

INNOVATION

The Future of AI Will Be About Less Data, Not More

by H. James Wilson, Paul R. Daugherty and Chase Davenport

JANUARY 14, 2019



YANGLEEPHOTO/MULTI-BITS/GETTY IMAGES

Companies considering how to invest in AI capabilities should first understand that over the coming five years applications and machines will become less artificial and more intelligent. They will rely less on bottom-up big data and more on top-down reasoning that more closely resembles the way humans approach problems and tasks. This general reasoning ability will enable AI to be more broadly applied than ever, creating opportunities for early adopters even in businesses and activities to which it previously seemed unsuited.

In the recent past, AI advanced through deep learning and machine learning, building up systems from the bottom by training them on mountains of data. For instance, driverless vehicles are trained on as many traffic situations as possible. But these data-hungry neural networks, as they are called, have serious limitations. They especially have trouble handling “edge” cases—situations where little data exists. A driverless car that can handle crosswalks, pedestrians, and traffic has trouble processing anomalies like children dressed in unusual Halloween costumes, weaving across the street after dusk.

Many systems are also easily stumped. The iPhone X’s facial recognition system [doesn’t recognize “morning faces”](#)—a user’s puffy, haggard look on first awakening. Neural networks have beaten chess champions and triumphed at the ancient Japanese game of Go but turn an image upside down or slightly alter it and the network may misidentify it. Or it may provide [“high confidence” identifications](#) of unrecognizable objects.

Data-hungry systems also face business and ethical constraints. Not every company has the volume of data necessary to build unique capabilities using neural networks. Using huge amounts of citizens’ data also raises privacy issues likely to lead to more government action like the European Union’s General Data Protection Regulation (GDPR), which imposes stringent requirements on the use of individuals’ personal data. Further, these systems are black boxes—it’s not clear how they use input data to arrive at outputs like actions or decisions. This leaves them open to manipulation by bad actors (like the Russians in the 2016 U.S. presidential election), and when something goes embarrassingly wrong organizations are hard put to explain why.

In the future, however, we will have top-down systems that don’t require as much data and are faster, more flexible, and, like humans, more innately intelligent. A number of companies and organizations are already putting these more natural systems to work. To craft a vision of where AI is heading in the next several years, and then plan investments and tests accordingly, companies should look for developments in four areas:

More efficient robot reasoning. When robots have a conceptual understanding of the world, as humans do, it is easier to teach them things, using far less data. Vicarious, a Union City, CA, startup whose investors include Mark Zuckerberg, Jeff Bezos, and Marc Benioff, is working to develop artificial [general intelligence for robots](#), enabling them to generalize from few examples.

Consider those jumbles of letters and numerals that websites use to determine whether you’re a human or a robot. Called *CAPTCHAs* (Completely Automated Public Turing tests to tell Computers and Humans Apart), they are easy for humans to solve and hard for computers. Drawing on computational neuroscience, researchers at Vicarious have developed a model that can [break through CAPTCHAs](#) at a far higher rate than deep neural networks and with 300-fold more data efficiency. To parse CAPTCHAs with almost 67% accuracy, the Vicarious model required only five training examples per character, while a state-of-the-art deep neural network required a 50,000-fold larger training set of actual CAPTCHA strings. Such models, with their ability to train faster and

generalize more broadly than AI approaches commonly used today, are putting us on a path toward robots that have a human-like conceptual understanding of the world.

Ready expertise. Modeling what a human expert would do in the face of high uncertainty and little data, top-down artificial intelligence can beat data-hungry approaches for designing and controlling many varieties of factory equipment. Siemens is using top-down AI to control the highly complex combustion process in gas turbines, where air and gas flow into a chamber, ignite, and burn at temperatures as high as 1,600 degrees Celsius. The volume of emissions created and ultimately how long the turbine will continue to operate depends on the interplay of numerous factors, from the quality of the gas to air flow and internal and external temperature.

