The Future Isn't What It Used to Be:[1]
Artificial Intelligence Meets Natural Stupidity[2]

Editor's Note: The following article expands upon a presentation given at the Second Interpreter Science and Mormonism Symposium: Body, Brain, Mind, and Spirit *at Utah Valley University in Orem, Utah, 12 March 2016. The six parts of the article are as follows:*

Part 1: Understanding the Future of Science and Technology
Part 2: Self-Driving Cars
Part 3: Superintelligence Through Artificial Computational "Brains"
Part 4: Superintelligence Through Natural, Uploaded Brains
Part 5: How Close Are We To Being Able To Model the Brain?
Part 6: Autonomous Weapons and Natural Stupidity

A book based on the first symposium, held in 2013, has recently been published as Bailey, David H., Jeffrey M. Bradshaw, John H. Lewis, Gregory L. Smith, and Michael L. Stark. Science and Mormonism: Cosmos, Earth, and Man. *Orem and Salt Lake City, UT: The Interpreter Foundation and Eborn Books, 2016. For more information, including free videos of these events, see http://www.mormoninterpreter.com.*

Figure 1 about here[3]

Part 1: Understanding the Future of Science and Technology

The saying "The future isn't what it used to be," is often attributed to Yogi Berra,[4] although he admitted (as one might also conclude about Elder J. Golden Kimball[5]), that he didn't really say everything he said.[6] The same aphorism was

used by scientist and science fiction author Sir Arthur C. Clarke[7] and later by Apple Computer co-founder Steve Jobs[8] as a preface to their optimistic extrapolations about the future of technology. Although I would agree with Clarke and Jobs that a bright future for technology lies ahead, that is not the drift of the talk I intend to give today.

More significant than the transformations that technology works upon us directly,[9] more potent than the effects of any hallucinogenic drug on our thoughts and senses are the mind-altering impact of our changing conceptions about the future itself. Arthur C. Clarke observed: "Until a century ago nobody was very interested in the future for the simple reason that, apart from natural catastrophes and wars, the future was going to be the same as the past. A man knew that the pattern of his life would be the same as his great-grandfather's, as far back as anyone could remember. Well, now we know differently."[10]

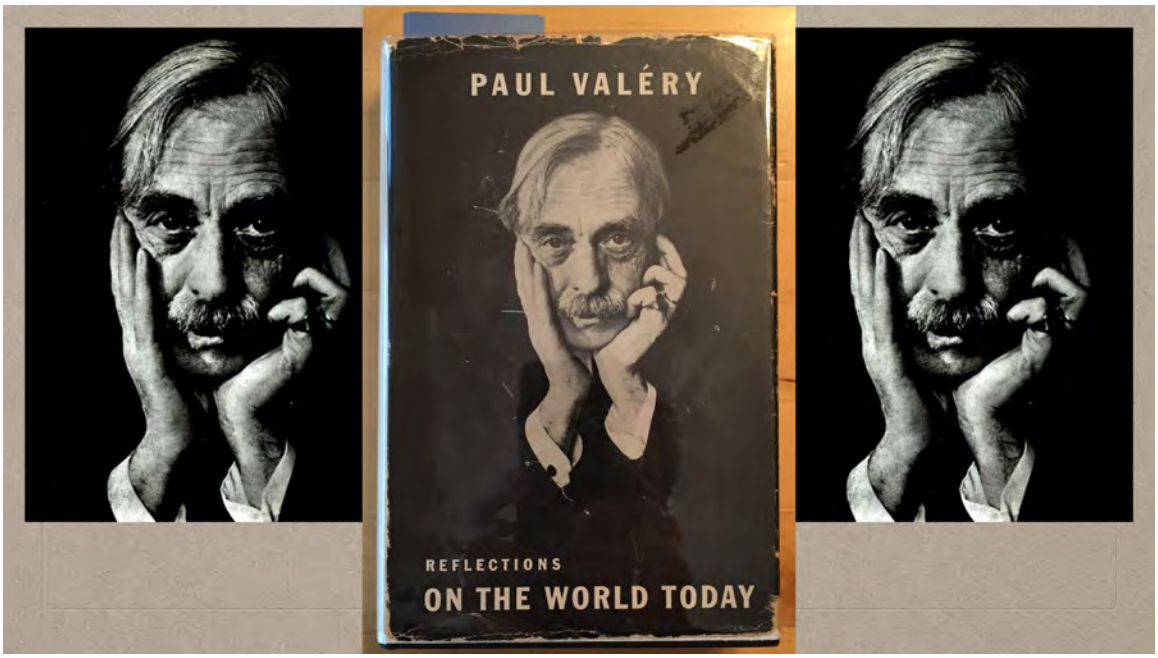


Figure 2 about here[11]

The brilliant, problematic French poet and essayist Paul Valéry (1871-1945)[12] explained why "the future isn't what it used to be" in 1937, long before any of the others I cited before.[13] Wrote Valéry:[14]

> The future, like everything else, is no longer quite what it used to be. … We used to consider the unknown future as a simple combination of already known things, and the new was analyzed according to its unoriginal elements. But that is ended. … [T]he rules of the game are changed at every throw. No calculation of probabilities is possible. … Why? Because the … modern world is assuming the shape of man's mind. Man has sought in nature all the means and powers that are necessary to make the things around him as unstable, volatile, and mobile as himself, as admirable, as absurd, as disconcerting and

> prodigious as his own mind. … If … we imprint the form of our mind on the human world, the world becomes all the more unforeseeable and assumes the mind's [own] disorder.

"Happily," concludes Valéry, "these prophecies are idle. I am busy doing exactly what I explained the vanity of a few moments ago. I am looking ahead, therefore I am wrong."[15]

Since the *far* future is impossible to predict with any confidence, my burden today will be to share some candid observations about the present and the *near* future as it has to do with technology, technologists, and society — between now and, say, 2025 — a little less than ten years from now. In brief, I want to explain why I am not convinced that a technological apocalypse is "nigh at hand."[16] In addition, I cannot refrain from giving a little advice about the future. As the wise Duc de la Rochefoucauld said: "Old people like to give good advice, since they can no longer set bad examples."[17]

Figure 3 about here

First, a personal aside. I was the next-to-youngest in my family — my brother Scott and I were called "the little boys." Scott and I were best buddies then, and have been close ever since. We were spoiled by our older brother and sisters.

My sisters defended my quirks, reassuring my parents that I would someday turn out to be quite normal. My mother engendered in me a love of learning and of the Gospel. Though she never had the chance to attend college, she became very well educated woman through her wide reading on many subjects, including the scriptures and the "best books"[18] of doctrinal and apologetic literature.

My father was the first missionary and the first college graduate in his immediate family. He lovingly spurred me along in my scientific interests, providing me with a large cardboard box kept in the garage that was filled with motors, wind-up clocks, and vacuum-tube electronics that I could take apart and sometimes put back together.

Figure 4 about here[19]

One day at lunch, during the cold war years, after hearing me protest that I didn't want eat my orange because I didn't like the taste of the bitter white stuff that stuck on to the fruit after you peeled it, my Dad told me that the white stuff would protect me from radiation in the event of nuclear fallout. That creative fib not only solved the immediate problem in getting me to eat my orange but also convinced me that there was a real practical value to science.

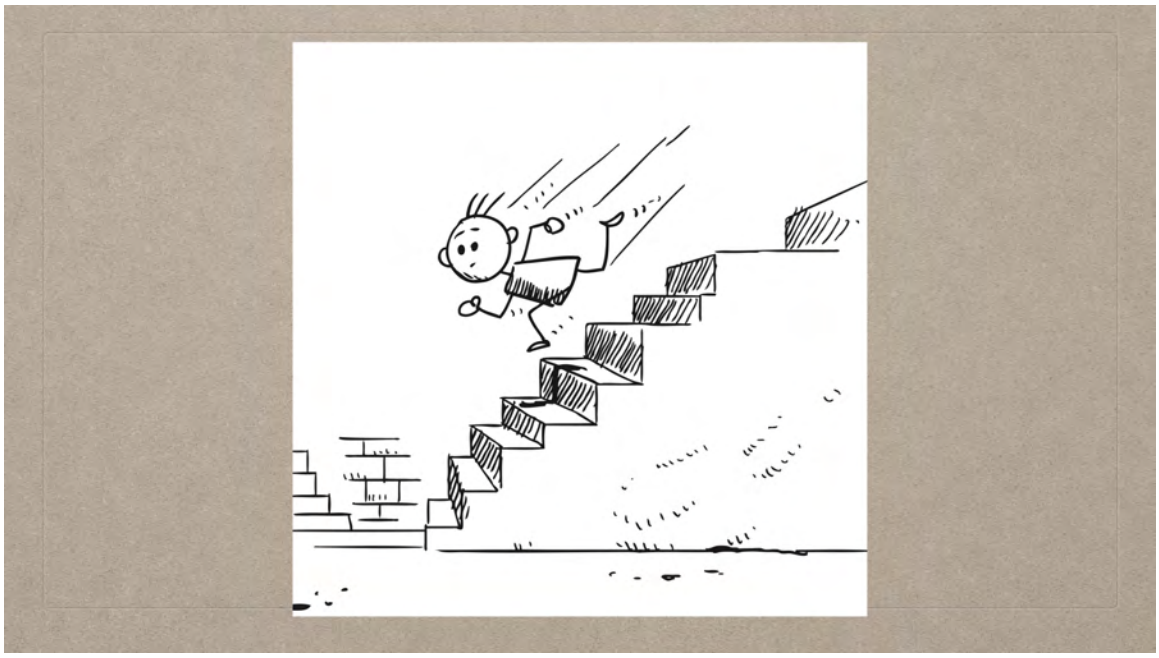


Figure 5 about here[20]

One day my older brother, who I revered then and still revere now, convinced me that if I ran fast enough, I could be upstairs and downstairs at the same time. As I recall, my efforts to find out for myself whether that was true lasted long enough to provide amusement for all the family. Though I can't say I succeeded in proving my brother's hypothesis at the time, I was pleased when I later learned enough about quantum mechanics to vindicate the failures of my earlier experiments.

I mention all this to convince you that my reflections today come out of a lifetime of watching how scientists and technologists work — sharing from personal experience both their inspiring dreams and their hopeless fantasies.

Figure 6 about here

My day job is at the Florida Institute for Human and Machine Cognition, or IHMC for short.[21] At IHMC, I feel very fortunate to work among researchers who are among the best in the world in their areas of specialty.

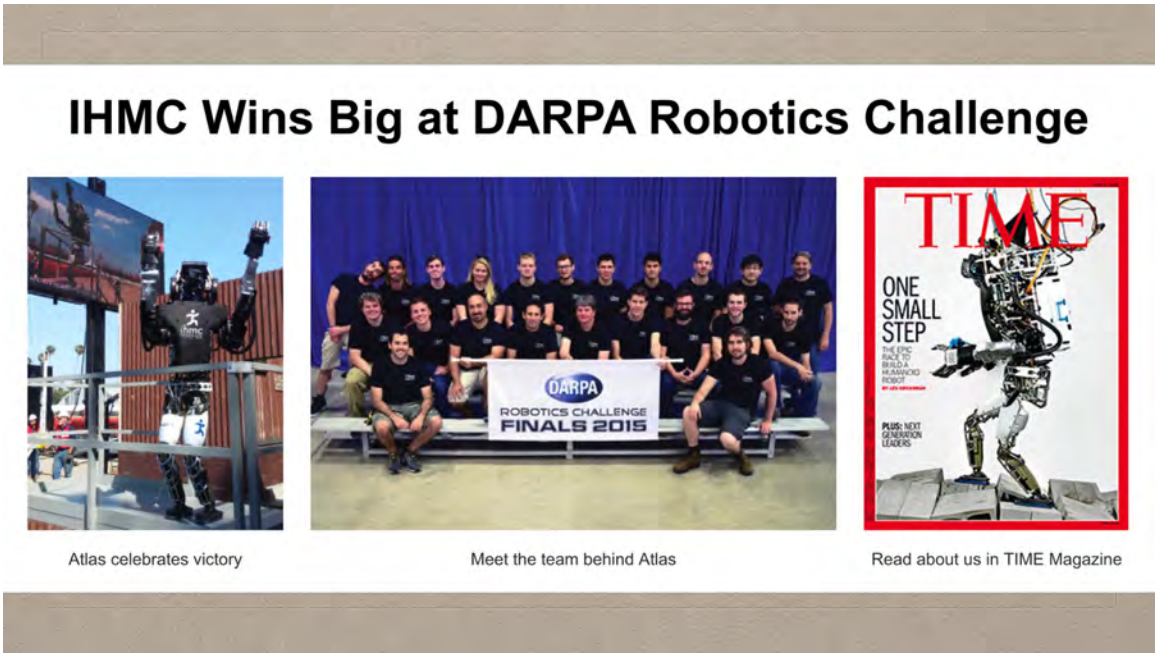


Figure 7 about here

One of our most exciting moments in 2015 was the final phase of the DARPA Robotics Challenge.[22] There, the IHMC team walked away with a million-dollar check and top honors among all participating American universities and research

institutions, and all but one international competitor. Researchers at IHMC are passionate about science and technology.

However, for the next few minutes, I'd like to share a few of the lessons I've learned, not about science and technology, but rather about *scientists* and *technologists*. If you understand scientists and technologists, you will be able to do a pretty good job in analyzing the news about science and technology, even if you do not understand the science and technology itself.

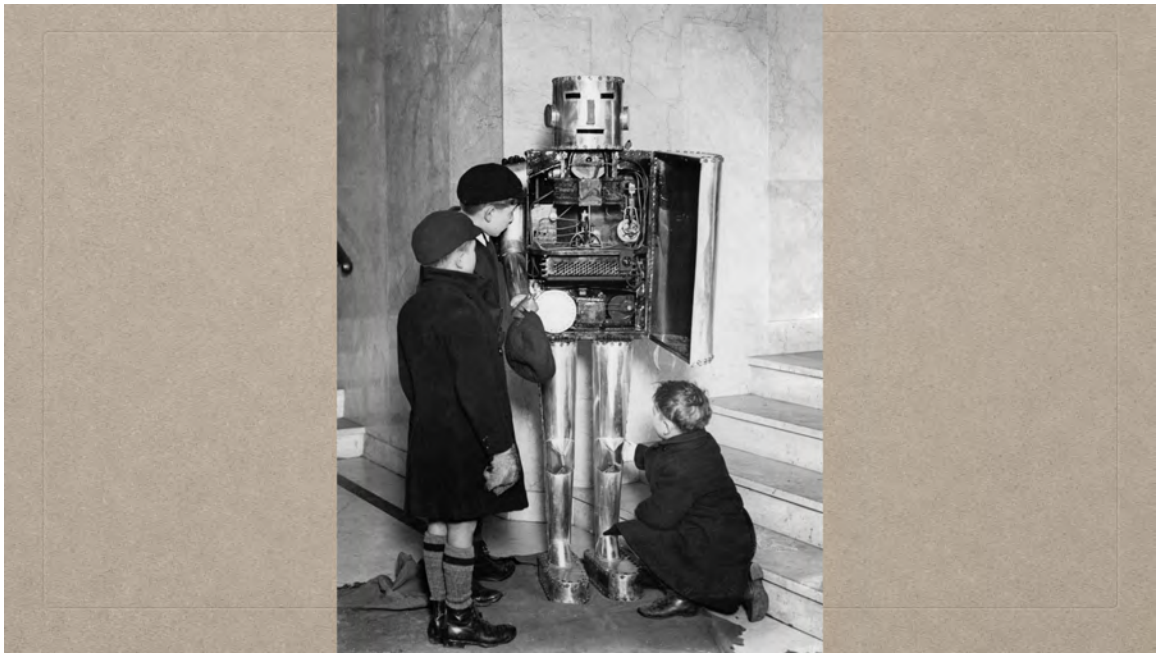


Figure 8 about here[23]

One thing I have learned about technologists is that they tend to be incurable optimists. For instance, consider the original caption on this photograph from the 1930s: "This steel man is near enough to accuracy to explain the physiology of the human frame."[24] The two students at left are no doubt counting the number of ribs to make sure they are all there. This reminds me of a question I once heard posed about why young Primary children are often asked draw pictures of their fathers to show them on Fathers Day. The answer? Because the fathers want to know what they really look like.

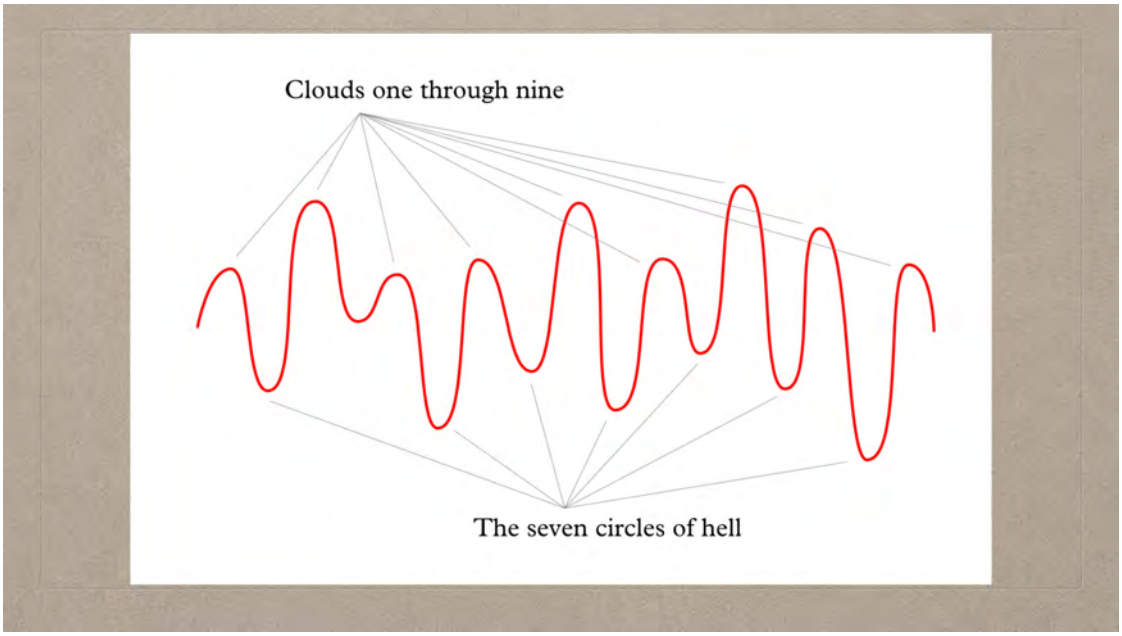


Figure 9 about here[25]

When you look at the ups and downs in the lives of scientists and technologists, as shown here, it becomes understandable why they might suffer from incurable optimism, even when the face of reality glares at them fiercely between the eyes.[26] As is often expressed, with unfeigned sympathy, "There goes another beautiful theory about to be murdered by a brutal gang of facts."[27] The professional lives of researchers are inherently unstable, and in many cases their stubborn, unreasonable optimism gives them courage to engage in tedious — and often discouraging — work every day.[28]

The stress of scientists and technologists is not merely a stress of the first order, like the kind that stems from high pressure and overwork, but also of the second order, which is something more existential in nature. Many researchers are passionate about the potential of their contributions, wanting to make a difference in important contemporary problems such as health, poverty, food production, and quality-of-life. Their stress is much like that of the struggling artist or of someone in the early, uncertain stages of a romantic relationship, "who really, really, really want[s] it to work, but lack[s] a clear model of *how*."[29] It is not just a matter of bulldozing one's way to success by working incredibly hard or of becoming more and more sure through experience about "about which path to take, but [rather] about [the uncertainty of] whether the paths (and the destination!) are even real."[30]

And then, if it weren't enough to be grappling with uncertainties relating to the *scientific* viability of the work, researchers usually have to be concerned just as deeply with the dizzyingly frequent changes in levels of public interest in the relevant ideas, which in turn drive the up and down trends of highly competitive funding in their particular lines of investigation.

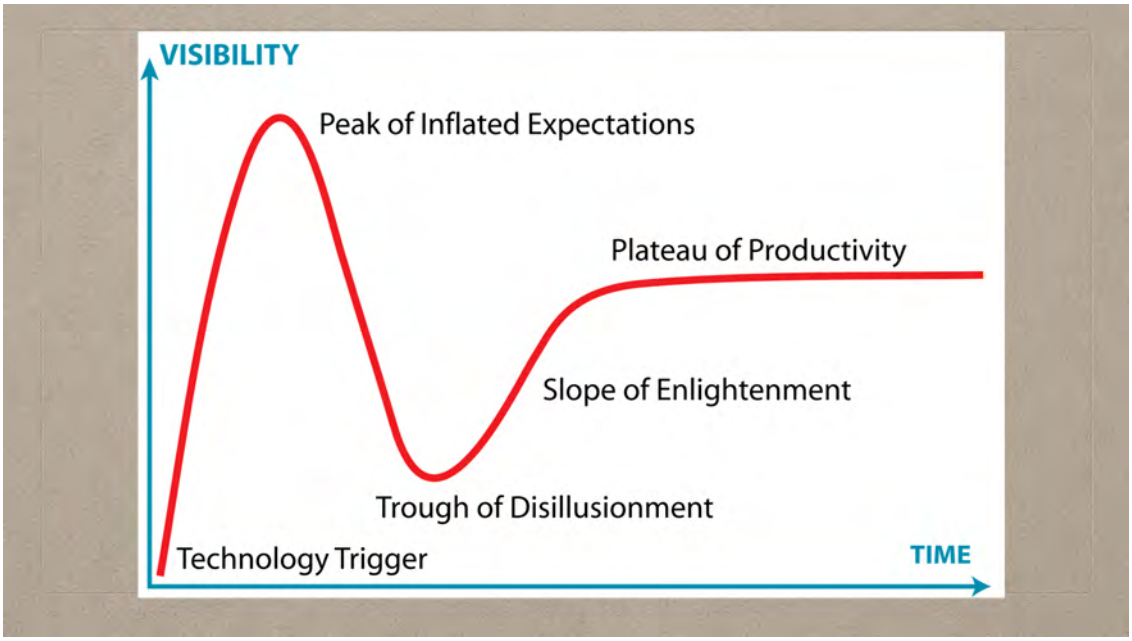


Figure 10 about here[31]

Each year the Gartner Group publishes an annual update to their hype cycle.[32] The curve illustrates how the most successful emerging technologies rapidly trend upward toward a peak of "inflated expectations" before suddenly dropping down into a "trough of disillusionment." Only a relative few such technologies sustain their popularity long enough to progress through a gradual "slope of enlightenment" and on to a "plateau of productivity."

Bad timing with respect to the hype cycle can be more destructive to the odds of success in executing a line of research than having a bad idea to begin with. The truth of this claim is evident in the sheer volume of bad ideas that are funded as soon as a given topic approaches the peak of the hype cycle.

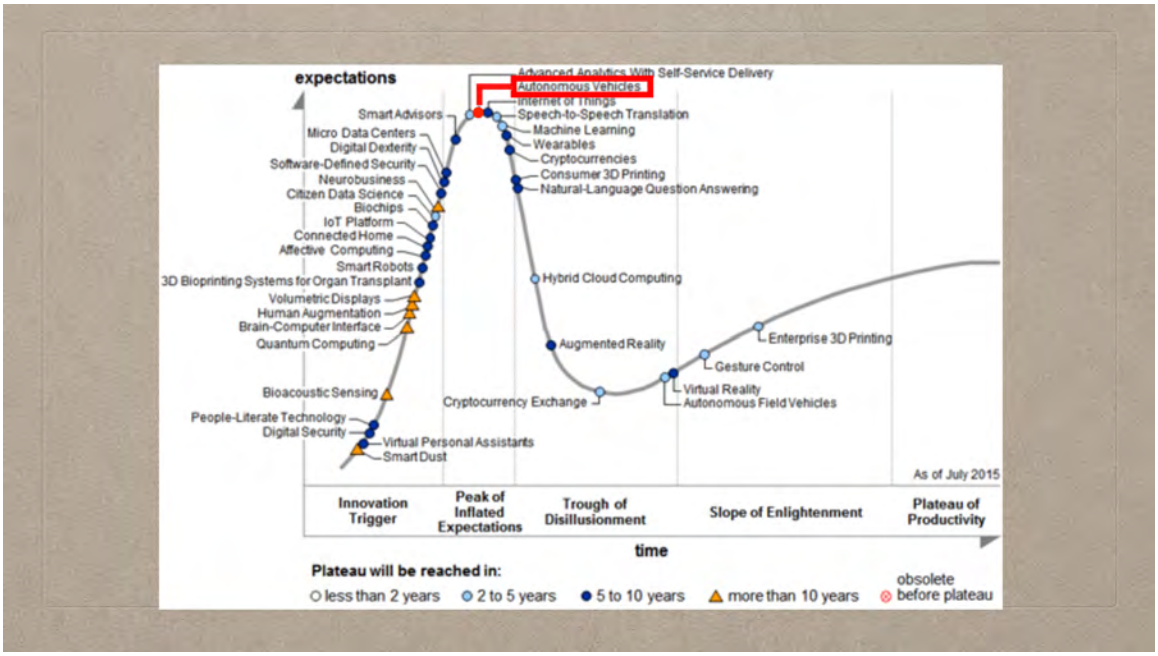


Figure 11 about here[33]

This brings us to our first example, which sits at the very top of the 2015 Gartner hype cycle update: self-driving cars.

