COGNITIVE COMPUTING: AN EMERGING HUB IN IT ECOSYSTEMS

DATA MANAGEMENT'S NEW IMPERATIVE

BY STEVE ARDIRE AND CHARLES ROE

Cognitive Scale

DATAVERSITY™

CONTENTS

1. EXECUTIVE SUMMARY

Will the "programmable era" of computers be replaced by Cognitive Computing systems which can learn from interactions and reason through dynamic experience just like humans?

With rapidly increasing volumes of Big Data, there is a compelling need for smarter machines to organize data faster, make better sense of it, discover insights, then learn, adapt, and improve over time without direct programming.

Cognitive Systems learn, adapt, hypothesize, and recommend in real time, but Cognitive Computing is not about replacing humans with machines. It is about harnessing the combined strengths of man and machine as partners solving complex problems that adapt to ever-changing factors and new information.

This report focuses on the growth of Cognitive Computing technologies as a key emerging hub within the Data Management industry. It will discuss the underpinnings of the recent groundswell of interest in Cognitive Technologies; the development of Cognitive Analytics, Machine Learning, Deep Learning, and Artificial Intelligence; and the latest developments in the field, providing an overall view of what Cognitive Computing means to a variety of enterprises.

The foundation of the report is an analysis of a DATAVERSITY™ survey on Cognitive Computing conducted in May and June of 2014. Some highlights of the report include:

- 53.4% of respondents believe that Cognitive System technologies need to provide more clarity in terms of business perspectives.
- 16.7% of respondents either are not aware of technologies like IBM's Watson, Siri, and Google Now or don't find them applicable to a discussion of Cognitive Computing.

- A majority of respondents had their own definitions of Cognitive Computing and added their comments to the definition we presented within the paper.
- Some of the most needed resources for eventual enterprise integration of Cognitive Systems include better education of benefits, more case studies, easier to use tool sets, and vendor demonstrations.
- More than a third of respondents said that they were still unclear about their organization's plans for Cognitive Systems implementation due to a lack of understanding about how to present the business case.
- 26% remarked that their organizations are early adopters of emerging technologies such as Cognitive Computing, NoSQL, and Big Data.
- Almost 60% said that one of the primary drawbacks to current integration with Cognitive Systems is the lack of knowledge and skills among existing IT staff, especially in terms of Data Scientists and Machine Learning experts.
- Business Intelligence/Cognitive Analytics was the top choice (81.8%) for how Cognitive Computing can help the enterprise.

2. RESEARCH AND DEMOGRAPHICS

The purpose of the 2014 online DATAVERSITY survey on Cognitive Computing (CC) was to gain a more comprehensive perspective on the topic from a wide range of Data Management industry practitioners. The findings are an essential aspect of this report. Instead of collecting disparate research on CC into a grand overview of the field, the report has allowed DATAVERSITY to gain firsthand insight into where CC currently exists within the enterprise and within the imaginations of those who work in the data industry.

What does Cognitive Computing mean to industry experts? How much development is happening at this time? How does it fit into current models of Enterprise Data? What sort of plans are there for implementing Cognitive Technologies? What tools and resources are needed? What are the possible future impacts of such technologies? These and other questions were asked of over 152 participants.

The survey contained seventeen questions broken into a number of subdivisions. Some questions included an open-ended format to allow respondents to give more comprehensive answers:

- Demographics (four questions)
- Familiarity and definition of Cognitive Computing (two questions)
- Tools and resources needed (two questions)
- Data Acquisition/Modeling/ Architecture (two questions)
- Cognitive Computing implementation (three questions)
- What CC can do to improve current methods/business practices (two questions)
- Future impacts (two questions)
- Nine of the seventeen questions had an open-ended comment choice available, and of those two of the questions were only

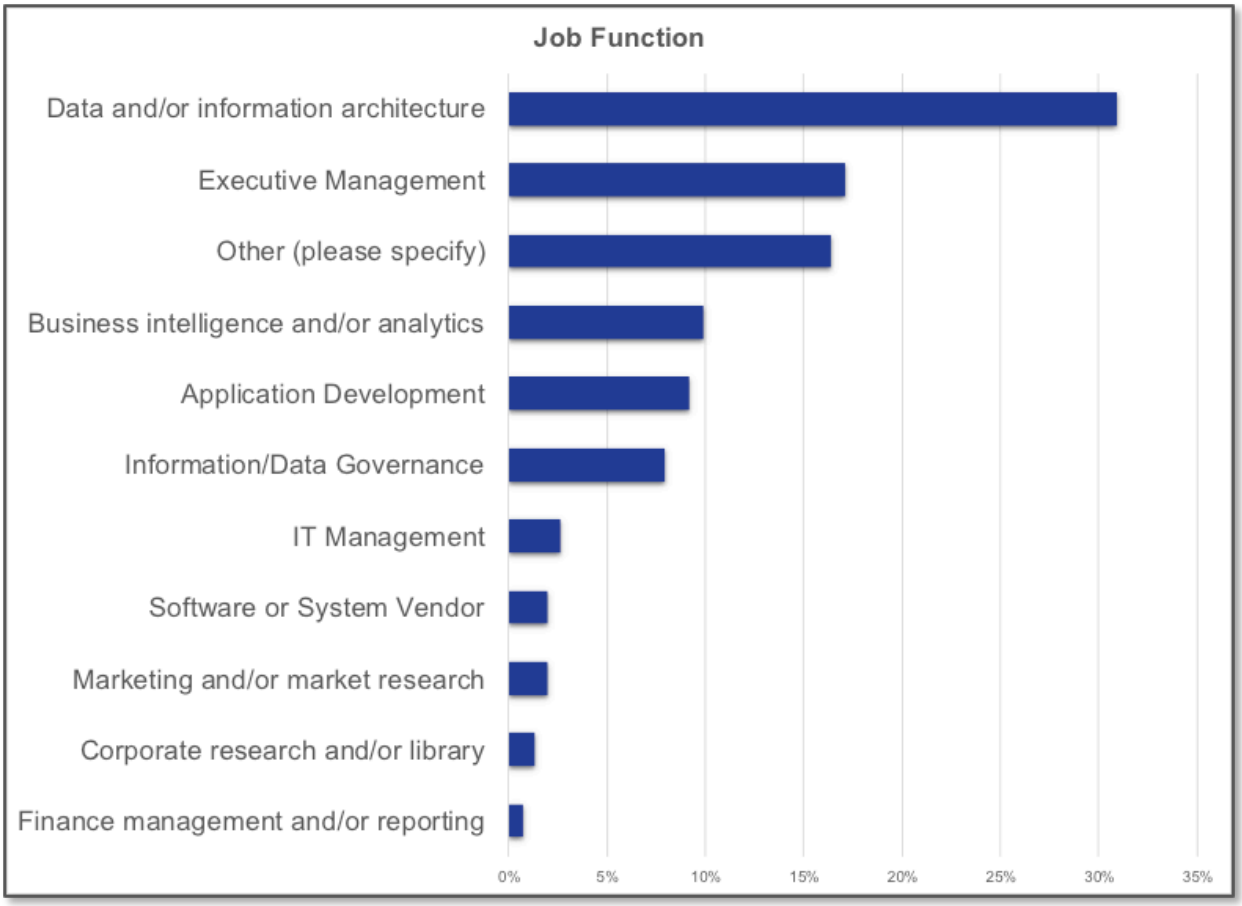


Figure 1

open-ended format. The first demographics question asked about job function [Figure 1]. The top three answers were:

- Data and/or Information Architecture: 30.9%
- Executive Management: 17.1%
- Business Intelligence and/or Analytics: 9.9%

Application Development (9.2%) and Information/Data Governance (7.9%) were also near the top of the list. There was also an Other choice (16.4%) which included answers relating to education, aerospace, hardware development, consultancy, and research.

The second demographics question asked about the respondent's industry [Figure 2]. The answers provided a comprehensive snapshot of the industry and included a wide range of responses. The top five were:

- Technology: 21.2%
- Consulting: 17.2%
- Insurance: 8.6%
- Education: 7.3%
- Government: 6.6%

The final demographics question focused on

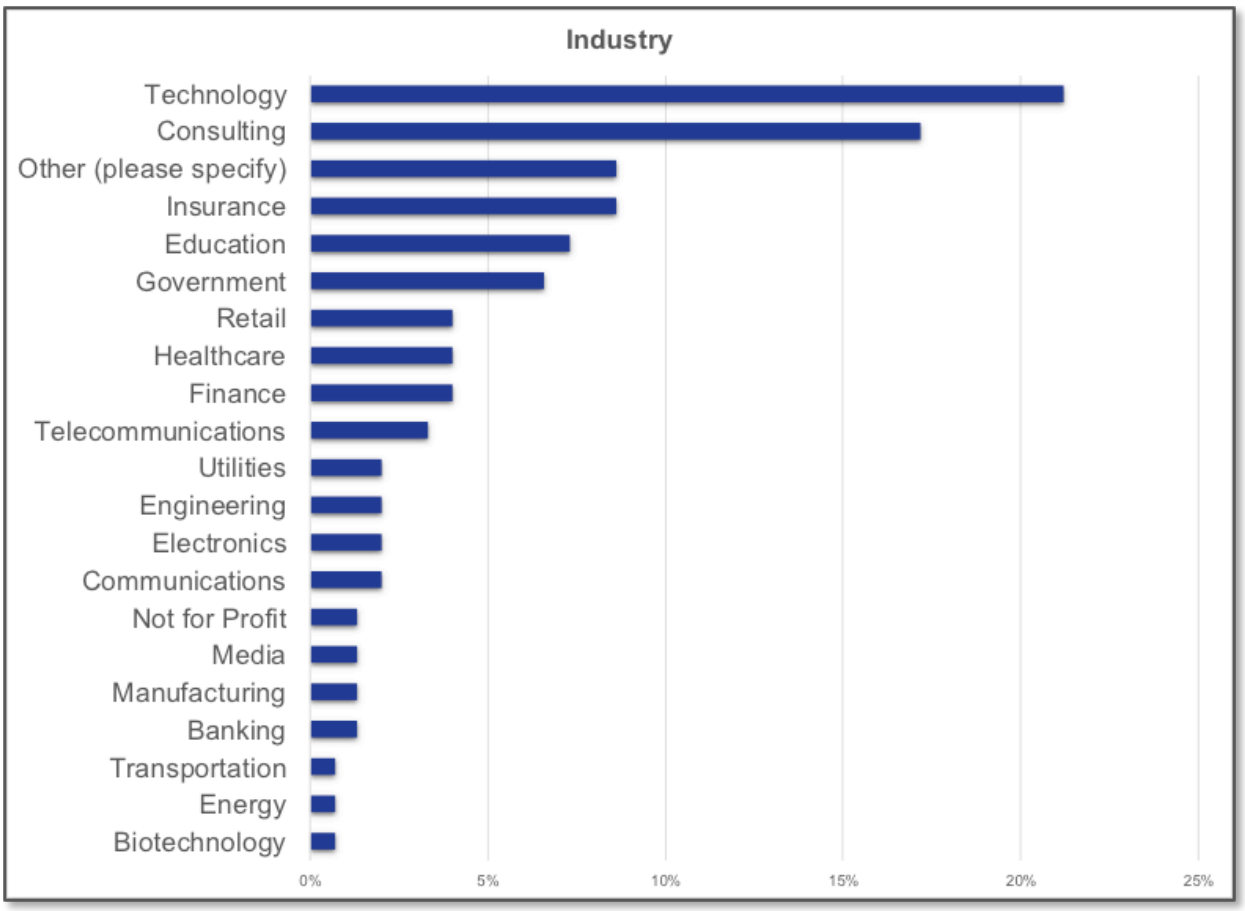


Figure 2

company size. It asked about the number of employees in the respondent's organization. The answers were comparatively close for all sizes, with the exception of 5,001 – 10,000 at only 4.6%. The top three answers were:

- Less than 10: 20.5%
- 1,001 – 5,000: 19.9%
- 11 – 100: 15.9%

The survey demonstrated a strong and varied sampling of multiple industries, with more than 30 involved. It also had a good range of company sizes and position types, so that all-in-all it provides a good snapshot of the Data Management industry as a whole.

