# Computing, cognition and the future of knowing

## How humans and machines are forging a new age of understanding

IBM

**Dr. John E. Kelly III**
Senior Vice President,
IBM Research and Solutions Portfolio

**It's not surprising that the public's imagination has been ignited by Artificial Intelligence since the term was first coined in 1955. In the ensuing 60 years, we have been alternately captivated by its promise, wary of its potential for abuse and frustrated by its slow development.**

But like so many advanced technologies that were conceived before their time, Artificial Intelligence has come to be widely misunderstood — co-opted by Hollywood, mischaracterized by the media, portrayed as everything from savior to scourge of humanity. Those of us engaged in serious information science and in its application in the real world of business and society understand the enormous potential of intelligent systems. The future of such technology — which we believe will be cognitive, not "artificial" — has very different characteristics from those generally attributed to AI, spawning different kinds of technological, scientific and societal challenges and opportunities, with different requirements for governance, policy and management.

Cognitive computing refers to systems that learn at scale, reason with purpose and interact with humans naturally. Rather than being explicitly programmed, they 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.

Those systems have been deterministic; cognitive systems are probabilistic. They generate not just answers to numerical problems, but hypotheses, reasoned arguments and recommendations about more complex — and meaningful — bodies of data.

What's more, cognitive systems can make sense of the 80 percent of the world's data that computer scientists call "unstructured." This enables them to keep pace with the volume, complexity and unpredictability of information and systems in the modern world.

None of this involves either sentience or autonomy on the part of machines. Rather, it consists of augmenting the human ability to understand — and act upon — the complex systems of our society. This augmented intelligence is the necessary next step in our ability to harness technology in the pursuit of knowledge, to further our expertise and to improve the human condition. That is why it represents not just a new technology, but the dawn of a new era of technology, business and society: the Cognitive Era.

The success of cognitive computing will not be measured by Turing tests or a computer's ability to mimic humans. It will be measured in more practical ways, like return on investment, new market opportunities, diseases cured and lives saved.



**Here at IBM, we have been working on the foundations of cognitive computing technology for decades, combining more than a dozen disciplines of advanced computer science with 100 years of business expertise. Now we are seeing first hand its potential to transform businesses, governments and society.**

We have seen it turn big data from obstacle to opportunity, help physicians make early diagnoses for childhood disease and suggest creative solutions for building smarter cities. And we believe that this technology represents our best — perhaps our only — chance to help tackle some of the most enduring systemic issues facing our planet, from cancer and climate change to an increasingly complex global economy.

## The history of computing and the rise of cognitive

To understand the future of cognitive computing, it's important to place it in historical context. To date, there have been two distinct eras of computing — the Tabulating Era and the Programming Era — and IBM has played a central role in the development of both. We believe cognitive computing is the third and most transformational phase in computing's evolution. (Image 1)

**The Tabulating Era** (1900s — 1940s)
The birth of computing consisted of single-purpose mechanical systems that counted, using punched cards to input and store data, and to eventually instruct the machine what to do (albeit in a primitive way). These tabulation machines were essentially calculators that supported the scaling of both business and society, helping us to organize, understand, and manage everything from population growth to the advancement of a global economy.

Image 1

The Tabulating Era (1900s – 1940s) | The Programming Era (1950s – present) | The Cognitive Era (2011 – )

**The Programming Era** (1950s — present)
The shift from mechanical tabulators to electronic systems began during World War II, driven by military and scientific needs. Following the war, digital "computers" evolved rapidly and moved into businesses and governments. They performed if/then logical operations and loops, with instructions coded in software. Originally built around vacuum tubes, they were given a huge boost by the invention of the transistor and the microprocessor, which came to demonstrate "Moore's Law," doubling in capacity and speed every 18 months for six decades. Everything we now know as a computing device — from the mainframe to the personal computer, to the smartphone and tablet — is a programmable computer.

**The Cognitive Era** (2011 — )
The potential for something beyond programmable systems was foreseen as far back as 1960, when computing pioneer J.C.R. Licklider wrote his seminal paper "Man-Computer Symbiosis." Much of modern computing is based on Licklider's research and insights:

*"Man-computer symbiosis is an expected development in cooperative interaction between men and electronic computers. It will involve very close coupling between the human and the electronic members of the partnership. The main aims are:*

1. *to let computers facilitate formulative thinking as they now facilitate the solution of formulated problems, and*
2. *to enable men and computers to cooperate in making decisions and controlling complex situations without inflexible dependence on predetermined programs…*

*Preliminary analyses indicate that the symbiotic partnership will perform intellectual operations much more effectively than man alone can perform them."* [1]

*— J.C.R. Licklider, "Man-Computer Symbiosis," March 1960*

Licklider knew that cognitive computing would be a necessary and natural evolution of programmable computing, even if he didn't yet know how it would be accomplished. Fifty years later, massively parallel computing and the accumulation of oceans of structured and unstructured data would lay the groundwork for cognitive computing.