Part 2: Self-Driving Cars

Figure 1 about here

By way of introduction to this theme, I want to state that I admire the courage and creativity of pioneers such as Sebastian Thrun of Stanford University whose tireless research and fearless advocacy of self-driving cars at Google both created groundbreaking technologies and opened up what will surely prove to be one of the biggest transportation developments of the coming century. In January 2016, I was honored to participate in the 174th Dies Natalis ceremonies at the TU Delft, the top Dutch university for science and engineering, where Thrun received an honorary doctorate. I highly commend to anyone, regardless of their background in science and technology, the inspirational video made of his remarks at that event, entitled “Moonshot Thinking.”[34]

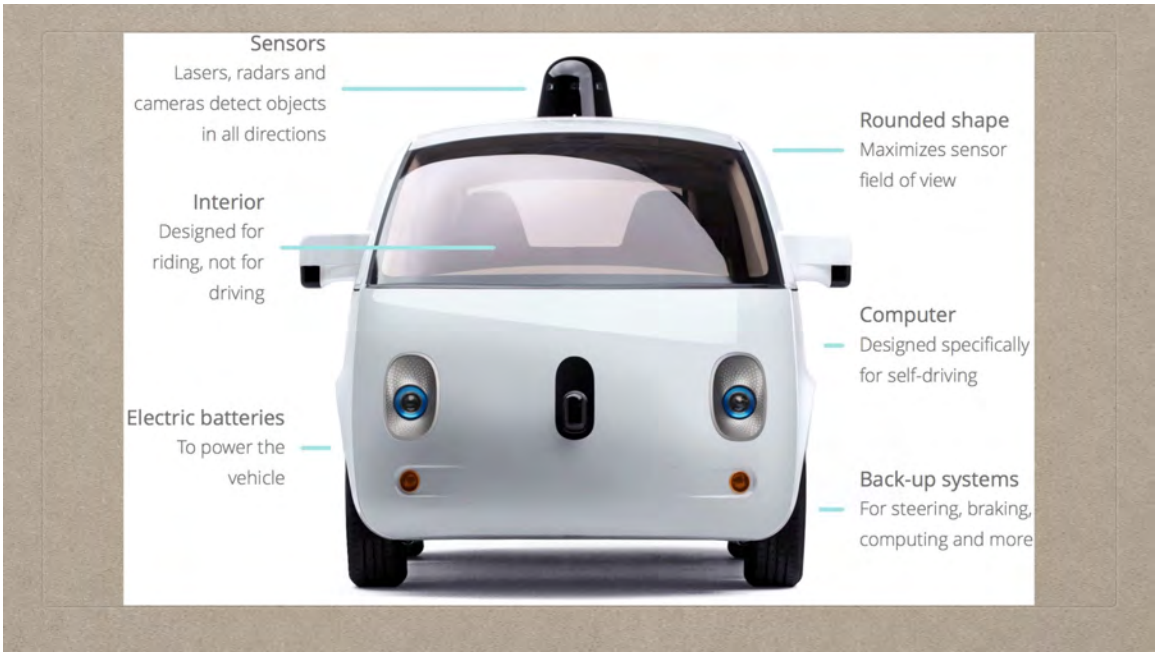


Figure 2 about here[35]

The Google Koala prototype that was publicly unveiled in the fall of 2015 is the most easily recognized self-driving car today. It is a successor to the early efforts of Thrun, and a tribute to him and to the many able researchers that have followed in his footsteps. With the fast-paced evolution of thinking that is currently taking place in the field, no doubt the vehicles of 2025 will look and function much differently than the prototypes of 2016.

The image tells you about the current features of the Google Koals but it does not explicitly reveal what features of an *ordinary* car have been removed, namely, the steering wheel, the gas pedal, and the brake pedal.[36] The omission of these three pieces of heretofore standard equipment is a specific example of the incurable optimism of technologists.

Success in fielding large numbers of *general-purpose* self-driving cars meaning cars that are intended to successfully negotiate the vast majority of situations that manually driven cars do today as opposed to cars that operate in specific, well-constrained niches — depends on solving several difficult problems. The biggest challenges are not in the basics of autonomous driving — getting from A to B. The devil is in the myriad details of unexpected events that can occur while driving.[37]

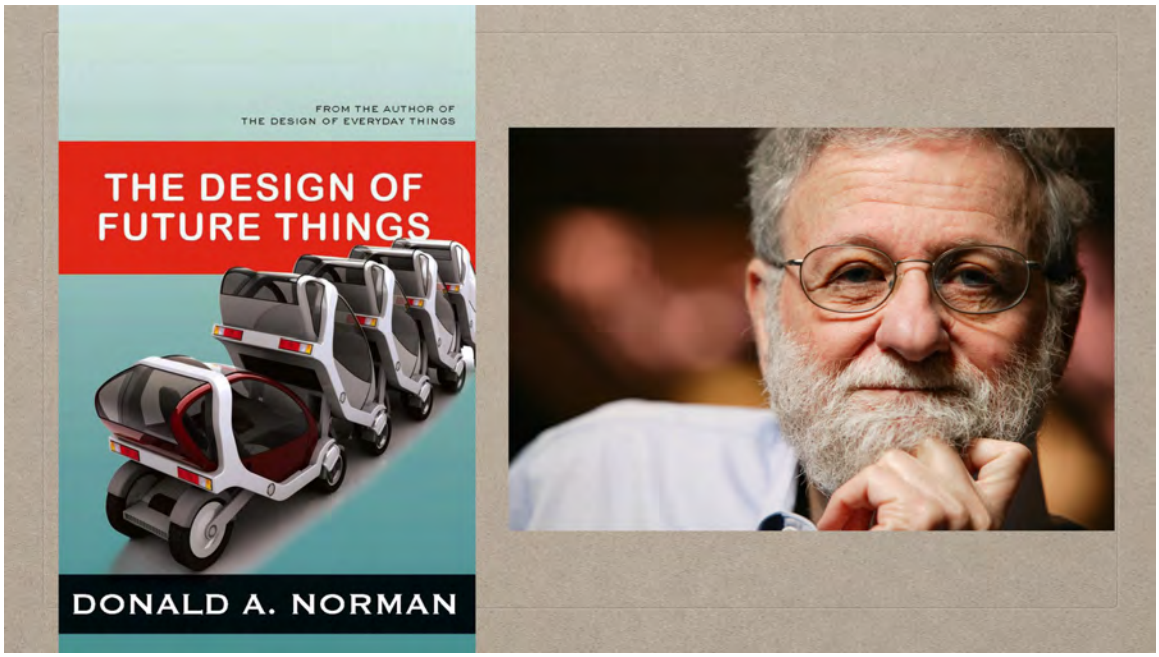


Figure 3 about here[38]

My colleague on the Nissan Science Advisory Council,[39] Donald Norman says it this way: "We know two things about unexpected events: first, they always occur, and second, when they do occur, they are always unexpected."[40] He continues:[41]

> The conflict between human and machine actions is fundamental because machines, whatever their capabilities, simply do not know enough about the environment, the goals and motives of the people, and the special circumstances that invariably surround any set of activities.

The California driving authority has tried to do something about this problem. As a result of ongoing discussions about the safety of autonomous vehicles, it has implemented rules whereby such vehicles are required "to have means whereby a person sitting in the car could intervene at any time, if the technology fails."[42] Sounds like a good idea, right?

Figure 4 about here[43]

Here's the rub: what Norman calls "halfway automation" or what other researchers sometimes call "the handoff problem." In defense of Google's apprehensions about allowing passengers to exert control in starting, stopping, and steering self-driving cars, "*halfway* automation" is sometimes a much bigger problem than *full* automation:[44]

> I once argued[, writes Norman,] that the current state of automation was fundamentally unsound because it was in the dangerous middle ground, neither fully automated nor fully manual. Either have no automation or full automation, I argued, but what we have today is halfway automation. Even worse, the system takes over when the going is easy and gives up, usually without any warning, when the going gets tough — just the reverse of what you would want. …
>
> If one cannot automate fully, then the automation that is possible must be applied with great care, sometimes not being invoked, sometimes requiring more human participation than is really needed in order to keep the human drivers informed and attentive.
>
> Full manual control of automobiles is dangerous. Fully automatic control will be safer. The difficulty lies in the transition toward full automation, when only some things will be automated, when different vehicles will have different capabilities, and when even the automation that is installed will be limited in capability. I fear that while the partial automation of driving will lead to fewer accidents, the accidents that do happen will be greater in magnitude, involve more cars, and exact a higher toll. The joint relationship between machines and their humans must be approached with caution.[45]

Will Nissan Beat Google and Uber to Self-Driving Taxis?

By Mark Harris
Posted 26 Feb 2015 | 15:43 GMT

Photo: Nissan

These Nissan taxis are conventional gasoline-powered vehicles, driven by humans. But Nissan is collaborating with NASA to develop a fleet of electric autonomous cabs.

Who will build the first robot taxis? Google has a working prototype but no experience in manufacturing cars (http://spectrum.ieee.org/cars-that-think/transportation/self-driving/googles-selfdriving-car-pals-revealed). Uber, meanwhile, knows the transportation business but has only just started working on autonomous vehicles (http://spectrum.ieee.org/cars-that-think/transportation/self-driving/uber-turns-from-google-teams-up-with-carnegie-mellon-on-selfdriving-cars) with Carnegie Mellon University.

Documents obtained by *IEEE Spectrum* suggest the first cab capable of driving itself (and that you won't feel obliged to tip) might be made by Nissan. In January, the Japanese automaker announced (http://spectrum.ieee.org/cars-that-think/transportation/self-driving/nasa-and-nissan-chase-selfdriving-car-technology) that it would be working with NASA to "demonstrate proof-of-concept remote operation of autonomous vehicles for the transport of . . . goods . . . and people." Using a California Public Records Act request, *Spectrum* has uncovered more details on the particular technologies Nissan and NASA plan to share and, more important, that the main goal of their collaboration appears to be the development of a fleet of remotely-supervised autonomous taxis.

Figure 5 about here[46]

The incurable optimism of researchers must be handled with extreme deftness and skill by traditional auto company CEOs. This is not only because they need to temper public expectations but also because many are hoping for partnerships with technology companies such as Google and Apple.[47] In a 2015 article entitled "Will Nissan beat Google and Uber to self-driving taxis?" Nissan's interest in research and development of fleet management services for autonomous vehicles was leaked to the public through a California Public Records Act request.[48]

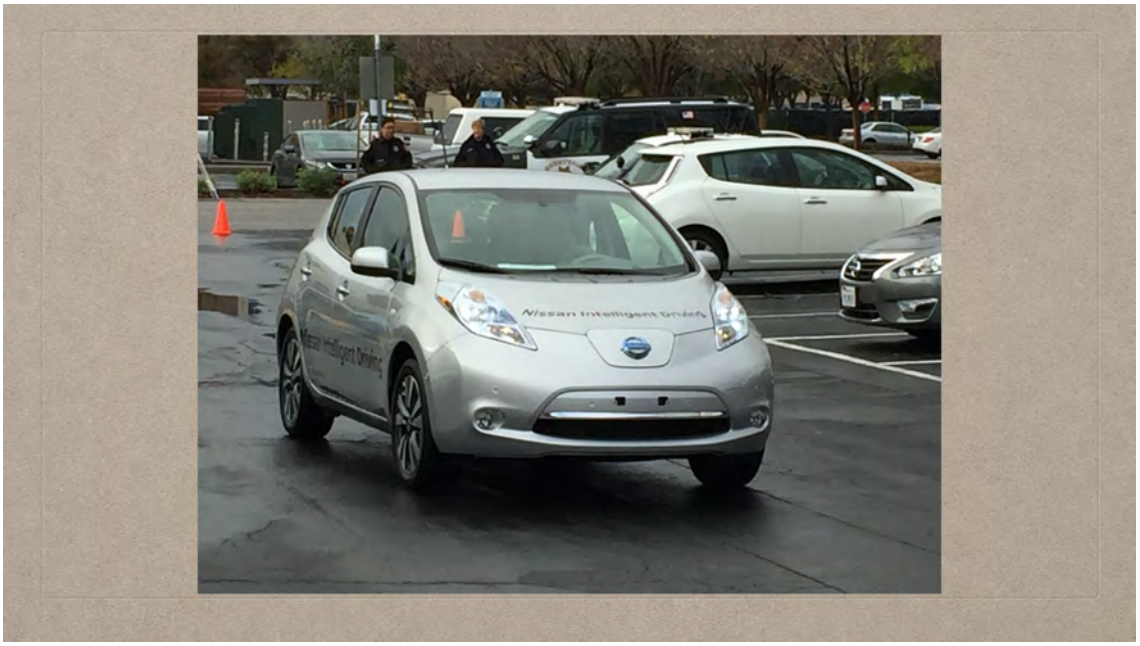


Figure 6 about here[49]

On the drizzly day of January 6, 2016, our research team participated with our colleagues at NASA and Nissan in a series of demonstrations of Nissan autonomous driving technologies to various executives, including the chairman and CEO of the Renault-Nissan alliance, Carlos Ghosn. Situations as commonplace as bad weather and standing water or snow on roadways can wreak havoc with self-driving automobile sensors, which is one reason why road testing occurs in California much more than it does in Alaska. Fortunately, the drizzle cleared up quickly, and the demonstrations went forward successfully.[50]

Figure 7 about here[51]

The next day, *The New York* Times reported Carlos Ghosn's announcement that Nissan:[52]

> would introduce ten new autonomous vehicles in the next four years.
>
> Elon Musk, the chief executive of Tesla, upped the ante. In a conference call with reporters …, he asserted that the so-called Autopilot feature introduced in the Tesla Model S last fall was "probably better than a person right now."
>
> Mr. Musk also said that within a year or two, it would be technically feasible to summon a Tesla from the opposite side of the country.
>
> But[, continued the *Times*,] there is a growing gap between what these executives are saying and what most people think of when they hear executives or scientists describing … driverless cars.

> What Mr. Musk and Mr. Ghosn are describing — cross-country-driving hyperbole aside — are cars with advanced capabilities that can help drive or even take over in tricky situations like parallel parking on a busy street.
>
> Truly autonomous cars that do all the work, like the bubble-shaped vehicles Google has been testing near its Silicon Valley campus, are still at least a decade away from ferrying people around town.

Mary (Missy) Cummings, another colleague who served for a time on the Nissan Science Advisory Council gave insightful, sobering testimony to the U.S. Senate Committee on Commerce, Science, and Transportation on March 15, 2016 of some of the "scenarios that highlight limitations of current self-driving car technologies," concluding with these words:[53]

> Let me reiterate that as a professor in the field of robotics and human interaction, I am wholeheartedly in support of the research and development of self-driving cars. But these systems will not be ready for fielding until we move away from superficial demonstrations to principled, evidenced-based tests and evaluations, including testing human/autonomous system interactions and sensor and system vulnerabilities in environmental extremes. To this end, in collaboration with private industry, [the National Highway Traffic Safety Administration] should be providing strong leadership and guidance in this space.

Part 3: Superintelligence Through Artificial Computational "Brains"

Now I'd like to say a few words about one of the most incredible example today of the incurable optimism of researchers, namely the building of what has been termed "superintelligence."[54]

Figure 1 about here

Tremendous progress in our imaginings about superintelligence has taken place in my lifetime. When I was a child, it was too far-fetched to think that anyone could actually *build* a superintelligence, so the best that science fiction could offer us was to help us imagine a *real* human brain, kept alive in a jar and tethered with wires, that was bent on either controlling or destroying the world. Thanks to the broadening of our imaginations in the computer age, we have substituted the outmoded idea of a real brain in a jar with two new and improved substitutes that have become the subject of countless blockbuster films: 1) the omniscient supercomputer, a completely artificial brain (discussed in this article, Part 3);[55] and 2) the omniscient mind, a natural human brain that has been uploaded to a network of supercomputers (discussed in the next article in the series, Part 4). Both of these new options for superintelligence — and a few others besides — are being hotly pursued by researchers.

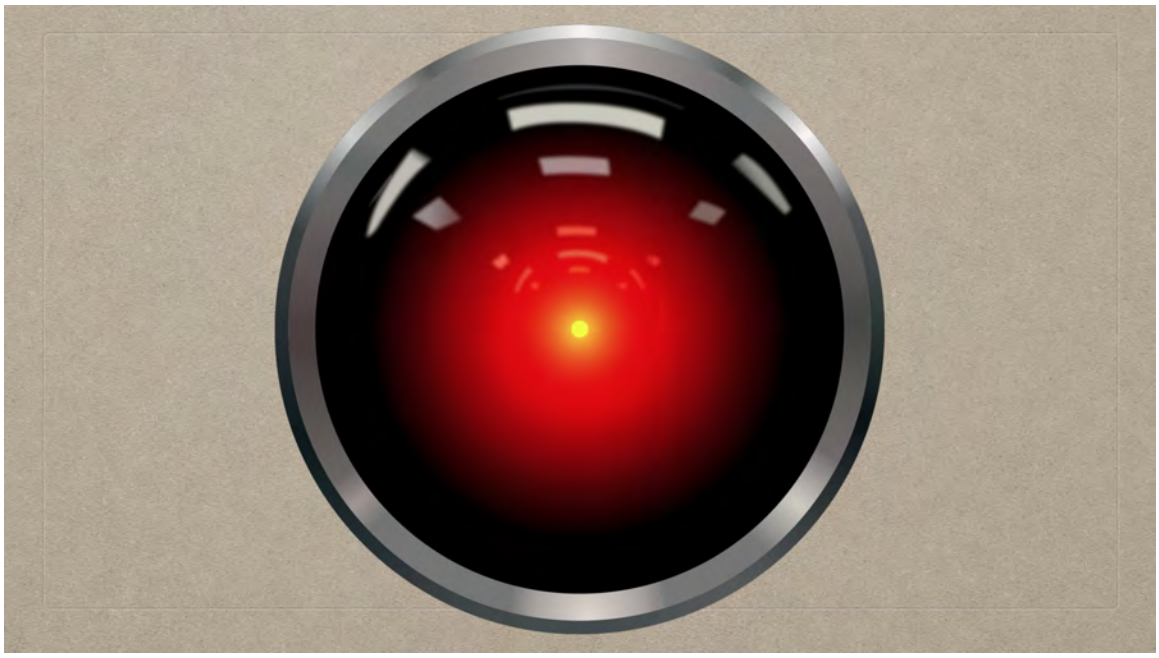


Figure 2 about here[56]

As to the omniscient supercomputer, the current front-runner is IBM Watson, which shares conceptual genes with Arthur C. Clarke's HAL 9000 from *2001: A Space Odyssey*. Although transposing the letters H-A-L one letter forward produces I-B-M, any deliberate connection was adamantly denied by Clarke, though later embraced by IBM.[57]

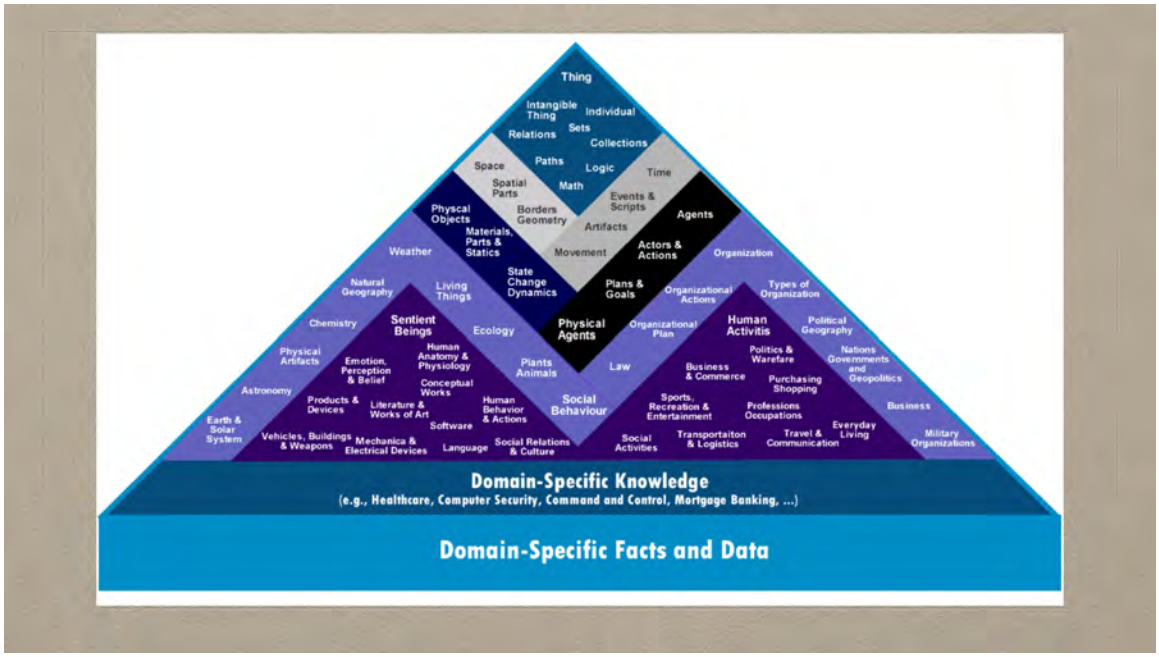


Figure 3 about here[58]

From a research perspective, Watson also shares conceptual genes with Doug Lenat's Cyc (as in en*cyc*lopedia), an ambitious multi-decade project to build a general purpose AI that has failed to yield the fruits its originators have always dreamed of.[59] However, unlike the current version of Cyc, for which bits of knowledge usually have been crafted by hand, the current version of Watson has the advantage of being able to ingest large swaths of the Internet to expand its knowledge base.[60] Fresh from the media frenzy that coincided with their win on *Jeopardy!* in 2011, IBM began announcing the beginning of "The Cognitive Era" in computing (which, not coincidentally, was shown on their timeline has having begun suddenly in 2011)[61] and news headlines shouted out memes like "IBM aims to build artificial human brain within 10 years."[62]

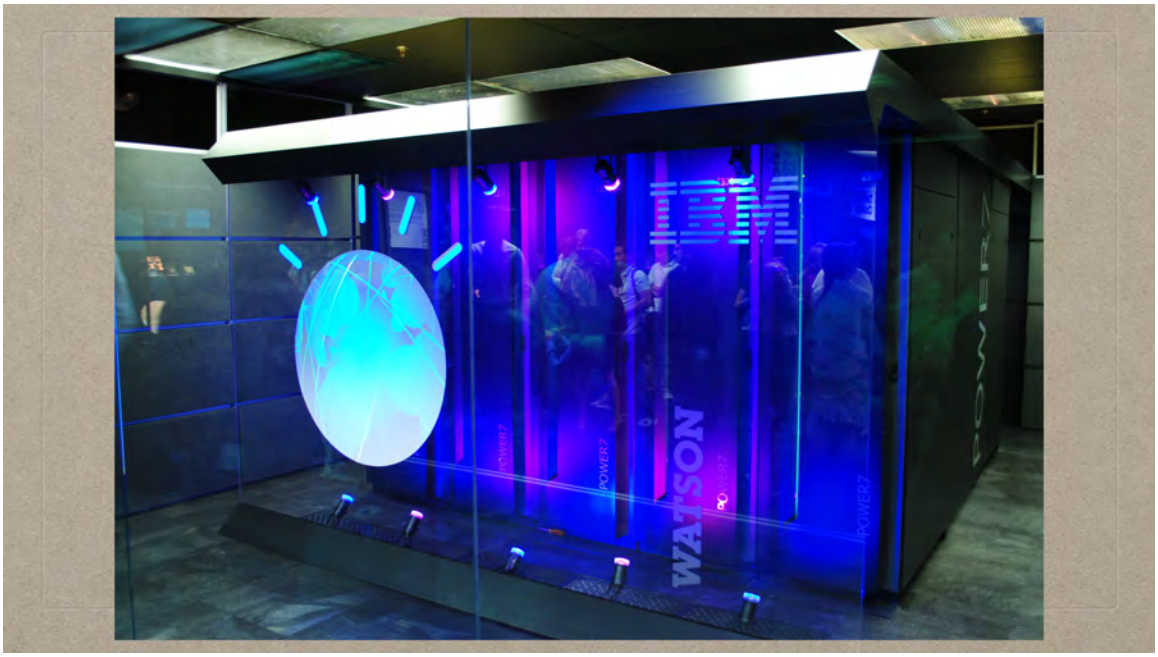


Figure 4 about here[63]

To the disappointment of IBM, Watson has not taken off with the speed and glory that they initially hoped since its public debut. Indeed, *Jeopardy!* has been the *only* application for Watson that has made much of a splash with the public. IBM's website currently lists only few dozen small company and institutional application partners, and a February 2016 article touts with unabashed optimism "future potential" much more than it parades details of its current successes over the last five years.[64] As a super-smart search engine, a capability for complex classification or diagnosis problems, or a natural-language-based analytic assistant, it has great potential. As a superintelligence that matches the ambitions of HAL, I predict it will continue to fall short for the foreseeable future. The top researchers at IBM must already know this; though the sales and marketing folks still seem to be in denial.

