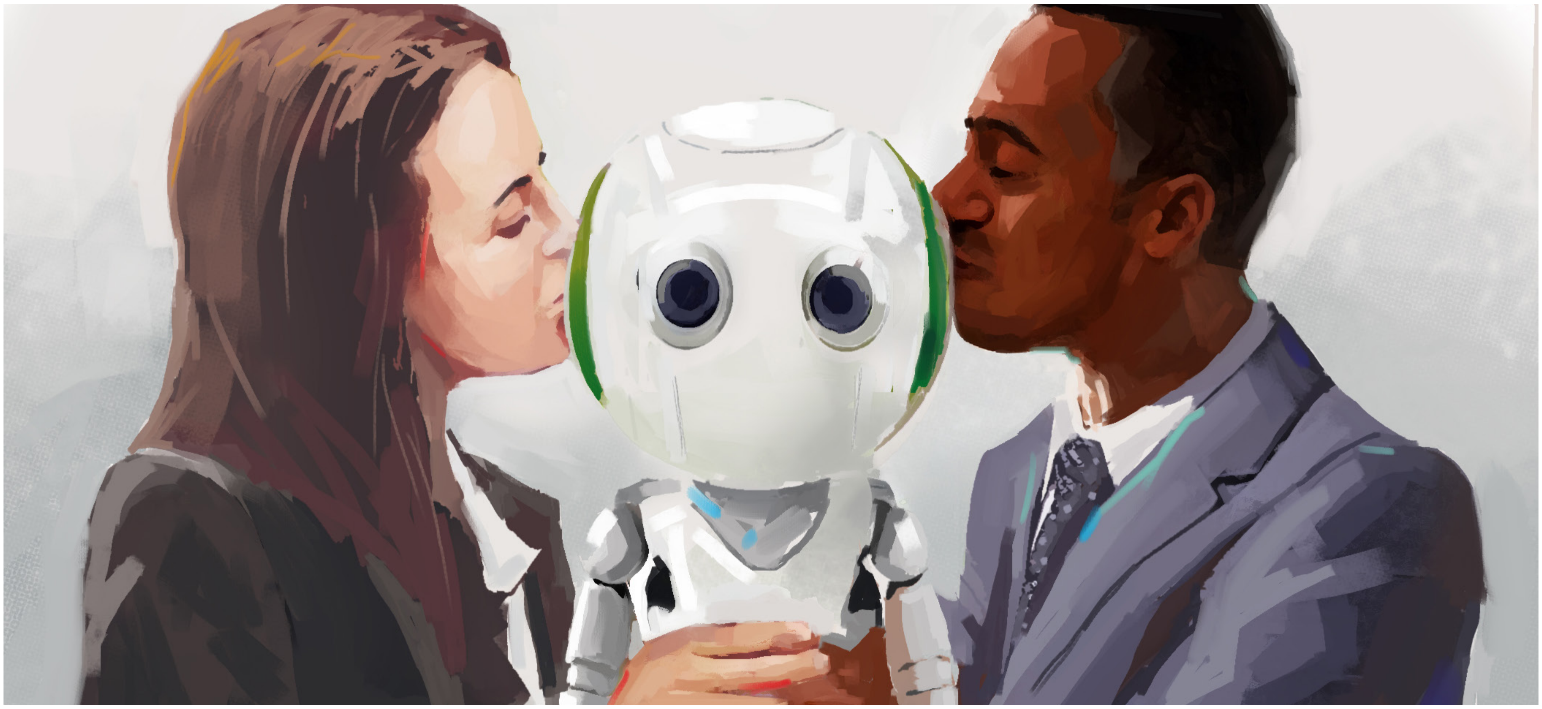

Deloitte.
University Press

Time to move

From interest to adoption of cognitive technology

By Thomas H. Davenport and Paul J. Roma

WHILE cognitive computing, often referred to as artificial intelligence (AI), is hardly new, the recent level of interest in it is astounding. The combination of vendor marketing, concerns about job losses, and even discussion of “robot overlords” have prompted massive interest in the media. There is also plenty of substance behind the hype. Cognitive technologies offer the possibility of increased productivity, better knowledge-based interactions with customers, and the ability to solve problems that are too complex for human brains.

