Transforming the Analytic Landscape

Cognitive Computing: A Competitive Advantage

By Michael L. Gonzales, PhD
Director of Research for Advanced Analytics, Prolifics
August, 2015

Research sponsored by: (Date Cover Style) [Insert Sponsor LOGO]

About the Author

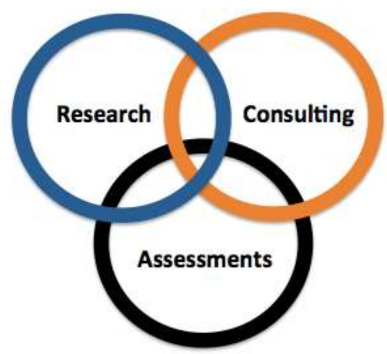
Michael L. Gonzales, Ph.D., is an active practitioner in the IT space with over 30 years of industry experience serving in roles of chief architect and senior solutions strategist. He specializes in the formulation of business analytics for competitive advantage in global organizations. Recent engagements include companies in the top global 100.

Dr. Gonzales holds a Ph.D. from the University of Texas with a concentration in Information and Decision Science. He has presented and published his research at leading IT international conferences, including: Decision Sciences Institute, Americans Conference on Information Systems, and Hawaii International Conference on Systems Science. His research streams include analytics against extremely large data and success factors for IT-enabled competitive advantage.

Dr. Gonzales is a successful author, industry speaker and is currently the Director of Research for Advanced Analytics for a leading IT consulting firm, Prolifics, Inc.

About Eckerson Group

Eckerson Group is a research and consulting firm that helps business and analytics leaders use data and technology to drive better insights and actions. The firm helps companies develop strategies and roadmaps that maximize their investment in data and analytics. Its consultants and researchers each have more than 20 years of experience in the field and are uniquely qualified to help business and technical leaders succeed with business intelligence and analytics, big data management, data governance, performance management, and the internet of things.



About This Series

Executive Summary

Many believe that we've entered a new era of technical progress that forces companies on a path of continuous innovations. Cognitive computing is considered not only part of this new era, but potentially the core technical driver due to its ability to improve human decision-making through automation and augmentation.

This paper will define cognitive computing, describe why cognitive computing is an IT-enabling competitive advantage for companies, and outline the necessary steps that IT must take if they hope to leverage the technology.

A Disruptive Climate

In the early 80's Microsoft announced to the world that computing power was moving from mainframes to desktops and companies that want to be industry leaders must make that move.[1] Apple seems to be willing to continue a long running success of introducing disruptive innovations even at the expense of their own products. The iPhone cannibalized iPod sales and the iPad is eating away at iMac.[2] Google has been buying up companies in seemingly diverse industries: from smart thermostat to robotics to artificial intelligent companies. The objective is to corner the market on skilled resources in order to create the next generation of human-computer interaction.[3]

These are examples of disruptive strategies that companies execute with the goal to reshape or otherwise transform entire industries and the market places in which they compete. As companies redefine the market on their terms, they essentially throw the entire space into a period of disruption. Many analysts concede that technical progress is the catalyst for more innovations, which perpetuates disruption and a fierce competitive business climate.

So what do disruptive strategies have to do with Cognitive Computing? Futurists[4], vendors, and leading companies believe that cognitive technology is a disruptive force in an information-intense age. Put simply: Cognitive Computing enables companies to identify and implement innovative, potentially game-changing products and services.

---

[1] Hagel, John et al, Shaping Strategy in a World of Constant Disruption, HBR, October, 2008.
[2] Fox, Justin, Apple Versus the Strategy Professors, HBR, January, 2013.
[3] Rowinski, Dan, Google's Game of Moneyball In the Age of Artificial Intelligence, January, 2014.
[4] Mayes, Randall, The Future, The Futurist, November-December, 2014.

What is Cognitive Computing?

There has been a growing interest in Cognitive Computing over the past several years. But that does not mean that the techniques used are all new. In fact many have been evolving and maturing for literally decades. These techniques have been the focus of academic researchers and to some degree private sector data scientists. And now they have the attention of C-level executives across the globe. Not because of the cool technology, but because they now recognize the need in this new era of competitive advantage.