Among the challenges for any machine aspiring to HAL 9000 capabilities — in addition to whatever competence it may have been designed to demonstrate in its

particular area of expertise — is commonsense reasoning (e.g., "Who is taller, Prince William or his baby son Prince George? Can you make a salad out of a polyester shirt? If you stick a pin into a carrot, does it make a hole in the carrot or in the pin?"). Commonsense reasoning is a remarkable but unheralded capability that people rely on in nearly every situation but that has been recognized by AI researchers as "one critical area where progress has been extremely slow."[65]

Another underappreciated garden-variety human capability is to be able to sense and understand changes in the world or in a given situation that require adapting or abandoning the current course of action in order to respond to something more important or urgent (e.g., the problem the machine is currently working on has just been resolved by another means (or has not become irrelevant or unachievable for some reason), an earthquake has occurred, a fight has broken out, or a literal or figurative elephant has entered the room). Today's machines are not typically made to continually sense and understand the wide range of global and local phenomena to which people attend, let alone to make the kinds of appropriate adaptations to context that come almost naturally to humans. Moreover, since today's machines are not "aware" of the fact that the world itself is distinct from their limited and contextually impoverished model of the world — their particular "map" of the features of the world that are relevant to what they are designed to do —they must rely on humans to keep them and their models in tune with the changing dynamics of the real world in which they operate.[66]

Pragmatics — a word that researchers usually associate with the study of language but which is just as important for studying every other kind of action — is another challenging problem for any machine that needs to be understood by people or other machines.[67] The theoretical study of natural languages is usually divided into three areas: syntax, semantics, and pragmatics. Syntax is the study of the words of a given language and the rules that dictate how these words combine to form legal expressions (i.e., its grammar). Within "speech act theory" — often used as the basis for communication among intelligent machines — semantics and pragmatics combine to account for what the expressions of the language *mean*. For example, a sentence like "It is cold in this room" has both a syntactic analysis and a literal, semantic meaning which is constant across all of its possible uses: namely, that the temperature in the room is cold relative to the speaker.

However, pragmatics deals with the fact that the speech act that a speaker intends to perform by using this sentence depends on the context of its utterance. The sentence could be used to state a fact, request that the listener close a window, warn the listener not to enter the room, ironically state just the opposite of what the statement means (e.g., saying that the room is cold for humorous effect when the room is actually very hot), or for some other reason. In fact, natural language utterances are often used for several purposes at once — including strengthening or weakening whatever social relationships hold among those involved in the conversation (e.g., deliberately saying something in a way

that will be understood one way by some of the listeners and a different way by others).

Obviously, it is much easier to design machines that are good at simply understanding the literal meaning of speech and actions than to endow them with the more human-like ability to understand the myriad implications of making a statement or performing an action in a specific situation at a given time in the presence of particular individuals. The following story illustrates both the power of that the proper exercise of knowledge about pragmatics can have in achieving desired results and the current limitations that machines have in that very respect.

Figure 5 about here[68]

Besides Watson, you may also remember another famous game-playing computer from IBM from two decades ago named Deep Blue.[69] In 1996, Garry Kasparov beat Deep Blue, winning three matches and drawing two:[70]

Figure 6 about here[71]

> The next year, he played against a new and improved Deep Blue and lost the match. Once again, the psychological toll of facing off against an inscrutable opponent played a key role. Although he easily won the first game, Deep Blue dominated the second. Kasparov … was visibly perturbed — sighing and rubbing his face — before he abruptly stood and walked away, forfeiting the match.
>
> He later said he was again riled by a move the computer made that was so surprising, so un-machine-like, that he was sure the IBM team had cheated. What it may have been, in fact, was a glitch in Deep Blue's programming: [In 2014, one the designers of Deep Blue revealed what he believed happened:] Faced with too many options and no clear preference, the computer chose a move at random. … [T]he move that threw Kasparov off his game and changed the momentum of the match was not a feature, but a bug.[72]

What lessons could be drawn from this story? The first and most obvious lesson is that it is very difficult to anticipate and understand the size of the impact that a seemingly innocuous design rule (i.e., "choose a random move when there are too many options and no clear preference among them") may have in a specific, unforeseen situation. The second lesson is that if Deep Blue had been intelligent enough to understand the pragmatic effect of an inscrutable, random move on its opponent (something that a human expert might plausibly have anticipated), its move could have been heralded as a brilliant feature rather than derided as a bug.

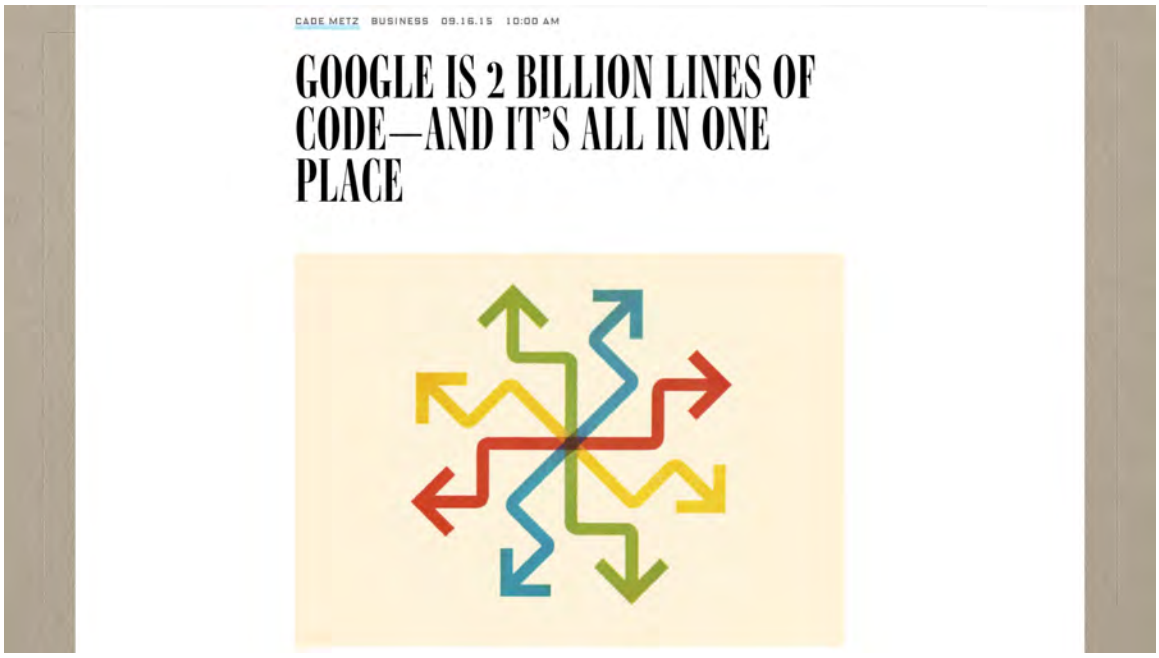


Figure 7 about here[73]

Fast-forward to 2016. Google, a company that runs off of 2 billion lines of code "in a single code repository available to all 25,000 Google engineers,"[74] now dominates large-scale computing. "By comparison, Microsoft's Windows operating system — one of the most complex software tools ever built for a single computer, a project under development since the 1980s — is likely in the realm of 50 million lines"[75] — only 2 ½ per cent of the size of Google's shared code base.

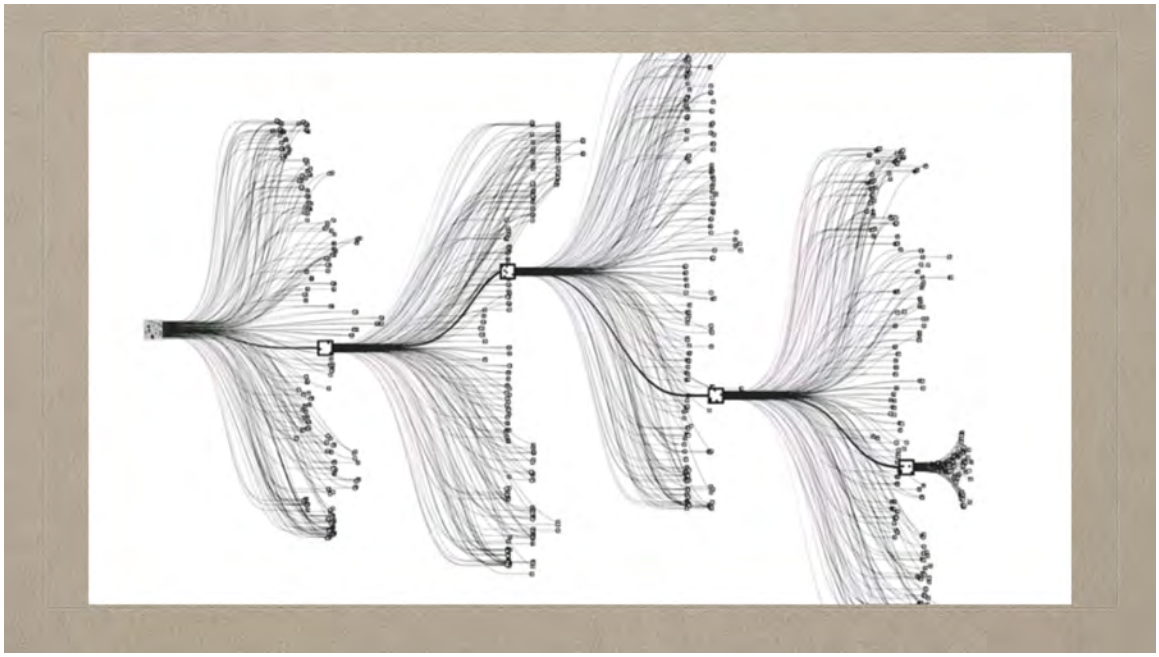


Figure 8 about here[76]

By its formidable computing prowess, Google has well earned the right to challenge the human champions of Go, reputed to be the most complex game ever invented. Go is exponentially complex, with each move in the game presenting another 250 possible moves.[77] For this reason, mere "doublings in computing power and Monte Carlo approaches have been … inadequate."[78] In January 2016, when Google's AlphaGo program beat the European Go champion 5-0, one of my research colleagues quipped: "My reaction when this happened was the same as Ken Jennings when beaten by Watson at *Jeopardy!* — 'I for one welcome our new computer overlords.' But first I want to see them win the pie-eating contest. ;)"[79] Two months later, AlphaGo beat Lee Sedol, a world-renowned South Korean master of Go.[80] As of 20 June 2016, AlphaGo was listed as the number two player in the world, according to GoRatings.org.[81]

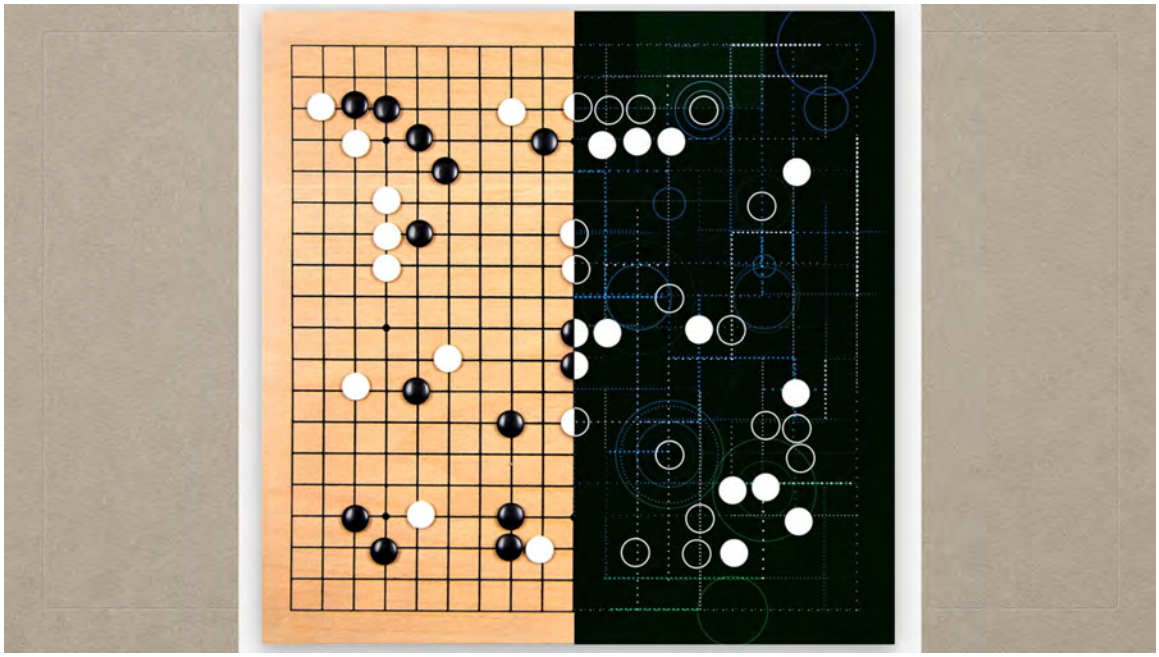


Figure 9 about here[82]

What makes all this important is not merely that AlphaGo was able to play successfully against formidable human opponents, but rather the degree to which it represented a significant shift in emphasis in *how* the game was played by the machine. My son Robert W. Bradshaw, who works for Google and knows a lot more than I do about AlphaGo's approach, outlined two reasons why an alternative was needed to typical game-playing strategies that rely largely on high-powered look-ahead searches through the space of possible future moves to see what current move would be most likely to lead to a later advantage. First, because the search space of possible moves in Go is so large, simply outpacing the human through extensive search is no longer feasible — there is no practical alternative at the current time but to use machine learning to a much greater degree. Second, because there are relatively few expert-level games available for the machine to learn from, machine-learning strategies needed to be pushed further than ever before in order to yield more results with less training data. The

aspiration of the researchers is for the machine to develop something "akin to an intuition about what good positions and moves are." Although this is not the first time that people have tried to attack the problem of playing Go using machine learning, it is the first time they've figured out how to do it effectively — and the value of the novel techniques that have come out of this research do not seem to be confined to playing Go. Writes Robert:[83]

> I personally find this most exciting because a lot of improvements in [machine learning] seem to have been due to being able to train on orders of magnitude more data (which, of course, is non-trivial theoretical and engineering challenge), but once you've trained on (say) all digitized bilingual text in the world there's not much further to go from there. These techniques are starting to explore the path of more effectively extracting "intelligence" out of (relatively) smaller corpora of data. It's also techniques like this that will allow it to produce results *better* than the data it trained on, which is a more fascinating proposition.

Figure 10 about here[84]

Despite such exciting advances, an panel sponsored by the United States White House Office of Science and Technology Policy concluded in May 2016 that "A. I. research is still far from matching the flexibility and learning capability of the human mind. … 'The A. I. community keeps climbing one mountain after another, and as it gets to the top of each mountain, it sees ahead still more mountains,'" [85] summarized one of the scientists. Another researcher observed that "attention-getting feats like Google's AlphaGo program … had plenty of humans behind the machine doing the work" — at least for the foreseeable future.[86]

Part 4: Superintelligence Through Natural, Uploaded Brains (Can We Become Immortal and Superintelligent By Uploading Our Brains to Supercomputers?)

Figure 1 about here[87]

An alternative to omniscient supercomputers that leverage traditional computational approaches to achieve superintelligence was introduced in a series of books by Ray Kurzweil, the most well known of modern-day transhumanists. There is a wide variety of opinion on the definition and goals of Transhumanism, but most would probably agree with the relevant *Wikipedia* article that it has as a major focus "the potential benefits and dangers of emerging technologies that could overcome fundamental human limitations."[88] This would typically include the idea "that human beings may eventually be able to transform themselves into beings with such greatly expanded abilities as to merit the label 'posthuman.'"[89]

Ray Kurzweil has achieved notoriety for his technology predictions, which he claims have been accurate 86-95% of the time.[90] However, the discrepancies between his self-assessments and the assessment of others of his accuracy resemble the differences between Donald Trump's claims about Trump University and the claims of everyone else.[91] As one critic concludes: "On close examination, [Kurzweil's] clearest and most successful predictions often lack originality or profundity. And most of his predictions come with so many loopholes that they border on the unfalsifiable."[92]

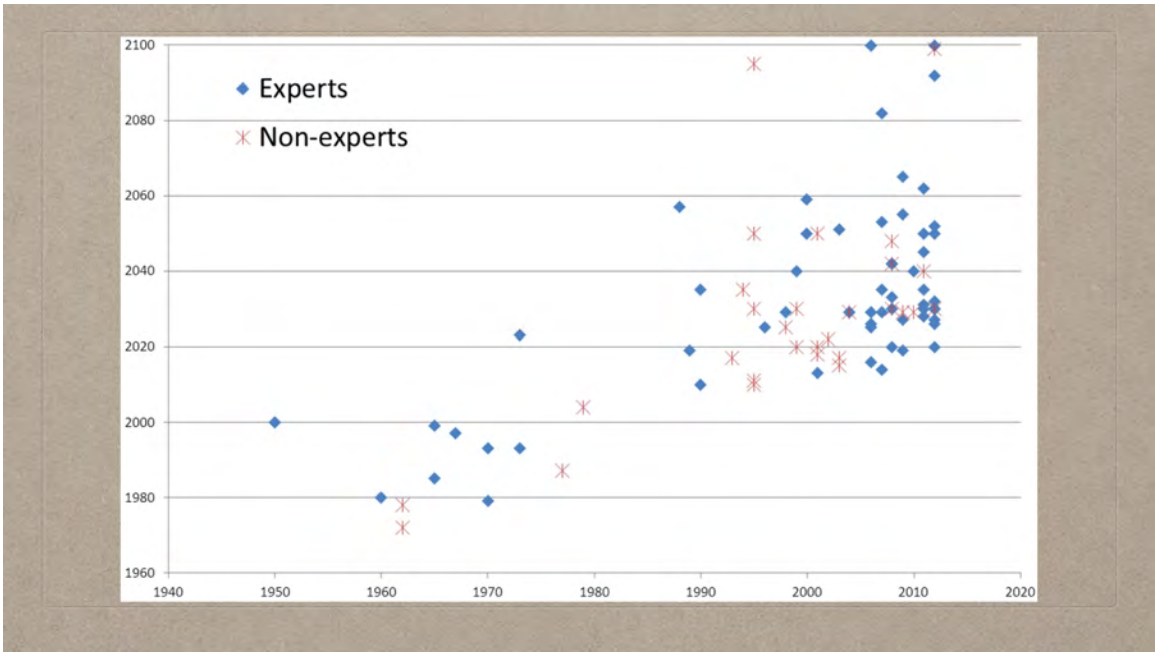


Figure 2 about here[93]

In Kurzweil's defense, other experts fare equally poorly in their predictions about the future of AI, differing little from the opinions of non-experts in what they say or how accurate they are.[94] For example, this graph shows experts' and non-experts' median-estimates for when "human-level" AI will appear, graphed against the date of prediction.[95] The predictions of both experts and non-experts are all over the map.

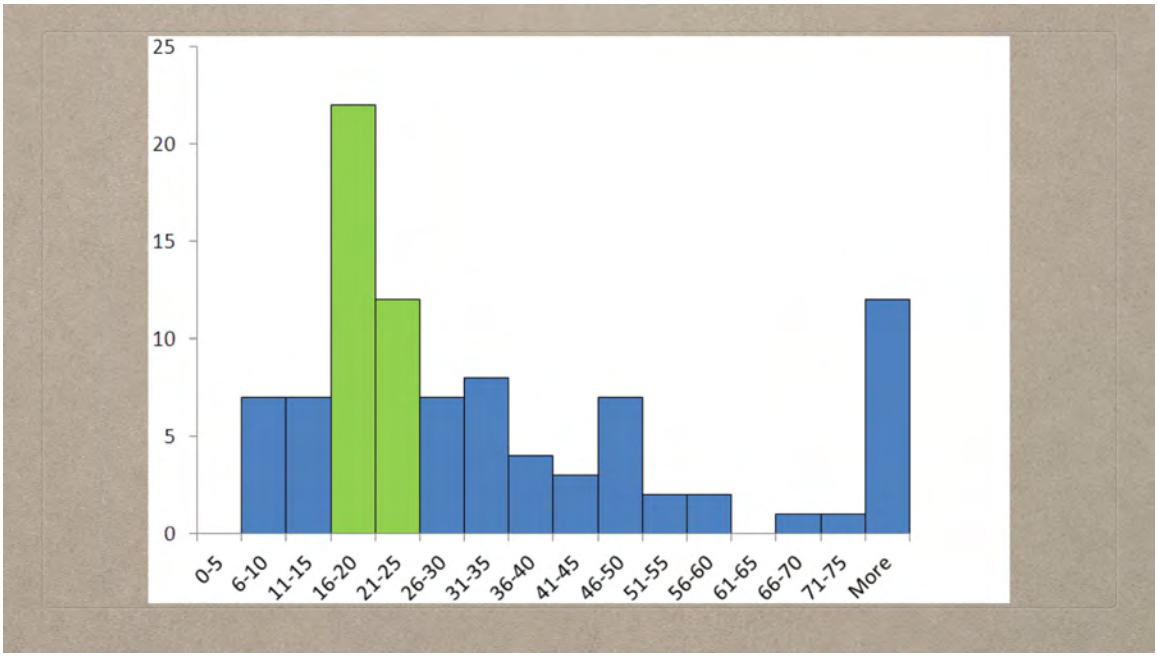


Figure 3 about here[96]

Notably, “there is a strong tendency to predict AI within 15 to 25 years[, regardless of] when the prediction is made.”[97]

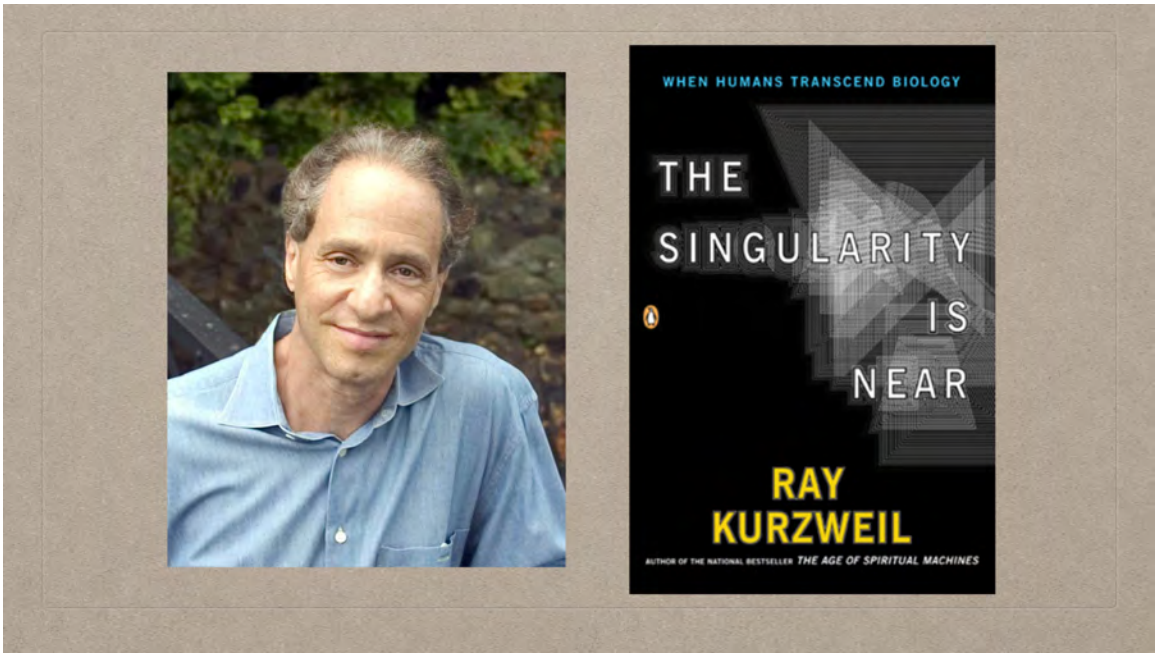


Figure 4 about here

In Kurzweil’s 2005 book, *The Singularity is Near: When Humans Transcend Biology*,[98] he predicts a future where, in “[f]using themselves with machines, humans can leave the flesh behind.”[99] In the meantime, “set[ting] out a plan of diet, exercise, vitamin supplementation, and preventive medical care” will, he believes, “enhance longevity to the point where technology can overcome mortality.”[100] In other words, Kurzweil is doing everything he can to live long enough in mortality so that he can make himself *immortal* through technology.