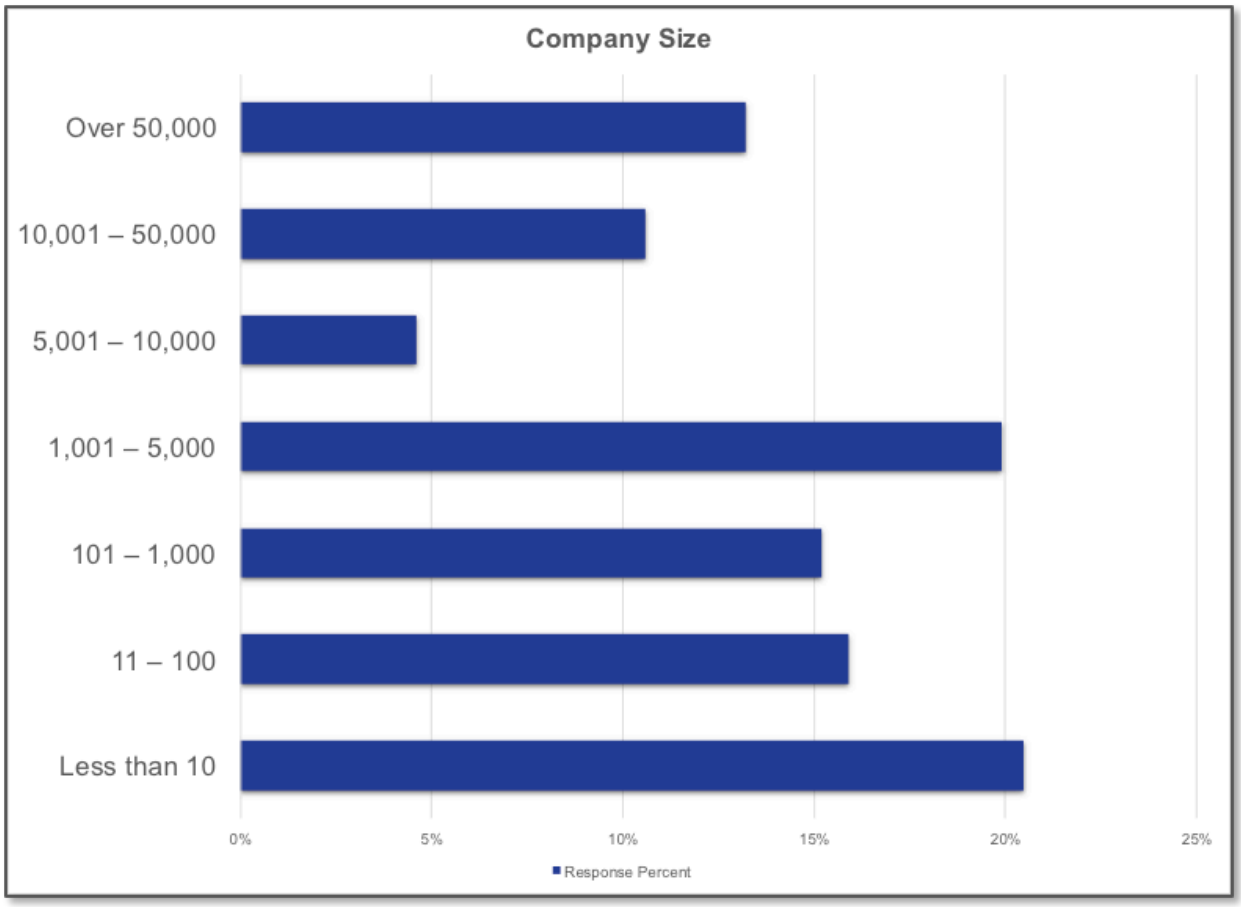

Company Size
Over 50,000
10,001 – 50,000
5,001 – 10,000
1,001 – 5,000
101 – 1,000
11 – 100
Less than 10
0%
5%
10%
15%
20%
25%
Response Percent

Figure 3

3. INTRODUCTION – NAVIGATING A BRAVE NEW WORLD

The interest in Cognitive Computing within various media channels is clear – it is one of the newest buzz words being parsed around the Web. Thought leaders and trailblazers within Data Management and related fields are writing about it, debating it, beginning to hold conferences on it, and developing technologies for it. Some think it is only a re-couched term for Artificial Intelligence (AI). Some decry it as a newly-packaged big player rollout of cobbled together components, all expertly branded with fancy modules like behavioral profiling, marketing automation, contextual navigation, Natural Language Processing, and Predictive Analytics – all meant to empty the bank accounts of data-driven enterprises.

The roots of CC reach back more than 40 years to the days when AI was first being developed in the laboratories of places like Stanford University, the Dartmouth Conference of 1956, and in the imaginations and research of such people as Alan Turing, Norbert Weiner, Marvin Minsky, Alan Newell, Edward Feigenbaum, and numerous others. The development of AI has held the imaginations of people worldwide for generations through the great works of science fiction by Isaac Asimov and Arthur C. Clarke, among others. The more recent work of IBM with precursors like Big Blue and the now oft-mentioned Watson systems have made big news. Humans have long dreamed of "thinking machines" to help simplify our lives, do the most tedious work for us, and bring us insights that our brains cannot deal with – whether due to time or sheer computational requirements.

Until recently, such dreams were unrealized. Now, thanks to decade's long research revolutionary advancements in computing power and parallel processing, deeper understandings of human learning behavior and neurobiology, and the development of more sophisticated algorithms, we stand at the threshold of a new era of human-computer interaction. As Rajeev

Ronanki and David Steier of Deloitte consulting wrote:

> "For the first time in computing history, it's possible for machines to learn from experience and penetrate the complexity of data to identify associations. The field is called cognitive analytics – inspired by how the human brain processes information, draws conclusions, and codifies instincts and experience into learning.
>
> Instead of depending on predefined rules and structured queries to uncover answers, cognitive analytics relies on technology systems to generate hypotheses, drawing from a wide variety of potentially relevant information and connections. Unlike in traditional analysis, the more data fed to a Machine Learning system, the more it can learn, resulting in higher-quality insights."[1]

Cognitive Computing is still in its infancy, yet it has already demonstrated a level of sophistication in regards to its ability to assimilate and analyze heretofore incongruous data streams. As more players get involved in its development, as more resources and tools are created to aid in its application, and as more enterprises begin to understand its considerable value in dealing with Big Data, Smart Data, and Unstructured Data (as well as traditional structured data assets), its impact will become ever more evident.

Investments in further Cognitive Computing R&D are pouring in. Venture capitalists are putting their money into numerous startups within the AI-Cognitive space, developing their own projects, and pushing the envelope ever farther with new technologies and new ideas. A short list (there are countless others not listed here) of recent CC investments and research includes:

- IBM recently announced a $1 billion investment in multiple Watson technology ecosystem initiatives from analytics to Cloud development, medical resources to education advancement.
- The healthcare website Modernizing Medicine, an online repository of medical data, now has more than 3,700 providers and 14 million patient visits in its cognitive data system, giving medical professionals the opportunity to gain deeper insights into possible avenues of patient care.
- The Chinese search-engine titan Baidu plans to invest $300 million over the next few years on Deep Learning and Big Data research. The company also hired Andrew Ng, an Artificial Intelligence expert and former head of Google's Deep Learning project.
- A prominent Hong Kong venture capital firm just named an AI tool known as VITAL to its Board of Directors with the goal of finding better investments through more innovative decision making.

- Alex Zunger, chief theorist at the University of Colorado's Center for Inverse Design, is pioneering a Cognitive System that allows materials scientists to use "inverse design" to identify better materials for manufacturing and other sectors. The system pinpoints the desired properties of a material and then ascertains the necessary atomic structure for it, rather than the other way around.
- Through the innovative algorithmic models of Jim Gao, a young Google engineer, Google is using Cognitive Computing to significantly improve the efficiency of its data centers.
- Intel recently paid $30 million for the personal assistant platform Ginger from Israel's Ginger Software. The purchase includes Natural Language Processing (NLP) tools and application assets, along with a number of Ginger's engineers.
- MIT researchers are working on a Cognitive System algorithm that is learning how to understand what is happening in videos with the hope of eventually tagging and indexing vast online video collections. The startup Emerald Logic has created a FACET (Fast Collection Evolution Technology) that tests tens of thousands of algorithms for a particular data set to discover the most predictive and thus valuable ones.
- Researchers at the Allen Institute and University of Washington have been developing LEVAN (Learn Everything about Anything), a Cognitive System that can teach itself essentially "everything it needs to know" by examining search engines through Natural Language Processing and Machine Learning techniques.
- Engineers at Sandia Laboratories are currently working on what they call a neuromorphic computer system that mimics the human brain both in terms of parallel processing power and efficiency.

A CLARIFICATION OF TERMINOLOGY

Before delving into a deeper analysis of Cognitive Computing, a shared definition should be established. Our current working definition of Cognitive Computing is:

"Natural Language Processing of structured, unstructured, streaming-in Big Data, or Smart Data layers with Machine Learning for reasoning and learning to generate textual patterns and associations that enable humans to connect the dots faster and smarter for more informed decisions to drive better outcomes."

The survey presented this definition to respondents in one of the earliest questions and asked for their reaction. The question [Figure 4] contained three possible answers, along with an open-ended question for further elaboration. The top answer was "Yes, this is very good" at 48.6%, with 15.5% responding "don't know/not applicable." [2]

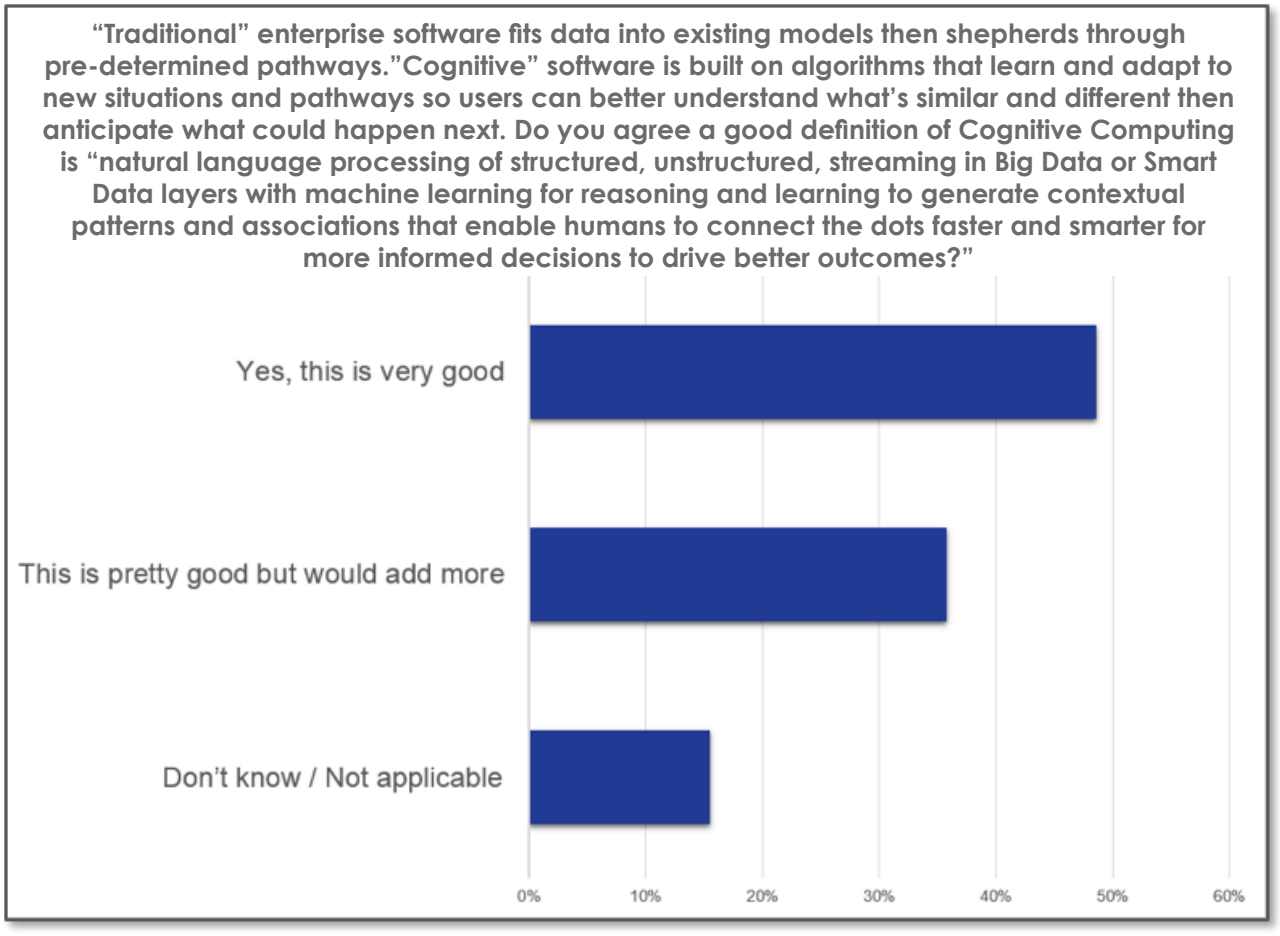


Figure 4

37% of respondents gave comments. Some of the most poignant were:

- "It's a bit long, but perhaps required to cover the entire scope. I don't think Cognitive Computing needs to just connect the dots for humans, but also for intelligent software agents that could take some actions on behalf of their human users."
- "Not sure I agree with attempt to re-brand AI."
- "Cognitive Computing should investigate and understand human cognition and cognitive process in the information usage. The process mentioned in the definition may be usable for the purpose, but it is not the definition of the Cognitive Computing itself in my perspective."
- "Use of NLP, graphs, and Machine Learning to aid search and inference problems."
- "Very good (more than pretty good), but shouldn't this also include something about intelligent, guided search?"
- "I wouldn't say it is just Natural Language Processing – for instance I want to be able to make sense of user interactions within applications, and derive intelligence from this. This aspect doesn't have much to do with Natural Language Processing."
- "Why just language? Why not include the cognitive processing of images, videos, gestures, maps, drawings, graffiti, smell etc? They do say that non-verbal cues are very important."
- "The algorithms are predefined, bounded, and biased by those who designed them. As a result they lack out of the box thinking. The systems don't 'think' and it is pretentious to call this cognitive. It is algorithmic computing."
- "Cognitive Computing applies to the tuning and automatic configuration of executing algorithms. That is, it applies as an aid to software systems themselves, not just humans."
- "Besides providing more meaningful information, Cognitive Computing can also help in building smart/automated self-learning devices."
- "We use 'cognitive' term in simulation for completely different reason & not related to natural language - it's for logic/models based on how people reason and make decisions (and how they learn and adapt)."

