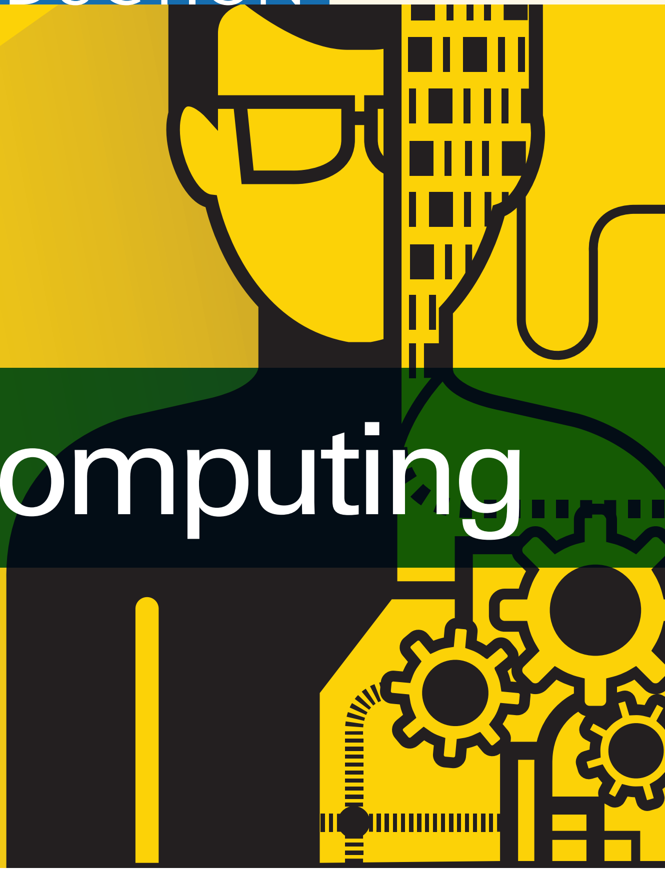
# Cognitive Computing



**Haluk Demirkan,** *University of Washington Tacoma*
**Seth Earley,** *Earley Information Science*
**Robert R. Harmon,** *Portland State University*

There is no doubt that computers are increasingly capable of doing things that in the past, only humans could do. Today, smart machines are becoming more like humans through their ability to recognize voices, process natural language, and learn. They learn by interacting with the physical world through devices that enable them to see, hear, smell, and touch, as well as through mobility and motor control. In some cases, they do a much faster and better job than humans at recognizing patterns, performing rule-based analysis on very large amounts of data, and solving both structured and unstructured problems.

## Smart Machines

*Cognitive computing* refers to smart systems that learn at scale, reason with purpose, and interact with humans and other smart systems naturally. Rather than being explicitly programmed, these systems learn and reason from their interactions with us and from their experiences with their environment. They are made possible by advances in a number of scientific fields over the past half-century, and are different in important ways from the information systems that preceded them.

Many researchers, practitioners, and educators are working diligently on this topic. Collectively, enterprises are spending billions of dollars to make these cognitive systems smarter. We already interact with services enabled by cognitive computing, such as Apple Siri, IBM Watson, Microsoft Cortana, Google Go, and Amazon Echo.

But what is this hype about "smart machines becoming the most disruptive in the history of IT"? In fact, this is not a new thing. One of the enabling technologies, artificial intelligence, was defined by John McCarthy at a conference held at Dartmouth in 1956 as "the science and engineering of making intelligent machines, especially intelligent computer programs."[1] It has taken 60 years to begin to realize this vision through a combination of advances in processing power,

 Published by the IEEE Computer Society 

development of new algorithms, and the explosion of data with which to train machines and enable intelligent functionality.

Cognitive computing—smart machines—has great potential to reduce costs, increase efficiency, and improve outcomes by performing routine tasks and analyzing large amounts of data. The development and adoption of cognitive computing is a process and the result of a combination of machine learning algorithms and traditional knowledge engineering extended by applying breakthroughs in emerging technology. The overall goal of cognitive computing is to increase the productivity and creativity (decision making, connectivity, innovation, and augmentation) of individuals and organizations.[2]

## In This Issue

This special issue of *IT Professional* seeks to provide readers with an overview of the current topics and practices related to cognitive computing. In addition, it looks into the future for IT professionals as these technologies become indispensable and essential enablers of product and service development and the creation of new markets.

The authors' contributions to this collection of articles have implications for cognitive computing that go beyond the immediate application settings on which they report. They also showcase the application of an array of research methods, including surveys, experiments, and design science. We next look at each of the articles to identify the main thrust of the authors' investigations and the relevant findings for theory and practice.

### Enterprise Cognitive Computing

This issue opens with "Enterprise Cognitive Computing Applications: Opportunities and Challenges," by Monideepa Tarafdar, Cynthia M. Beath, and Jeanne W. Ross. In this article, the authors provide an overview of cognitive computing applications for the enterprise. In particular, they provide a classification of opportunities for developing enterprise cognitive computing (ECC) applications and describe challenges in implementing them. In the 33 user organizations—which represented a broad range of industries distributed across North America, Europe, and the Asia-Pacific region—the authors studied a total of 51 initiatives and use cases of ECC applications, about 70 percent of which were either in production or in a working proof-of-concept stage.