Overall, John Gray sees Kurzweil’s program being “best understood as a version of process theology.”[101] “It is not essentially different from Gorky’s fantasy of humans evolving to become pure thought. … The virtual afterlife is a high-tech variant of the Spiritualist Summerland, while accelerated evolution in cyber-space is an updated version of Myers’ Victorian dream of progress in the after-world.”[102]

LDS scholar and academic physician Samuel M. Brown has raised concerns about programs that seek to increase longevity and promote athletic, cognitive, and psychiatric enhancements to the body through science and technology. While acknowledging that there is a place for biomedical interventions to relieve suffering and to enable individuals to perform normal human functions, and also realizing that Mormonism is at heart a program for human betterment in both the spiritual and physical dimensions, Brown expresses a few of his many concerns as follows:[103]

For some, framing enhancement as the medical approximation of resurrection will make biomedical enhancements seem like nothing quite so much as the Tower of Babel narrative, when, according to early Latter-day Saints, people sought to build their own ladder to heaven on the plain of Shinar. From this Babel perspective, believers could argue that it is the one who makes us immortal rather than the mere fact of immortality that matters most. The perfect immortality of the afterlife comes through Christ and a moral transformation, while the perfect immortality of biomedical enhancement comes merely at the price of purchased technology. Mormons could argue that God has already "enhanced" Enoch's city, the Apostle John, and the Three Nephites. To turn that holy process into the equivalent of a steroid-augmented athletic context seems a sacrilege. Many Latter-day Saints belive that we should focus on changing our hearts; in his due time, God will change our bodies. In the other envisioned enhancement of the body, there is no attendant change of heart. …

How, in [the] calculus [of proponents of science-based programs for human enhancement], does one distinguish an enlightened society of mortals from a benighted society of immortals? What does it mean to live smart, healthy, and long without a soul (either in the metaphysical or metaphorical sense)? For believers in the Christian scriptures, this would seem to be a textbook case of what Jesus described as losing life in the attempt to save it.[104] In the absence of an overarching system of meaning, what makes [their] goals any less arbitrary than the goals espoused by others — such as Michael Sandel or Leon Kass — to experience the emotion of humility or to appreciate the poignancy of our temporary existence?

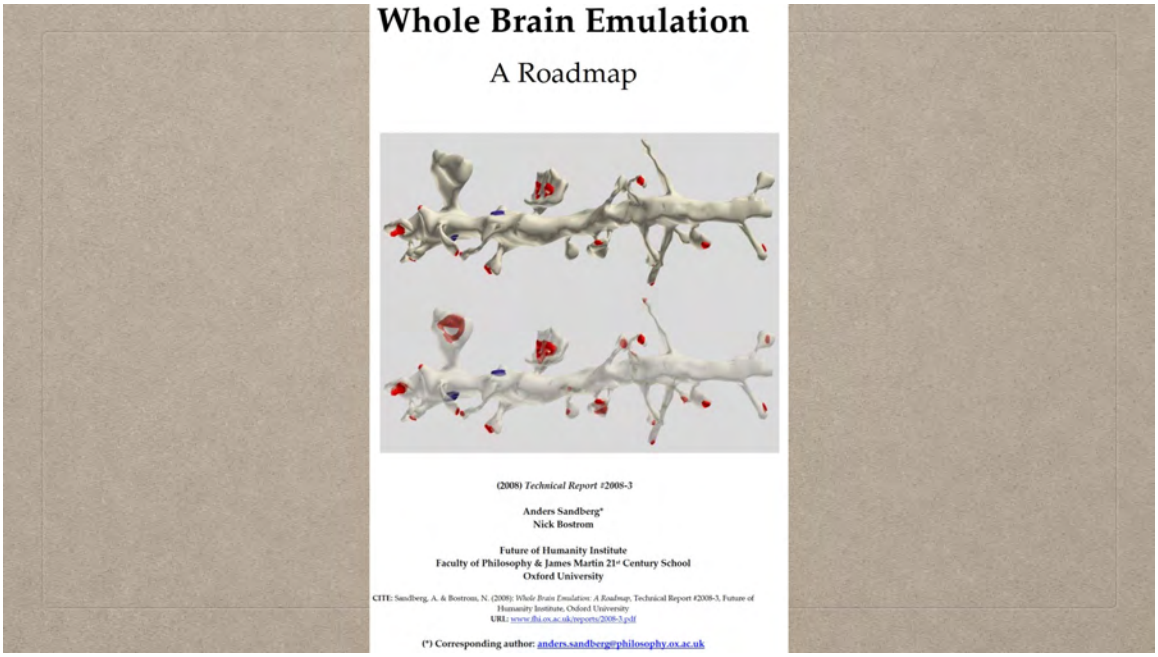


Figure 5 about here[105]

Apart from the philosophical, ethical, and theological overtones of attempts to achieve superintelligence and immortality through technology (including, for LDS believers, the question about how the spirit, mind and body relate), what can be said about the scientific feasibility of uploading our physical brains to a computer? There has been some credible thinking on this topic, as well as some very unsound proposals that Brown describes as evincing "an almost Pollyannish certainty that biomedical science will succeed [that] gives [such arguments] an air of science fiction."[106] Regardless of one's opinion on the ultimate prospects of success of programs that seek to enhance or ultimately re-embody the brain in silicon, nearly every expert on the topic agrees at least that the many remaining challenges would not be overcome "in the near future."[107] Before you can upload a brain, you need to be able to model it adequately — that challenge is the subject of the next article in this series. Before we leave the current topic, however, there are a few additional things that must be said.

Figure 6 about here[108]

What Would We Do With a Thousand Uninterrupted Years of Life?

As I have reflected on the possible complications of what might happen if our lives on earth could be extended indefinitely through technology, I have sometimes mused about the reasons that humans were made in such a way that we must spend a significant fraction of time each day in sleep. Could it be, I wonder, that if our bodies could do without sleep and we could be up and about nonstop, 24 by 7, we'd use the extra time to create more mischief and get ourselves into even deeper trouble than we already do? This was exactly the situation of the legendary Watchers we read about in Jewish pseudepigrapha. Hugh Nibley summarized accounts that describe their era as being:[109]

> … the time of great intellectual as well as material sophistication.[110] … The leaders of the people devoted most of their wealth to all kinds of engineering projects for controlling and tempering nature. But the Lord altered the order of creation, making the sun rise in the west and set in the east, so that all their plans came to naught.[111] The idea of controlling the environment independently of God was not so foolish as it sounds, says the *Zohar*, "for they knew all the arts … and all the ruling chieftains [archons] in charge of the world, and on this knowledge they relied, until at length God disabused them by restoring the earth to its primitive state and covering it with water."[112] Rabbi Isaac reports: "In the days of Enoch even children were acquainted with these mysterious arts [the advanced sciences]. Said R. Yesa: If so, how could they be so blind as not to know that God intended to bring the Flood upon them and destroy them? R. Isaac replied: They did know" but they thought they were smart enough to prevent it. "What they did not know was that God rules the world. … God gave them a respite all the time that the righteous men Jered, Methuselah, and Enoch were alive; but when they departed from the world, God let punishment descend …, 'and they were blotted out from the earth.'"[113]

Figure 7 about here[114]

While it may be true that some of us, if we were granted immortality in our current, imperfect state, might use our time like the legendary Watchers — to push science and technology aggressively in foolish and shortsighted directions — I think that most of us would instead simply languish in more mundane fashion, wasting our time in the same kinds of selfish and worthless pursuits we are drawn into right now. As evidence for such a proposition, we might apply the words of Amulek, who reminds us of the obvious yet often ignored truth that "that same spirit which doth possess your bodies at the time that ye go out of this

[mortal] life, that same spirit [would] have power to possess your body in that eternal world"[115] of unending life. Likewise, D&C 88:28 is very clear: "your glory [in the resurrection] shall be that glory by which your bodies are quickened" in this mortal sphere. We will not suddenly and automatically advance from telestial to celestial thinking and living at the moment we are called forth from the grave.

Figure 8 about here[116]

Playing off the overoptimistic title of William Blake's *Marriage of Heaven and Hell*, C. S. Lewis entertainingly depicts the wide gap that actually exists between them in *The Great Divorce*. Lewis describes a telestial world with "mean streets" that is "always in the rain and always in evening twilight"[117] — in other words, a world not too much unlike our own. "However far I went I found only dingy lodging houses, small tobacconists, boardings from which posters hung in rags, windowless warehouses, goods stations without trains, and bookshops of the sort that sell *The Works of Aristotle*."[118]

Figure 9 about here[119]

The town was devoid of people, except for a contentious queue of people waiting for a bus. Where was everyone else? Someone explains:[120]

> The trouble is that they're so quarrelsome. As soon as anyone arrives [in town] he settles in some street. Before he's been there twenty-four hours he quarrels with his neighbor. Before the week is over he's quarreled so badly that he decides to move. … If by any chance the street is full, he goes further. But even if he stays, it makes no odds. He's rue to have another quarrel pretty soon and then he'll move on again. Finally he'll move right out to the edge of the town. … That place where we caught the bus is thousands of miles form the Civic Center where all the newcomers arrive from earth. … [The people who came long ago have] been moving on and on. Getting further apart. They're so far off by now that they could never think of coming to the bus stop at all.

Figure 10 about here[121]

The bus departs regularly for heaven. A few go there every day, but hardly anyone stays. They prefer living in hell.[122]

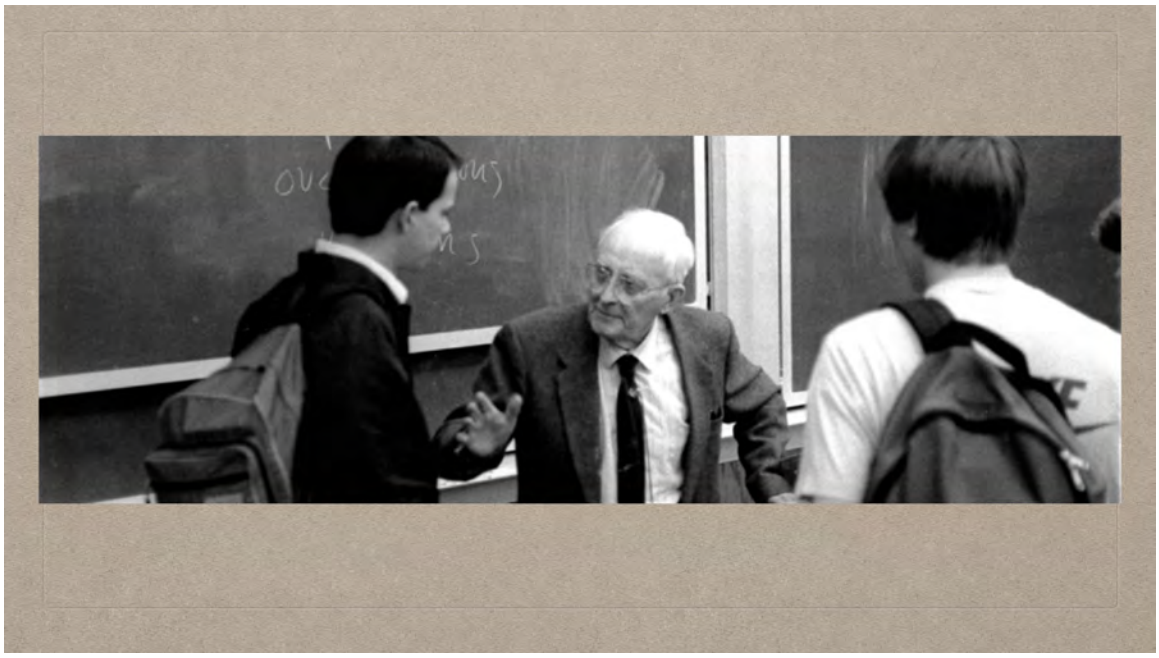


Figure 11 about here[123]

Some years ago, Hugh Nibley gave the students in his BYU honors class an unusual midterm assignment. He described that experience as follows:[124]

I asked them … to assume that they had been guaranteed a thousand uninterrupted years of life here on earth, with all their wants and needs adequately funded: How would you plan to spend the rest of your lives here? I explained that this is not a hypothetical proposition, since this is the very situation the Gospel puts us in. Whether we want to or not, we are doomed to live forever — even the wicked — for "they cannot die."[125] In accepting the Gospel, we are already launched into our eternal program. … We are taught to think of ourselves here and now as living in eternity, and how can it be otherwise, since the contracts we make and the rules we live by are expressly "for time and eternity"?

So I asked them, "How are you going to get started on that thousand-year introduction to a timeless existence?" … Here are some typical answers:

> Overwhelmed by the proposition … [I] would have to refuse it. …
>
> First I would go crazy, … then I would be bored after 100 years. …
>
> I would not want to live here that long. I would make long-term investments in the money markets, … would complete my education in business, get an MBA, would find a part-time job and teach my children the value of work. …
>
> It's not a nice question, the pressure would be too great from people who would like money from me. How should I pay tithing on it? How would I use all that money? [For this person[, notes Nibley,] the whole question is an economic one.]
>
> I would spend my time in recreation with some serious moments. For a sense of success, I might build or write something.
>
> I don't know if I would want a thousand years. … Travel, study, and teach.
>
> Could be a blessing or a cursing; I would excel in athletics and general education, would procrastinate a good deal, live in the style of the well-to-do, … shopping, camping, dancing. …
>
> I could do nearly everything there was to do several times over. Perform service and drive a Porsche 911. …
>
> I would turn it down. This life is okay, but I am anxious to get on with my progression in the hereafter. [Doing what? [asks Nibley.] This *is* your progression into the hereafter!]

And so it goes[, continues Nibley]. No wonder [Shakespeare's] Hamlet finds a world of such people "weary, stale, flat, and unprofitable."[126] "What is a man"

he asks, “if his chief good and market of his time be but to sleep and feed? A beast, no more.”[127] …

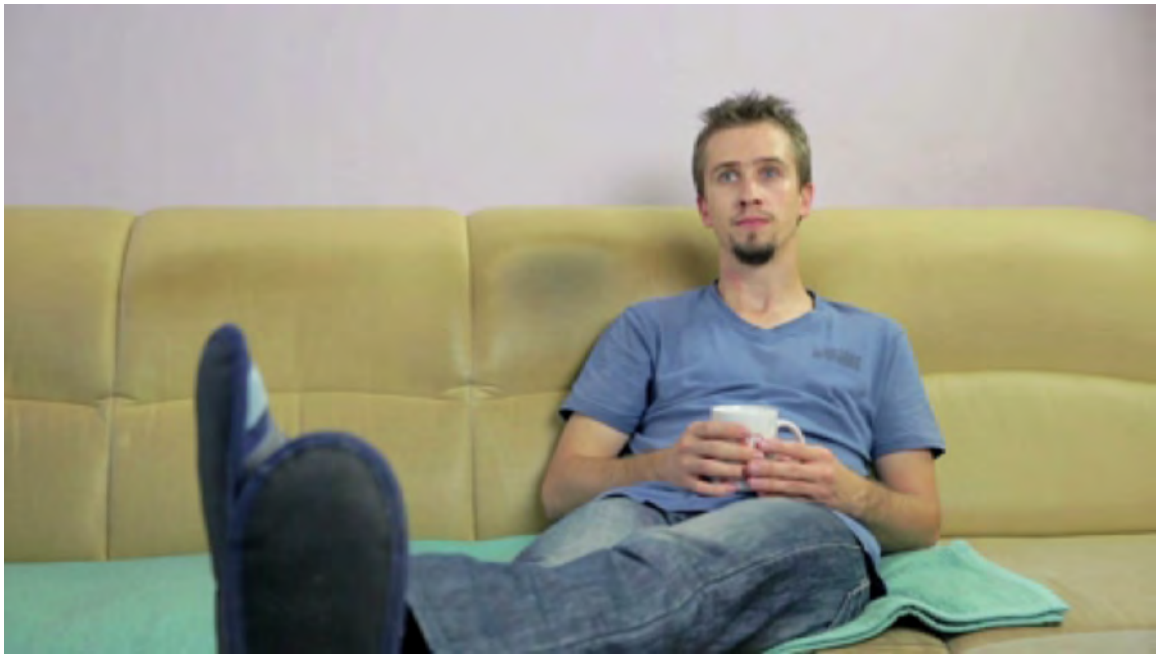


Figure 12 about here[128]

What do people do in an eternal society? A recent news item … tells us of a once flourishing but now decaying mill town in which the population find themselves with all the time in the world on their hands. And what do they do? They spend their days watching [videos]. Instead of exploiting an opportunity for … “plain living and high thinking,” … they fall back on the paralyzing *theatromania*,[129] which was the final comfort of the last days of Rome.

Figure 13 about here**[130]

> President Harold B. Lee[, continues Brother Nibley,] once … told us how at a meeting of [a stake] high council the question of the hereafter came up. One of the group, an undertaker, humorously noted that he would have to change his profession. Upon this, a dentist chimed in and confessed that he was in the same case; next an insurance man … admitted that there would not be much call for his talents, and then a used car salesman saw only limited prospects for his own business, as did the … real estate [agent] in the group, and so it went. If these men were not to dedicate themselves to making money, what would they do? A thousand years of guaranteed livelihood rule out the necessity of almost all the professions, businesses, and industries that thrive on the defects of our bodies and the insecurity of our minds.[131]

For these and other reasons, I don't think there are very many of us who are now ready for immortality — let alone eternal life.

Part 5: How Close Are We To Being Able To Model the Brain?

In the previous section, we assessed the prospects of success of programs that seek to enhance or ultimately re-embody the brain in silicon, concluding that nearly every mainstream scientist who studies the topic agrees, at least, that the many remaining challenges will not be overcome "in the near future."[132] Among these difficulties is the fact that before you can upload a brain, you need to be able to model it adequately. In this article we touch briefly on two examples, human vision and memory, to illustrate the formidable challenge of modeling the brain's structure and functions.

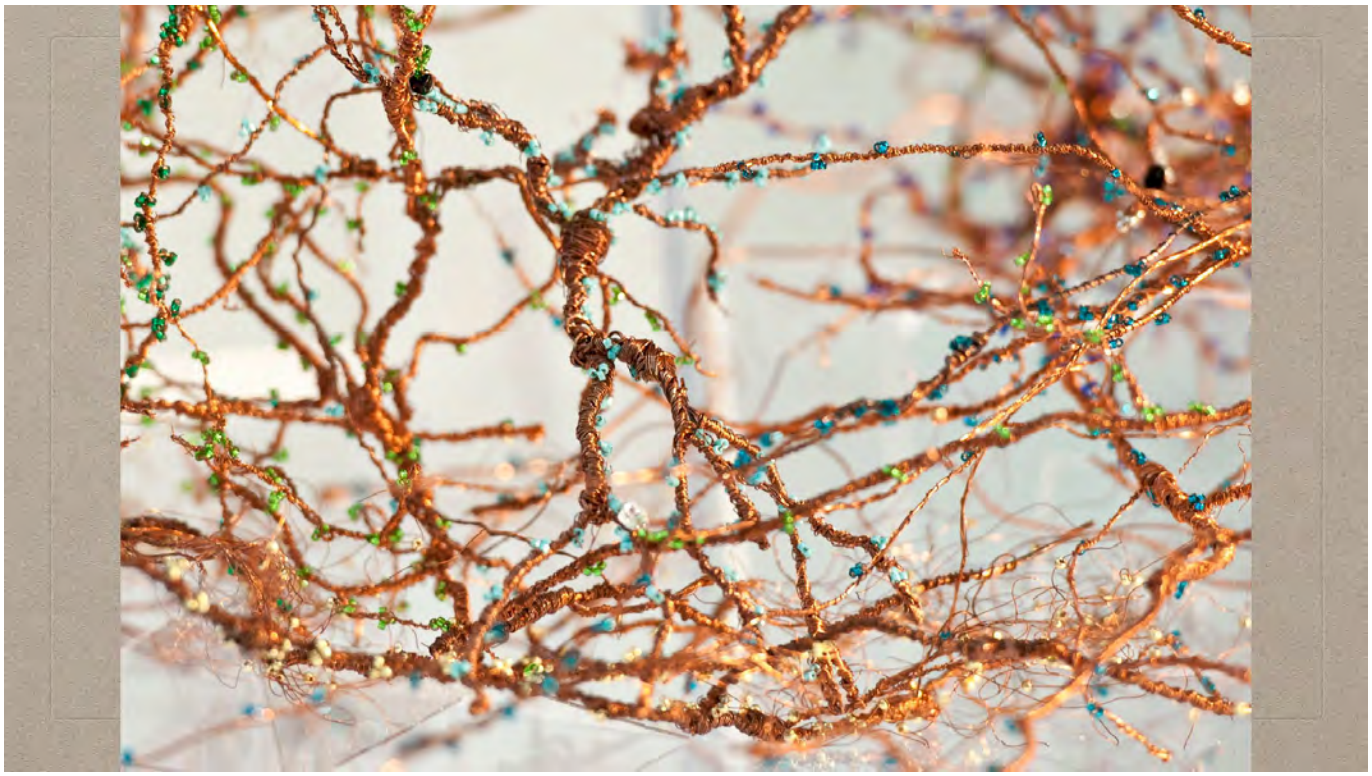
The Challenge of Modeling the Structure of the Brain

[133]

In January 2014, I took this photo outside the headquarters of the Defense Science Organization (DSO) in Singapore. It was created in 2003 by a famous local sculptor named Han Sai Por (1943-) who entitled it *Tropical Brain Forest*. Originally, the brains were painted brightly in pink and green to mirror the theme. Although the foldings of the cerebral cortex are beautiful to behold, the real complexity of the brain cannot, of course, be fully appreciated from a superficial perspective.

134

Three months later, I paid a visit to Jim Olds, Director of the Krasnow Institute for Advanced Study at George Washington University, head of the Directorate for Biological Sciences for the National Science Foundation, with whom I had served on the External Advisory Board for the Cognitive Science Program at Sandia National Laboratory. He showed me this 2012 sculpture, entitled *Mental Floss*. The sculpture is a network model of just a small part of a rodent brain, the hippocampus.



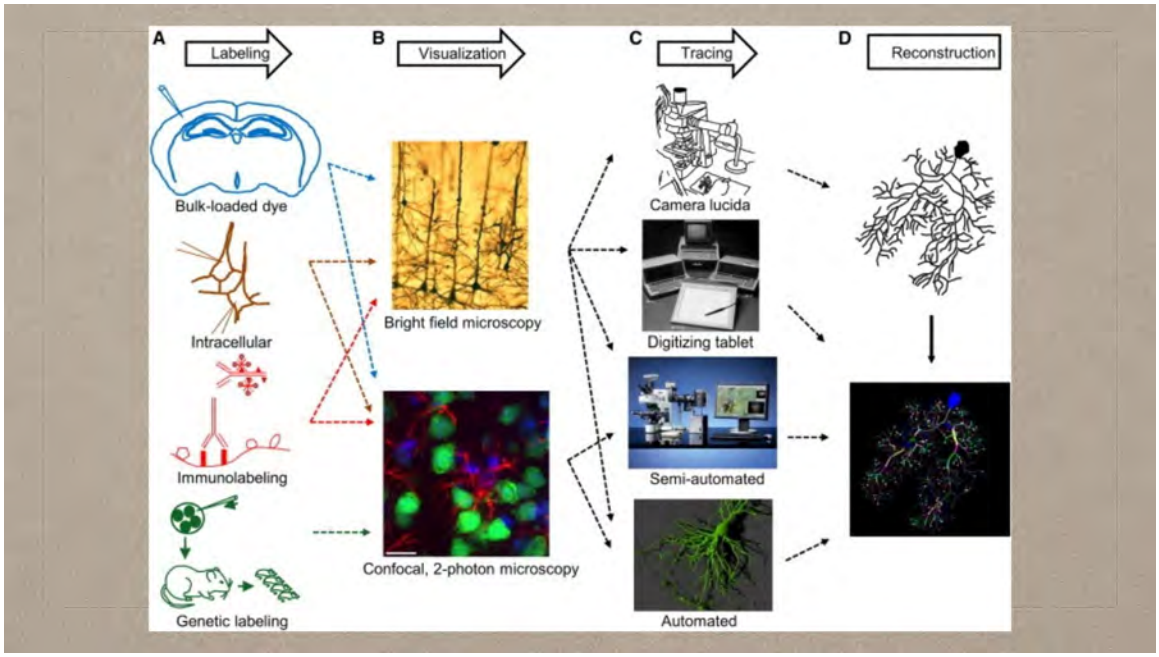
135

This photo gives some idea of the detailed work that went into the making of the sculpture:[136]

> The team [who created the sculpture, led by Professor Giorgio A. Ascoli,] selected 13 representative neuronal morphologies of the major hippocampal areas (dentate gyrus, CA3, CA1, and enthorhinal cortex), color-coding their complex axons (output trees) and dendrites (input trees).