## The world's first cognitive system

In February 2011, the world was introduced to Watson, IBM's cognitive computing system that defeated Ken Jennings and Brad Rutter at *Jeopardy!*. It was the first widely seen demonstration of cognitive computing, and it marked the end of the so-called AI winter. The programmable systems that had revolutionized life over the previous six decades could not have made sense of the messy, unstructured data required to play *Jeopardy!*. Watson's ability to answer subtle, complex, pun-laden questions made clear that a new era of computing was at hand.

Since *Jeopardy!*, Watson has tackled increasingly complex data sets and developed understanding, reasoning and learning that go far beyond deciphering. Indeed, the goal of cognitive computing is to illuminate aspects of our world that were previously invisible — patterns and insight in unstructured data, in particular — allowing us to make more informed decisions about more consequential matters. The true potential of the Cognitive Era will be realized by combining the data analytics and statistical reasoning of machines with uniquely human qualities, such as self-directed goals, common sense and ethical values.

This is what Watson was built to do, and is in fact already doing. Banks are analyzing customer requests and financial data to surface insights to help them make investment recommendations. Companies in heavily regulated industries are querying the system to keep up with ever-changing legislation and standards of compliance. And oncologists are testing ways in which cognitive systems can help interpret cancer patients' clinical information and identify individualized, evidence-based treatment options that leverage specialists' experience and research.

What is the experience of this like for the professionals involved? World-renowned oncologist Dr. Larry Norton of the Memorial Sloan Kettering Cancer Center, which is working with Watson to help physicians personalize cancer treatments, says, "Computer science is going to evolve rapidly, and medicine will evolve with it. This is coevolution. We'll help each other. I envision situations where myself [*sic*], the patient, the computer, my nurse and my graduate fellow are all in the examination room interacting with one another." [2]

We first saw hints of this symbiosis after Watson's chess-playing predecessor, Deep Blue, defeated World Chess Champion Garry Kasparov in 1997. Following that demonstration, Kasparov would go on to participate in new "freestyle" chess leagues, in which players are free to use any computer programs they like. In these leagues, some players compete unassisted. Some rely entirely on computer programs. But those players who combine computer input with their intuition and competitive instincts are the most successful. [3]

"Teams of human plus machine dominated even the strongest computers. Human strategic guidance combined with the tactical acuity of a computer was overwhelming. We [people] could concentrate on strategic planning instead of spending so much time on calculations. Human creativity was even more paramount under these conditions." [4]

— Garry Kasparov

## The technical path forward and the science of what's possible

While Licklider helped inform a philosophical approach to cognitive computing, there was little he could do to articulate the technical path forward. That path is still being defined, constantly adjusting as the world outside our computer labs evolves. In particular, we are keenly aware of how data is shaping the future. Gartner estimates that the world's information will grow by 800 percent in the next five years, and that 80 percent of that data will be unstructured. It includes everything humans have recorded in **language** (from textbooks to poems), captured in **images** (CAT scans to family photos) and recorded in **sounds**. It is the data hidden in aromas, tastes, textures and vibrations. It comes from our own activities, and from a planet being pervasively instrumented.

In a global economy and society where value increasingly comes from information, knowledge and services, this data represents the most abundant, valuable and complex raw material in the world. And until now, we have not had the means to mine it effectively.

Programmable systems are based on rules that shepherd data through a series of predetermined processes to arrive at outcomes. While they are powerful and complex, they are deterministic — thriving on structured data, but incapable of processing qualitative or unpredictable input. This rigidity limits their usefulness in addressing many aspects of a complex, emergent world, where ambiguity and uncertainty abound.

Cognitive systems are probabilistic, meaning they are designed to adapt and make sense of the complexity and unpredictability of unstructured information. They can "read" text, "see" images and "hear" natural speech. And they interpret that information, organize it and offer explanations of what it means, along with the rationale for their conclusions. They do not offer definitive answers. In fact, they do not "know" the answer. Rather, they are designed to weigh information and ideas from multiple sources, to reason, and then offer hypotheses for consideration. A cognitive system assigns a confidence level to each potential insight or answer.