Current Definition

While there seems to be a broad interest in evolving the definition of cognitive computing it is best to start with the simplest: Cognitive computing attempts to think like, and interact with, humans; essentially emulating the decision process humans undertake when considering complex problems in a changing landscape of data. The objective is to:

- Augment human cognition by interacting naturally with humans in a process of exploration and discovery
- Automate the decision making process; thus, eliminating human intervention
- Apply self-learning to support continuous improvement

Cognitive computing complements and supplements the capacity of humans to make sense of highly complex data, expose insights, consider alternative means for addressing problems, and arriving at optimum solutions. It can be argued that artificial intelligence meets business intelligence…literally.

Unlike analytics that deals with deterministic models, Cognitive computing addresses problems characterized by ambiguity and uncertainty. In order to achieve this level of computing, these systems must be:

- Adaptive – they must learn as information changes and as goals and requirements evolve. They must feed on dynamic data in real-time or near real-time.
- Interactive – they must interact with users as well as other processes, devices, and services.
- Iterative – they must aid in defining a problem by asking questions or finding additional inputs if the problem is ambiguous or incomplete as well as “remember” previous interaction.
- Contextual – they must understand, identify, and extract contextual elements such as meaning, syntax, time, location, etc. And draw upon structured and unstructured data.

Better Decision-Making Process

The human cognition science community is one group that continues to study and dissect the decision-making process. What we present in this section is one aspect of this research that makes a strong connection between the application of CC and the value of making more effective and efficient decisions. To that end, presented in Figure 1 is a decision ladder with widely accepted steps taken for decision making.

Figure 1. Decision Ladder

When you consider all 8 steps humans commonly take to make a decision, traditional Business Intelligence (BI), Data Warehousing (DW), and even predictive analytics provide little support in the decision steps. Operational systems capture the events (Step 1) and BI/DW environments present the information (Step 2 and Step 3). Predictive analysis based on deterministic data provides some additional information (Step 4), but a human (user, analyst, etc.) really carries the burden for decision making from Step 4 through Step 8. Although this human and computer system balance may be effective in some cases, it doesn't lend itself to the modern challenge we've discussed earlier: Growing Big Data and Shorter Decision Cycles.

On the other hand, consider the scenario where a computer system can make a decision quickly if it recognizes the event (Step 3). If so, it can simply jump from Step 3 to Step 6 (shown with the red arrow) because the system knows what task must be executed for recognized events and goes straight to its planning and execution tasks. Essentially, the computer system can eliminate the need for human intervention. This type of scenario is perfect for automating.

Alternatively, all decision processes slow down when the event is not recognizable (Step 3). That's when more information is needed to better understand the consequences of the event and then select and compare options to address the situation, and then finally select a single option and execute. Steps 5 and 6 require significant human participation. These unrecognizable events are unique but are common as Big Data becomes larger and more complex. While this situation is impossible to automate, it does represent a perfect opportunity for cognitive techniques to augment the human (user, analyst, executive, etc.) decision capability.

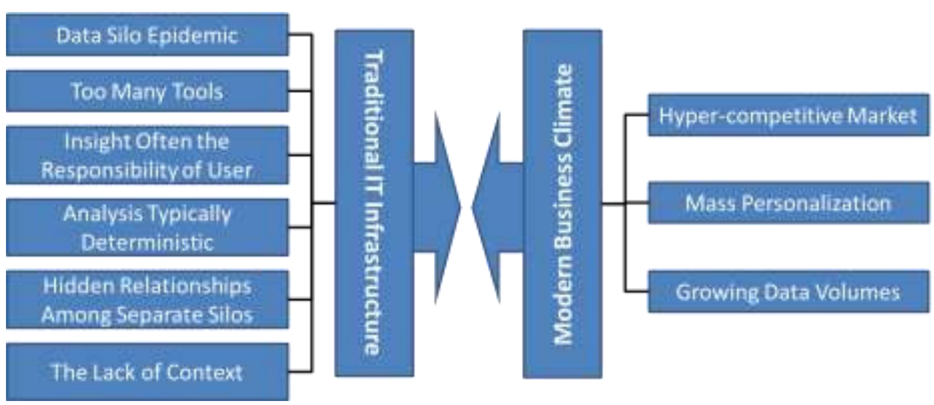
When considering the scope of cognitive computing, a great place to start is to decide if your use-case is a unique event (unrecognizable) or a known event.