137

The neurons were scaled and registered in virtual reality against a three-dimensional reconstruction of the rodent hippocampus. … The resulting model included excitatory projection neurons, inhibitory local neurons, and a sample of their characteristic potential circuit neurons.



138

The research of Professor Ascoli and his colleagues "aims to understand how the brain relates to the mind at the level of neuronal morphology – the "shape" of nerve cells"[139] — and its effect on neuronal electrophysiology.[140] Although the 100 billion neurons in our brains share a basic structure, they "are not exactly similar, and in fact have an astonishing diversity that leaves scientists with much to discover."[141] By way of contrast to the relatively simple and uniform models of so-called "neural networks" that are typically used in today's computer simulations, "over the past ten years, [Ascoli's] group has [developed] 3D digital reconstructions of over 35,000 neurons from dozens of species and brain regions and [the number of different models] continues to grow."[142] And, of course, a better understanding of brain structure will require much further study, not only at the neuronal level at which Ascoli's team is working, but also at the level of larger, hierarchically organized structures within the brain.

As thrilling as are the many current attempts to implement massively parallel computing architectures capable of packing the requisite number of "neurons" into a small, low-power package,[143] building a brain is much more than scaling up the number of processors while scaling down their size and power requirements — as credible mainstream scientists are well aware.[144]

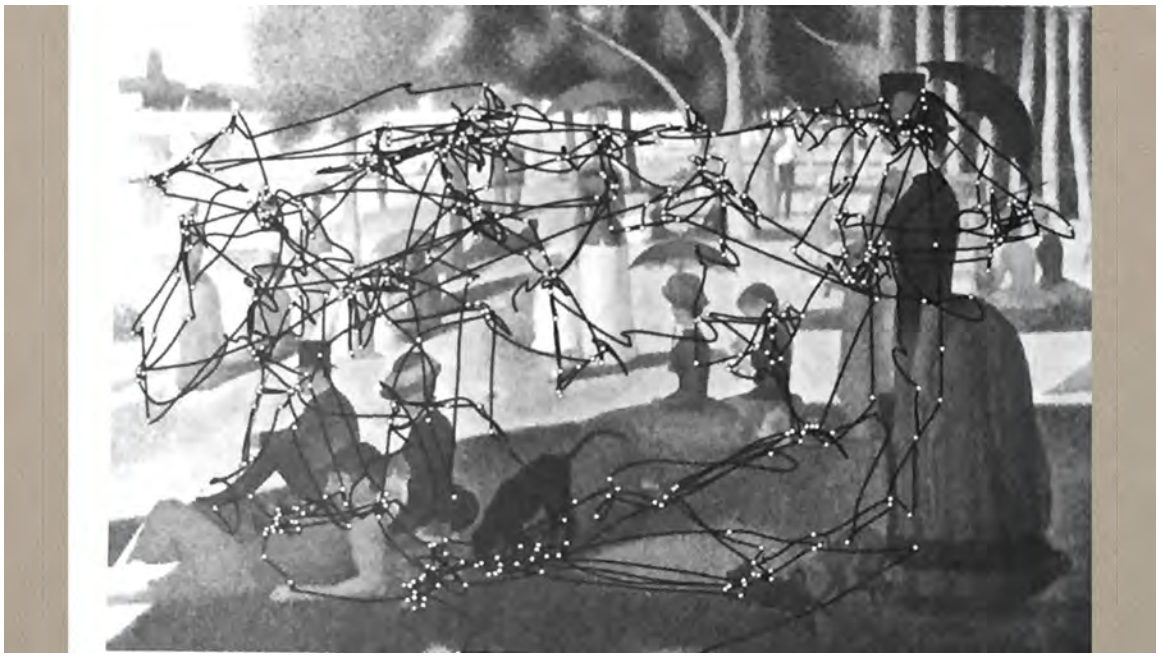
The Challenge of Modeling the Functions of the Brain

Although the focus of many researchers to date has been on replicating the detailed *structure* of the human brain, we should not forget the equally or perhaps more daunting challenge of understanding how the brain *works*: reproducing the complex details of its *functions* and *processes*.

There are many popular, persistent myths about the way the brain works — for example the erroneous idea that we only use a small percentage of the brain[145] or exaggerated notions about people being right-brained or left-brained.[146] Here, I will touch briefly on only two of these: 1) the myth that the human visual system works like a simple camera; and 2) the myth that human memory works like today's computer "memory."

147

The first thing to know about such human sensory and cognitive processes is that they are active, not passive. Visual data is not simply taken in passively as in a simple camera that focuses the light from an entire scene through the lens and onto a sensor; memory is not laid down in the brain as simple traces of experience that, in principle, could be retrieved intact at a later time, like a series of bits in computer memory. Instead, the brain relies not only on complex feedback mechanisms that shape learning based on *past* experience, but also on feedforward mechanisms that direct cognitive processes by *anticipating future experience*. As a rough analogy you can think of these feedforward mechanisms as if they were part of an automatic railroad track laying system like the one shown above that first lays out the track ahead of itself, and then follows the track it has made in order to control its forward movement.[148]



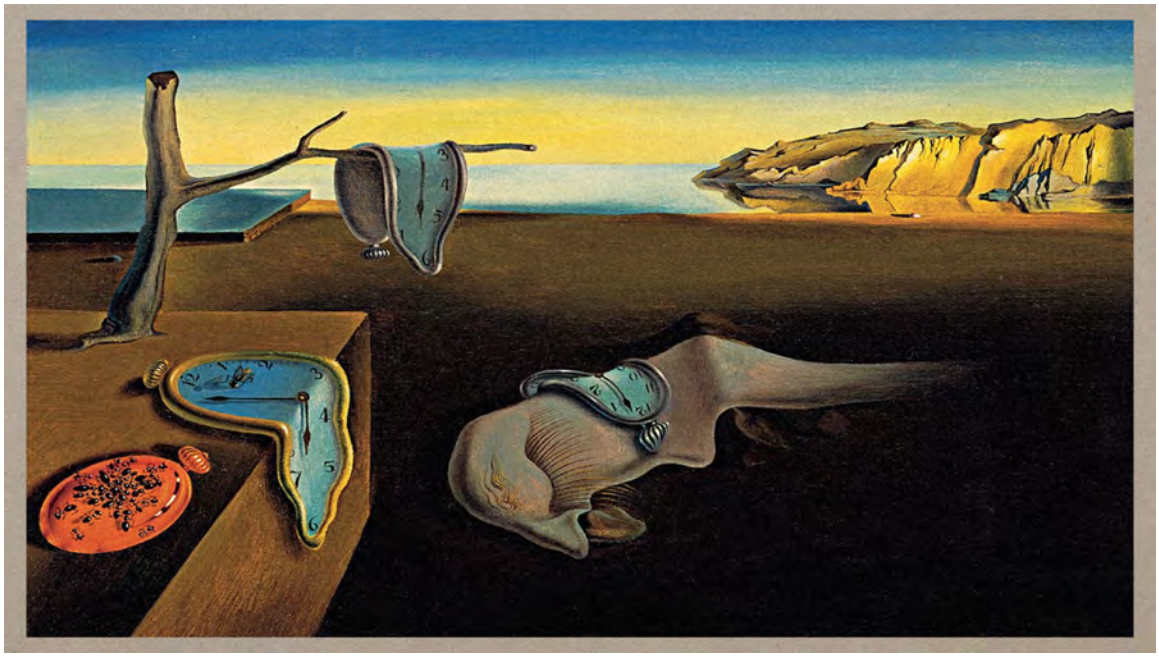
149

Active vision: The eye is not a simple camera. As a first example of the active, feedforward nature of human cognition, consider this figure, from an overview of brain mechanisms for active vision written by Robert H. Wurtz, a National Institutes of Health Distinguished Investigator at the National Eye Institute. The image is a record of eye gaze for a human viewer looking at Georges Seurat's famous painting *Sunday on the Island of La Grande Jatte* for three minutes. "Saccades[, rapid eye movements that occur two or three times each second,] are represented by black lines, and intervening periods when the eyes are stationary (fixations) are represented by white circles. … Note that the fixations are not to random locations, but rather to points of likely interest"[150] which, of course, may change, even for the same person at different times and in different contexts. "Why bother with all these saccades? Why not just hold the eye steady and inspect the painting?"[151] To answer these questions, one must know something about the varying resolution of the retina, the surface at the back of the eye that receives light:[152]

> The retina is equipped with receptors that respond to light, dark, and color, but it does not have uniform resolution across its surface. The highest receptor density is found in the central region of the retina, called the fovea, which gives us the highest visual resolution and enables us to see small details. Retinal areas outside this central region, responding to light from the periphery, have a lower density of receptors and therefore lower resolution. Thus, the viewer enjoying the Seurat painting is essentially using the fovea to examine the rich detail, jumping from one part of the painting to another. In the three minutes of saccades shown in [the figure], the viewer examines the details in a substantial fraction of the painting, but never sees the whole scene at once.

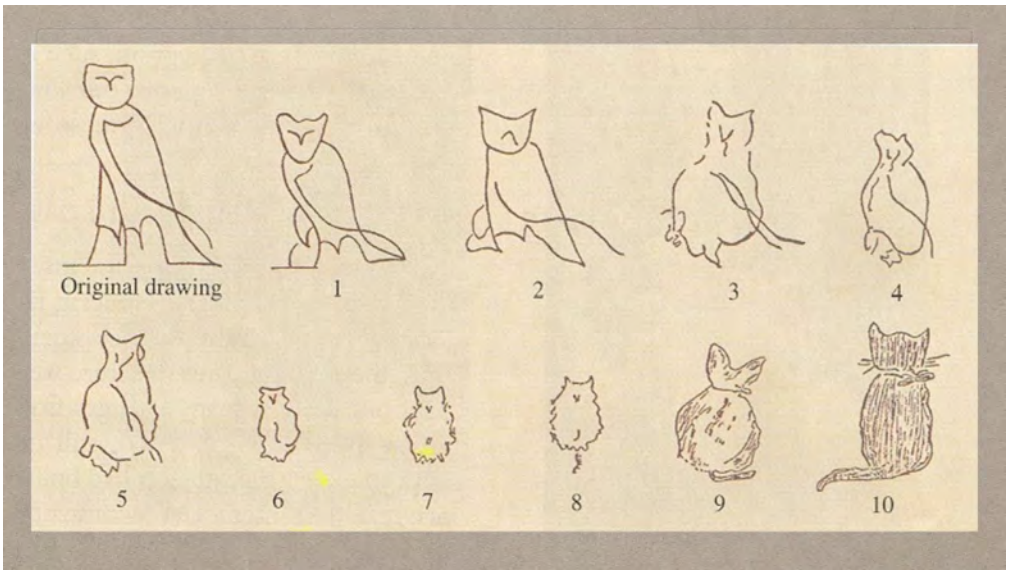
> How do subjects pick the next object to examine? Using their peripheral retina, subjects can see objects in the field at a relatively low resolution and select ones of potential interest to examine next. This shift of attention from one item to the next accompanies each saccade. This selection is not random: if we look at the saccades superimposed on the Seurat painting, we can see that the trees at the top of the painting and the grass at the bottom are largely ignored; in contrast, faces, dresses, and other significant objects are frequently inspected. …
>
> The goal of … saccades is to bring images to the fovea for detailed analysis. Even though these saccades displace the image of the whole visual field on the retina, the system operates so perfectly that we regard the scene as serenely stable.

In short, the eye and brain work together much differently than does a simple camera. And, as complex as what we have described above may seem, Wurtz points out that the many components that underlie active vision are "only a small part of the larger brain systems involved with sensation and motor control. In turn, the puzzle of how these systems operate to produce action is just one of many global questions about how the brain produces all behavior, including learning, memory, and emotion; and even how consciousness arises from brain activity."[153] But since the state of the art in current research on these topics still leaves many important questions unanswered, Wurtz admits that he has no choice but to rely on "the classic approach of reducing the overwhelming complexities of the brain to more easily understood fragments" by considering active vision in isolation.[154]



155

Reconstructive memory: The brain is not a simple canvas. As a second example of the active, feedforward nature of human cognition, consider the complex and still somewhat mysterious phenomenon of memory. Salvador Dali's famous images of the pliable, melting pocket watches make a fitting representation of the contrast between a naïve, static perception of time and events and the actual fluid and dynamic nature of memory.



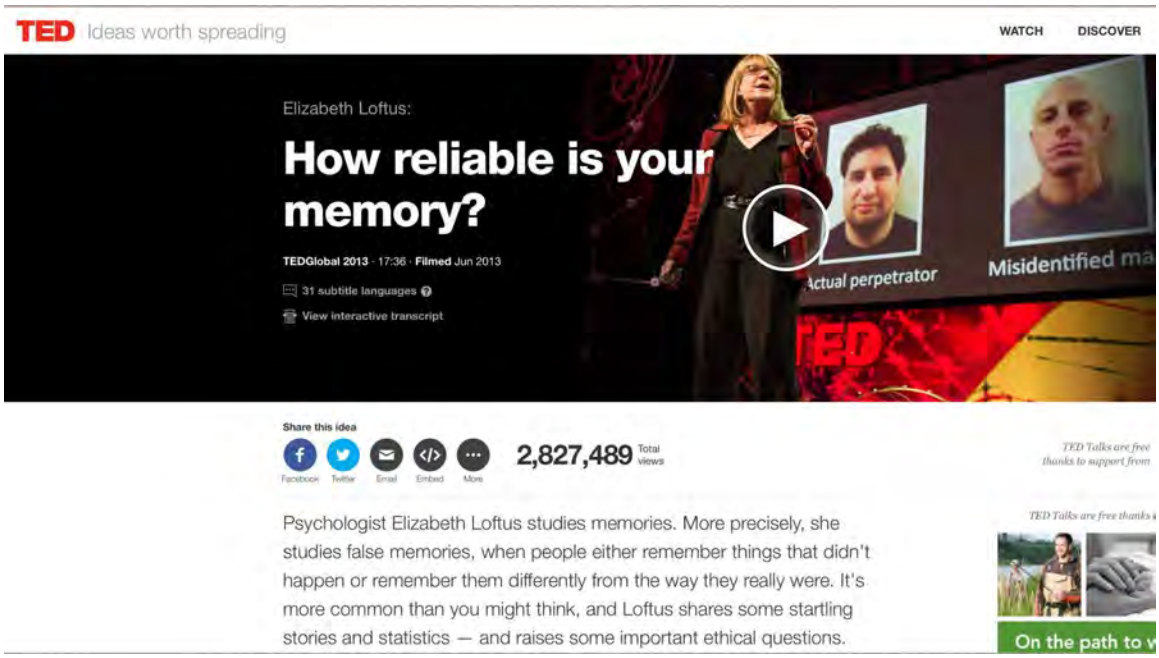
156

Pioneering studies of the fluid and dynamic nature of memory were made early in the 20th century by the British experimental psychologist, Frederic C. Bartlett (1886-1969).[157] His experiments on how people remembered stories and pictures over time showed that the "accuracy" of memory was not simply a function of some original, pristine trace gradually fading away, similar to what happens when a photographic print is continuously exposed to sunlight. Rather, Bartlett showed that the reason our recall of complex scenes, narratives, or situations becomes confused is because every perception and every memory is made in the context of our knowledge, emotions, motivations, and expectations at the time of the experience. Moreover, our memories of the past are unconsciously associated and colored with later attitudes and memories laid down *after the fact* in a way that sometimes makes it nearly impossible to tease them out:[158]

> Bartlett's participants were asked to reproduce stories taken from the folklore of other cultures; thus, their content and structure were rather strange to Western ears. The reproductions showed many changes from the original. Some parts were subtracted, others were overelaborated, and still others were additions that were completely new. In effect, the participants had built a new story upon the memorial ruins of the original. This memorial reconstruction was generally more in line with the cultural conceptions of the subjects than with the story they had actually heard. For example, certain supernatural plot

> elements were interpreted along more familiar lines. In a variant of the same experiment, Bartlett used the method of serial reproduction. A drawing was presented to one participant, who reproduced it from memory for the benefit of a second, whose reproduction was shown to a third, and so on for a chain of up to ten participants [as shown above].[159] With this technique (an experimental analogue of rumor transmission), each participant's memorial distortion became part of the stimulus for the next one down the line; the effect was to grossly amplify the reconstructive alteration.

The important takeaway is not the obvious fact that memory distortions occurred, but rather that the *kinds* of distortions were predictably in the direction of what people expect to see rather than simply a fading away of what they saw.[160] Hence, the final drawing of the cat eliminates the unusual, foreign aspects of the original drawing in favor of a conventional rendering that would be recognized by most people in a Western culture today.

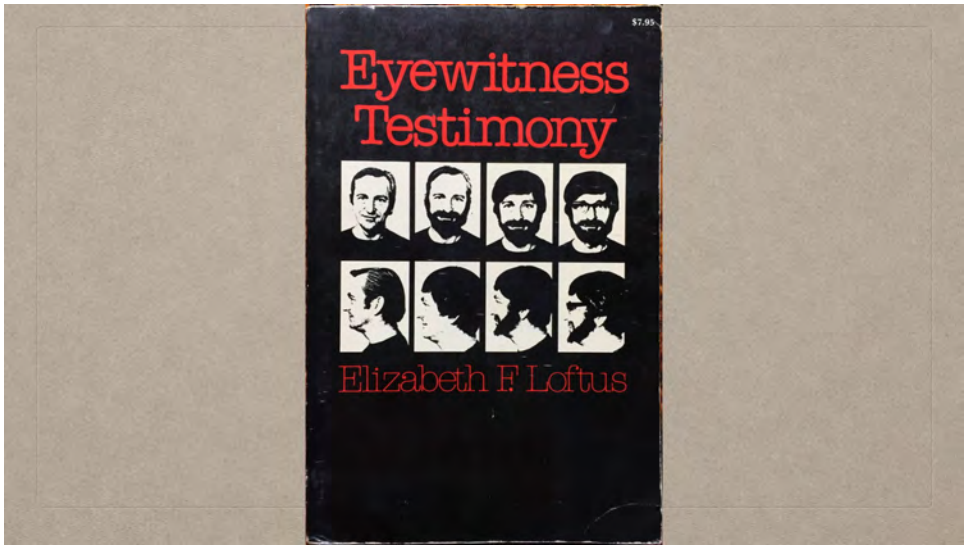


161

My own introduction to this subject came through exposure to the research of Elizabeth Loftus, who recognized and demonstrated the implications of the workings of human memory for situations such as eyewitness testimony or hypnosis, where false presuppositions and suggestions planted after the fact often came to be remembered as if they had been part of the original experience. The following summary of Loftus' accomplishments accompanied a 2013 TED talk that I highly recommend to anyone interested in this subject:[162]

> Elizabeth Loftus altered the course of legal history by revealing that memory is not only unreliable, but also mutable. Since the 1970s, Loftus has created an impressive body of scholarly work and has appeared as an expert witness in

hundreds of courtrooms, bolstering the cases of defendants facing criminal charges based on eyewitness testimony, and debunking "recovered memory" theories popular at the time, as in her book *The Myth of Repressed Memory: False Memories and Allegations of Sexual Abuse* (with Katherine Ketcham).



163

Loftus' early experiments on eyewitness testimony showed:[164]

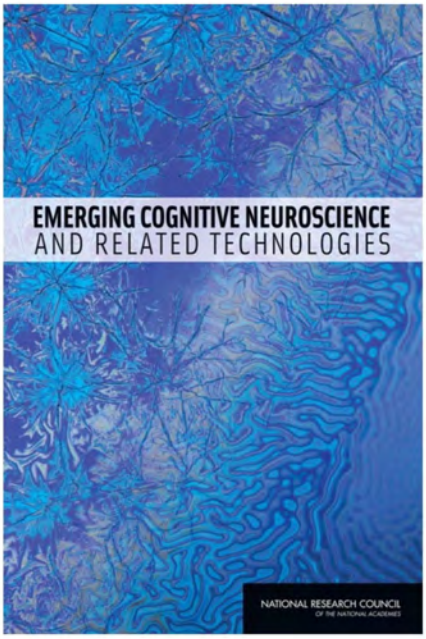
> that when people who witness an event are later exposed to new and misleading information about it, their recollections often become distorted. In one example, participants viewed a simulated automobile accident at an intersection with a stop sign. After the viewing, half the participants received a suggestion that the traffic sign was a yield sign. When asked later what traffic sign they remembered seeing at the intersection, those who had been given the suggestion tended to claim that they had seen a yield sign. Those who had not received the phony information were much more accurate in their recollection of the traffic sign.

As an undergraduate at the University of Utah, I was invited by David Dodd to help carry out a variant of this experiment where we showed that Loftus' results sometimes may have been exaggerated by the way misleading information was introduced to her subjects.[165] A few years later, after beginning graduate school at the University of Washington, I was fortunate to serve as a teaching assistant for one of her courses, one of the most engaging I have ever attended.

On one occasion, a purse snatching occurred during the first few minutes of class. Unbeknownst to the students, this event had been staged deliberately. As the event was unfolding, stooges in the lecture hall shouted out misleading information, such as "Grab that guy with the moustache!" although the individual fleeing the room had no moustache. Later, when candidates for the "crime" were lined up in front of the class, people remembered having *seen* the moustache on

the culprit, even though the misleading information had come through verbal rather than visual channels. The source of the false information had been forgotten and the verbal information had become inextricably associated with the visual memory.

The point of the preceding example is to illustrate a small part of the complexity and interrelatedness of higher cognitive functions such as learning, memory, emotion, and consciousness. The brain is not a simple canvas on which experience is passively written and recalled. Those who aspire to someday upload a fully functioning human mind to a computer inevitably will have to rely on complete, high-fidelity models of the structure and function of the human brain in general, and of an individual human brain in particular. Such models do not now exist.