"Cognitive Computing applies to the tuning and automatic configuration of executing algorithms. That is, it applies as an aid to software systems themselves, not just humans."

- "I think that's too restrictive (i.e., you're describing a higher level of Cognitive Computing) - I think you need something that focuses on learning, and includes but doesn't require neuromorphic processors."

The definition listed above, along with the additions of the respondents, contain enough terminology to fill up a book on the subject. To add clarity, some other descriptions of CC include:

> "Cognitive Computing systems get better over time as they build knowledge and learn a domain - its language and terminology, its processes and its preferred methods of interacting. Unlike expert systems of the past which required rules to be hard coded into a system by a human expert, cognitive computers can process natural language and unstructured data and learn by experience, much in the same way humans do. While they'll have deep domain expertise, instead of replacing human experts, cognitive computers will act as a decision support system and help them make better decisions based on the best available data, whether in healthcare, finance or customer service."[3]

> "Cognitive systems were conceived of and are being built to deal with this new immense amount of information. They have a number of characteristics different from today's computers. One is that they learn patterns and trends. They no longer require reprogramming by humans for all the tasks we want them to do. Secondly, cognitive systems interact with people in a much more natural way. They understand our human language, they recognize our behaviors and they fit more seamlessly into our work-life balance. We can talk to them, they will understand our mannerisms, our behaviors – and that will shift dramatically how humans and computers interact."[4]

> "Cognitive systems will extract insights from data sources from which we acquire almost no insight today, such as population-wide health care records, or from new sources of information, such as sensors monitoring pollution in delicate marine environments. Such systems will still sometimes be programmed by people using *if A, then B logic*, but programmers won't have to anticipate every procedure and every rule that will be required. Instead, computers will be equipped with interpretive capabilities that will make it possible for them to learn from the data and evolve over time as they gain new knowledge or as the demands on them change. The goal isn't to replicate human brains, though. This isn't about replacing human thinking

> with machine thinking. Rather, in the era of cognitive systems, humans and machines will collaborate to produce better results - each bringing their own superior skills to the partnership. The machines will be more rational and analytic - and, of course, possess encyclopedic memories and tremendous computational abilities. People will provide judgment, intuition, empathy, a moral compass and human creativity."[5]

Some of the other major players in this discussion include Artificial Intelligence, Machine Learning (along with Deep Learning), and Natural Language Processing. The original definition of AI, coined by John McCarthy in 1955 and recently updated in 2007, still works as a taxonomic umbrella which other terms have been developed under:

> "It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."

Natural Language Processing (NLP) also has a number of different definitions, but one that works for our purposes is:

> "Natural Language Processing is a theoretically motivated range of computational techniques for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications."[6]

Some of those tasks or applications include machine translation, parsing and tagging, sentence understanding, question answering, information extraction, language interpretation, sentiment analysis, speech recognition, word and topic segmentation, word sense disambiguation, natural language generation, et al.

Tied into NLP is Machine Learning (ML). It is a foundational concept that is essential to the success of Cognitive Systems, for without ML the entire structure of the Cognitive Systems in question wouldn't function:

> "Machine Learning is the modern science of finding patterns and making predictions from data based on work in multivariate statistics, data mining, pattern recognition, and advanced/ predictive analytics. Machine Learning methods are particularly effective in situations where deep and predictive insights need to be uncovered from data sets that are large, diverse and fast changing — Big Data. Across these types of data, Machine Learning easily outperforms traditional methods on accuracy, scale, and speed. For example, when detecting fraud in

> the millisecond it takes to swipe a credit card, Machine Learning rules not only on information associated with the transaction, such as value and location, but also by leveraging historical and social network data for accurate evaluation of potential fraud."[7]

In a recently published white paper on ML, Jason Brownlee gave a simple definition which he then elaborated on throughout the paper:

> "Machine Learning is the training of a model from data that generalizes a decision against a performance measure."[8]

The white paper discusses many of ML's essential concepts, including testing and training datasets, induction, over and under learning, and modeling. The substructure of ML is based on algorithms. Without well-developed algorithms that allow for the desired results, the datasets don't matter. Deep Learning (DL) is an subset of ML that takes it to further detail: "Deep Learning is a collection of statistical Machine Learning techniques used to learn feature hierarchies, often based on artificial neural networks,"[9] and:

> "It's about teaching machines to think more hierarchically or more contextually – to see a picture of a mole, for example, and down from recognizing the features that comprise an animal to recognizing the specific features that make it a mole. With text, the process might be teaching machines to recognize how words are related to one another and how they fit together to form phrases or express ideas."[10]

CONVERTING CONCEPT INTO PRACTICE: COGNITIVE COMPUTING APPLICATIONS

Cognitive Computing – and all the concepts, practices, innovations, and advanced analytics within its domain – is working to give enterprises the ability to gain valuable and necessary insights from all of their data assets and information streams, no matter the structure, no matter the size, no matter how seemingly incompatible all that data may seem. Advances in Cognitive Computing --along with the sophistication of new Machine Learning algorithms, evolution of NLP, and new offerings by an ever-expanding vendor marketplace -- are making the jump into the cognitive sphere more readily achievable for enterprises of all sizes.

An enterprise may have only a few terabytes or many petabytes of data to analyze, but if its analytical systems are substandard, the results will still be unsatisfactory and most likely damaging. Enterprises want a complete end-to-end perspective of their information landscape. They want an Information Architecture that allows them to see the entire picture of their data assets across the entire data lifecycle. They want effective frameworks that allow for full integration of their Business Intelligence and Analytics, ETL, Data Governance, Enterprise Data Models, Metadata Management, Data Quality, business process and decision making systems, and data acquisition. Enterprises want a clear idea about what is going on with their data, how to better interact with it, how to get the most operative analysis they can, and all at the best cost. Cognitive Computing has the potential to streamline such processes while also protecting the bottom line.

"There's a compelling need for smarter machines to organize and actuate data faster so therefore we see Cognitive Computing technologies as emerging hub within the Data Management industry. "

FAMILIARITY WITHIN AN UNFAMILIAR WORLD

It is difficult to stay abreast of changes within a practitioner's field of expertise, whether that field is vascular medicine, mortgage investment, or something as new as Cognitive Computing. To gauge where our respondents were in terms of their familiarity with some of the most well-known CC developments and with Cognitive Computing as a whole, the first question we asked (after the demographics information) centered on some of the most recognized CC offerings [Figure 5]: "Are you familiar with IBM's Watson, Siri, Google Now?"

The top two answers, not surprisingly, were:

- Yes, and see them as important pieces for Cognitive Computing's growth: 40%
- Yes, but have not thought of them as representing Cognitive Computing: 30%

It is good that close to 70% of respondents have at least heard of these technologies and sees them as important, though not necessarily as part of CC. But more importantly, 16.7% of respondents either do not know about such technologies or do not find them applicable to this discussion. Certainly, Watson is a major player for IBM in the Cognitive space, and both Siri and Google Now are important end-user applications that employ NLP technologies. If such a significant percentage of data professionals are not generally aware of such familiar offerings, how will that effect the future development and adoption of CC technologies?

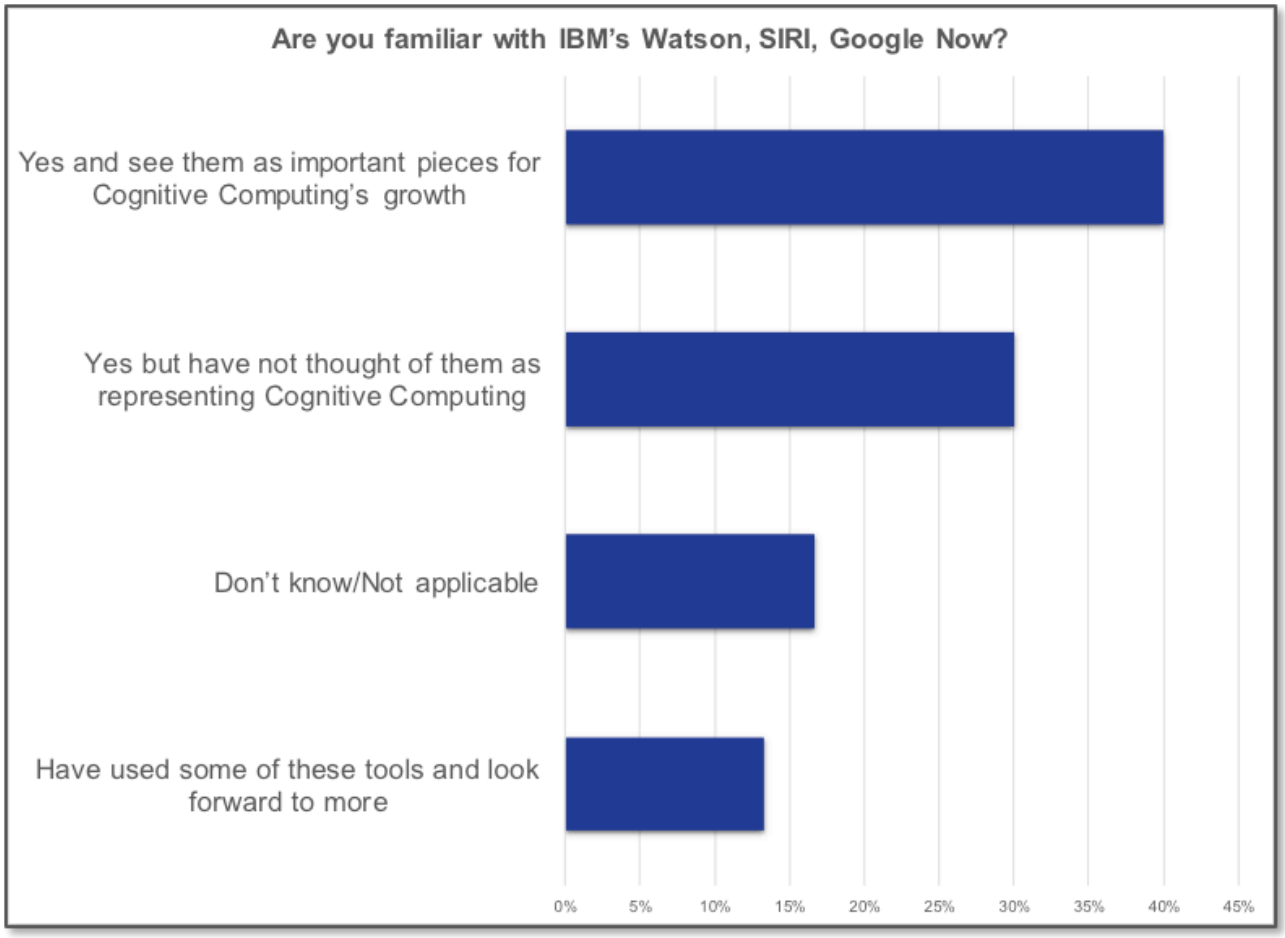


Figure 5

4. WHAT IS NEEDED – TOOLS AND RESOURCES AND DATA ACQUISITION/MODELING/ARCHITECTURE

Whenever a new technology enters the industry (any industry), a number of reactions occur: there are detractors who say it's just a fad, there are traditionalists who decry it as inconsequential or valueless and try to ignore it, and there are thought leaders and gurus who dive right in and learn everything they can about it. In terms of enterprise reactions, the larger organizations often come in a bit late (though not always), then use their massive resources to create huge, packaged offerings that drive slower enterprises into the fray. Early adopters and smaller startups either don't have the resources to deal with new technologies or are at the forefront of the wave, which may either make or break them.

Some of the most prevalent trends from the past decade include Cloud Computing with "as-a-Service" packages, Semantic Technologies, Big Data, NoSQL, and Data Science. Cognitive Computing, though its roots go back decades, is the newest of these and has really only begun to enter board room, server room, and conference hall discussions.

As with these other trends, a major struggle for enterprises wishing to start preliminary CC integration within their particular data environments is a lack of tools, resources, training, and education. Beyond those hurdles, integration must begin with data acquisition, then move into the realms of data warehousing, ETL, databases, BI and analytics, and eventually up into the application engagement layers of an organization's Information Architecture. Therefore, to begin this analysis of CC as it pertains to Data Management, we will first focus on necessary tools and resources and end this section with a discussion of modeling and architecture.