While there have been several “AI winters” and “AI springs” over the past 50 years, there is reason to be confident that the flowering this AI spring is changing the garden permanently. However, in order to get value from this impressive set of technologies, organizations should move rapidly from interest to adoption. In many cases, the technology is ready for immediate use, and many large, established companies are moving ahead with production applications.

The current term for AI is “cognitive technology,” suggesting an ability to perceive and learn that approaches—and sometimes exceeds—that of humans. On both the vendor and user sides of information technology, cognitive is likely the most exciting technology of our time. It is on track to become the most influential technology since the rapid growth of the Internet around the turn of the last century.

There are still some obstacles to cognitive technologies assuming their place at the head of the technology table, but determined organizations can overcome them. It's not clear when or if cognitive technologies will ever reach the level of "artificial general intelligence" at which they are better than humans at every kind of thinking we do (also known as "The Singularity"). But it is already having dramatic and, generally, beneficial impacts on businesses and organizations.

Why the rapid rise?

Why has this rapid growth in visibility and interest for cognitive technologies taken place, and what are the implications of it? There are both demand and supply factors underlying this trend. On the demand side, there are strategic, tactical, and operational aspects. From a strategic standpoint, companies always want profitable growth, and increased productivity is a common way to achieve it. But there hasn't been much productivity improvement in advanced economies over the past several years (only 1.3 percent average annual growth from 2007 to 2015, and decreasing productivity in the first two quarters of 2016),[1] and companies are anxious to learn whether cognitive technologies can finally spur productivity growth.

Also, on the demand side, there are many situations today in which a traditional human approach to analytics and decision-making is simply impossible. These decisions need to be made with too much data and in too short a time for humans to employ their own brains in the process. Digital advertising, medical diagnosis, predictive maintenance or industrial equipment, and many other realms of business today are impossible to execute well without some form of cognitive technologies.

Finally, in terms of the simple operational aspects of demand, while human labor has many strengths, it remains expensive (even when outsourced) and difficult to manage relative to machines. Of course, in many cases smart machines will work alongside smart humans, and that is the best combination in many work settings. But there's no doubt that the siren call of automation is propelling cognitive technology forward.

On the supply side, we now have both software and hardware that is well suited to performing cognitive tasks. Both proprietary and open source software is widely available to perform various types of machine cognition. Several prominent vendors have all made available open source machine learning libraries. And some of the world's largest IT companies are providing proprietary offerings.

Cognitive capabilities are available as standalone software, and, increasingly, as embedded capabilities within other types of software. IBM has placed a big bet on Watson as a standalone software offering, although as a series of APIs (application program interfaces) it can also be combined with other programs. Salesforce.com recently announced "Einstein," a set of cognitive capabilities that are embedded within its programs for sales, marketing, and service. Oracle has announced a series of "adaptive intelligence apps" that will interface with the company's existing cloud-based software and learn from user behavior. It's likely that virtually every major software vendor will embed cognitive capabilities in their transactional systems before long.

Data scientists, however, tell us that the supply side factors that really make a difference for this generation of AI are data and processing power. Neural networks, for example, have been available since the 1950s. But current versions of them—some of which are called "deep learning" because they have multiple layers of features or variables to make a decision about something—require massive amounts of data to learn on and massive amounts of computing power to solve the complex problems they address. Fortunately, Moore's Law has made the requisite processing power available. And in many cases, there are data sources at the ready for training purposes. The ImageNet database, for example—a crowdsourced research database used for training cognitive technologies to recognize images—has over 14 million images that a deep learning system can sink its teeth into.

The availability of almost unlimited computing capability in the cloud means that researchers and application developers can readily obtain the horsepower they need to crunch data with cognitive tools. And relatively new types of processors like graphics processing units (GPUs) are particularly well suited to addressing some cognitive problems. GPUs in the cloud provide virtually unlimited processing power for many cognitive applications. There are also emerging computational infrastructures that combine multiple processors in a mesh to enable the use of more complex cognitive algorithms.