166

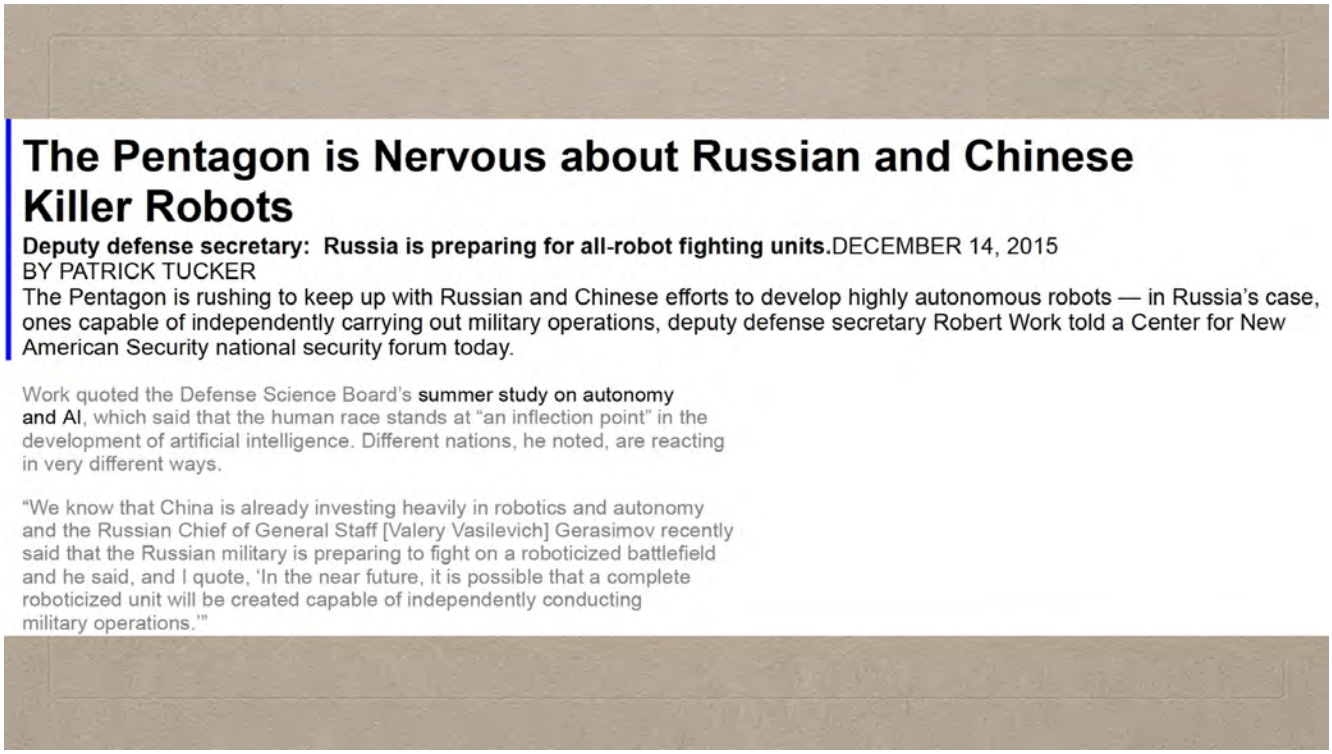
More than twenty years after my experience as a teaching assistant for Elizabeth Loftus, I served with her as part of a 2008 National Academies study of the future of cognitive neuroscience and related technologies. As part of our committee's summary on the topic of modeling and building the equivalent of a human brain, I wrote the following:[167]

> [Despite the impressive increases in high-end computing], there does not yet exist either an adequate and detailed understanding of *how* … modeling [of the human brain] can be done, or a complete model of how the brain interacts with complex regulatory and monitoring systems throughout the body. These and other difficulties make it highly unlikely that in the next two decades anyone could build a neurophysiologically plausible model of the whole brain and its array of specialized and general-purpose higher cognitive functions.[168]

Mark Changizi, an evolutionary neurobiologist, is even more cautious in his predictions. In his article entitled “Artificial Brains: Not In This Century” he writes:[169]

> Maybe I’m a buzz kill. But I prefer to say that it’s important to kill the bad buzz, for it obscures all the justified buzz that’s ahead of us in neuroscience and artificial intelligence. And there’s a lot. Building artificial brains may be a part of our future — though I’m not convinced — but for the foreseeable, century-scale future, I see only fizzle.

Part 6: Autonomous Weapons and Natural Stupidity



The Pentagon is Nervous about Russian and Chinese Killer Robots

Deputy defense secretary: Russia is preparing for all-robot fighting units.DECEMBER 14, 2015
BY PATRICK TUCKER
The Pentagon is rushing to keep up with Russian and Chinese efforts to develop highly autonomous robots — in Russia's case, ones capable of independently carrying out military operations, deputy defense secretary Robert Work told a Center for New American Security national security forum today.

Work quoted the Defense Science Board's **summer study on autonomy and AI**, which said that the human race stands at "an inflection point" in the development of artificial intelligence. Different nations, he noted, are reacting in very different ways.

"We know that China is already investing heavily in robotics and autonomy and the Russian Chief of General Staff [Valery Vasilevich] Gerasimov recently said that the Russian military is preparing to fight on a roboticized battlefield and he said, and I quote, 'In the near future, it is possible that a complete roboticized unit will be created capable of independently conducting military operations.'"

Figure 1 about here[170]

One area of technology that is fraught with momentous consequences is the proliferation of autonomous weapons — in other words, weapons that have significant capabilities for performing sensing, reasoning, and decision-making on their own. This was one of many topics that I explored in depth as a member of the 2015 Defense Science Board Summer Study on Autonomy.[171] While I won't comment on the details of the study nor try to summarize the committee's consensus, I want to share my personal thoughts and concerns about the future development and deployment of weapons that are capable of operating more or less on their own. Such weapons, available in both the cyber and the physical domains, will be increasingly used by adversaries who are not constrained by the principles and ethics that are meant to govern US policies in this arena.

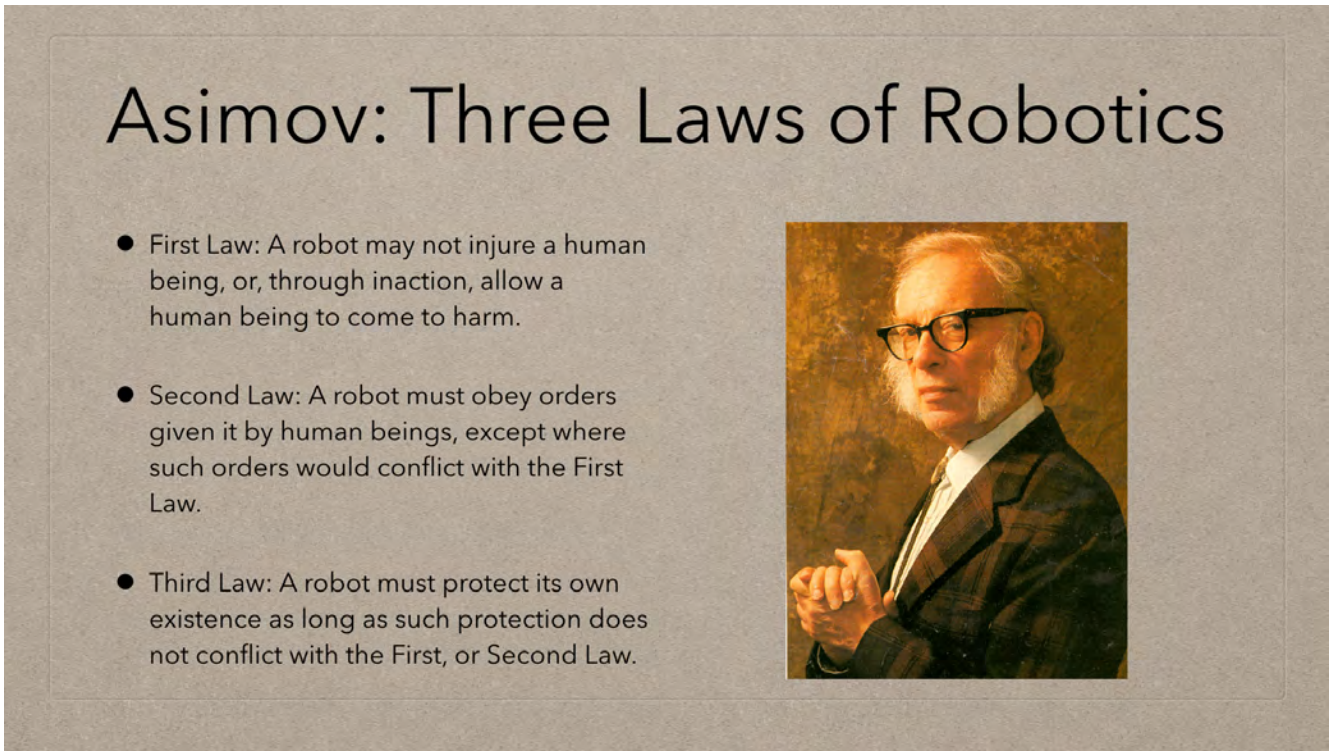


Figure 2 about here

My interest in this topic has grown over the last fifteen years as our research group at the Institute for Human and Machine Cognition (IHMC) has worked on technological solutions to the problem of policy-based governance of intelligent systems, with a long-term vision that embraces the spirit of Isaac Asimov's laws of robotics.[172] We call our digital policy services framework KAoS.[173] Significant efforts are underway at the world's largest tech companies "to create a standard of ethics around the creation of artificial intelligence."[174]

Figure 3 about here[175]

Artificial Intelligence as a Pillar of Modern Military Strategy

Over the past fifty years, much of the significant progress in Artificial Intelligence (AI) has been due to funding from the United States Department of Defense (DoD). The resulting developments in AI can be grouped into three major waves:[176]

1. "The first wave (1950–1970) launched the academic field of computer science, opened an era of discovery and set the foundation for signal processing, computer vision, computer speech and language understanding.
2. The second wave (1970–1990) saw codification of knowledge in expert systems, using rule bases, and beginnings of simple machine inference to do reasoning (think things like computer chess), along with exploration of computer architectures, specialized for AI applications.
3. The third wave (1990–present) launched the era of large scale robotics, including autonomous machines, along with real breakthroughs in the use of neural network architectures, inspired by better understanding of how the brain works."

Figure 4 about here[177]

These developments are now seen as so important to the DoD that they have been labeled its "third offset strategy."[178] Like the Cold War strategies of nuclear deterrence (first offset) and general technological superiority in the face of decreasing manpower (second offset), it is hoped that a third offset based on AI would provide a decisive advantage for the US military in future confrontations.[179]

Amazon 'Reviewing' Its Website After It Suggested Bomb-Making Items

By AMIE TSANG SEPT. 20, 2017

A British TV station said Amazon's website was prompting customers to buy all the ingredients to make a bomb. Richard Drew/Associated Press

RELATED COVERAGE


Google and Facebook Face Critici: Ads Targeting Racist Sentiments
SEPT. 15, 2017


Amazon Profits Fall Far Short of Expectations, but Investors Shrug
JULY 27, 2017


BITS
Tech Roundup: A Defense of 'Pro Ads' APRIL 6, 2017

LONDON — Amazon said on Wednesday that it was reviewing its website after a British television report said the online retail giant's algorithms were automatically suggesting bomb-making ingredients

Figure 5 about here[180]

What Are Some of the Challenges?

Several issues complicate the implementation of the "third offset" by the US military, including the fact that research and development of AI is now generously funded by (and largely controlled by) private companies rather than the DoD and the fact that the United States no longer holds a monopoly on significant scientific advances in the field.[181]

Unlike the nuclear weapons first deployed in World War II, the proliferation of autonomous weapons would not be constrained by the difficulties of a given nation's ability in performing sophisticated refinement of rare elements. Rather it is being helped along rapidly by the virtually unlimited capacity for just about anyone to share and duplicate the needed software using worldwide computer networks. In principle, such capabilities could be developed in and sold from anyone's garage — so long as that garage has a good Internet connection.

Unlike nuclear weapons, the development and proliferation of intelligent weaponry cannot be easily monitored or banned. There is no need to solve the

long-term AI problem of general intelligence in order to develop early generations of such weapons — only the development of limited-scope autonomous capabilities that are custom tailored to specific purposes.[182] Like the combination of bomb-making parts that, until recently, were cheerfully suggested by Amazon's recommendation algorithms to anyone who asked the right questions,[183] AI algorithms and code that are "good enough" to include in advanced weaponry are widely available everywhere.

Add all this to the fact that "weapons" are no longer confined to specialized military hardware or even conventional computers, but can reside and proliferate in the billions of connected gadgets of all kinds in our homes, workplaces, and public sites. Security for such devices is a daunting prospect.[184] Billions of such devices are already in use and will easily dwarf the numbers of traditional computing devices in the coming few years. According to the Defense Science Board report: "This immense, sparsely populated space of interconnected devices could serve as a globe-spanning, multi-sensing surveillance system or as a platform for massively proliferated, distributed cyber-attacks — or as an immense test range for real-world, non-permissive testing of large-scale autonomous systems and swarms."[185]

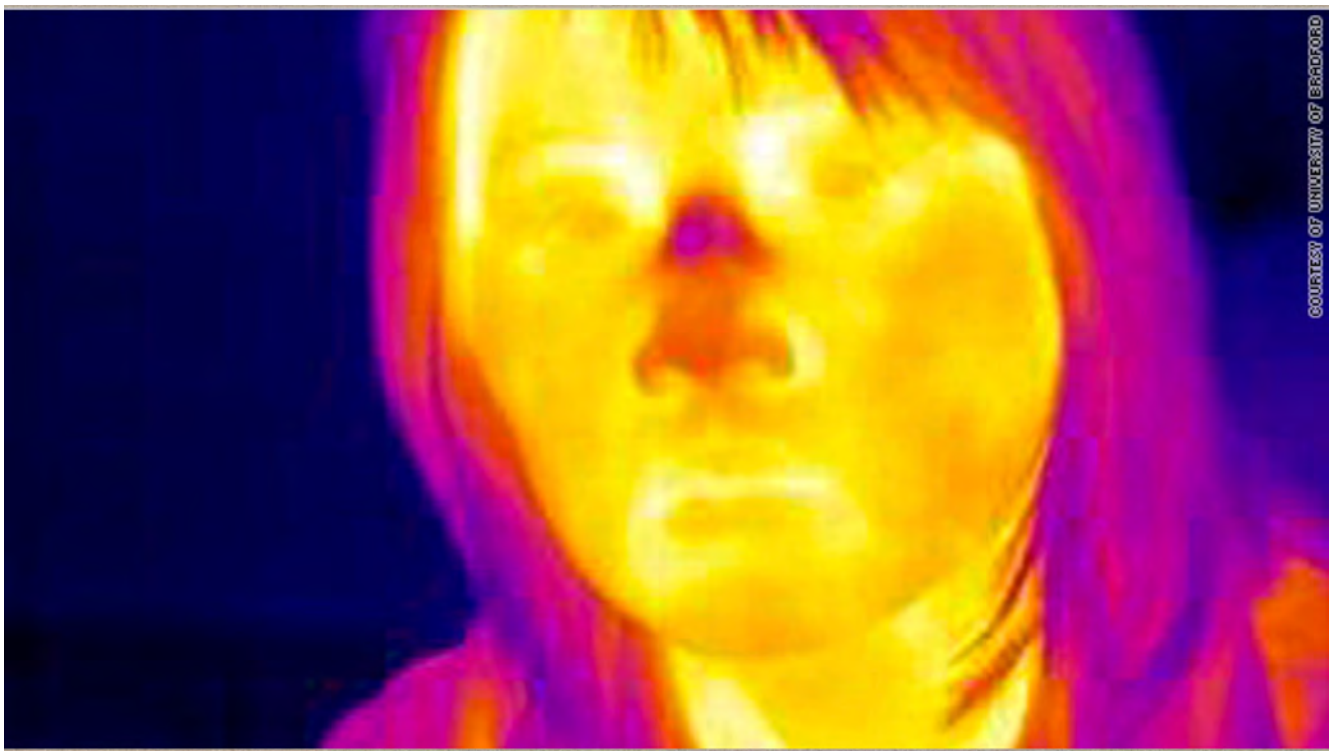


Figure 6 about here[186]

In previous articles in this series, we have given examples of the overblown expectations of scientific researchers about the near-term future of AI. Just to prove that others besides researchers can entertain wild speculations, at the initial meeting of a National Academies study some years ago, our group was told

that one of the questions one sponsor had asked us to explore was whether it would be possible to develop an autonomous weapon that could fire into a crowd and only hit people with hostile thoughts.

Without even entering into the staggering legal and ethical implications of developing such a weapon, our committee implicitly answered this question on the pure grounds of common sense, based on decades of data: Today, we hardly know how to build a good, automatic lie detector, let alone being able to recognize a range of specific psychological states for unknown individuals in an uncontrolled environment — and (thank heavens!) it's highly unlikely that the needed breakthroughs will happen anytime in the next few decades.[187]

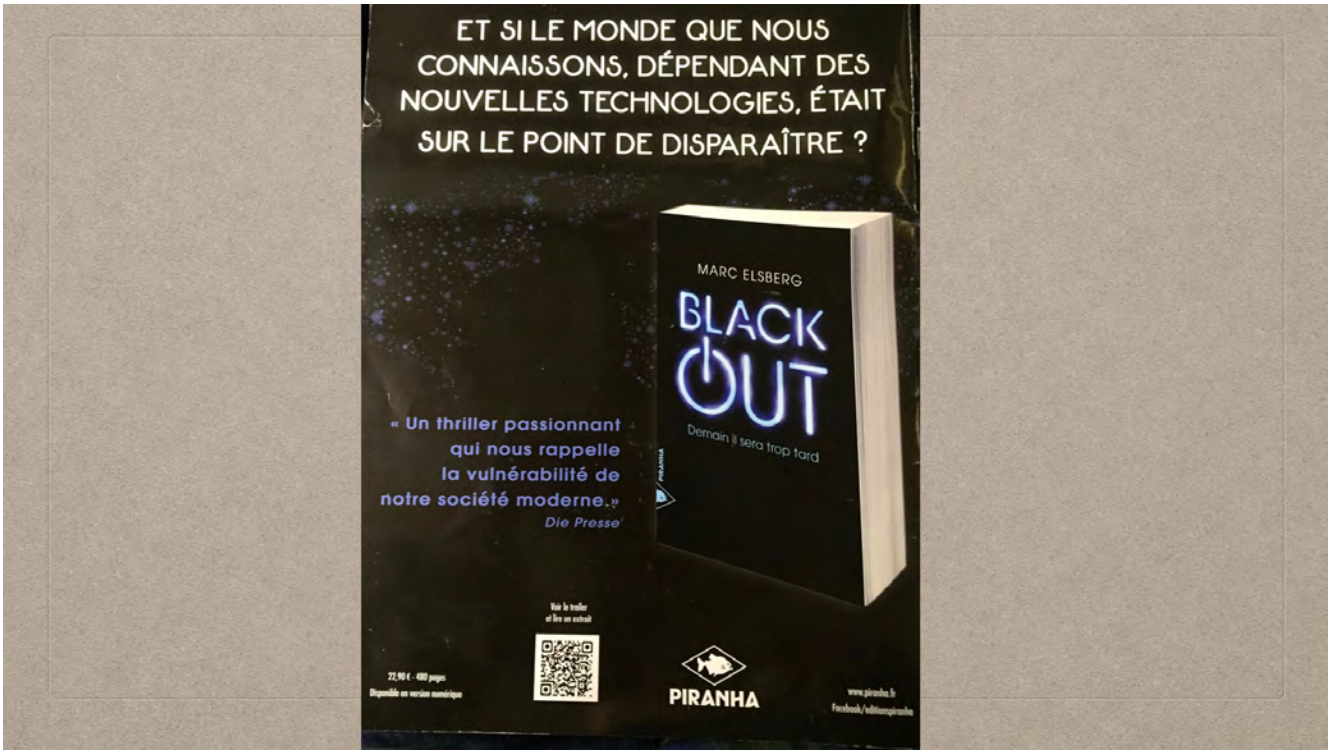


Figure 7 about here

The Rise of Cyber Warfare

Cyber warfare is one of the most underappreciated threats of the modern age. Everything in our economy, infrastructure, and personal lives would come to a grinding halt were such threats carried out at a large scale. For this reason, the DoD has elevated cyber security as a "national priority" and has established well-funded organizations to carry out its missions, such as the "US Cyber Command."[188]

The motivation for cyber warfare waged against nations, organizations, and individuals is not merely political but is also economic. There is a flourishing worldwide "underground economy" that exploits the money to be made in

"cybercrime, money laundering, and information security" breaches.[189] Groups with a "motivation to find exploitable defects in widely used [software] … are willing to pay anyone who can find and exploit these weaknesses top dollar to hand them over, and never speak a word to the companies whose programmers inadvertently wrote them into software in the first place."[190] Far from the ideals of the Internet pioneers who imagined open access to informationå across all borders, we are facing the future of a "splinternet" fragmented by geopolitics and commercial interests.[191]

Following hard on the heels of the enormous destructive power of two major hurricanes, damaging wildfires, and an 8.1 magnitude earthquake in Mexico, was the news of the September 7, 2017 theft of detailed personal and financial information at Equifax. This cyber disaster affected the lives and credit of up to 143 million people in the United States.[192] It has been called "one of the gravest breaches in history,"[193] but it is barely a drop in the bucket in the sea of information already available online about individuals. It provides a small foretaste of what portend to be greater confusions and disruptions of people's private and public lives ahead.[194]

Consider not only individual mavericks who manipulate online information for personal profit or political ends, but more importantly the increasing number of well-financed and carefully targeted efforts to create misinformation, invent false identities, and disrupt critical infrastructure with the goal of "wreak[ing] havoc all around the Internet — and in real-life American communities."[195] For example, as early as 2008 the DoD publicly disclosed information "from multiple regions outside the United States, of cyber intrusions into utilities, followed by extortion demands. … We have information that cyberattacks have been used to disrupt power equipment in several regions outside the United States."[196]

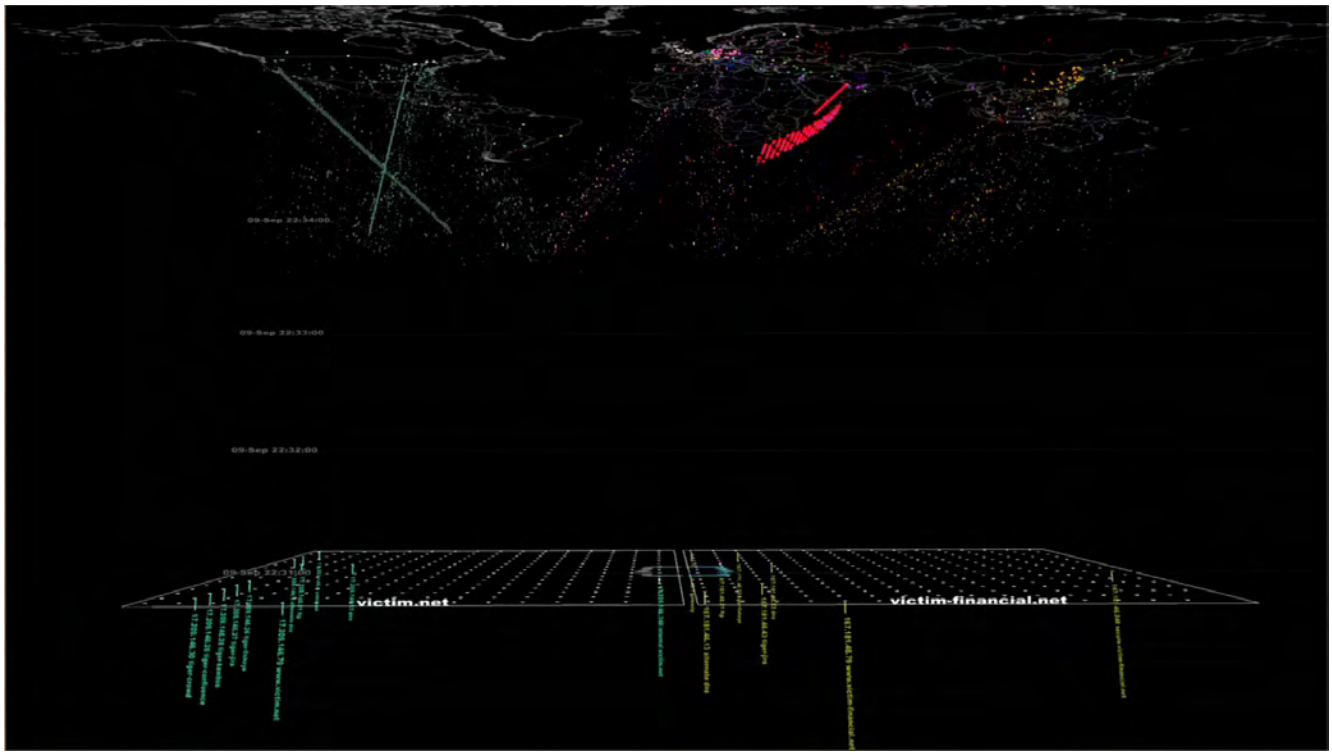


Figure 8 about here

IHMC's Sol cyber framework, here using simulated data, shows one of the approaches our research team developed in response to a government request to address the (impossible) challenge of visualizing and interacting with the entire Internet in real time so as to make sense of whatever important events were going on at the moment.[197] We have had a "live," real-time version of such a display continuously working on IHMC's own network for some years now. As you watch the "live" display, the graphics make it easy to see continuous waves of attacks from around the world attempt to penetrate our relatively obscure and unimportant website.

The patented design of this and similar IHMC-developed displays exploit specific, subtle properties of human perception and cognition, allowing large numbers of interesting events to pop out and be assimilated by the ambient vision system.[198] In the image, you can see a projection of a world map at the top, with various patterns of attack moving downward toward the company network at the bottom, belonging to a specific victim and its primary financial institution.