SURVEY RESULTS AND STATISTICS

The survey dealt with the issue of tools and resources through two different questions. The first was a simple 'select all that apply':

"If you're still unclear as to how it fits into your strategy, how does Cognitive Computing and its associated technologies need to develop to become useful tools for your organization? (Select all that apply)" [Figure 6]The top two answers in Figure 6 highlight what many organizations require for any new IT ecosystem offering. The first, "provide clarity on business perspective purpose and effectiveness," (53.4%) clearly demonstrates that for enterprises to even begin to consider implementing a CC strategy, they need to understand the business value. The second choice, "showcase more wins in areas that matter to us – stress results not science," (41.4%) is a central issue within the Cognitive space at this time. There are numerous academic and technical papers on the subject and they are not that hard to find. The real work that needs to be done for enterprises to get on board, though, is for marketplace vendors to create better and more accessible materials that expound the benefits of CC for non-technical users. The current lack of appraisable demos, easily perused white papers, marketing bundles, and business-oriented forums, make it difficult for those not directly aligned within the AI-CC landscape to gain a clear perspective. This is further highlighted by the fact that 23.3% of respondents answered, "don't know/not applicable" for this question.

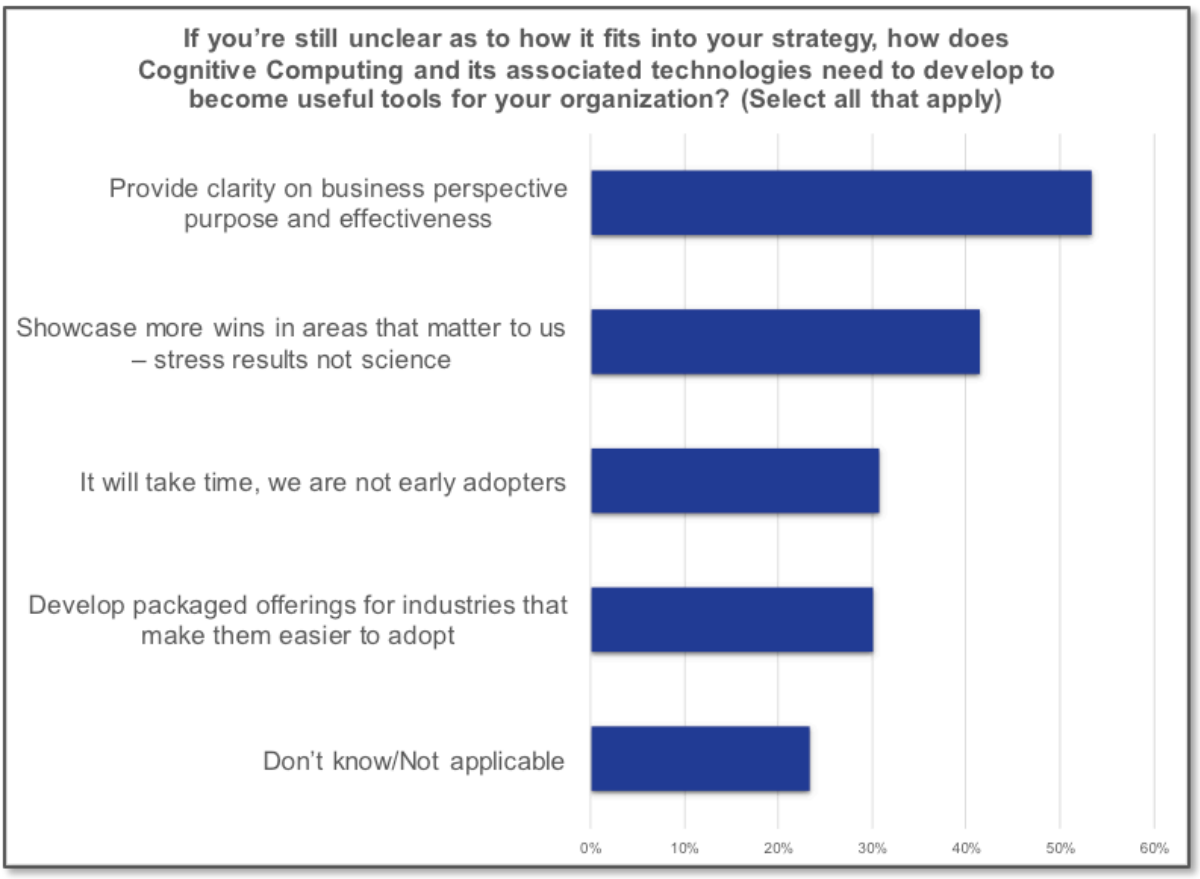


Figure 6

The open-ended question that followed directly after gives more insight into this issue:

- "If you are tracking Cognitive Computing developments and beginning to develop a roadmap, what resources would be most valuable for you to have in order to better evaluate solutions? (Open-Ended Comment Question)"

A total of 45.8% of respondents answered the question. Their answers ran the gamut from use cases to solution ontologies, Web resources to reference architectures, business SMEs with knowledge of the trend to proof-of-concept evaluations, workload requirements to implementation guidelines, and credible, easy-to-access educational resources. A short list of some responses follows:

- "Access to worked examples or proof of concept opportunities that will help demonstrate the 'art of the possible'."
- "The lack of broad information about the potential and shortcomings of cognitive technology is a hurdle for many newcomers. A catalog of known solutions and strategies, like a pattern catalog, would be very helpful. Ideally supported by a solution ontology."
- "Central, up-to-date, well-organized clearinghouse of current Web resources."
- "Reference architectures and citations or examples of successful applications."
- "Real solutions to show the use and advantages of Cognitive Computing."
- "More how, less wow! Need more guidance/instruction for how to develop/build this capability from end-to-end. All I see is 'pieces/parts,' need the 'integrated picture' with guidance on how to do each phase and how each phase links to the others."
- "Ecosystem framework and both vendor and academic information consolidated in a way that helps keep up with new developments but puts them in perspective."

"More how, less wow! Need more guidance/instruction for how to develop/ build this capability from end-to-end. All I see is 'pieces/parts,' need the 'integrated picture' with guidance on how to do each phase and how each phase links to the others."

- "Detailed case studies using open source software and strong ROI details."
- "Information on case studies, pilot projects and initial offerings. Also, information on new developments using languages other than English."
- "An overview of the different companies and their approaches and definitions to what Cognitive Computing is – to understand the differences in approach to what different companies are referencing."
- "Workload requirements and workload distribution across mobile devices, enterprise, and Cloud."
- "A very concise definition of what Cognitive Computing includes and does not include. In today's lexically challenged Data Management arena many terms are used, mis-used, and abused to the extent that they have become meaningless. Cognitive Computing is heading there without a very concise definition that is used consistently."
- "Simple self-run Demos that I could use to present to leadership."
- "Success rates in terms of: implementation of technology and achievement of strategic business objectives."
- "Comparative evaluation of developments including solutions that work 'out of the box' vs solutions that require considerable consultancy services to work. Definitely compare on platforms, connectors/integrations to existing infrastructure, integration with Cloud etc."
- "Existing open source (or inexpensive) code (Machine Learning) that we can easily integrate into our existing applications - don't want to have to redevelop these algorithms/ tools - just want to use and leverage them."
- The next two questions in this section inquired with more detail about the issue of tools and resources, with specific focus on Data Modeling. Any enterprise considering a CC integration within its current IT infrastructure is inevitably going to ask about how CC works within current modeling structures. Data Modeling has long been an essential element to an organization's Information Architecture. Even with the advent of Big Data and NoSQL

"An overview of the different companies and their approaches and definitions to what Cognitive Computing is – to understand the differences in approach to what different companies are referencing."

technologies, Data Modeling is still an indispensable piece of the IT landscape and is not going to disappear anytime soon. A few important points to consider:

- 'Traditional' enterprise software fits data into existing models then shepherds through pre-determined pathways, so unless problems are not subject to much change, decision-making and outcomes are suspect.
- 'Cognitive' software is built on algorithms that learn and adapt to new situations and pathways so users can better understand what's similar, what's different, and anticipate what could happen next.
- The real power of Cognitive Computing comes when it is embedded in IT ecosystems and applications where highly complex decisions and constant change make it the optimal choice to augment human decision making.
- Cognitive Computing systems have their greatest impact when they complement the work being done by humans.

So what does all this mean to Data Modeling as a whole? The two questions asked in the survey dealt with Enterprise Data Models and the expected retention of Modelers/Architects heading forward. Both questions had a standard question format and an open-ended comment section.

"Does your organization have an Enterprise Data Model, and if so, how would Cognitive Computing technologies further enhance this?" [Figure 7]The top two answers for the first question were "don't know/not applicable" at 53.3% and "Yes, partially completed" at 37.2%. Once again, the responses show a definitive lack of understanding around the potential benefits of CC:

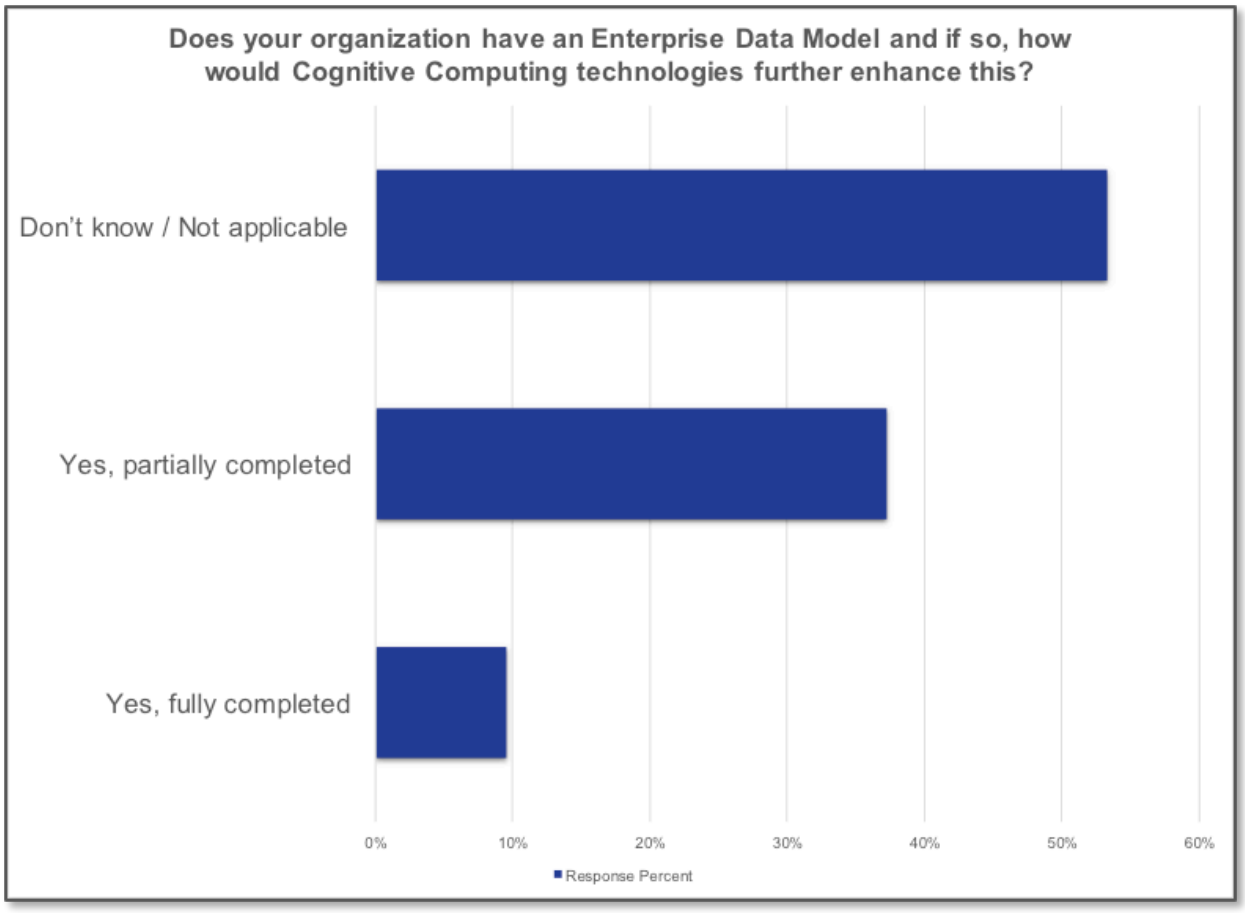


Figure 7

Some of the many open-ended responses to this question include:

- "Most of my clients, including the most recent ones, have partial enterprise models, often not well organized. Cognitive Computing should help produce the models."
- "In general Cognitive Computing is fluid – in order to adhere to an enterprise model it would require confining the fluidity to [a] specific type of outcome that would meet the model criteria but is doable and useful when implemented appropriately."
- "We need a cognitive engine to act as the backend that can help create dynamic interfaces and feedback loops from human interactions with data + information."
- "It would lead us to morph the data model into more of a knowledge map, or ontology."

"In general Cognitive Computing is fluid – in order to adhere to an enterprise model it would require confining the fluidity to [a] specific type of outcome that would meet the model criteria but is doable and useful when implemented appropriately."

- "Company as a whole has an enterprise data model; my business unit IT group lacks an enterprise data model – we need to tackle that before considering Cognitive Computing."
- "Cognitive Computing would help if it emphasized a single, organization wide, comprehensive data architecture for the organization, with data models being drawn from that architecture for the intended audience."
- The final survey question in this section asked about the future for Data Architects and Modelers:
- "With Cognitive Computing technologies becoming a hot topic and with massive amounts of data to be analyzed in the coming years, do you foresee you will have more, less, or the same number of specialist Data Architects and Modelers in your organization three years from now? Why?" [Figure 8]

The top response said that there will be more Architects/Modelers in the future at 57.1%, while only 4.8% of respondents said there would be less.