Why Now?

If it's not new, why is there such an elevated interest in cognitive computing now? The advent of Big Data is a primary driver as the world goes digital. From the Internet of Things to social media to smart phones, the activity generated is dynamic, uncertain, and voluminous. Videos, music, pictures, and natural language are data that stresses the infrastructures of most organizations, and more so the capability to consume and digest the information, identifying insights, and formulating responses.

We now face unprecedented data volumes coupled with the market competition for mass personalization, access, and availability. Organizations can no longer simply throw traditional technologies and techniques at the modern problems and expect to succeed. Refer to Figure 2. Traditional computing techniques are reaching the limits of what can be done to make all this data useful for the average knowledge worker. Cognitive techniques, however, are uniquely qualified to effectively leverage Big Data and implement the innovative applications necessary to compete.

Figure 2. Converging Conditions



And here resides the connection between disruptive strategies, Big Data, and the value of cognitive computing. It is a circular, self-perpetuating relationship: As more disruptive strategies are executed, more Big Data is created, and the more business has a need for cognitive techniques.

Extending the Analytic Ecosystem

Detailed discussions about shaping strategies are beyond the scope of this paper, however, we will detail the steps that IT must undertake to ensure their organization can embrace the innovation-enabling capability of cognitive computing and the effective consumption of Big Data. Consider the objectives discussed earlier:

- Augment human cognition by interacting naturally with humans in a process of exploration and discovery
- Automate the decision making process; thus, eliminating human intervention
- Self-learn with usage

The overall application value of cognitive computing is in its ability to augment the human decision-making process by interacting with humans in an exploratory and discovery function or by automating the decision process and thus eliminating the need for human intervention. Self-learning is a constant between either augmentation or automation of decision-making. Consequently, companies must understand the decision-making process in order to better assess the value and where and how to implement cognitive applications.

A Rich History of Evolution

The roots of cognitive computing date back to the late 19th century when researchers like Charles Babbage referred to the analytical engine. We fast forward to the 1950's when John McCarthy coined the term, Artificial Intelligence, as the science and engineering of making intelligent machines[5], and Alan Turing, Arthur Samuel, and Tom Mitchell's work on machine learning. It is these two broad disciplines, AI and Machine Learning, on which modern cognitive computing is based.

It represents the culmination of more than 60 years of computer science research and commercialization. It is an umbrella term that encompasses many new as well as established techniques, technologies, and concepts. For example, new hardware like the neurosynaptic chip introduced in 2014. While other techniques like neural networks were initially introduced in the 40's and have a relatively long history of success as learning algorithms that adapt to changes in data.[6]

The broad range of cognitive techniques drives the scope of opportunities and related applications it spawns, everything from taking and understanding images to understanding written or spoken language to emulating human decision-making processes. Six features/functionalities used to compare and contrast traditional and cognitive analytics are outlined in Table 1.

---

[5] John McCarthy, Stanford University, 1955, revised in 2007.
[6] Warren McCulloch and Walter Pitts, 1943.

Table 1. Toward Cognitive Computing[7]

Traditional Analytics	Cognitive Analytics
Search	Discovery
Deterministic	Probabilistic
Enterprise Data	Big Data
Machine Language	Natural Language
Simple Outputs	Intelligent Options
Right Answer	Best Solution

For most of our computing history we have provided functionality to search data sets, build enterprise (mostly structured) databases, interface with computers using machine language (computer programming), and provide simple output for standard reports, ad hoc analysis, and dashboards. Ultimately traditional computing sought the right answer for a deterministic (non-random) problem. But traditional methodologies simply are not sufficient in today's business climate.

Cognitive Computing is dramatically different across these six features. It does not focus on simply searching data sets based on enterprise, structured data, but attempts to uncover, discover, and otherwise expose insight and information contained within huge, structured and unstructured data sets (Big Data). And when you interact with a cognitive system, the objective is to do so using natural language as opposed to computer languages. This functionality supports one of the key objectives which is to enable the subject matter expert to interact with a cognitive application without the need for technical assistance either from analysts or IT.