This combination of hardware, software, and data has already yielded impressive levels of capability. In image recognition, for example, a deep learning algorithm running on GPU processors first beat human-coded software for computer vision in recognizing ImageNet images in 2012. By 2015, it had exceeded not only the best human-written software, but the best human eye in image recognition.[2]

Venture capital, of course, is stoking the cognitive engine. VC funding for various forms of cognitive startups increased from $282 million in 2011 to almost $2.4 billion in 2015.[3] Many of these startups will undoubtedly be acquired by larger IT firms that sell to large enterprises.

Rapid cognitive take-up

All the planets on both the supply and demand sides seem to be aligned for organizations to employ cognitive technology in their businesses, and research suggests they are doing so. In a 2016 survey of 450 US executives from companies with revenues over $1 billion,[4] 38 percent of respondents agreed that their companies were "investing in cognitive computing or artificial intelligence to improve business operations, customer service, marketing, or other areas." Almost as many, 36 percent, agreed that their organizations "use machine learning [a branch of cognitive technology] to automatically generate analytical models." And 50 percent of the respondents agreed with two statements that correspond with the use of cognitive tools: "We integrate 'big data' with 'small data' to get business results," and "We mix data science with traditional analytics to get business results." The survey suggests that many large organizations are jumping enthusiastically onto the cognitive bandwagon.

Many of the leading companies we speak to and consult are moving beyond the early days of cognitive technology in business, in which the primary focus was pilot projects or proofs of concept. These firms are now pursuing production applications. Some are transformative, addressing the core objectives or mission of the organization. Others are "low hanging fruit" applications that apply cognitive tools to smaller but useful business issues.

Technologies like "robotic process automation" may not have all the desired cognitive capabilities (they generally can't learn, for example), but they can be implemented rapidly and at low cost to perform many digital tasks. More companies could adopt them now, and would benefit from doing so.

The most aggressive users are pursuing both transformative and "low hanging fruit" applications at once. At one major cancer hospital, for example, IBM's Watson is being trained to diagnose and treat cancer. And this is not all; the hospital's collection department leverages cognitive technologies to rank the patients most likely to realize revenue, while family members of their patients are provided recommendations for lodging and restaurants based on treatment and patient information, radically changing the hospital-patient relationship for the better.

While it is entirely possible that the high expectations for cognitive technologies in business could lead to what Gartner Inc. calls a "trough of disillusionment,"[5] any dissatisfaction is likely to be only temporary at worst. Some early use cases were perhaps not appropriate for the technology applied to them, but understanding of the differences and applications of alternative cognitive tools is growing rapidly. These technologies have the ability now to solve pressing problems in business and other types of institutions. They can make scarce knowledge more widely available, and make fast and accurate decisions. They work at all hours of the day and night. They can greatly augment our human capabilities for perceiving our environments and undertaking decisions and actions based on them. We cannot afford to be complacent as we plan our own futures in the workplace, but we also cannot afford to ignore the opportunities that cognitive technologies present.

ENDNOTES

1. Ben Leubsdorf, "Productivity slump threatens economy's long-term growth," *Wall Street Journal*, http://www.wsj.com/articles/u-s-productivity-dropped-at-0-5-pace-in-the-second-quarter-1470746092.
2. Jen-Hsun Hwang, "Accelerating AI with GPUs: A new computing model," NVidia blog post, January 12, 2016, https://blogs.nvidia.com/blog/2016/01/12/accelerating-ai-artificial-intelligence-gpus/.
3. CB Insights, "Artificial intelligence explodes: New deal activity record for AI startups," https://www.cbinsights.com/blog/artificial-intelligence-funding-trends/.
4. Survey of analytics trends, August 2016, Deloitte Consulting LLP.
5. Gartner Inc., "Gartner hype cycle," accessed February 22, 2017, http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp.

Deloitte.
University Press

Follow @DU_Press

Sign up for Deloitte University Press updates at www.dupress.deloitte.com.

About Deloitte University Press
Deloitte University Press publishes original articles, reports and periodicals that provide insights for businesses, the public sector and NGOs. Our goal is to draw upon research and experience from throughout our professional services organization, and that of coauthors in academia and business, to advance the conversation on a broad spectrum of topics of interest to executives and government leaders.

Deloitte University Press is an imprint of Deloitte Development LLC.