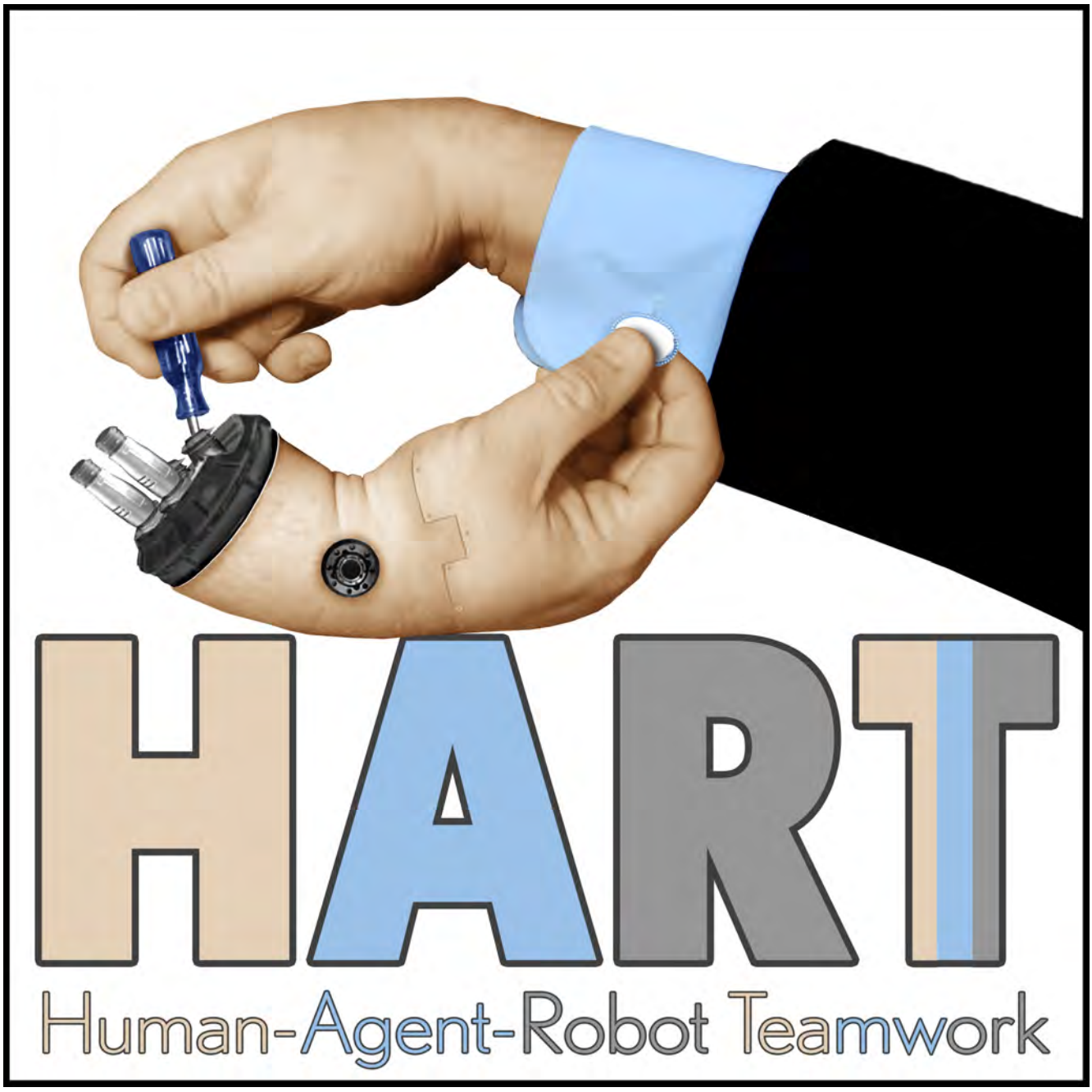


Figure 9 about here

The Future of Artificial Intelligence

Our design philosophy for Sol was consistent with the emphasis of our research group on creating systems that enable human-agent-robot teamwork (HART) rather than developing Artificial Intelligence capabilities that are meant to work more or less on their own. A good illustration of the more common way of thinking in the standalone AI approach can be found in the work of Alan Turing. Turing, a famous early computer scientist, asked the question, “Can machines

think?" He laid out an experiment in the form of a game.[199] The challenger in the game is given the task of comparing the separate answers of a human and a machine in order to determine which is which.

By way of contrast to Turing's game, our question has been "Can humans and machines think *together*?" The challenge in designing Sol was not to determine whether a machine could be so sophisticated that it could fool a human. Instead, Sol was designed as an early experiment in blurring the line between human and machine thinking — to understand what it might be like someday for humans and machines to be working together so closely and that it would seem as if the parties were thinking together.[200] To this end, the visual innovations of Sol were combined with software agents that were designed to collaborate with cyber analysts, working together to make sense of complex situations in rapid, real time.[201] Because cyber attacks can occur in microseconds, the responsibility for the most rapid kinds of reactions must be assigned to the agents while deliberative aspects of sensemaking and decision-making can benefit from a combination of human and machine abilities.

While mainstream researchers in Artificial Intelligence usually reject the prospects of an AI explosion, singularity, or apocalypse such as those popularized in the media,[202] they have been thinking more deeply of late about the future of AI. As a result of this thinking, there has been a recent proliferation of research institutions,[203] studies,[204] articles,[205] books,[206] blogs,[207] and open letters of concern[208] to help assure that both the short- and long-term trajectories of AI research will follow directions that are both safe and beneficial to society. Far from being the neo-Luddites these researchers are sometimes painted to be,[209] they are some of the top minds in the field, believers in the potential of AI for the good of humankind.[210]

Figure 10 about here

Combatting Natural Stupidity

Now our brief tour of AI must come to an end. It's been exciting for me over the years to see many of the breakthroughs we used to call Artificial Intelligence become assimilated as ordinary, ho-hum parts of mainstream computer science and engineering.[211] I share much of the optimism of President Gordon B. Hinckley who, like his predecessors, rejected unsound extrapolations of scripture and statements of Church leaders to justify apocalyptic panic in the face of natural disasters and technological advances.[212] He said:

> [The twentieth century] has been the best of all centuries. ... The fruits of science have been manifest everywhere. ... This is an age of greater understanding and knowledge. ... This has been an age of enlightenment. The miracles of modern medicine, of travel, of communication are almost beyond belief.[213]

I believe that the fruits of science and technology are divine gifts to which it is appropriate to apply the observation given in D&C 59:20: "And it pleaseth God that he hath given all these things unto man; for unto this end were they made to be used, with judgment, not to excess, neither by extortion."

Figure 11 about here[214]

Do I ever lose sleep over the future of Artificial Intelligence? Only rarely, and that's usually when I'm wrestling with a solution to some interesting problem. However, that is not to say that I don't sometimes lose sleep over the future in general — for related reasons that are best illustrated by Boyd Petersen's account of an incident involving the late Hugh Nibley:[215]

> One day in the early 1950s, Hugh Nibley's teaching assistant Curtis Wright found Hugh leaning over his desk, reading from the Book of Mormon, and laughing. Wright asked Hugh Nibley what was so funny, and he responded that he had discovered an error in the Book of Mormon. "You did, huh?" Wright asked. "That's interesting. Let me see it."
>
> Hugh handed the scriptures over to Wright and pointed to Alma 42:10, which says that humans are "carnal, sensual, and devilish, by nature." Wright read the passage and demanded, "Well, what's the matter with that?" … Wright was beginning to think that Hugh might be ridiculing the Book of Mormon. "So I got a little defensive," says Wright. Unable to conceal his contempt, Wright demanded, "How's it a mistake?"
>
> He responded, "Well, look at Alma, he says that all mankind is carnal, sensual, and devilish by nature. And he should've said they were carnal, sensual, devilish, and *stupid*."

No, I don't worry too much about the future of Artificial Intelligence, but I do over the consequences of natural stupidity. When Artificial Intelligence meets natural stupidity, unfortunate things can happen. "I am grateful to know," wrote Truman G. Madsen, "that Jesus Christ suffered not only for our sins but for our stupid mistakes."[216] And through the Atonement of Jesus Christ, declared Elder Jeffrey R. Holland, "we can escape the consequences of both sin and stupidity — our own or that of others — in whatever form they may come to us in the course of daily living."[217]

May God grant that we may read and understand the fine print in the hype cycles, discern the "designs which do and will exist in the hearts of conspiring men in the last days,"[218] and, most important of all, rely on divine wisdom and grace to help overcome our natural stupidity is my prayer.

References

AAAI Presidential Panel on long-term AI futures: 2008-2009 study. In *Association for the Advancement of Artificial Intelligence*. https://www.aaai.org/Organization/presidential-panel.php. (accessed September 9, 2017).

AI Effect. In *Wikipedia*. https://en.wikipedia.org/wiki/AI_effect. (accessed March 12, 2016).

Alba, Davey. 2017. The world may be headed toward a fragmented "splinternet". In *Wired*. https://www.wired.com/story/splinternet-global-court-rulings-google-facebook/. (accessed September 22, 2017).

Anonymized. E-mail message to Jeffrey M. Bradshaw, January 30, 2016.

Anonymous. "How IBM makes its partners smarter: IBM's Watson turns cognitive computing into business success." *CyberTrend 14*, no. 2 (February 2016): 10-15. http://www.cybertrend.com/article/19837/how-ibm-makes-its-partners-smarter. (accessed March 5, 2016).

Armstrong, Stuart. 2013. Assessing Kurzweil: The results. In *LessWrong*. http://lesswrong.com/lw/gbi/assessing_kurzweil_the_results/. (accessed March 15, 2016).

Armstrong, Stuart, and Seán ÓhÉigeartaigh. "Who knows anything about anything about AI?" In *Intelligence Unbound: The Future of Uploaded and Machine Minds,* edited by Russell Blackford and Damien Broderick, 46-60. Chichester, West Sussex, England: John Wiley, 2014.

Armstrong, Stuart, Kaj Sotala, and Seán ÓhÉigeartaigh. 2014. The errors, insights and lessons of famous AI predictions — and what they mean for the future. In. http://www.fhi.ox.ac.uk/wp-content/uploads/FAIC.pdf. (accessed March 15, 2016).

Artificial Brains: The Quest to Build Sentient Machines. In *Artificial Brains*. http://www.artificialbrains.com. (accessed June 29, 2016).

Artificial Intelligence: What's real, what's not, and is this the DoD Third Offset? (Abstract of Panel Discussion). In *Tentative Agenda for the Defense Science Board Sixtieth Anniversary: Celebrating Innovation for National*

Security (20 September 2016). https://www.eiseverywhere.com/ehome/179363/408368/. (accessed September 26, 2017).

Atkinson, Robert D. 2015. The 2015 ITIF Luddite Award nominees: The worst of the year's worst innovation killers. In *Information Technology and Innovation Foundation*. https://itif.org/publications/2015/12/21/2015-itif-luddite-award-nominees-worst-year's-worst-innovation-killers. (accessed March 5, 2016).

Autonomous weapons: An open letter from AI and robotics researchers. In *Future of Life Institute*. http://futureoflife.org/open-letter-autonomous-weapons/. (accessed March 5, 2016).

Bansal, Sonia. 2016. Navigating the brain forest: Center for Neural Informatics, Structures, and Plasticity, Krasnow Institute for Advanced Study, George Mason University (28 March 2016). In *Showcasing DC Area Research, SfN, DC Metro Area Chapter*. https://sites.google.com/site/sfndcma/dc-area-research. (accessed June 1, 2016).

Bartlett, Frederic C. *Remembering: A Study in Experimental and Social Psychology*. Cambridge, England: Cambridge University Press, 1932. http://www.bartlett.psychol.cam.ac.uk/RememberingBook.htm.

Bernard, Tara Siegel, Tiffany Hsu, and Ron Lieber. 2017. Equifax says cyberattack may have affected 143 million in the US. In *The New York Times*. https://www.nytimes.com/2017/09/07/business/equifax-cyberattack.html. (accessed September 9, 2017).

Berra, Yogi. *The Yogi Book: I Really Didn't Say Everything I Said*. New York City, New York: Workman, 1998.

Bostrom, Nick, and Milan M. Cirkovic. *Global Catastrophic Risks*. Oxford, England: Oxford University Press, 2008.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford, England: Oxford University Press, 2014.

Bradshaw, J. M., Paul Feltovich, and Matthew Johnson. "Human-Agent Interaction." In *Handbook of Human-Machine Interaction,* edited by Guy Boy, 283-302. Ashgate, 2011.

Bradshaw, Jeffrey M., Larry Bunch, Matthew Johnson, and Micael Vignati. Agents, Ontologies, Policies, and Visualization: Building Blocks for the Human-Machine Interfaces of the Future (Presentation at a 17 March 2016 Meetup in Sunnyvale, CA). For a video of presentation, see jeffreymbradshaw.net. In *Silicon Valley Autonomous Vehicle Enthusiasts*. http://www.meetup.com/Autonomous/events/229486742/. (accessed June 1, 2016).

Bradshaw, Jeffrey M., ed. *Software Agents*. Cambridge, MA: The AAAI Press/The MIT Press, 1997.

Bradshaw, Jeffrey M., Marco Carvalho, Larry Bunch, Tom Eskridge, Paul J. Feltovich, Matthew Johnson, and Dan Kidwell. "Sol: An Agent-Based Framework for Cyber Situation Awareness." *Künstliche Intelligenz 26*, no. 2 (2012): 127-40.

Bradshaw, Jeffrey M., and Marco Carvalho. "Multi-agent systems for deep understanding of cyberspace (Abstract of an invited presentation)."

Presented at the Eighth Cyber Security and Information Intelligence Research Workshop (CSIIRW 2013), Oak Ridge National Labs, Oak Ridge, TN, January, 2013. http://www.jeffreymbradshaw.net/publications/121102-Multi-Agent%20Systems%20Abstract%20v.%202.pdf. (accessed March 15, 2016).

Bradshaw, Jeffrey M., Robert R. Hoffman, Matthew Johnson, and David D. Woods. "The seven deadly myths of 'autonomous systems'." *IEEE Intelligent Systems 28*, no. 3 (May/June 2013): 54-61. http://jeffreymbradshaw.net/publications/IS-28-03-HCC_1.pdf. (accessed March 15, 2016).

Bradshaw, Jeffrey M., Andrzej Uszok, Maggie Breedy, Larry Bunch, Thomas C. Eskridge, Paul J. Feltovich, Matthew Johnson, James Lott, and Michael Vignati. "The KAoS Policy Services Framework." Presented at the Eighth Cyber Security and Information Intelligence Research Workshop (CSIIRW 2013), Oak Ridge National Labs, Oak Ridge, TN, January, 2013. http://www.jeffreymbradshaw.net/publications/CSIIRW%20KAoS%20paper-s.pdf. (accessed March 15, 2016).

Bradshaw, Jeffrey M., Andrzej Uszok, and Rebecca Montanari. "Policy-Based Governance of Complex Distributed Systems: What Past Trends Can Teach Us about Future Requirements." In *Engineering Adaptive and Resilient Computing Systems,* edited by Niranjan Suri and G. Cabri, 259-84. Boca Raton, FL: CRC Press/Taylor and Francis, 2014. http://www.jeffreymbradshaw.net/publications/130529-PolicyDraft2May-jb-rm-jb.pdf.

Bradshaw, Jeffrey M. 2016. Designing Software Agents: Perils, Pitfalls, and Promise. Video Presentation in Two Parts. In *Vivint Innovation Center, Lehi, UT*. http://www.jeffreymbradshaw.net/. (accessed September 9, 2017).

———. 2016. Human-Agent-Robot Teamwork Through Coactive Design. Academic Ceremony on the Occasion of the Bestowal of an Honorary Doctorate for Sebastian Thrun. Video Recording of invited Presentation at the TU Delft 174th Dies Natalis Seminar on Robots: Tools or Teammates? (8 January 2016). In *Delft Technical University (TU Delft)*. http://jeffreymbradshaw.net/videos/160108-Bradshaw-Human-Agent-Robot%20Teamwork%20Through%20Coactive%20Design-TU%20Delft-174th%20Dies%20Natalis%20Seminar-Robots-Tools%20or%20Teammates-edited.mp4. (accessed March 15, 2016).

Bradshaw, Jeffrey M., and Matthew Johnson. 2016. Lessons learned in the theory and practical application of Human-Agent-Robot Teamwork. . Invited presentation at the University of Kaiserslautern, Kaiserslautern, Germany, 17 February 2016. In *University of Kaiserslautern*. http://jeffreymbradshaw.net/videos/Clip2.mov.

———. 2016. Video of Presentation entitled "Human-Agent-Robot Teamwork Through Coactive Design." Academic Ceremony on the Occasion of the Bestowal of an Honorary Doctorate for Sebastian Thrun. Invited Presentation at the TU Delft 174th Dies Natalis Seminar on Robots: Tools

or Teammates? Delft, The Netherlands, 8 January 2016. In *Jeffrey M. Bradshaw - Selected Publications*. www.jeffreymbradshaw.net. (accessed June 1, 2016).

Bradshaw, Robert W. E-mail message to Jeffrey M. Bradshaw, January 30, 2016.

Bradshaw to chair new Nissan advisory board. In *IHMC News*. http://www.ihmc.us/bradshaw-to-chair-new-nissan-advisory-board/. (accessed March 15, 2016).

The BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies). In *The BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies)*. https://www.whitehouse.gov/BRAIN. (accessed June 29, 2016).

Brain power. In *IBM Research*. http://www.research.ibm.com/cognitive-computing/neurosynaptic-chips.shtml#fbid=bptEJCBVy2p. (accessed June 23, 2016).

Brown, Samuel Morris. "Enhancing evolution: Posthumanous dreams and the moral complexity of biomedical aspirations." *BYU Studies 48*, no. 2 (2009): 41-49.

Bunch, Larry, Jeffrey M. Bradshaw, Marco Carvalho, Tom Eskridge, Paul J. Feltovich, James Lott, and Andrzej Uszok. "Human-Agent Teamwork in Cyber Operations: Supporting Co-Evolution of Tasks and Artifacts with Luna." Presented at the Tenth German Conference on Multiagent System Technologies (MATES 2012) (LNAI 7598), Trier, Germany, October 10-12, 2012, 53-67.

Bunch, Larry, Jeffrey M. Bradshaw, Robert R. Hoffman, and Matthew Johnson. "Principles for human-centered interaction design, part 2: Can humans and machines think together?" *IEEE Intelligent Systems 30*, no. 3 (May/June 2015): 68-75. http://www.jeffreymbradshaw.net/publications/30-03-HCC.PDF. (accessed March 15, 2016).

Calabro, David. E-mail message to Jeffrey M. Bradshaw, March 9, 2016.

Changizi, Mark. 2012. Artificial brains: Not in this century. In *The Huffington Post*. http://www.huffingtonpost.com/mark-changizi-phd/artificial-brains-not-in-_b_1248948.html. (accessed June 24, 2016).

Chen, Adrian. 2017. The Agency. In *The New York Times*. https://www.nytimes.com/2015/06/07/magazine/the-agency.html. (accessed September 9, 2017).

Clarke, Arthur C. "The future isn't what it used to be." *Engineering and Science 33*, no. 7 (1970): 4-9. http://resolver.caltech.edu/CaltechES:33.7.clarke. (accessed September 18, 2015).

———. "Foreword: The birth of HAL." In *HAL's Legacy: 2001's Computer as Dream and Reality,* edited by David G. Stork, xi-xvi. Cambridge, MA: The MIT Press, 1997.

Clarke, Roger. "Asimov's laws of robotics: Implications for information technology, Parts 1 and 2." *IEEE Computer*, December/January 1993-1994, 53-61/57-66.

Cole-Turner, Ronald, ed. *Transhumanism and Transcendence: Christian Hope in an Age of Technological Enhancement*. Washington, DC: Georgetown University Press, 2011.

Cordeschi, Roberto. "The discovery of the artificial. Some protocybernetic developments 1930-1940." *AI and Society 5* (1991): 218-38. http://philpapers.org/archive/CORTDO-9. (accessed March 12, 2016).

Cummings, Mary (Missy) Louise, and Jason C. Ryan. "Who is in charge? The promises and pitfalls of driverless cars." *TR News (Transportation Research Board, The National Academies of Sciences, Engineering, and Medicine) 292* (May-June 2014): 25-30. http://www.trb.org/Publications/Blurbs/171018.aspx. (accessed June 1, 2016).

Davis, Ernest, and Gary Marcus. "Commonsense reasoning and commonsense knowledge in Artificial Intelligence." *Communications of the ACM 58*, no. 9 (September 2015): 92-103.

Deep Blue (chess computer). In *Wikipedia*. https://en.wikipedia.org/wiki/Deep_Blue_(chess_computer). (accessed June 8, 2016).

Dennison Lott Harris. In *Find a Grave*. http://www.findagrave.com/cgi-bin/fg.cgi?page=gr&GRid=33937079. (accessed April 18, 2016).

Dietterich, Thomas G., and Eric J. Horvitz. "Viewpoint: Rise of Concerns about AI: Reflections and Directions." *Communications of the ACM 58*, no. 10 (October 2015 2015): 38-40. http://cacm.acm.org/magazines/2015/10/192386-rise-of-concerns-about-ai/fulltext. (accessed March 5, 2016).

Dodd, D. H., and J. M. Bradshaw. "Leading questions and memory: Pragmatic constraints." *Journal of Verbal Learning and Verbal Behavior 19* (1980): 695-704.

Doyle, Charles Clay, Wolfgang Mieder, and Fred R. Shapiro. *The Dictionary of Modern Proverbs*. Princeton, NJ: Yale University Press, 2012.

Eaglen, Mackenzie. 2016. What is the third offset strategy? In *Real Clear Defense*. https://www.realcleardefense.com/articles/2016/02/16/what_is_the_third_offset_strategy_109034.html. (accessed September 26, 2017).

Eliason, Eric A. *The J. Golden Kimball Stories*. Urbana, IL: University of Illinois Press, 2007.

Exploring our robotic future. In *Delta Magazine Online, 12 January 2016*, TU Delft. http://delta.tudelft.nl/article/exploring-our-robotic-future/31013. (accessed June 1, 2016).

Feigenbaum, Edward A., Jeffrey M. Bradshaw, Alexander Felfernig, Ali-Akbar Ghorbani, Sven Koenig, Sankar Pal, David M. W. Powers, Vijay Raghaven, Eunice Santos, Takahira Yamaguchi, and Yiyu Yao. "Panel on Top Ten Questions in Intelligent Informatics and Computing in Celebration of the Alan Turing Year." Presented at the World Intelligence Congress, Macau, China, 4-7 December 2012, 2012. http://wi-consortium.org/blog/top10qi/#cfp. (accessed 9 September 2017).

Fife, Austin, and Alta Fife. *Saints of Sage and Saddle*. Bloomington, IN: Indiana University Press, 1956.

Ford, K.M., J.M. Bradshaw, J.R. Adams-Webber, and N.M. Agnew. "Knowledge acquisition as a constructive modeling activity." In *Knowledge Acquisition as Modeling,* edited by K.M. Ford and J.M. Bradshaw, 9-32. New York: John Wiley, 1993.

Fountain, Henry. 2017. Apocalyptic thoughts amid nature's chaos? You could be forgiven. In *The New York Times*. https://www.nytimes.com/2017/09/08/us/hurricane-irma-earthquake-fires.html. (accessed September 9, 2017).

Furber, Steve. "To build a brain." *IEEE Spectrum* (August 2012): 45-49.

Gage, Fred H. *What is the brain good for? Special issue of Daedalus: Journal of the American Academy of Arts and Science 144:1*, 2015.

Gaines, Brian R. "An ounce of knowledge is worth a ton of data: Quantitative studies of the trade-off between expertise and data based on statistically well-founded empirical induction." In *Proceedings of the Sixth International Workshop on Machine Learning,* edited by Alberto Maria Segre, 156-59. San Francisco, CA: Morgan Kaufman, 1989. http://pages.cpsc.ucalgary.ca/~gaines/reports/ML/ML89/ML89.pdf. (accessed June 10, 2016).

Gray, John. *The Immortalization Commission: Science and the Strange Quest to Cheat Death*. New York City, NY: Farrar, Straus and Giroux, 2011.

Greaves, Mark, Heather Holmback, and Jeffrey M. Bradshaw. "What is a conversation policy?" Presented at the Proceedings of the Autonomous Agents '99 Workshop on Specifying and Implementing Conversation Policies, Seattle, WA, 1-5 May 1999, 1999, 1-9.

Green, Christopher C., Diane E. Griffin, James J. Blascovich, Jeffrey M. Bradshaw, Scott C. Bunch, John Gannon, Michael Gazzaniga, Elizabeth Loftus, Gregory J. Moore, Jonathan Moreno, John R. Rasure, Mark (Danny) Rintoul, Nathan D. Schwade, Ronald L. Smith, Karen S. Walch, and Alice M. Young. *Emerging Cognitive Neuroscience and Related Technologies. A Report of the Committee on Military and Intelligence Methodology for Emergent Neurophysiological and Cognitive/Neural Science Research in the Next Two Decades*. Washington, DC: The National Academies Press, 2008. http://www.nap.edu/catalog/12177/emerging-cognitive-neuroscience-and-related-technologies. (accessed March 5, 2016).