A mistake Watson made during *Jeopardy!* illustrates this. At the end of the first day of competition, the category for "Final Jeopardy" was "US Cities," and the clue was, "Its largest airport was named for a World War II hero; its second largest, for a World War II battle." The answer is Chicago (O'Hare and Midway). Watson guessed, "What is Toronto?????" There are many reasons why Watson was confused by this question, including its grammatical structure, the presence of a city in Illinois named Toronto and the Toronto Blue Jays playing baseball in the American League.

As a result, Watson's confidence level was incredibly low: 14 percent. If this had been a regular *Jeopardy!* clue, where contestants must buzz in, rather than one in "Final Jeopardy," that confidence level would have caused Watson not to press its buzzer. Watson knew what it didn't know as indicated by the five question marks shown in Image 2.

Image 2

## However, cognitive systems can learn from their mistakes. Large-scale machine learning is the process by which cognitive systems improve with training and use.

After ingesting a corpus of knowledge, curated by experts on any given subject, cognitive systems are trained by being fed a series of question-and-answer pairs. This machine "knowledge" is then enhanced as humans interact with the system, providing feedback on the accuracy of the system's responses.



When Watson played *Jeopardy!*, it did one thing — natural language Q&A, based on five technologies. Today, Q&A is only one of many Watson capabilities available as an application programming interface. Since then, we have developed more than two dozen new APIs, powered by 50 different cognitive technologies. This is a critical distinction between the technical approach to cognitive computing and other current approaches to Artificial Intelligence. Cognitive computing is not a single discipline of computer science. It is the combination of multiple academic fields, from hardware architecture to algorithmic strategy to process design to industry expertise.

## Many products and services that we use every day — from search-engine advertising applications to facial recognition on social media sites to "smart" cars, phones and electric grids — are beginning to demonstrate aspects of Artificial Intelligence.

Most consist of purpose-built, narrowly focused applications, specific to a particular service. They use a few of the core capabilities of cognitive computing. Some use text mining. Others use image recognition with machine learning. Most are limited to the application for which they were conceived.

**Cognitive systems, in contrast, combine five core capabilities:**

1. **They create deeper human engagement:** Cognitive systems create more fully human interactions with people — based on the mode, form and quality each person prefers. They take advantage of the data available today to create a fine-grained picture of individuals — such as geolocation data, web interactions, transaction history, loyalty program patterns, electronic medical records and data from wearables — and add to that picture details that have been difficult or impossible to detect: tone, sentiment, emotional state, environmental conditions and the strength and nature of a person's relationships. They reason through the sum total of all this structured and unstructured data to find what really matters in engaging a person. By continuously learning, these engagements deliver greater and greater value, and become more natural, anticipatory and emotionally appropriate.

2. **They scale and elevate expertise:** Every industry's and profession's knowledge is expanding at a rate faster than any professional can keep up with — journals, new protocols, new legislation, new practices and entirely new fields. A clear example is found in healthcare, where it is estimated that in 1950, it took 50 years to double the world's medical knowledge; by 1980, seven years; and in 2015, less than three years. Meanwhile, each person will generate one million gigabytes of health-related data in his or her lifetime, the equivalent of about 300 million books.

   Cognitive systems are designed to help organizations keep pace, serving as a companion for professionals to enhance their performance. Because these systems master the language of professions — the language of medicine, or sales, or cuisine — they can both understand and teach complex expertise. This reduces the time required for a professional to become an expert. In addition, because these systems are taught by leading practitioners — whether in customer service, oncology diagnosis, case law or any other field — they make available to broad populations the know-how of the best.

3. **They infuse products and services with cognition:**
   Cognition enables new classes of products and services to sense, reason and learn about their users and the world around them. This allows for continuous improvement and adaptation, and for augmentation of their capabilities to deliver uses not previously imagined. We see this happening already with cars, medical devices, appliances and even toys. The Internet of Things is dramatically expanding the universe of digital products and services — and where code and data go, cognition can now follow.

4. **They enable cognitive processes and operations:**
   Cognition also transforms how a company operates. Business processes infused with cognitive capabilities capitalize on the phenomenon of data, from internal and external sources. This gives them heightened awareness of workflows, context and environment, leading to continuous learning, better forecasting and increased operational effectiveness — along with decision-making at the speed of today's data. This is good news in a world where, for example, an average billion-dollar company spends almost 1,000 person hours per week managing its suppliers.

5. **They enhance exploration and discovery:**
   Ultimately, the most powerful tool that cognitive businesses will possess is far better "headlights" into an increasingly volatile and complex future. Such headlights are becoming more important as leaders in all industries are compelled to place big bets — on drug development, on complex financial modeling, on materials science innovation, on launching a startup. By applying cognitive technologies to vast amounts of data, leaders can uncover patterns, opportunities and actionable hypotheses that would be virtually impossible to discover using traditional research or programmable systems alone.