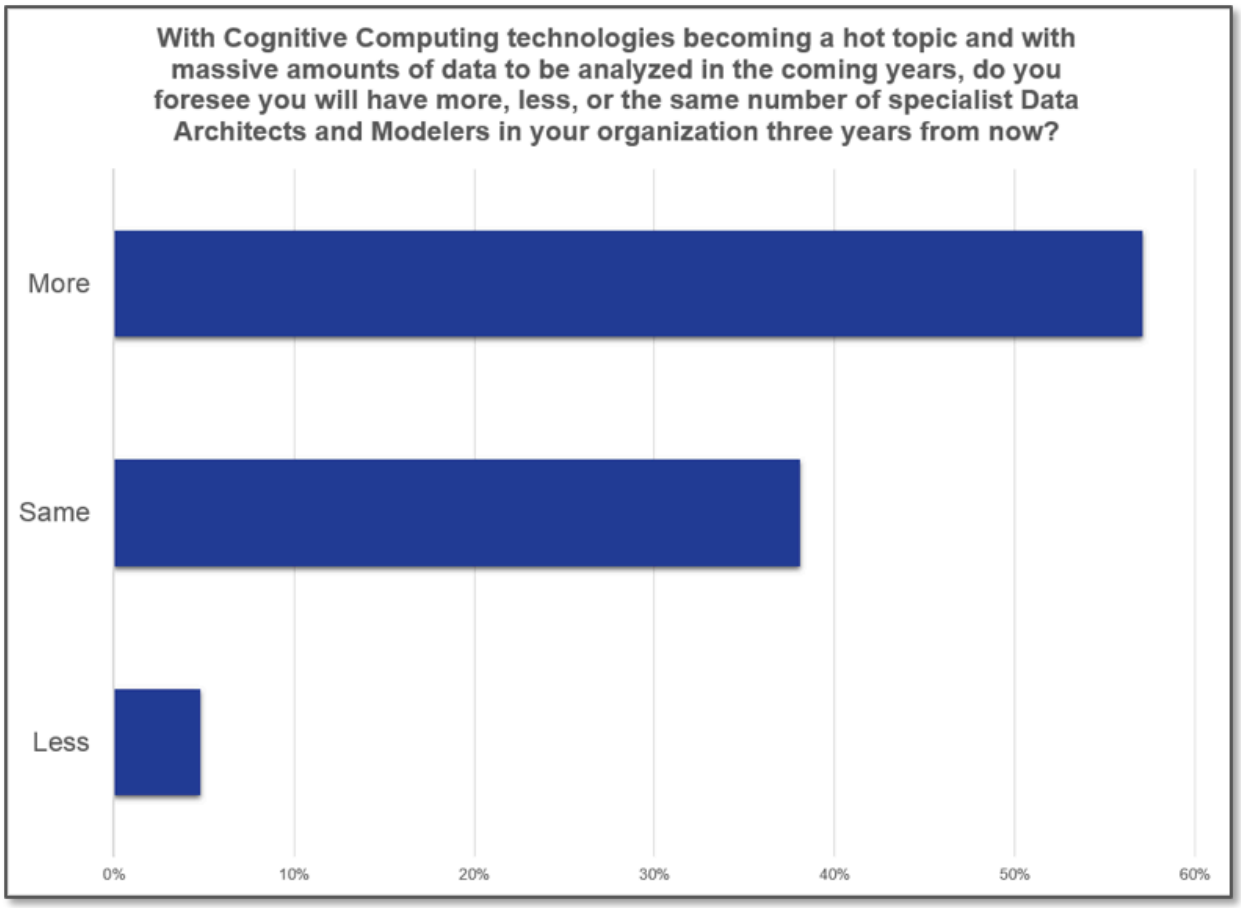


Figure 8

A few of the responses to the open-ended aspect of this question included:

- "This new technology supplements and fits a not-yet-addressed space. Current technologies are not replaced by these."
- "Organizations will begin to realize that they are spending money on staff that push data around from one spreadsheet or database to another. These data pushers will become redundant and Data Modelers will also become increasingly redundant."

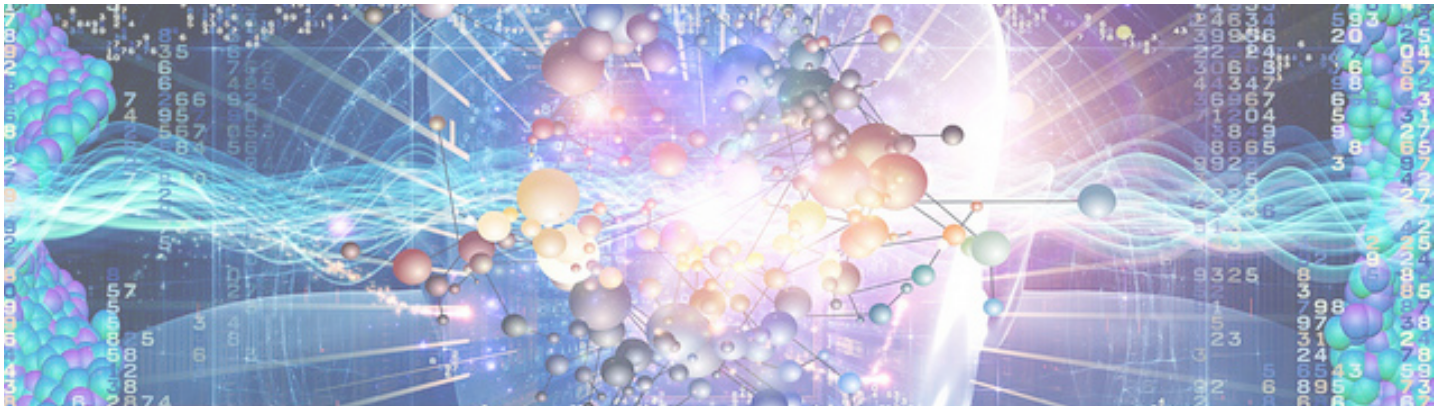
"What will increase is 'Data Scientists' and 'Ontologists.' Based on my experience, Data Modelers are so entrenched in their methodologies they can convert to concepts, classes, categories, and relationships."

- "There will be traditional roles and Big Data roles until it is sorted out which will take some time."
- "Data Architecture workload likely to remain constant, increase in sciences/analytics spaces."
- "What will increase is 'Data Scientists' and 'Ontologists.' Based on my experience, Data Modelers are so entrenched in their methodologies they can convert to concepts, classes, categories, and relationships."
- "Less traditional ER modeling and more document modeling."
- "I think the role of the traditional IA will evolve more towards enterprise solutions that focus on data integration and federation to provide users easy, reliable, and consistent access to information and not just raw data."

ANALYSIS OF RESULTS

The survey results indicate a good technical understanding of Cognitive Computing technologies, albeit with some different points of view. However, the results also demonstrate a muddled view of how to sell the business value of Cognitive Systems, therefore creating problems for implementing a sound CC strategy. Going forward, more compelling customer use cases, evaluation tools, CC conferences and webinars, and documentation that articulate the benefits of Cognitive Computing is needed.

5. COGNITIVE COMPUTING IMPLEMENTATION



It's no wonder that many organizations have a difficult time developing a roadmap for implementing something as new as CC given the limited general understanding of Cognitive Analytics, Machine Learning, Deep Learning, Reasoning, and next generation AI.

In short 'Traditional' enterprise software (e.g. enterprise Data Modeling, Data Warehousing) fits data into existing models then shepherds through pre-determined pathways.

'Cognitive' software is built on algorithms that learn and adapt to new situations and pathways so users can better understand what's similar and different then anticipate what could happen next and because 80 - 90% of new information is unstructured this plays right into sweet spot of Cognitive Computing.

SURVEY RESULTS AND STATISTICS

To investigate the current environment for Cognitive System implementation, the survey respondents were asked three questions. The first [Figure 9] was a straight forward question about their organization's plans for CC implementation. The question included an open-ended comment section so they could add more detail. The top answer for the question was "don't know/not applicable" at 34.6%. "Still unclear as to how it fits into our strategy without a clear enough understanding, so may just take time to build the business case" was a close second at 33.8%. The last choice, "have active plans or projects beginning now," accounted for 14% of responses, demonstrating yet again why considerably more education is needed on the potential of CC.

Figure 9

There were a number of comments in the open-ended section that aid in further underlining this issue:

- "We are active in this field and a member of OpenCog."
- "Demonstrations have been held here on Watson, Big Data and need for Enterprise Data Governance."
- "Using several dashboard applications across multiple channels."
- "Starting with eXist, XQuery and Stanford NLP. We will wait for open source to mature."
- "The lack of Spanish handling capabilities is a big draw back for these initiatives."
- "We are developing a new layer of 'data architecture' more human cognitive oriented [sic], collaborating with neuroscientists, cognitive psychologists, and data modelers."
- "Are working to create standard APIs that are vendor neutral that will allow 3rd parties to offer cognitive services."
- "We are watching the market to see where the early adopt[er]s are and if there is consistency in the use cases."

- "We are looking at Machine Learning to support some of the cognitive simulation models we have - but not natural language applications."

Such comments are once again across-the-board: some organizations are actively working on tracking such developments, some are waiting for more case studies, some are considering offerings but need more presentable business case statistics, and others have a very long way to go before even thinking of CC as a potential implementation.

The next question asked was only open-ended and meant to provide further clarity as a follow-up to Figure 9. It asked, "If you have active plans or Cognitive Computing projects beginning now, can you briefly describe the reason that your organization chose to use Cognitive Computing technologies?"

- "We have vast amounts of information that we need assistance in correlating relationships and cause/effects that are beyond human capacity."
- "We apply cognitive technology in many areas, for discovery of optimized solutions in the construction and design process, for decision support, for automated safeguarding of machinery, etc."
- "We believe that having a cognitive feedback loop to inform decisions, update business rules/flows is critical to maintaining/gaining a competitive advantage by maximizing the return on their existing IT Assets."
- "We are hired for our deliverability to the bottom line. The quicker we can plug in the variables and arrive at a set of cognitive benchmarks improves our desire[d] outcomes."
- "We have had exposure to semantic/cognitive technologies and understand the advantages that [they] can bring to machine automation."
- "We are a vendor of a form of AI and all of our customer apps are looking for a solution that could be defined as Cognitive Computing projects."
- "We want to build intelligent interfaces and applications on top of Cognitive Computing engines."

"We apply cognitive technology in many areas, for discovery of optimized solutions in the construction and design process, for decision support, for automated safeguarding of machinery, etc."

"The need to integrate many different types of data – particularly unstructured data – in order to better understand and address issues related to health care practice on a population basis, as well as to adapt to the major changes and challenges facing our industry today and in the foreseeable future."

- "The need to integrate many different types of data – particularly unstructured data – in order to better understand and address issues related to health care practice on a population basis, as well as to adapt to the major changes and challenges facing our industry today and in the foreseeable future."
- "Increase business responsiveness by taking greater advantage of Big Data Analytics in support of decision automation."

The final question on CC implementation was written as a series of three agreement statements with five choices for each statement ranging from Strongly Agree to Strongly Disagree. Each of the agreement statements had some similarity to the original implementation question discussed in Figure 8, but included a more detailed response in order to ascertain a clearer picture of possible integration and implementation plans:

Agreement Statement 1:	"Our organization is NOT considering any Cognitive Computing technologies now or in the future"
Statistics:	A total of 33.8% respondents answered Strongly Agree/ Agree, indicating that they were NOT considering CC. 25.9% of respondents were Neutral, and 40.1% responded Strongly Disagree/Disagree.
Commentary:	For organizations where business practices revolve mostly around Transactional Structured Data like data from ERP and traditional data warehouse applications, the upsides for Cognitive Computing Systems are more limited.

Agreement Statement 2:	"Our organization is currently evaluating or considering Cognitive Computing technologies in pilot projects"
Statistics:	30.6% (were in Agreement (either Strongly Agree or Agree) about current evaluations or pilot projects, with the exact same amount 30.6% being Neutral, and 48.5% in Disagreement.
Commentary:	Organizations that have implemented Big Data Hadoop architectures are now looking for smarter, highly-scalable architectures (like Cognitive Computing Systems) that transcend the ability to simply collect, join, consolidate, store, and manage web and enterprise Big Data, both structured and unstructured.

Agreement Statement 3:	"Our organization is an early adopter of Cognitive Computing technologies because they are a terrific complement to SQL, NoSQL, Big Data"
Statistics:	The results of this question had a sizeable percentage of respondents in Agreement (26.1%) considering previous discussions within this paper. The largest percentage was in Disagreement at 46.7%, and 26.8% remained Neutral.
Commentary:	Organizations need to unify massive volumes of dynamic and disparate Web and enterprise data, with contextual associations for predictive and prescriptive analysis to better understand customers intent and buying behavior, This will likely increase the need for Data Scientists, Information Architects, Ontologists, Taxonomists, etc.