A final differentiation is the outcomes. While traditional systems work with deterministic (non-random) data and algorithms in an attempt to arrive at the "right" answer, cognitive environments are best used against ambiguous, complex problems and data with the goal of providing the most intelligent options and the best possible solutions. For example, linear regression is an extremely popular predictive algorithm. But like many traditional analytic techniques it is built using historical data where the answer we want to predict is already known. This is described as a deterministic model. Comparatively, one technique of cognitive analytics is fed a broad array of data for which there is no single, right answer sought. Instead, the model learns from the incoming data and is prepared to attempt answering questions posed. Questions are given a range of potential responses, and the user is asked to rate the responses based on relevance and accuracy.

[7] Cognitive Computing and Watson, Data Science Day, May 2014.

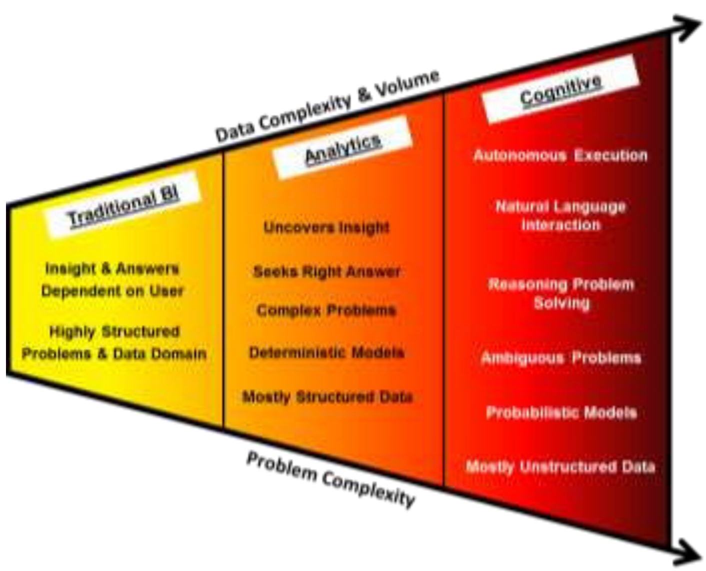
Cognitive Computing Compliments Not Supplants

An important point that executives must understand is that cognitive systems are not intended to supplant the existing analytic infrastructure. Cognitive tools should be viewed as a compliment to your current systems and therefore, extend your analytic landscape. For example, you don't invest in cognitive computing in order to supplant standard reporting and dashboards, or predictive analysis based on non-random data (deterministic), or even exploratory and what-if analysis based on the structured data stored in data warehouses and data marts. It is likely you have the necessary infrastructure to perform these tasks. You invest in cognitive systems when the challenges exceed these traditional systems; specifically:

1. data is far too large and the complexity is broad and random
2. problem being analyzed is ambiguous and complex
3. decision process is executed against volumes of repeatable events that can be automated for efficiency

Some argue that analytics and cognitive computing can be viewed as extensions to BI. But that definition over-estimates the analytic role and value of BI while over-simplifying the significance of the other two domains. Moreover, analytics and cognitive techniques are not synonymous. The way to view all three is as a continuum of analysis as shown in Figure 3.

Figure 3. Continuum of Analysis



For brevity, a single example from each domain will be used to illustrate two points: 1) the significant areas of difference between the domains, and 2) the role of each domain within the broader context of an analytic continuum.

Traditional BI
Consider traditional BI using a popular technique, Online Analytic Processing (OLAP). It requires a highly structured data domain either in the form of a star schema or multidimensional cube (MOLAP) or both. And the questions it supports are exclusive to the data domain built. If I want to answer sales questions by clients, products, and dates, then the star schema/cube will need to have those 3 dimensions (client, product, and date) as well as a sales fact table. A highly structured data domain designed to answer highly structured questions. Moreover, it is incumbent on the user to find any insight in the data domain built. OLAP, in and of itself, exposes no insight. It merely provides the technique, specifically statistical aggregation (average, sum, min, max, count), and the related data. It is the responsibility of the user to ask the right question(s) to get information sought and, hopefully, to expose insight. This dependency between OLAP and the user is why we historically wanted a subject matter expert to use the OLAP application. Only then could an organization maximize its chance of getting value from the investment.