The authors describe two unique capabilities that characterize ECC applications. The first is the processing and making sense of increasingly large and growing volumes of data. ECC applications can handle data that, although available directly to humans, can be overwhelming because of the sheer size of the corpus. Examples include medical literature on a given topic and case law of a country or state. Moreover, ECC applications can adjust and adapt their models based on new data to return increasingly accurate results. The second capability is the automation of tasks that typically require human interpretation, such as queries in a call center that can be interpreted by natural language processing (NLP) or image recognition that can identify individuals.

The authors also discovered that the largest number of use cases involved ECC applications that enhanced an organization's operational excellence. These included sophisticated search and retrieval from very large corpora of technical information, such as legal and accounting data, financial laws and regulations, and medical literature; predictive maintenance of machines; product classification; and fraud detection. ECC applications that intend to delight customers by either offering superior products and services or fostering customer loyalty and engagement formed the second largest number of use cases. The third group was ECC applications that help create a superior employee experience.

Finally, the authors found that four challenges are particularly important. To take advantage of ECC's ability to process massive amounts of data and improve efficiency and effectiveness, business leaders must choose the right tools, make sure needed data is available to those tools, consistently supervise the applications, and appropriately allocate responsibilities between humans and machines.

### Cognitive Compliance in Finance

This issue's theme continues with an article by Arvind Agarwal, Balaji Ganesan, Ankush Gupta, Nitisha Jain, Hima P. Karanam, Arun Kumar, Nishtha Madaan, Vitobha Munigala, and Srikanth G. Tamilselvam, "Cognitive Compliance for Financial Regulations." Financial institutions are faced with rapidly changing regulatory policies and an ever-growing number of regulations. Compliance tasks are further complicated by the complex language used in regulatory documents,[3] which forces banks to hire domain experts whose primary job is to identify relevant regulations and accordingly introduce or suggest changes to specific internal controls to remain compliant. This demand for skilled labor has led to growing operational costs in recent years, sometimes accounting for more than 10 percent of total operational expenses.[4] To address the volume, velocity, variety, and complexity of regulations, banks are increasingly seeking technological help.

The authors propose a system that aids compliance officers in understanding. They describe the architecture of Cogpliance, a cognitive compliance platform that uses machine learning, information retrieval, and NLP techniques coupled with a novel user experience design to provide an end-to-end system. The architecture consists of multiple phases—that is, data ingestion, preprocessing, data enrichment, data store, supported services, and applications.

The authors also describe two applications built on this architecture: regulatory change tracking and a knowledge-graph-driven question-answering system. They show through two case studies how these applications can aid in increased compliance. Practitioners can replicate these solutions by developing necessary components such as document ingestion, data enrichment, and data storage. Organizations can apply this architecture to improve the efficiency of compliance tasks.

### An Intelligent Academic Advisor

In the third article, entitled "Building a Cognitive Application Using Watson DeepQA," Christopher Asakiewicz, Edward A. Stohr, Shrey Mahajan, and Lalitkumar Pandey aim to address the gap between the technical and broad overview literature on cognitive computing by developing guidelines that can be used by nontechnical analysts and developers for the development of cognitive applications. They present a cognitive advising system that answers questions related to their university and its programs. This system uses IBM Watson's machine learning algorithms to identify the question category and provide an appropriate response. To implement the system, they developed a four-step development process as illustrated in their article.

Their guidelines were developed through their experience in developing a frequently asked questions (FAQs) application for a graduate analytics program. The system answers questions from a number of perspectives: potential applicants to the program, newly admitted students, current students, faculty members, and business professionals. One of their goals was to demonstrate how easy it can be to develop a cognitive application. The application addresses a relatively simple domain with a narrow corpus of knowledge to train the system.

Building cognitive business applications requires a combination of both business focus and technical skill. The business focus helps the development team to best identify those opportunities suitable for cognitive development. Technical skills (such as Python and Java programming) help in developing and deploying cognitive solutions. Having a comprehensive platform such as Bluemix on which to build cognitive applications, as well as a set of template applications (with associated code and documentation contained in GitHub repositories) that can be extended and enhanced, allowed the development team to get their prototype running quickly.

### Railway Cognitive Communication

Cheng Wu and Yiming Wang aim to demonstrate a railway cognitive communication model to assess the key issues of spectrum scarcity and uncertainty in motion in their article, "Cognitive Communication in Rail Transit: Awareness, Adaption, and Reasoning." With the continuous construction of high-speed rail in recent years, wireless communication technology plays an increasingly vital role in the daily operation, safety monitoring, and efficiency improvement of high-speed rail.

The authors tested their cognitive application model by using a network simulation of a wireless communication framework. They analyzed the main limitations of rail traffic wireless communication caused by the high-speed motion of trains. To solve the uncertainty of spectrum accessibility, they proposed adding a cognitive reasoning module with a Bayesian network as the core engine. They also discuss performance features related to spectrum accessibility.