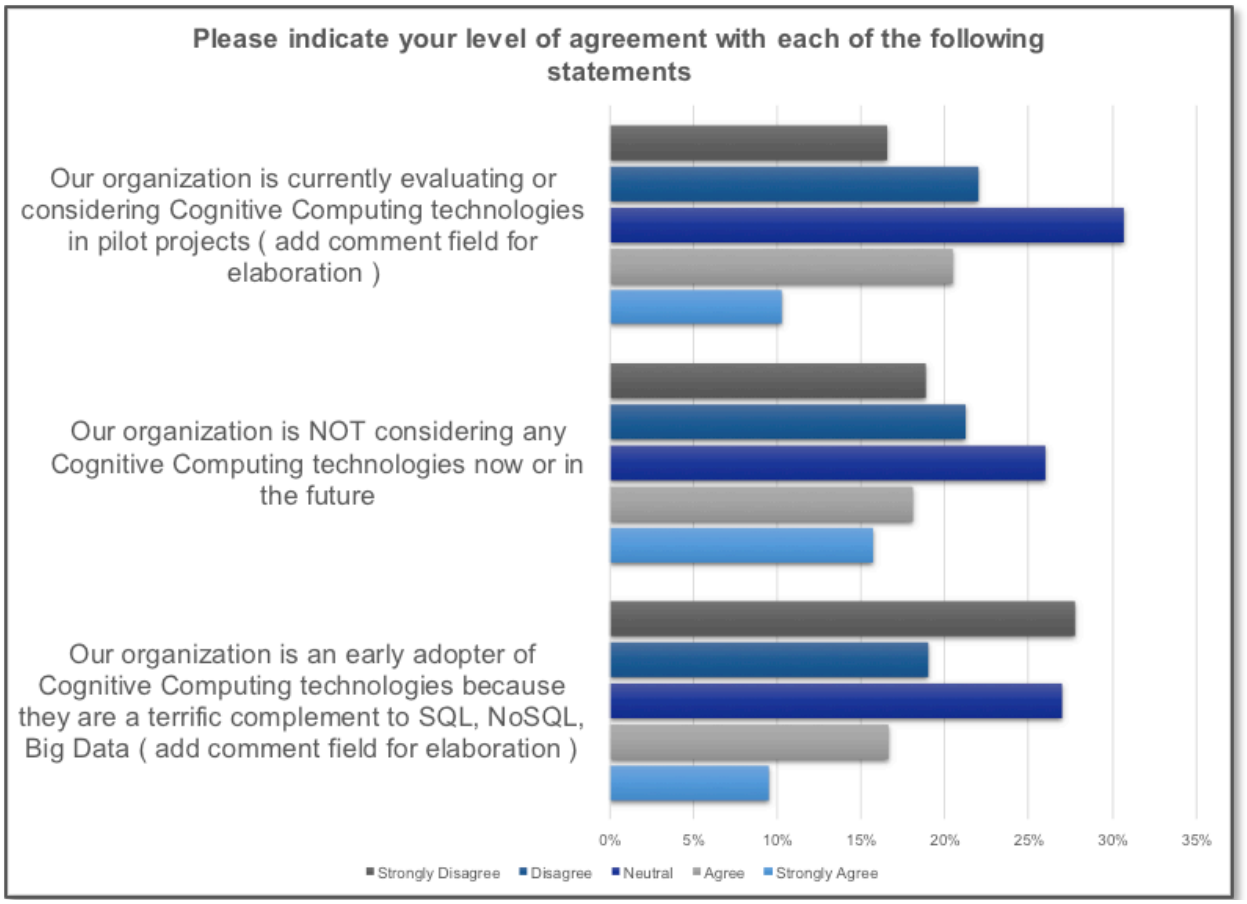
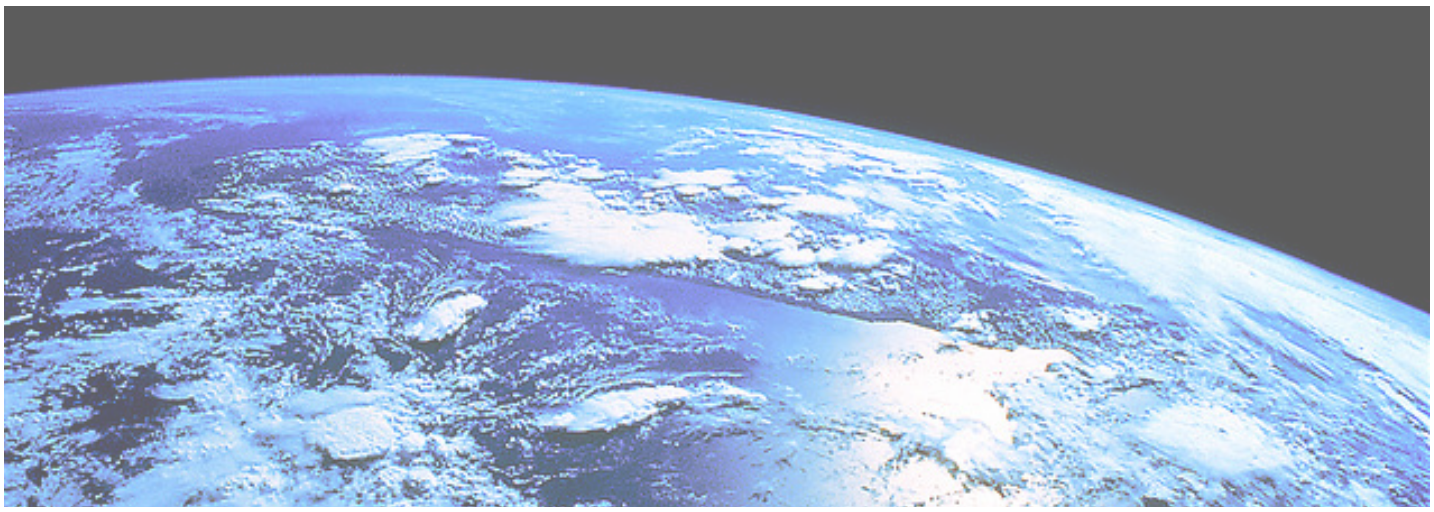

Please indicate your level of agreement with each of the following statements
Our organization is currently evaluating or considering Cognitive Computing technologies in pilot projects (add comment field for elaboration)
Our organization is NOT considering any Cognitive Computing technologies now or in the future
Our organization is an early adopter of Cognitive Computing technologies because they are a terrific complement to SQL, NoSQL, Big Data (add comment field for elaboration)
0%
5%
10%
15%
20%
25%
30%
35%
Strongly Disagree
Disagree
Neutral
Agree
Strongly Agree

Figure 10

6. WHAT CAN CC TECHNOLOGIES DO FOR THE ENTERPRISE?



Cognitive Computing enables humans to connect various data-driven dots faster and smarter for more informed decisions. CC systems learn (they are not programmed) with increasingly sophisticated Machine Learning algorithms via different methods (such as linear models, neural nets, SVMs, kernel methods, classification trees, Deep Learning, et al.), and numerous AI-Cognitive space players are pushing the envelope further. Some are big-name players that are always in the news, but many are small, pioneering startups that are just beginning to make themselves felt within the industry due to their innovative work. As Cognitive Computing moves into the mainstream, the essential question being asked by the C Suite is, how does CC integrate and complement the existing enterprise IT environment?

One of the primary recurring themes coming from CIOs, LOB heads, and other executives/managers in all areas of Data Management is actionable insights. Companies are drowning in data while simultaneously suffering from a paucity of insights that could change the decisions they make every day. Businesses need reliable insights from all of that data, or what is the point of paying to collect, store, and analyze it? It is critical to boil data down into something that can be acted upon in a reasonable time frame to either help companies generate more revenue, operate more efficiently, serve their customer base better, or in a perfect world, all three. The "sweet spot" for Cognitive Computing is shown in the graphic below [Figure 11]:

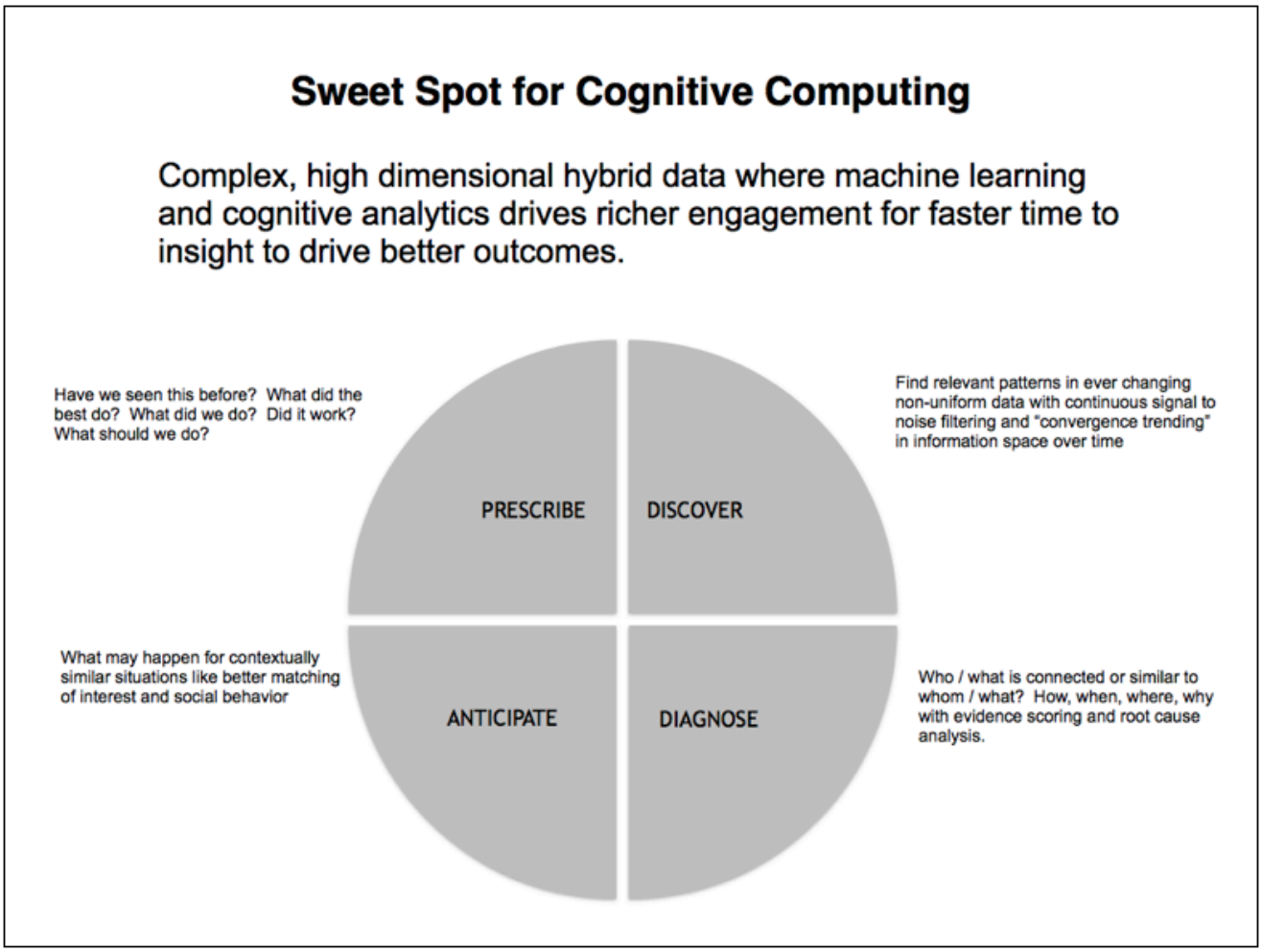


Figure 11

SURVEY RESULTS AND STATISTICS

The survey asked two specific questions about what CC needs to do in order to improve current methods and practices:

- "Based on your experience with SQL, NoSQL, and Big Data products, what can Cognitive Computing technologies do to improve current methods and business practices? (Select all that apply)"

The first question [Figure 12] had three possible responses, and they all came out quite similar in terms of percentages. The top choice at 56.9% was "shortage of Cognitive Computing skills/knowledge amongst existing IT staff especially for Data Scientists, Machine Learning experts." The second choice "we are still trying to sell the technical and business benefits of Big Data and NoSQL, and the Cognitive Computing business case is not well enough understood with lack of executive or business-level support" was a few percentage points behind at 51.2%, while the last choice, "too much to do, too little time," was close behind at 49.6%:

