level by applying machine learning algorithms, visualization, and natural language processing. Companies want their models to look beyond their internal assumptions about customers and products so that they are better prepared to respond to changing market dynamics.

If models are designed to continuously learn based on each new interaction, the accuracy will improve. For example, a mobile service provider uses analytical models to help customer service agents reduce the level of customer churn. These models analyze information about a particular customer and predict what action the company needs to take to retain this customer. The company's predictive models were updated infrequently and lacking in accuracy and sensitivity to competitive changes in the market. The company significantly improved its customer retention rate by designing a new model that is more prescriptive. The model is designed to be self-learning by feeding each new interaction back into the model, capturing changing market conditions. In addition, the model incorporates social analytics to understand the customer's interactions with and influence on others. These changes improved the model's capability to help drive accurate decision making regarding what the next best action should be to support the customer.

Models that are designed to adapt and change are beginning to be used by companies to predict when a machine is likely to fail so that corrective action can be taken before a catastrophic event occurs. For example, patterns identified in streams of machine data coming from sensors in a train can be used to build models that will anticipate equipment failure before it happens. By using adaptive learning, the model's accuracy can be continuously improved to provide a real-time warning of equipment failure in time for the company to take corrective action.

Answering Business Questions in New Ways

Cognitive computing can also be thought of as a set of enabling technologies that can be applied to a variety of business problems. Many vendors are providing Application Programming Interfaces (APIs) that can add cognitive capability to a variety of initiatives. For example, one emerging vendor called Expect Labs is adding the ability to take conversations and discussions and discover the key concepts and actions buried in spoken language. This same type of approach can be used to discover patterns of similarities across hundreds of pictures of faces to find a particular person.

This type of discovery of meanings inside data will have a dramatic impact on how business is conducted. We are already seeing what happens when businesses can analyze the content of social media conversations. A company can interact with customers on sites such as Twitter or Facebook to intervene when a customer is unhappy. If a business can fix a problem before the customer gets angry, the company can turn a bad situation into a positive engagement.

In addition to discovering patterns, companies need to be able to impart knowledge to employees with limited expertise. A sophisticated practitioner can use a cognitive computing system to input best practices that can be used by less knowledgeable professionals. This sets the foundation so that the new doctor or engineer can gain a faster understanding of most up to date processes. These systems are fed newer data overtime so that the depth of knowledge is expanded and refined. The promise of knowledge management was always difficult to achieve because it assumed that it would be possible to actively capture what experts knew. In contrast, by using a cognitive approach, a system can ingest written information that can be vetted by experts. In addition, this same system can be trained as new information and new best practices emerge. This new dynamic knowledge source can become a competitive differentiator for a business. Imagine, employees with only a few weeks of experience can have immediate access to the right answers at the time of engagement with customers.

Building Business Specific Solutions

In addition to the availability of APIs and cognitive services, an emerging set of applications is being developed. In Chapters 11, 12, and 13 you will find details about how cognitive computing applications are being designed to address data driven solutions in various industries. All of these solutions, whether we are looking at healthcare, metropolitan areas, or security and commerce, have common characteristics. The commonalities include:

- A huge amount of data in many different forms
- Industry-specific data (typically unstructured) that is constantly expanding

- A need to correlate a variety of data sources to determine context, patterns, and anomalies
- The requirement to find a way to match the data with deep expertise
- The need to analyze large amounts of data to support decision making, such as next best action
- The ability to have the systems learn and change as business conditions change

Cognitive systems are changing the way people interact with computing systems to help them find new ways of exploring and answering questions about the business. These systems will learn and interact to provide expert assistance to scientists, engineers, lawyers, and other professionals in a fraction of the time it now takes.

Making Cognitive Computing a Reality

What makes a cognitive computing approach different is that these systems are built to change. The system continues to change based on the ingestion of more data and the capability to identify patterns and linkages between elements. Therefore, companies can look for associations and links between data elements that they might not even know existed beforehand.

The results of creating these types of solutions can be profound. They enable a new level of engagement in which the business leader can have an intuitive interface between the system and the huge volume of data managed in the corpus. Even more important is that these systems are not static. As new data is added, the system learns and determines new ways of understanding situations. For example, new associations may suddenly appear that were not visible or present in the past. Perhaps there is an association between someone who buys books and takes a certain type of vacation. Perhaps there is a relationship between two drugs that can cause a never-before-seen interaction. There may be a new method of treating a serious condition based on a series of new research findings that were published only in the past month in an obscure journal.

The underlying value of a cognitive approach to interacting with technology is that it has the potential to change the way individuals in organizations think about information. How do we ask systems what the data we are seeing means? How can we interact with a system to provide insight when we don't know what direction to take or what question to ask?

It is becoming clear that we have scratched only the surface of the power of information managed in new ways to discover new ways to act and transform organizations.

How a Cognitive Application Can Change a Market

When industries are in transition with new competitive threats, it is impossible to simply build an application. Traditional applications are intended to automate processes and manage data. When a business is trying to transform a traditional industry such as travel or customer care, innovators need sophisticated technologies that allow leaders to discover new techniques and new knowledge. A travel company that can discover what customers want will have a differentiation. What if a travel company can know what the customer will buy even when the customer has no idea? What if a customer service representative can anticipate that the customer's problem is related to a partner's product within minutes rather than hours?

The new generation of solutions will look beyond codified practices and find the answers that are not obvious. Disrupters in every industry throughout the centuries have done precisely this—they have taken traditional approaches to solving problems and turned them upside down.

Summary

Cognitive computing is an emerging area that has the potential to change the way humans interact with machines. Creating a corpus of information that collects massive amounts of structured and unstructured information is game changing. Cognitive computing is not designed to be a back-office function such as existing systems of record that keep track of past transactions and interactions. Rather, cognitive computing is intended to provide businesses with solutions that help them move beyond silos of data to expand how they understand the world. In the coming decade, cognitive computing will transform the machine to a human interface and accelerate how we act to solve problems and facilitate change. Although systems of record and engagement won't disappear, cognitive discovery, support, and training systems will enable new and improved ways to understand and serve customers, and raise the performance levels of experts in virtually every knowledge-based industry.