If cognitive computing is to fulfill its true promise, the underlying platform must be broad and flexible enough to be applied by any company in any industry. And it must be able to be applied across industries. To do that requires a holistic approach to research and development, with the goal of creating a robust platform with a range of capabilities to support diverse applications from an ecosystem of developers.

This platform must encompass machine learning, reasoning, natural language processing, speech and vision, human-computer interaction, dialog and narrative generation and more. Many of these capabilities require specialized infrastructure that leverages high-performance computing, specialized hardware architectures and even new computing paradigms. Each grows from its own scientific or academic field. But these technologies must be developed in concert, with hardware, software, cloud platforms and applications that are built expressly to work together in support of cognitive solutions.

With Watson's rapid evolution, we are getting a glimpse of what is possible. One example is a cognitive medical imaging application that analyzes X-rays, MRIs and ultrasound images. It processes the natural language of medical journals, textbooks and articles. It uses machine learning to correct and improve its understanding. And it develops deep knowledge representations and reasoning that can help surface possible diagnoses. To do this requires specialized graphic processing units to support the large scale of data and human domain expertise to guide its learning and interpret its results.

The power of this new model can be applied to any domain. Oil and gas companies can combine seismic imaging data with analyses of hundreds of thousands of papers and reports, current events, economic data and weather forecasts to create risk and reward scenarios for exploratory drilling. Or, by analyzing test scores, attendance and information about student behavior on digital learning platforms, schools can create longitudinal student records and personalized education plans.

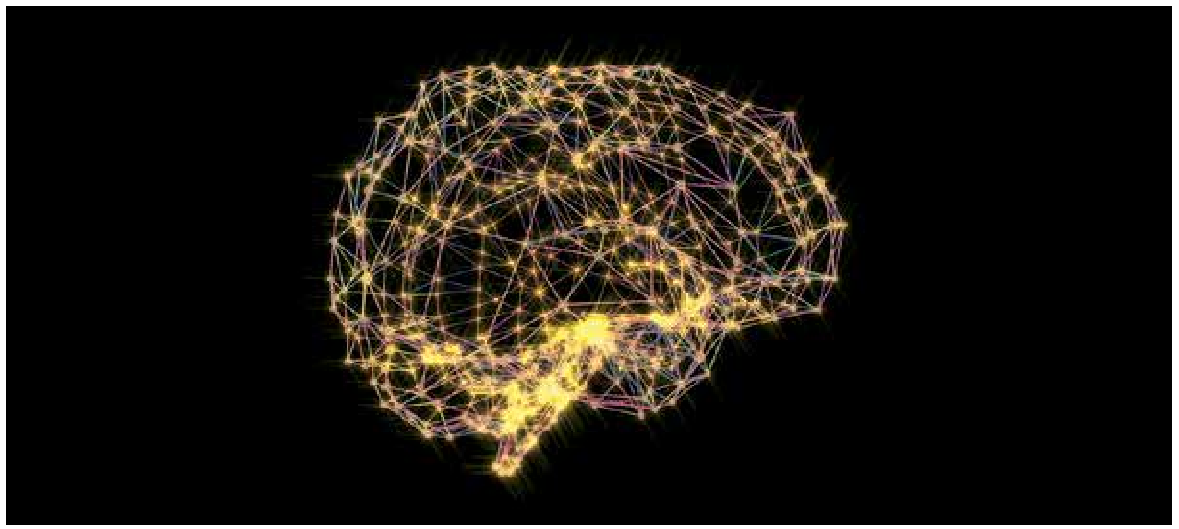
In one of the most promising near-term applications of cognitive computing, IBM is working with more than a dozen leading cancer institutes to accelerate the ability of clinicians to identify and personalize treatment options for patients. The program seeks to reduce from weeks to minutes the time it takes to translate DNA insights, understand an individual's genetic profile and gather relevant information from medical literature. The resulting analysis allows doctors to target the specific cancer-causing genetic mutations in each patient. Watson completes the genetic material and medical literature review process in only a few minutes, producing a report and data visualization of the patient's case and evidence-based insights on potential drugs that may be relevant to an individual patient's unique DNA profile. The clinician can then evaluate the evidence to determine whether a targeted therapy may be more effective than standard care for the patient.