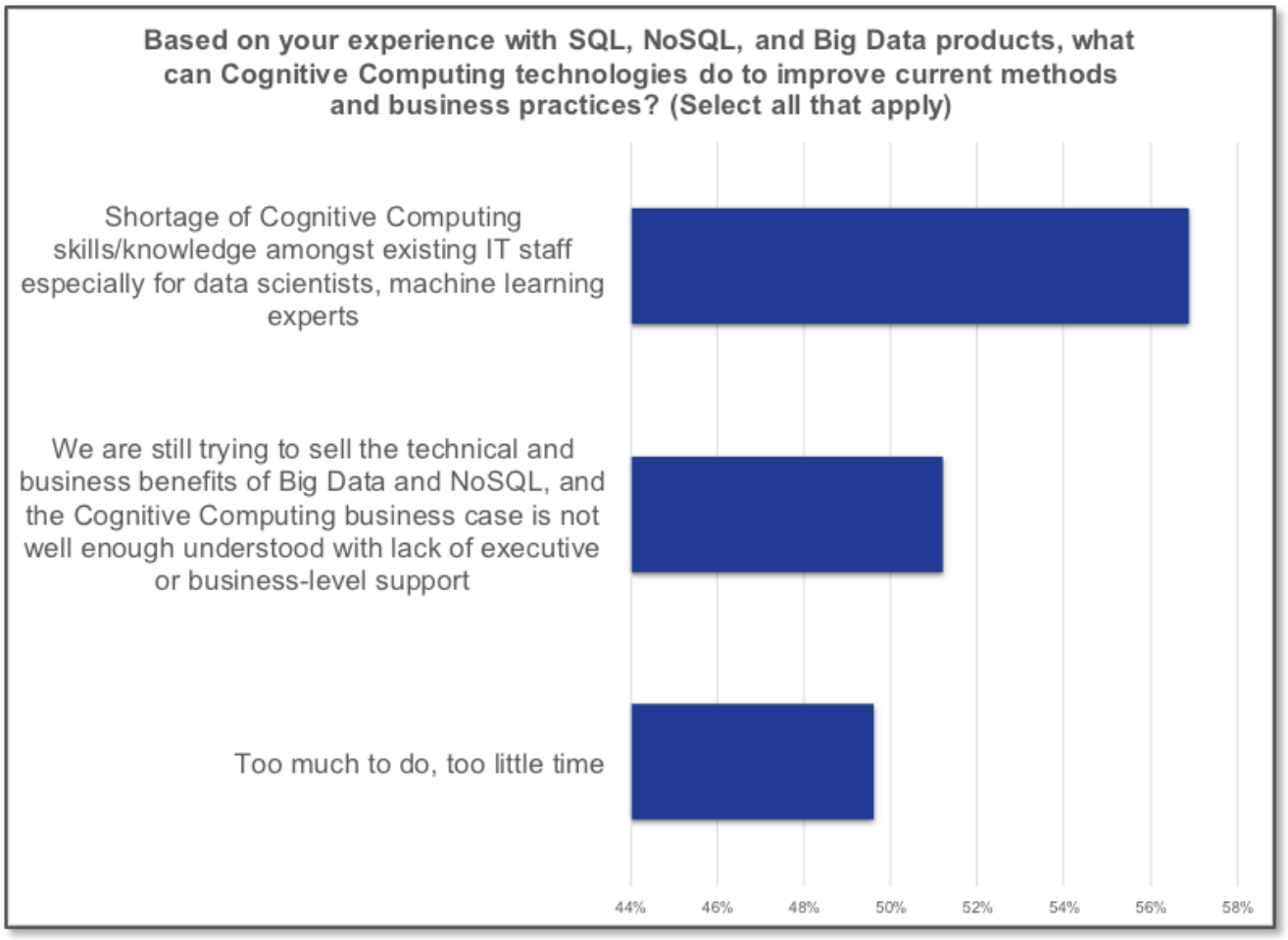


Figure 12

The second question [Figure 13] was written so that we could drill deeper into the issue of what respondents felt CC could potentially augment in terms of the existing enterprise IT environment.

- "As Cognitive Computing becomes more embedded in the IT ecosystem and applications where complex decision- and constant change-making is needed to augment human decision-making, what enterprise activities do you believe could be enhanced with the implementation of Cognitive Computing technologies? (Select All that Apply)"
- The question had seven possible answers. The respondents could select all that apply. The top three choices were:
- Business Intelligence with Cognitive Analytics: 81.8%
- Data Quality Management, Data Services, Data Governance: 58.7%
- Richer and Enhanced User Experience: 58.7%

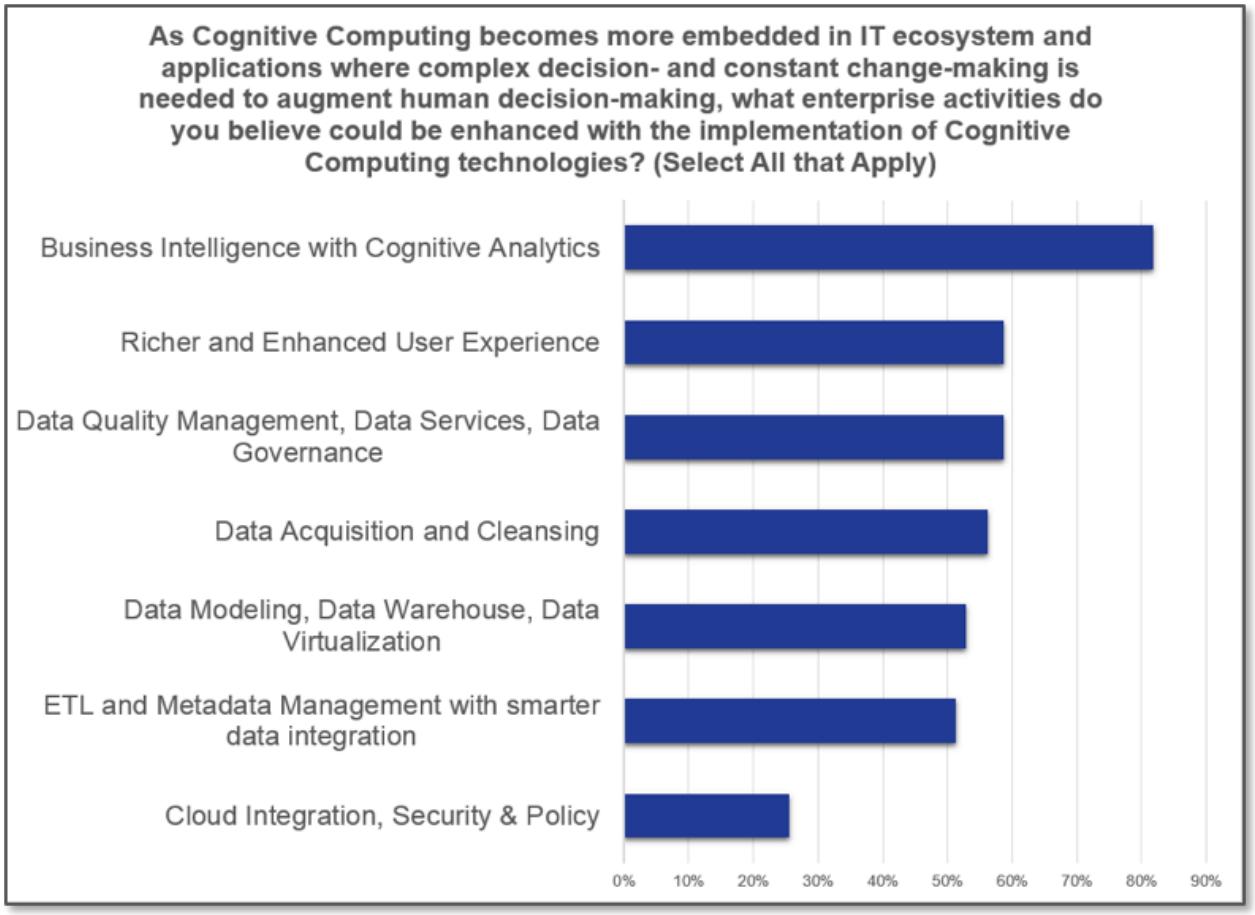


Figure 13

ANALYSIS OF RESULTS

Hadoop ecosystems will increasingly play a role in the fusion of Big Data with Cognitive Analytics and Machine Learning [Figure 14]. One of the great things about Hadoop is the openness and flexibility of HDFS where you can store practically any type of data. Hadoop is proving itself very useful in allowing users to analyze the entire iceberg of data without having advanced programming and statistical skills. The growth in use of customer interactions and machine data with the @ApacheSpark framework as the underlying data framework in Hadoop is starting to blow away traditional enterprise data warehouse approaches that are also more expensive.

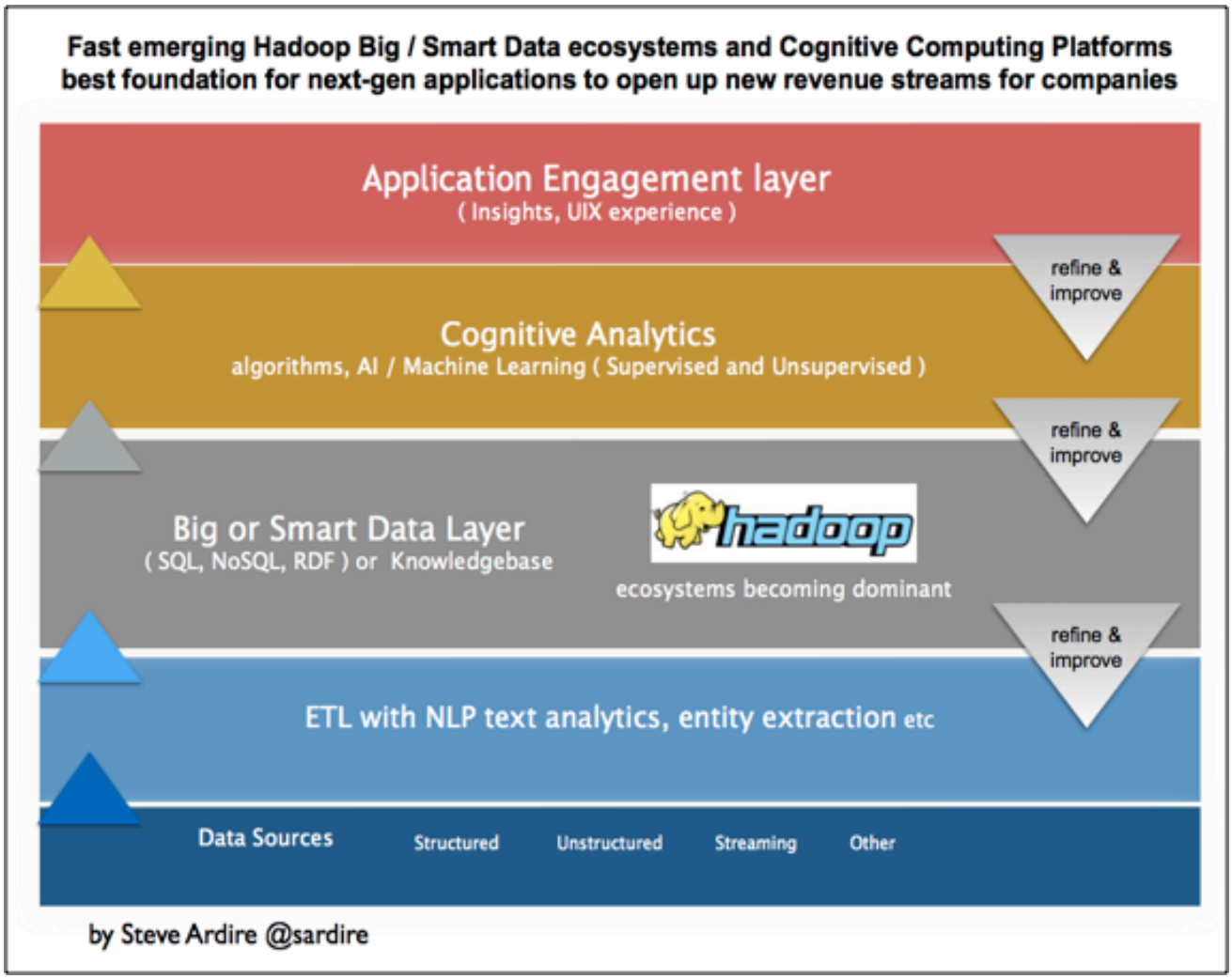

Fast emerging Hadoop Big / Smart Data ecosystems and Cognitive Computing Platforms best foundation for next-gen applications to open up new revenue streams for companies
Application Engagement layer
(Insights, UIX experience)
refine & improve
Cognitive Analytics
algorithms, AI / Machine Learning (Supervised and Unsupervised)
refine & improve
Big or Smart Data Layer
(SQL, NoSQL, RDF) or Knowledgebase
hadoop
ecosystems becoming dominant
refine & improve
ETL with NLP text analytics, entity extraction etc
Data Sources
Structured
Unstructured
Streaming
Other
by Steve Ardire @sardire

Figure 14

7. FUTURE IMPACTS

Cloud Computing, Big Data, NoSQL, Data Science, and Semantic Technologies are here to stay – they are not just new fads that will disappear into the data stream quagmire. They are important game-changing trends that competitive enterprises must put concentration and resources into if they want to stay aggressive in an ever-changing global technology market. To do so, enterprises need to educate their existing staff on these new technologies, hire Data Scientists, Data Architects, and other data experts. They must integrate emerging technologies like CC systems into their extant IT ecosystems, so that they have the ability to make real-time decisions and improve business process management while collecting, storing, and analyzing larger quantities and more diverse structures of data faster than ever before.

SURVEY RESULTS AND STATISTICS

To get a better idea of where current organizations sit within this expanding milieu of new technologies, the survey asked two prognostication questions. The first asked about a series of four technologies/trends and how they would impact the respondent's organization in 2015. The second asked about the same four technologies/trends in 2020. Respondents had three choices: "Major Influence," "Minor Influence," and "Not Applicable."