Analytics
Now let's examine analytics. While there are many dimensions to analytics we will use data mining to discuss the topic since it is likely the most widely familiar to readers. Key characteristics of analytics are outlined in Figure 5, *uncovers insight* being the first. This is a significant distinction between BI and analytics. Data mining techniques and technologies can literally expose previously unknown insight and report its findings to a user. For example, association algorithms, such as Apriori, offer automated, unguided capability to data discovery and return with association rules (things done together) of significance. Market basket analysis is a common application based on these algorithms. We've all read about products commonly purchased together, e.g. wine and cheese or milk and eggs. It is impossible to have a user read through 100 million point-of-sale tickets to determine what products are commonly purchased together. Automated, unguided data discovery is a real benefit and value of analysis.

Cognitive
For our discussion regarding the cognitive domain, we use an application that epitomizes many of the characteristics listed in Figure 5; specifically, Siri, the iPhone virtual assistant. Siri is the offspring of one of the largest AI projects funded by the Department of Defense in an effort to create a virtual assistant that could reason and learn.[8] When acquired by Apple and launched as part of the iPhone, Siri has not been able to live up to expectations. Nevertheless, it represents an era of cognitive computing designed to assist humans with deciphering the complexity of data and even offering potential solutions.

Siri's core strength as a user-facing interface is the ability to translate and understand a user query. This represents a leap forward when compared to BI or analytics. And not just natural language processing, but in terms of being able to execute on demand, self-directed, and self-learning (autonomous execution), dealing with ambiguous problems/questions, and offering up, albeit sometimes limited, options (reasoning). Many argue that Siri's primary limitation is the lack of sufficient sources of data. But what if a cognitive application similar but more robust than Siri had all the data it needed for a specific domain? Watson, a familiar brand to many, is now dispensing advice and guidance to doctors treating

[8] Bosker, Bianca, Siri Rising, Huffington Post, Jan., 2013.

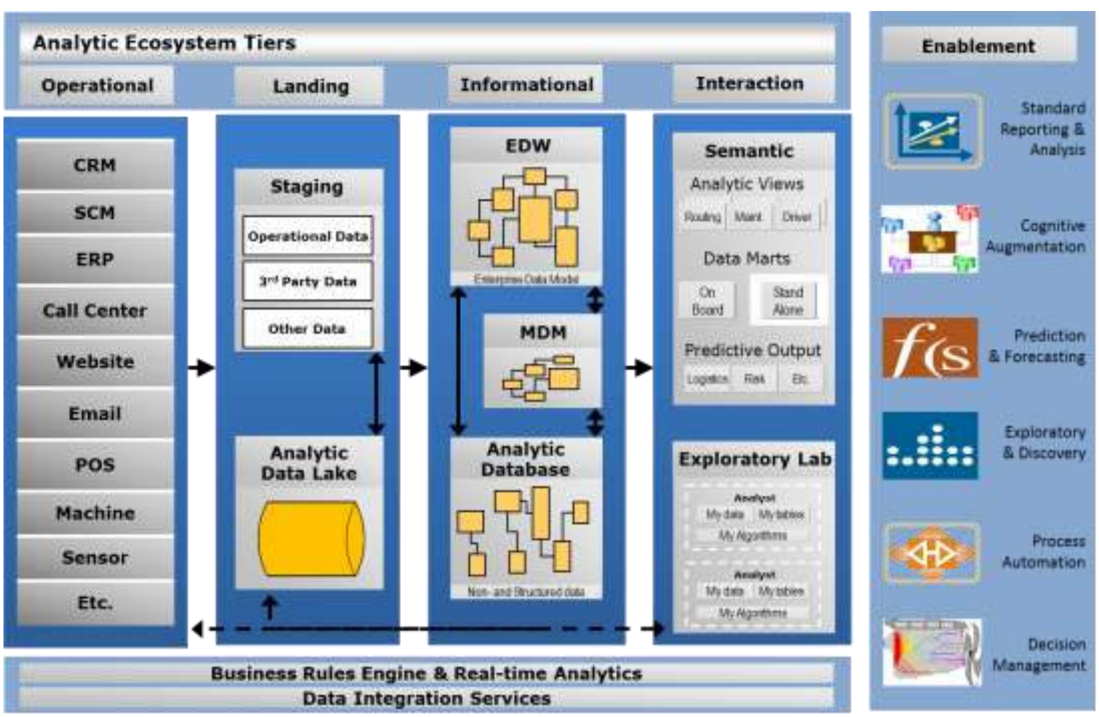
patients of New York's Memorial Sloan-Kettering Cancer Center.[9] Or Google's Now virtual assistant which has access to your Gmail account, location history, and additional information provided by Android devices.[10] Now is launched as providing the right information at the right time with information sources from more than 100 provides such as Fandango, Kayak, Spotify, and Pandora just to name a few.