Using bottom-up machine learning methods, the gas turbine would have to run for a century before producing enough data to begin training. Instead, [Siemens researchers](#) Volar Sterzing and Steffen Udluft used methods that required little data in the learning phase for the machines. The monitoring system that resulted makes fine adjustments that optimize how the turbines run in terms of emissions and wear, continuously seeking the best solution in real time, much like an expert knowledgeably twirling multiple knobs in concert.

Common sense. A variety of organizations are working to teach machines to navigate the world using common sense—to understand everyday objects and actions, communicate naturally, handle unforeseen situations, and learn from experiences. But what comes naturally to humans, without explicit training or data, is fiendishly difficult for machines. Says Oren Etzioni, the CEO of the Allen Institute for Artificial Intelligence (AI2), “No AI system currently deployed can reliably answer a broad range of simple questions, such as, ‘If I put my socks in a drawer, will they still be in there tomorrow?’ or ‘How can you tell if a milk carton is full?’”

To help define what it means for machines to have common sense, AI2 is developing a [portfolio of tasks](#) against which progress can be measured. DARPA is investing \$2 billion in AI research. In its [Machine Common Sense \(MCS\) program](#), researchers will create models that mimic core domains of human cognition, including “the domains of *objects* (intuitive physics), *places* (spatial navigation), and *agents* (intentional actors).” [Researchers](#) at Microsoft and McGill University have jointly developed a system that has shown great promise for untangling ambiguities in natural language, a problem that requires diverse forms of inference and knowledge.

Making better bets. Humans routinely, and often effortlessly, sort through probabilities and act on the likeliest, even with relatively little prior experience. Machines are now being taught to mimic such reasoning through the application of [Gaussian processes](#)—probabilistic models that can deal with extensive uncertainty, act on sparse data, and learn from experience. Alphabet, Google’s parent company, launched [Project Loon](#), designed to provide internet service to underserved regions of the world through a system of giant balloons hovering in the stratosphere. Their navigational systems employ Gaussian processes to predict where in the stratified and highly variable winds aloft the balloons need to go. Each balloon then moves into a layer of wind blowing in the right direction,

arranging themselves to form one large communication network. The balloons can not only make reasonably accurate predictions by analyzing past flight data but also analyze data during a flight and adjust their predictions accordingly.

Such Gaussian processes hold great promise. They don't require massive amounts of data to recognize patterns; the computations required for inference and learning are relatively easy, and if something goes wrong its cause can be traced, unlike the black boxes of neural networks.

Though all of these advances are relatively recent, they hark back to the very beginnings of AI in the 1950s, when a number of researchers began to pursue top-down models for mimicking human intelligence. But when progress proved elusive and the rich potential for bottom-up machine learning methods became apparent the top-down approach was largely abandoned. Today, however, through powerful new research and computational techniques top-down AI has been reborn. As its great promise begins to be fulfilled, smart companies will put their money where the mind is.

H. James Wilson is a managing director of Information Technology and Business Research at Accenture Research. Follow him on Twitter [@hjameswilson](#). He is a coauthor, with Paul Daugherty, of *Human + Machine: Reimagining Work in the Age of AI* (Harvard Business Review Press, March 2018).

Paul R. Daugherty is Accenture's chief technology and innovation officer. He is a coauthor, with H. James Wilson, of *Human + Machine: Reimagining Work in the Age of AI* (Harvard Business Review Press, 2018).

Chase Davenport is a San Francisco-based technology researcher.

Harvard Business Review Notice of Use Restrictions, May 2009

Harvard Business Review and Harvard Business Publishing Newsletter content on EBSCOhost is licensed for the private individual use of authorized EBSCOhost users. It is not intended for use as assigned course material in academic institutions nor as corporate learning or training materials in businesses. Academic licensees may not use this content in electronic reserves, electronic course packs, persistent linking from syllabi or by any other means of incorporating the content into course resources. Business licensees may not host this content on learning management systems or use persistent linking or other means to incorporate the content into learning management systems. Harvard Business Publishing will be pleased to grant permission to make this content available through such means. For rates and permission, contact permissions@harvardbusiness.org.