The initial question for 2015 had some telling results [Figure 15]:

- Only 33.6% of respondents see Big Data/NoSQL/Hadoop Ecosystems as a "Major Influence," while 50% say they are a "Minor Influence."
- 26% of respondents said Cognitive Computing/Machine Learning/ Intelligent User Interfaces were "Not Applicable," and 28.4% said they were a "Major Influence."
- The only technology/trend that had "Major Influence" as its highest response was Sematic Technologies/Unstructured Data at 43.8%.

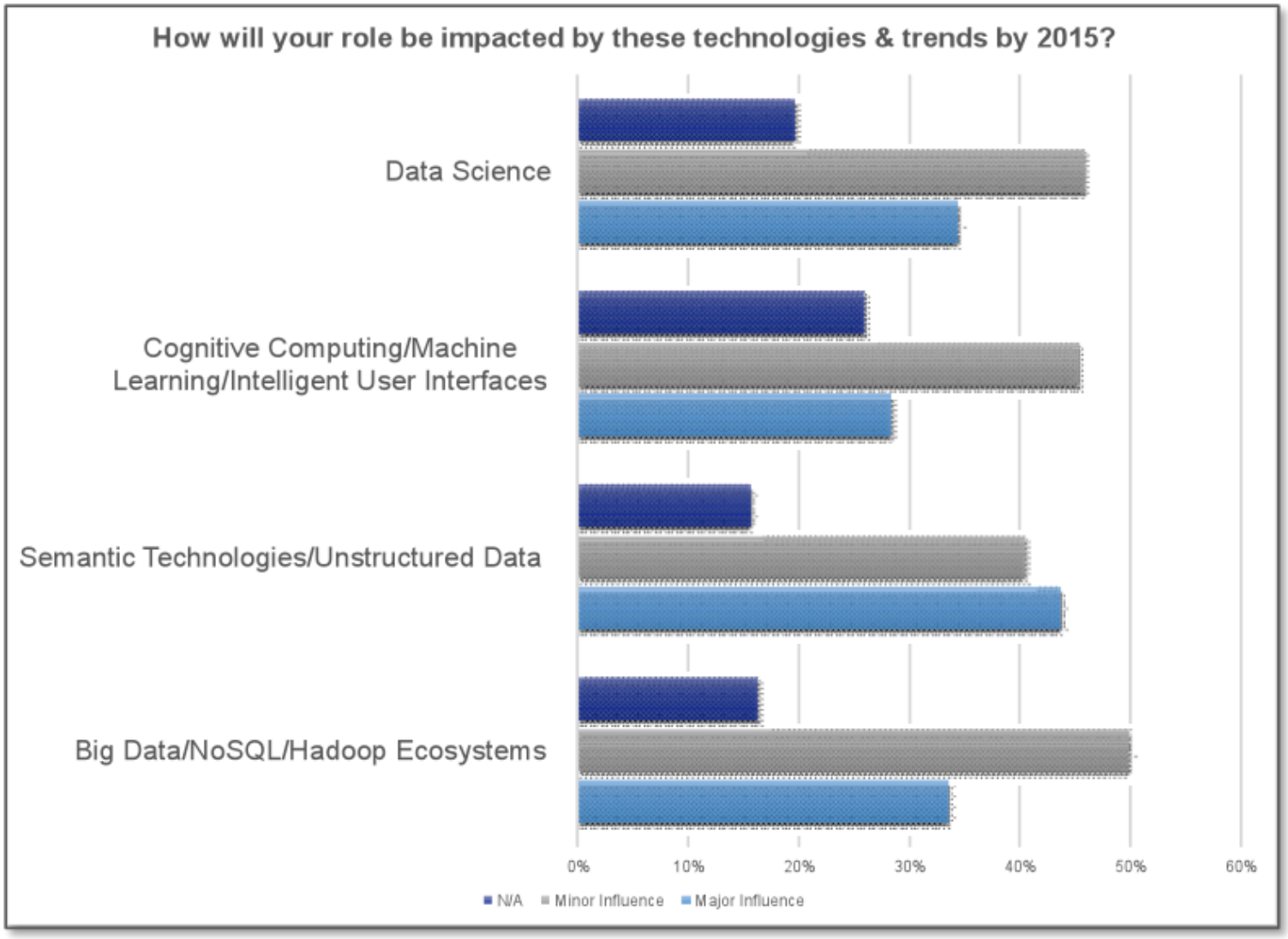


Figure 15

In comparison, respondents believe that all of the listed technologies will have a significant rise in influence by 2020, some by noteworthy percentage points [Figure 16]:

- Cognitive Computing/Machine Learning/Intelligent User Interfaces increased as a "Major Influence" to 63.9% (a 35.5% increase from the 2015 response).
- Big Data/NoSQL/Hadoop Ecosystems came in as a "Major Influence" in 2020 at 56.4% (a 28.4% increase).
- Data Science rose as a "Major Influence" by 27.2%, while Semantic Technologies/ Unstructured Data rose the least (15.2%), but still showed a substantial increase over five years.

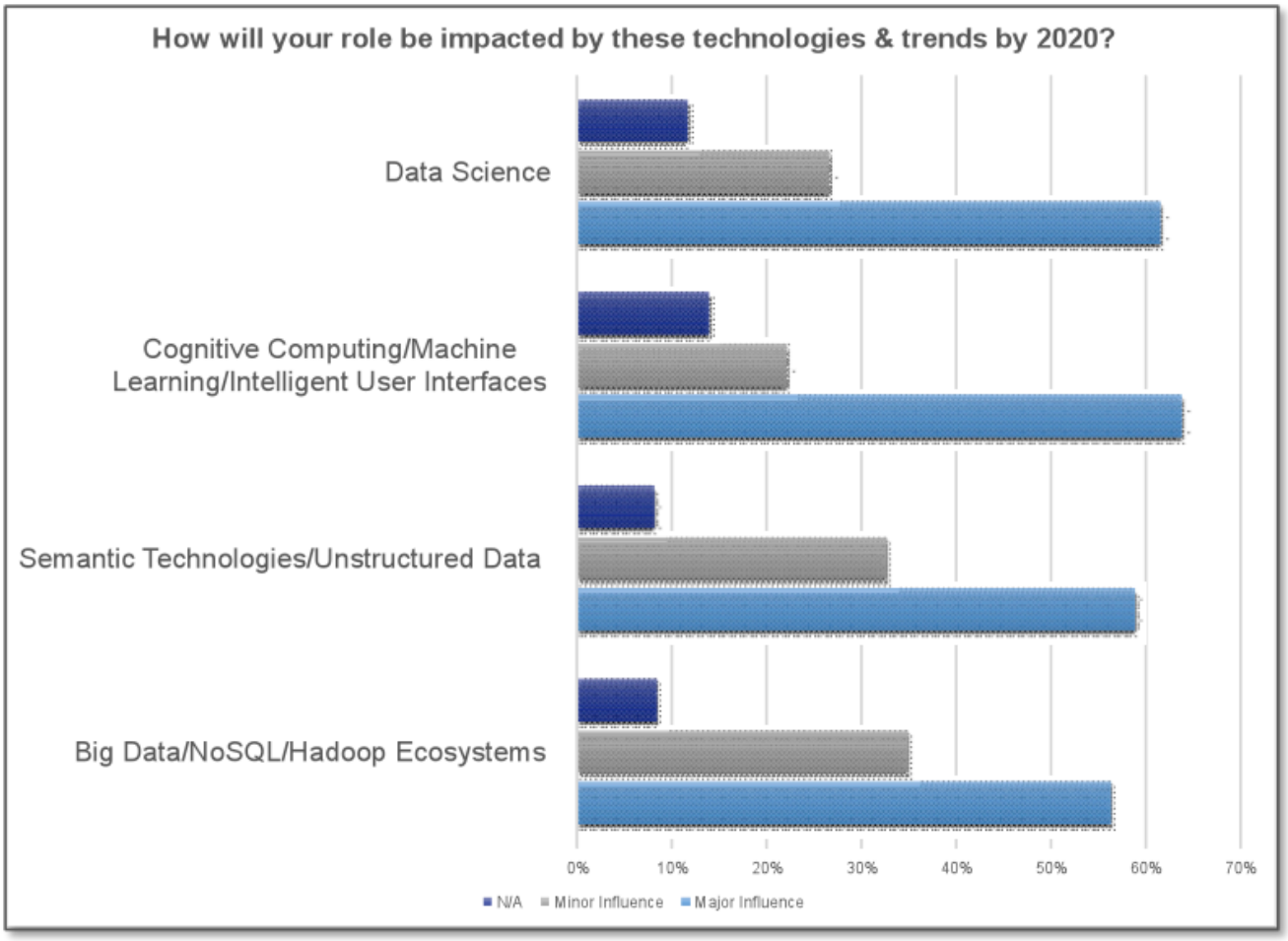

How will your role be impacted by these technologies & trends by 2020?
Data Science
Cognitive Computing/Machine Learning/Intelligent User Interfaces
Semantic Technologies/Unstructured Data
Big Data/NoSQL/Hadoop Ecosystems
0%
10%
20%
30%
40%
50%
60%
70%
N/A
Minor Influence
Major Influence

Figure 16

8. CONCLUSION

The brain is the best Cognitive System ever created. As human-made Cognitive Systems strive to approach the power of the human brain, they will help businesses think better, learn from interactions with data and with us, and adapt their behavior automatically based on new knowledge and experience. Cognitive systems will help us understand ourselves, our biases, and our reasoning. They won't make decisions for us. But they will help us make better decisions in an ever more complex world.

We hope this gave you a better understanding Cognitive Computing where Cognitive Analytics, Machine Learning, Deep Learning, Reasoning, and next generation AI is destined to become an emerging hub in IT Ecosystems.

9. REFERENCES

[1] Rajeev Ronanki and David Steier (2014). Human Brain Inspires New Cognitive Analytics. http://deloitte.wsj.com/cio/2014/05/13/human-brain-inspires-new-cognitive-analytics/

[2] Steve Ardire Interview, May 2014

[3] IBM Research – Cognitive Computing. http://www.research.ibm.com/cognitive-computing/#fbid=ezr0bSmCfZd

[4] Larry Greenemeier (2013). Will IBM's Watson Usher in a New Era of Cognitive Computing? http://www.scientificamerican.com/article/will-ibms-watson-usher-in-cognitive-computing/

[5] John E. Kelly III and Steve Hamm (2014) Smart Machines: IBM's Watson and the New Era of Cognitive Computing. http://www.cup.columbia.edu/static/cognitive

[6] Syracuse University. Natural Language Processing. http://surface.syr.edu/cgi/viewcontent.cgi?article=1043&context=istpub

[7] Martin Hack (2014). Use Data to Tell the Future: Understanding Machine Learning. http://www.wired.com/2014/03/use-data-tell-future-understanding-machine-learning/all/1

[8] Jason Brownlee (2014). Machine Learning Foundations: Master Machine Learning Definitions and Concepts. http://machinelearningmastery.com/machine-learning-foundations

[9] Data Robot (2014). A Primer on Deep Learning. http://www.datarobot.com/blog/a-primer-on-deep-learning/

[10] Derrick Harris (2013). The Gigaom guide to deep learning: Who's doing it, and why it matters. https://gigaom.com/2013/11/01/the-gigaom-guide-to-deep-learning-whos-doing-it-and-why-it-matters/

AUTHOR BIOS

Steve Ardire shape's serendipity by interrogating reality to connect & illuminate the dots that matter for 'select' software startups with cognitive computing and AI/machine learning as current focus. Work with founders / exec team to close seed & Series A funding, shape competitive business strategy, formulate strategic & product marketing, and tactical sales execution with high touch customer engagement to drive business model forward to success.

For past 20+ yrs advised / consulted with 30 software startups in US, Canada, Europe in cognitive computing, AI/machine learning, semantic technology, Big Data, cloud computing, predictive analytics, infoviz, DAM, plus much more.

Charles Roe, freelance writer & founder of CR Scribes, is backed with advanced degrees in English, History and a Cambridge degree in Language Instruction. He worked for 10 years as an instructor of English, History, Culture and Writing at the college level in the USA, Europe and Turkey. He grew up working for a family-owned business in the construction industry, has owned and operated a web design and hosting company, a photo studio, has written numerous academic papers and worked as a professional copyeditor/proofreader for close to 15 years. He spent many years after graduate school working in the high tech industry in tech support, as a database analyst for an ophthalmic software design company and a part-time server administrator. He writes on a variety of topics, including more technical topics, for a host of businesses. He writes creatively in his spare time.

ABOUT THE SPONSORS

Cognitive Scale

www.cognitivescale.com

Insights Fabric from Cognitive Scale is the first Cognitive Cloud platform that delivers Insights as a Service from all types of social, public, private and device data and context signals. It will help you master Big Data and uncover Dark Data to improve decision-making, personalize consumer experiences, and create profitable relationships.

www.dataversity.net

DATAVERSITY™ provides resources for information technology (IT) professionals, executives, and business managers to learn about the uses and management of data. Our worldwide community of practitioners, advisers, and customers participates in, and benefits from, DATAVERSITY's educational conferences, discussions, articles, blogs, webinars, certification, news feeds, and more. Members enjoy access to a deep knowledge base of presentations, research and training materials, plus discounts off many educational resources including webinars and conferences.