Modern Infrastructure Landscape

The high-level reference architecture in Figure 4 demonstrates a rich blend of technologies, techniques, and data architecture: from the traditional, warehouse-centric structures to advanced technologies and techniques, such as a data lake and exploratory labs. When you combine the proven, traditional design with leading and innovative technologies and techniques, you establish a robust foundation for a broad array of analytics. The framework is designed to offer a diverse landscape of analytic capability, including Cognitive Analysis. This is a strategic distinction and, consequently, a competitive advantage.

All the best-practices for modern information management and analytics were considered in a cohesive, integrated, and synchronized framework. By design, the architecture represents best-practices for analytics, including: real-time analytics and process automation, a Data Lake for Big Data, an Analytic Database for deep and scalable analytics, and an Exploratory Lab for advanced exploration and discovery leveraging a rich set of algorithms. There is, of course, the need for traditional information management as well as illustrated.

Figure 4. Analytic Ecosystem Tiers



[9, 10] Byrne, Richard, When Watson met Siri, VB News, July 16, 2014.

The reference architecture represents a logical environment that can be implemented in various ways. On one end of the spectrum is a best-of-breed approach which physically implements several different technologies to meet the requirements outlined in Figure 4. This can include in-house or cloud-based platforms or a combination of both. The challenge with implementing best-of-breed has always been the level of integration between the technologies and the skill necessary to implement and support a broad variety of tools. This is exacerbated when adding a cognitive computing layer to an already complex ecosystem. At the other end of the spectrum are purpose built appliances. An organization can choose to leverage a single appliance specifically designed to support a broad array of integrated analytic data management and analysis, such as Teradata's Aster technology.

Irrespective of how you implement the analytic ecosystem it must be designed to enable an organization to perform everything from standard reporting to cognitive augmentation to process automation. For brevity, each capability is defined in Table 2.

Table 2. Enablement

Enablement	Questions Answered
Standard Reporting & Analysis	What happened?
Cognitive Augmentation	What do you think is the best option?
Prediction & Forecasting	What will happen?
Exploratory & Discovery	What is happening?
Process Automation & Decision Management	What action must be taken right now?

Conclusion

The overall objective of this paper is to draw the following links:

- From the fierce competitive business climate to the need for innovation
- From the need for innovation to Big Data
- From Big Data to the innovation-enabling technologies and techniques of Cognitive Computing

As innovation drives more digital data, the need to consume and expose opportunities and risks within Big Data is critical. No traditional technology can address this issue effectively. Cognitive Computing is not only the tool for spotting opportunities in Big Data, but it is also the foundation of innovative products and services for many organizations.

Implementing a modern ecosystem is not without its challenges. First and foremost is the level of expertise that will be required. Many companies already struggle with basic big data technologies such as Hadoop and advanced analytics like data mining and statistics. Adding another layer of expertise necessary for leveraging cognitive computing may constrain its adoption. Nevertheless, the momentum and direction of a disruptive business climate is clear and organizations have no other choice but to address the implementation challenges of the ecosystem if they want to compete.

There seems to be a heightened level of marketing buzz surrounding this technology but many dimensions of cognitive computing are mature and have been evolving over decades. For instance, self-learning algorithms like neural networks or technology for natural language processing have been the focus of research and development for years and have found broad commercial application and acceptance. Whether the dimensions are mature or recently advanced, there are two key benefits offered to support better decision-making: Automation and Augmentation.

Whether your organization chooses to implement the latest dimensions of cognitive computing or rely on one or more of the mature components, the key to implementing this technology is to recognize that it compliments, supplements existing IT infrastructure and the applications it supports. Investments in BI, DW, or predictive and exploratory analytics remains intact, with specific roles, use cases, and user communities they support. Cognitive computing extends the analytic landscape, broadening the type of applications and use cases your company can address.