### Cognitive Gaming

In the final article of this issue, "Cognitive Gaming," Wei Cai, Yuanfang Chi, and Victor C.M. Leung describe the concept of cognitive gaming with a proposed enabling architecture. Video game virtual experiences are emerging as the earliest adopters of these types of cognitive capability.

After defining cognitive gaming and presenting an architectural framework, the authors discuss examples and opportunities in cognitive game content generation. Two cognitive game system optimization approaches with their inherent challenges are discussed. Building cognitive services specifically for game content is a challenging task, given that it requires complex mathematic models and large amounts of training data for learning algorithms. Collecting players' information for game content generation is identified as an emerging trend in cognitive gaming. This is an extension of personalization approaches used in targeted advertising, which tracks consumer behavior and delivers appropriate content. The rich information available across the web makes it possible for a machine to quickly derive game players' characteristics.

With games as dynamic human-computer interaction systems, player behaviors and performances in different game scenarios pose enormous challenges in game design and optimization. These issues become more complicated in multiplayer games because a variety of factors create more uncertainties. The self-adaptation feature of cognitive computing makes it a powerful approach to address these issues.

As the articles in this special issue illustrate, cognitive computing encompasses the processing and making sense of large volumes of data to enhance operational excellence, improve compliance with regulations, develop academic advising systems, optimize spectrum sharing, and optimize complex game system design. Although the articles here provide an overview of a range of cognitive computing applications, the discipline and its applications are growing rapidly. Cognitive computing is cross-disciplinary in nature and focuses on methodologies and systems that can implement autonomous computational intelligence in applications as varied as expert systems, robotics, autonomous vehicles, medical diagnostics, machine vision, translation, employee performance evaluations, planning and scheduling, marketing analytics, remote maintenance monitoring, and many others. We hope the articles presented here will motivate readers to continue the journey of discovery about what constitutes one of the most important domains for IT professionals. IT


## Acknowledgments

We know the importance of having to start somewhere to get new ideas moving, and of finding the appropriate collaborators to make initial steps and advances in new knowledge possible. We thank editor in chief San Murugesan for the vision he shared with us and for getting the discussion rolling. This issue received several submissions which underwent a two-cycle "review and revise" process before we were able to select the final articles. We would especially like to acknowledge the anonymous reviewers who so generously offered their time, effort, and helpful insights for us to make these hard choices and for helping us develop the final product. Finally, we thank the authors whose work was accepted and those whose research we were not able to publish in this edition. We hope the reviewers' comments will strengthen their future success. We look forward to the "next generation" of IT submissions to *IT Professional*.


## References


1. J. McCarthy, "What Is Artificial Intelligence?" Computer Science Dept., Stanford Univ., 12 Nov. 2007, pp. 1–15; stanford.io/2lSo373.
2. H. Demirkan, J.C. Spohrer, and J.J. Welser, "Digital Innovation and Strategic Transformation," *IT Professional*, vol. 18, no. 6, 2016, pp. 14–18.
3. T.D. Breaux and A.I. Antón, "Mining Rule Semantics to Understand Legislative Compliance," *Proc. 2005 ACM Workshop on Privacy in the Electronic Society*, 2005, pp. 51–54.
4. *The Cost of Compliance: Global Hedge Fund Survey*, AIMA, MFA, and KPMG, Oct. 2013; bit.ly/2s5oWwB.



***Haluk Demirkan*** *is a Milgard professor of service innovation and business analytics and director of the Center for Business Analytics at the Milgard School of Business, University of Washington Tacoma. His research interests* *include driving service innovation and transformation with analytics and cognitive computing for value co-creation and outcomes. Demirkan is a cofounder of and member of the board of directors for the International Society of Service Innovation Professionals, and has more than 150 publications. He received a PhD in information systems and operations management from the University of Florida. Contact him at haluk@uw.edu.*

***Seth Earley** is CEO of Earley Information Science (www.earley.com). His interests include knowledge strategy, information architecture, search-based applications, and accelerating the value of analytics solutions. Earley has worked with a diverse roster of Fortune 1000 companies, helping them achieve higher levels of operating performance; is founder of the Boston Knowledge Management Forum; and is a former adjunct professor at Northeastern University, where he taught graduate courses in knowledge management infrastructure and e-business strategy. Contact him at seth@earley.com.*

***Robert R. Harmon** is a professor of marketing and service innovation and Cameron Research Fellow at Portland State University. His research interests include cloud-enabled service innovation and the development and implementation of asymmetric business strategies. Harmon recently coauthored* Cloud as a Service: Understanding the Service Innovation Ecosystem *(Apress/Springer, 2016) with Enrique Castro-Leon of Intel. He is a member of the IEEE Computer Society and the International Society of Service Innovation Professionals, and serves as an associate editor in chief for* IT Professional. *He received a PhD in marketing and information systems from Arizona State University. Contact him at harmonr@pdx.edu.*