## Implications and obligations for the advance of cognitive science

The Cognitive Era is the next step in the application of science to understand nature and improve the human condition. In that sense, it is a new chapter of a familiar story, and the controversy surrounding Artificial Intelligence is merely the latest example of the age-old debate between those who believe in progress and those who fear it. Within the scientific community — as opposed to the media and popular entertainment — the verdict is in. There is broad agreement on the importance of pursuing a cognitive future, along with recognition of the need to develop the technology responsibly.

> "Technology creates possibilities and potential, but ultimately, the future we get will depend on the choices we make. Technology is not destiny. We shape our destiny." [5]
>
> — Erik Brynjolfsson, MIT

Specifically, we must continue to shape the effect of cognitive computing on work and employment. Like all technology, cognitive computing will change the nature of work done by people. It will help us perform some tasks faster and more accurately. It will make many processes cheaper and more efficient. It will also do some things better than humans, which has been the case since the dawn of civilization. What has always happened is that higher value is found in new skills, and humans and our institutions adapt and evolve. There is no reason to believe it will be different this time. Indeed, given the exponential growth in knowledge, discovery and opportunity opened up by a Cognitive Era, there is every reason



to believe that the work of humans will become ever-more interesting, challenging and valuable.

Equally important is the need for societal controls and safeguards. Again, such concerns are not unique to intelligent systems. Questions about security — both individual and institutional — attach themselves to every transformational technology, from automobiles to pharmaceuticals to mobile phones. These issues are already urgent, and will remain so as cognitive technologies develop. They are fueled especially by today's radical democratization of technology — driven by the rapid spread of networks and the cloud, and the accompanying reduction in costs.

We believe that the answer lies not in attempting to limit that democratization, but rather in embracing it while designing cognitive systems with privacy, security and human control integrated into their fabric.

## Paving the way for the next generation of human cognition

**In the end, all technology revolutions are propelled not just by discovery, but also by business and societal need. We pursue these new possibilities not because we can, but because we must.**

As with every revolutionary technology, our initial understanding will be limited — both by the world's complexity and by our own deeply ingrained biases and heuristics. However, for all these limitations, progress is imperative. Indeed, we pay a significant price for not knowing: not knowing what's wrong with a patient; not knowing which part of the population wants to buy a product; not knowing where to find critical natural resources; not knowing where the risks lie in every investor's portfolio.

"The greatest bar to wise action and the greatest source of fear is ignorance. A tiny candle gives misleading light and throws huge and ominous shadows. The sun at noon gives great light and throws no shadows. It is time to get this whole problem of men and machines under a blazing noonday beam. Computers will never rob man of his initiative or replace the need for his creative thinking. By freeing man from the more menial or repetitive forms of thinking, computers will actually increase the opportunities for the full use of human reason." [6]

— Thomas Watson Jr., former Chairman and CEO

At IBM, we believe that many of the perennial problems of the world can be solved — and we believe that cognitive computing is the tool that will help us accomplish this ambitious goal.

The much-hyped drama of "man vs. machine" is a distraction from that, and it rests on an exciting but misguided fiction. Cognitive systems are not our competitor, nor will they become so. Neither the science nor the economics support such fears. Cognitive systems, as they actually exist, are a tool to deepen the relationship that really matters — the relationship between humans and the world.

In so doing, we will pave the way for the next generation of human cognition, in which we think and reason in new and powerful ways. It's true that cognitive systems are machines that are inspired by the human brain. But it's also true that these machines will inspire the human brain, increase our capacity for reason and rewire the ways in which we learn. In the 21st century, knowing all the answers won't distinguish someone's intelligence — rather, the ability to ask better questions will be the mark of true genius.

1. Licklider, J.C.R. Man-Computer Symbiosis. IRE Transactions on Human Factors in Electronics. 1960. http://groups.csail.mit.edu/medg/people/psz/Licklider.html
2. Kelly, John E. and Steve Hamm. Smart Machines: IBM's Watson and the Era of Cognitive Computing. New York. Columbia University Press, 2014.
3. Kelly, Kevin. The Three Breakthroughs That Have Finally Unleashed AI on the World. Wired Magazine. October, 2014.
4. McAfee, Andrew. Did Garry Kasparov Stumble into a New Business Process Model? Harvard Business Review. February, 2010. https://hbr.org/2010/02/like-a-lot-of-people/
5. Brynjolfsson, Erik and Andrew McAfee. The Second Machine Age: Work, progress, and prosperity in a time of brilliant machines. New York. W.W. Norton & Company, 2014.
6. Watson, Thomas Jr. IBM Archives. https://www-03.ibm.com/ibm/history/exhibits/watsonjr/watsonjr_quoted.html


IBM




IBM Global Services
Route 100
Somers, NY 10589
USA.

Produced in the United States of America
October 2015