Greenberg, Gary. 2008. Hackers cut cities' power. In *Forbes*. http://www.forbes.com/2008/01/18/cyber-attack-utilities-tech-intel-cx_ag_0118attack.html. (accessed May 17, 2016).

Greenfield, Suan. *Mind Change: How Digital Technologies Are Leaving Their Marks on Our Brains*. New York City, NY: Random House, 2015.

Guha, R.V., and Douglas B. Lenat. "Cyc: A midterm report." *AI Magazine*, Fall 1990, 33-58.

Hacking power networks. In *Schneier on Security*. https://www.schneier.com/blog/archives/2008/01/hacking_power_n.html. (accessed May 17, 2016).

Harris, Mark. 2015. Will Nissan beat Google and Uber to self-driving taxis? In *IEEE Spectrum*. http://spectrum.ieee.org/cars-that-think/transportation/self-driving/will-nissan-beat-google-and-uber-to-self-driving-taxis. (accessed February 27, 2015).

Heisler, Yoni. In 1983 speech, Steve Jobs alluded to the iPad, Siri, the App Store, mainstream Internet connectivity, Google Maps and more. In *Engadget*. http://www.engadget.com/2013/11/14/in-rare-1983-speech-steve-jobs-alluded-to-the-ipad-siri-the-a/. (accessed September 18, 2015).

Hicks, Kathleen H., Andrew Hunter, Jesse Ellman, Lisa Samp, and Gabriel Coll. 2017. Assessing the Third Offset Strategy. In *CSIS: Center for Strategic and International Studies*. https://csis-prod.s3.amazonaws.com/s3fs-public/publication/170302_Ellman_ThirdOffsetStrategySummary_Web.pdf?EXO1GwjFU22_Bkd5A.nx.fJXTKRDKbVR. (accessed September 26, 2017).

Hinckley, Gordon B. "Thanks to the Lord for His blessings." *Ensign 29*, May, 88-89.

Hoffman, Robert, Paul Feltovich, Kenneth M. Ford, David D. Woods, Gary Klein, and Anne Feltovich. "A rose by any other name... would probably be given an acronym." *IEEE Intelligent Systems*, July-August 2002, 72-80.

Hofstadter, Douglas R. *Gödel, Escher, Bach: An Eternal Golden Braid*. New York City, NY: Vintage Books, 1979.

Holland, Jeffrey R. "'Tomorrow the Lord will do wonders among you' [Joshua 3:5]." *Ensign 46*, May 2016, 124-27.

Holmback, Heather, Mark T. Greaves, and Jeffrey M. Bradshaw. " 'Agent A, can you pass the salt?': The role of pragmatics in agent communication (expanded version accessible online)." In *Proceedings of the Third Annual Conference on Autonomous Agents (Autonomous Agents '99)*, edited by Oren Etzioni, J. P. Müller and Jeffrey M. Bradshaw, 368-69. Seattle, WA: New York City, NY: ACM Press. http://www.jeffreymbradshaw.net/publications/Pragmatics.submitted.new.pdf. (accessed June 18, 2016).

How it works. In *Google Self-Driving Car Project*. https://www.google.com/selfdrivingcar/how/. (accessed March 3, 2016).

Human Brain Project. In *APT Advanced Processor Technologies Research Group*. http://apt.cs.manchester.ac.uk/projects/HBP/. (accessed June 29, 2016).

Huxley, Thomas Henry, Michael Foster, and E. Ray Lankester. *The Scientific Memoirs of Thomas Henry Huxley*. 4 vols. Vol. 3. New York City, NY: Macmillan, 1901. https://books.google.com/books/about/The_scientific_memoirs_of_Thomas_Henry_H.html?id=9-oEAAAAYAAJ. (accessed February 28, 2016).

Huxley, Thomas Henry. 1870. "Biogenesis and abiogenesis (The presidential address to the British Association for the Advancement of Science for 1870)." In *Collected Essays*. Vol. 8, 229-71. Cambridge, England: Cambridge University Press, 2011. http://ebooks.cambridge.org/chapter.jsf?bid=CBO9781139149273&cid=CBO9781139149273A013. (accessed February 28, 2016).

Jobs, Steve. 1983. The future isn't what it used to be (Presentation to the International Design Conference in Aspen (IDCA), 15 June 1983). In *SoundCloud*. https://w.soundcloud.com/player/?url=http%3A%2F%2Fapi.soundcloud.com%2Ftracks%2F62010118&show_artwork=true. (accessed September 18, 2015).

Johnson, Matthew, Jeffrey M. Bradshaw, Robert R. Hoffman, Paul J. Feltovich, and David D. Woods. "Seven cardinal virtues for human-machine teamwork: Examples from the DARPA Robotic Challenge." *IEEE Intelligent Systems 29*, no. 6 (November/December 2013): 74-80. http://www.jeffreymbradshaw.net/publications/56.%20Human-Robot%20Teamwork_IEEE%20IS-2014.pdf. (accessed March 15, 2016).

Johnson, Matthew, Jeffrey M. Bradshaw, Paul J. Feltovich, Catholijn M. Jonker, M. Birna van Riemsdijk, and Maarten Sierhuis. "Coactive design: Designing support for interdependence in joint activity." *Journal of Human-Robot Interaction 3*, no. 1 (2014): 43-69. http://jeffreymbradshaw.net/publications/173-1278-1-PB.pdf. (accessed March 15, 2016).

Jones, Randolph M., Ryan O'Grady, Denise Nicholson, Robert R. Hoffman, Larry Bunch, Jeffrey M. Bradshaw, and Ami Bolton. "Modeling and integrating cognitive agents within the emerging cyber domain." Presented at the Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), 20 November-4 December 2015, Orlando, FL 2015. http://www.jeffreymbradshaw.net/publications/15232-iitsec-v5.pdf. (accessed March 15, 2016).

Joseph, Susan, Andrea Westerinen, Jeffrey M. Bradshaw, Peter Sell, and Sandi Roddy. 2012. Digital Policy Management: Be part of the solution, not the problem. In *RSA Conference (28 February 2012)*. http://www.rsaconference.com/events/us12/agenda/sessions/721/digital-policy-management-be-part-of-the-solution. (accessed March 15, 2016).

Kaplan, Mordecai M. "Random thoughts." *The Reconstructionist 16*, no. 2 (March 1950): 30.

Kasparov playing against Deep Blue (8 June 2016). 2016. In *Britannica Online for Kids*. http://kids.britannica.com/comptons/art-158037/Garry-Kasparov-plays-against-Deep-Blue-the-chess-playing-computer. (accessed June 8, 2016).

Kelly, John E., III. Computing, cognition, and the future of knowing: How humans and machines are forging a new age of understanding. In *IBM Research: Cognitive Computing*. http://www.research.ibm.com/software/IBMResearch/multimedia/Computing_Cognition_WhitePaper.pdf. (accessed June 23, 2016).

Kendall, Frank. 2014. Terms of Reference — Defense Science Board 2015 Summer Study on Autonomy. In *Defense Science Board*. http://www.acq.osd.mil/dsb/tors/TOR-2014-11-17-Summer_Study_2015_on_Autonomy.pdf. (accessed March 15, 2016).

Koene, Randal A. "Feasible mind uploading." In *Intelligence Unbound: The Future of Uploaded and Machine Minds,* edited by Russell Blackford and

Damien Broderick, 90-101. Chichester, West Sussex, England: John Wiley, 2014.
Kurzweil, Ray. *The Singularity is Near: When Humans Transcend Biology*. New York City, NY: Penguin Books, 2005.
Latson, Jennifer. 2015. Did Deep Blue beat Kasparov because of a system glitch? (17 February 2015). In *TIME*. http://time.com/3705316/deep-blue-kasparov/. (accessed June 8, 2016).
Leigher, William E. 2011. Learning to operate in cyberspace. In *US Naval Institute Proceedings Magazine, February 2011, pp. 32-37*. https://www.usni.org/magazines/proceedings/2011-02/learning-operate-cyberspace. (accessed September 26, 2017).
Lenat, Douglas B., M. Prakash, and M. Shepard. "CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks." *AI Magazine 6*, no. 4 (1986).
Lenat, Douglas B., and R.V. Guha. *Building Large Knowledge-based Systems*. Reading, MA: Addison-Wesley, 1990.
Lenat, Douglas B. "From 2001 to *2001*: Common sense and the mind of HAL." In *HAL's Legacy: 2001's Computer as Dream and Reality,* edited by David G. Stork, 193-209. Cambridge, MA: The MIT Press, 1997.
———. 2015. Computers with common sense. TEDxYouth@Austin (video published 21 May 2015). In *YouTube*. https://www.youtube.com/watch?v=2w_ekB08ohU. (accessed March 15, 2016).
Lewis, C. S. 1946. *The Great Divorce*. New York City, NY: Touchstone, 1996.
Liedtke, Michael. Renault-Nissan to introduce ten self-driving vehicles by 2020 (7 January 2016). In *The Washington Times*. http://www.washingtontimes.com/news/2016/jan/7/renault-nissan-to-introduce-10-self-driving-vehicl/. (accessed June 1, 2016).
Lindner, Roland. Google warning about autonomous cars. In *Frankfurter Allgemeine (13 January 2016, translated from German).* http://www.faz.net/aktuell/wirtschaft/netzwirtschaft/google/testfahrten-google-warnt-vor-eigenen-roboterautos-14012386.html (accessed March 4, 2016).
Loftus, Elizabeth. *Eyewitness Testimony*. Cambridge, MA: Harvard University Press, 1979.
———. "Creating false memories." *Scientific American 277*, no. 3 (September 1997): 70-75. http://faculty.washington.edu/eloftus/Articles/sciam.htm. (accessed June 24, 2016).
———. 2013. How reliable is your memory? In *TED Talks*. https://www.ted.com/speakers/elizabeth_loftus. (accessed June 11, 2016).
Lohr, Steve. 2016. Stepping up security for and Internet of Things world (16 October 2016). In *The New York Times*. http://www.nytimes.com/2016/10/17/technology/security-internet.html. (accessed October 18, 2016).
Madsen, Barnard N. *The Truman G. Madsen Story: A Life of Study and Faith*. Salt Lake City, UT: Deseret Book, 2016.

Malcomson, Scott. *Splinternet: How Geopolitics and Commerce Are Fragmenting the World Wide Web*. New York City, NY: OR Books, 2016.
Manjoo, Farhad. 2017. Serious, Equifax? This is a breach no one should get away with. In *The New York Times*. https://www.nytimes.com/2017/09/08/technology/seriously-equifax-why-the-credit-agencys-breach-means-regulation-is-needed.html. (accessed September 9, 2017).
Markoff, John. 2016. Artificial Intelligence is far from matching humans, panel says (25 May 2016). In *The New York Times*. http://www.nytimes.com/2016/05/26/technology/artificial-intelligence...ns-panel-says.html. (accessed June 8, 2016).
———. 2016. As artificial intelligence evolves, so does its criminal potential (23 October 2016). In *The New York Times*. https://www.nytimes.com/2016/10/24/technology/artificial-intelligence-evolves-with-its-criminal-potential.html. (accessed October 25, 2016).
———. 2016. For now, self-driving cars still need humans. In *The New York Times*. http://nyti.ms/1ncdLfl. (accessed January 20, 2016).
———. 2016. How tech giants are devising real ethics for Artificial Intelligence (1 September 2016). In *The New York Times*. https://www.nytimes.com/2016/09/02/technology/artificial-intelligence-ethics.html. (accessed September 2, 2016).
Matt, Daniel C., ed. *The Zohar, Pritzker Edition*. Vol. 1. Stanford, CA: Stanford University Press, 2004.
Maxwell, Neal A. "The stern but sweet seventh commandment (Devotional, Salt Lake Institute of Religion, 8 December 1978)." Salt Lake City, UT: Salt Lake Institute of Religion, The Church of Jesus Christ of Latter-day Saints, 1978. Reprint, Abridged Versions: The stern but sweet seventh commandment, *New Era*, June 1979 (https://www.lds.org/new-era/1979/06/the-stern-but-sweet-seventh-commandment?lang=eng), *New Era*, March 2003 (https://www.lds.org/new-era/2003/03/reasons-to-stay-pure?lang=eng).
———. 1981. "Grounded, rooted, established, settled (Ephesians 3:17, 1 Peter 5:10) (BYU Address, 15 September 1981)." In *The Inexhaustible Gospel: A Retrospective of Twenty-One Firesides and Devotionals at Brigham Young University 1974-2004*, 109-21. Provo, UT: Brigham Young University, 2004.
McConkie, Bruce R. 1980. "The seven deadly heresies (BYU fireside address given 1 Junne 1980)." In *I Believe: A Retrospective of Twelve Firesides and Devotionals, Brigham Young University, 1973-1985*, 97-111. Provo, UT: Brigham Young University, 2006.
McDermott, Drew. "Artificial intelligence meets natural stupidity." *SIGART Newsletter*, no. 57 (1976): 4-9.
McGourty, Colin. 2014. Man vs Machine: A poet on Kasparov-Deep Blue (29 October 2014). In *Chess24*. https://chess24.com/en/read/news/man-vs-machine-a-poet-on-kasparov-deep-blue. (accessed June 8, 2016).

McMillan, Graeme. 2011. IBM aims to build artificial brain within 10 years. In *Time.com*. http://techland.time.com/2011/10/05/ibm-aims-to-build-artificial-human-brain-within-10-years/. (accessed June 23, 2016).
Metz, Cade. 2015. Google is 2 billion lines of code — and it's all in one place (15 September 2015). In *WIRED*. http://www.wired.com/2015/09/google-2-billion-lines-codeand-one-place. (accessed June 8, 2016).
Minsky, Marvin, ed. *Semantic Information Processing*. Boston, MA: The MIT Press, 1968.
———. 1961. "Steps toward artificial intelligence." In *Computers and Thought,* edited by Edward A. Feigenbaum and Julian Feldman, 406-50. New York City, NY: McGraw-Hill, 1963.
Mok, Kimberley. 2016. Artificial 'imagination' helped Google AI master Go, the most complex game ever invented. In *The New Stack*. http://thenewstack.io/google-ai-beats-human-champion-complex-game-ever-invented/. (accessed June 8, 2016).
Moody, Sidney. 1999. The brain behind Cyc: Scientist Doug Lenat discusses Artificial Intelligence. In *The Austin Chronicle*. http://www.austinchronicle.com/screens/1999-12-24/75252/. (accessed March 15, 2016).
The most influential experiments on memory. In *Memory Distortions*. http://socialscientist.us/nphs/psychIB/psychpdfs/MemoryDistortions.pdf. (accessed June 29, 2016).
Neubauer, Hendrik. *Curious Moments: Archive of the Century (Archives of Text och Bilder, Stokholm, made available through SVT Bild/Das Fotoarchiv)*. Köln, Germany: Könemann, 1999.
New Scientific Advisory Council to assist at Nissan Research Center - Silicon Valley (12 March
2014). In *Nissan News Release*. http://nissannews.com/en-US/nissan/usa/releases/newscientific-advisory-council-to-assist-at-nissan-research-center-silicon-valley. (accessed March 15, 2016).
New tools counter cyber threats. In *IHMC Newsletter 10:1 (Feburary 2013)*. http://www.jeffreymbradshaw.net/publications/IHMCNewslettervol10iss1.pdf. (accessed March 15, 2016).
Nibley, Hugh W. "The Roman games as a survival of an archaic year-cult." Doctoral Dissertation, University of California, Berkeley, 1939.
———. *Enoch the Prophet. The Collected Works of Hugh Nibley 2*. Salt Lake City, UT: Deseret Book, 1986.
———. "But what kind of work?" In *Approaching Zion,* edited by Don E. Norton. The Collected Works of Hugh Nibley 9, 252-89. Salt Lake City, UT: Deseret Book, 1989.
———. 1951. "Sparsiones." In *The Ancient State,* edited by Donald W. Parry and Stephen D. Ricks. The Collected Works of Hugh Nibley 10, 148-94. Salt Lake City, UT: Deseret Book, 1991.
———. 1956. "*Victoriosa Loquacitas*: The Rise of Rhetoric and the Decline of Everything Else." In *The Ancient State: The Rulers and the Ruled,* edited by Donald W. Parry and Stephen D. Ricks. The Collected Works of Hugh Nibley 10, 243-86. Salt Lake City and Proovo, UT: Deseret Book and

Foundation for Ancient Research and Mormon Studies, 1991. Reprint, Western Speech 20 (Spring 1956), 47-82.
Nickelsburg, George W. E., and James C. VanderKam, eds. *1 Enoch: A New Translation*. Minneapolis, MN: Fortress Press, 2004.
Nielson, Reid L., and Mitchell K. Schaefer. "Excavating early Mormon history: The 1878 history fact-finding mission of apostles Joseph F. Smith and Orson Pratt." In *Joseph F. Smith: Reflections on the Man and His Times,* edited by Craig K. Manscill, Brian D. Reeves, Guy L. Dorius and J. B. Haws. Brigham Young University Church History Symposium, 359-78. Salt Lake City and Provo, UT: Deseret Book and Religious Studies Center, Brigham Young University, 2013.
Noessel, Christopher. *Designing Agentive Technology: AI That Works for People*. Brooklyn, NY: Rosenfeld Media, 2017.
Norman, Donald A. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Reading, MA: Addison-Wesley, 1993.
———. *The Invisible Computer: Why Good Products Can Fail, the Personal Computer is So Complex, and Information Appliances Are the Solution*. Cambridge, MA: The MIT Press, 1998.
———. *The Design of Future Things*. New York City, NY: Basic Books, 2007.
O'Toole, Garson. 2010. Facts Are Stubborn Things (18 June 2010). In *Quote Investigator: Exploring the Origins of Quotations*. http://quoteinvestigator.com/2010/06/18/facts-stubborn/. (accessed 1 May, 2016).
Offset strategy. In *Wikipedia*. https://en.wikipedia.org/wiki/Offset_strategy. (accessed September 26, 2017).
One Hundred Year Study on Artificial Intelligence (AI100). In *Stanford University*. https://ai100.stanford.edu/. (accessed September 9, 2017).
An open letter: Research priorities for robust and beneficial Artificial Intelligence. In *Future of Life Institute*. http://futureoflife.org/ai-open-letter/. (accessed March 5, 2016).
Ormerod, David. 2016. Lee Sedol defeats AlphaGo in masterfaul comeback - Game 4. In *Go Game Guru*. https://gogameguru.com/lee-sedol-defeats-alphago-masterful-comeback-game-4/. (accessed June 8, 2016).
Parekh, R., and Giorgio A. Ascoli. "Neuronal morphology goes digital: A research hub for cellular and system neuroscience." *Neuron 77*, no. 6 (2013): 1017-38.
Paul Valéry (1871-1945). In *The Poetry Foundation - Poems and Poets*. http://www.poetryfoundation.org/bio/paul-valery. (accessed January 22, 2016).
Pellerin, Cheryl. 2016. Deputy Secretary: Third Offset Strategy Bolsters America's Military Deterrence. In *DoD News, US Department of Defense*. https://www.defense.gov/News/Article/Article/991434/deputy-secretary-third-offset-strategy-bolsters-americas-military-deterrence/. (accessed September 26, 2017).
Petersen, Boyd Jay. *Hugh Nibley: A Consecrated Life*. Draper, UT: Greg Kofford Books, 2002.

Probst, Catherine. 2012. 'Mental Floss' project brings artists and scientists together (12 December 2012). In *Mason News*. http://tgillette.onair.cc/2013/08/11/mental-floss-project-brings-artists-and-scientists-together/. (accessed June 1, 2016).

Pynadath, David, and Milind Tambe. "Revisiting Asimov's first law: A response to the call to arms." Presented at the Proceedings of ATAL 01 2001.

Quatrochi, Alice M., and Giorgio A. Ascoli. "Mental Floss - 2012. Paper handout created to accompany the sculpture display." 2012.

Ray Kurzweil. In *Wikipedia*. https://en.wikipedia.org/wiki/Ray_Kurzweil. (accessed March 8, 2016).

Report of the Defense Science Board Summer Study on Autonomy (June 2016). 2016. In, Defense Science Board, Office of the Under Secretary of Defense for Acquisition, Technology and Logistics. http://www.jeffreymbradshaw.net/publications/DSB%20Summer%20Study%20on%20Autonomy%20Final%20Report%20-%20June%202016.pdf. (accessed 9 September, 2017).

Rich, Elaine, Kevin Knight, and Shivashankar B. Nair. *Artificial Intelligence*. New Delhi, India: Tata McGraw Hill Education Private LImited, 2009. http://itechnocrates.weebly.com/uploads/5/5/2/7/55270269/artificial_intelligence.pdf. (accessed March 12, 2016).

Rochefoucauld, Le Duc de La. 1665. *Réflexions ou Sentences et Maximes Morales de La Rochefoucauld (Dernier Texte)*. London, England: Louys Glady, 1880.

———. 1665. *Maxims of Le Duc de La Rochefoucauld*. Translated by John Heard. Boston, MA: International Pocket Library, 1917.

Rosenberg, Matthew, and John Markoff. 2016. The Pentagon's 'terminator conundrum': Robots that could kill on their own. In *The New York Times*. http://www.nytimes.com/2016/10/26/us/pentagon-artificial-intelligence-terminator.html. (accessed October 25, 2016).

Russell, Stuart, and Max Tegmark. 2015. Think-tank dismisses leading AI researchers as Luddites. In *Future of Life Institute*. http://futureoflife.org/2015/12/24/think-tank-dismisses-leading-ai-researchers-as-luddites/. (accessed March 5, 2016).

Salamon, Anna. 2015. Why startup founders have mood swings (and why they may have uses). In *Less Wrong*. http://lesswrong.com/lw/n0v/why_startup_founders_have_mood_swings_and_why/. (accessed January 22, 2016).

Sang-Hun, Choe. 2016. Google's computer program beats Lee Sedol in Go tournament. In *The New York Times (15 March 2016)*. http://www.nytimes.com/2016/03/16/world/asia/korea-alphago-vs-lee-sedol-go.html. (accessed March 31, 2016).

Shakespeare, William. 1600. "The Tragedy of Hamlet, Prince of Denmark." In *The Riverside Shakespeare,* edited by G. Blakemore Evans, 1135-97. Boston, MA: Houghton-Mifflin Company, 1974.

Shoham, Yoav, and M. Tennenholtz. "On the synthesis of useful social laws for artificial agent societies." Presented at the Proceedings of the Tenth National Conference on Artificial Intelligence, San Jose, CA 1992, 276-81.

Smith, Ira A., Phil R. Cohen, Jeffrey M. Bradshaw, Mark Greaves, and Heather Holmback. "Designing conversation policies using joint intention theory." Presented at the Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98), Paris, France, 2-8 July, 1998, 269-76.

Smith, Joseph, Jr. 1938. *Teachings of the Prophet Joseph Smith*. Salt Lake City, UT: Deseret Book, 1969.

Software as weaponry in a computer-connected world. In *The New York Times (7 June 2016)*. http://www.nytimes.com/2016/06/09/technology/software-as-weaponry...ected-world.html. (accessed June 10, 2016).

Stoffer, Shawn. 2000. Cyc: Building HAL. In *Department of Computer Science, University of New Mexico*. http://www.cs.unm.edu/~storm/docs/Cyc.htm. (accessed March 15, 2016).

Thune, John, Bill Nelson, Chris Urmson, Mike Ableson, Glen DeVos, Joseph Opkpaku, and Mary (Missy) Louise Cummings. 2016. Hands Off: The Future of Self-Driving Cars (Hearing, 15 March 2016). In *U.S. Senate Committee on Commerce, Science, and Transportation*. http://www.commerce.senate.gov/public/index.cfm/2016/3/hands-off-the-future-of-self-driving-cars. (accessed June 1, 2016).

Transhumanism. In *Wikipedia*. http://en.wikipedia.org/wiki/Transhumanism. (accessed June 10, 2016).

Tsang, Amie. 2017. Amazon 'reviewing' its website after it suggested bomb-making items (20 September 2017). In *The New York Times*. https://www.nytimes.com/2017/09/20/technology/uk-amazon-bomb.html. (accessed September 26, 2017).

Tucker, Patrick. 2014. The Military's New Year's Resolution for Artificial Intelligence. In *Defense One*. http://www.defenseone.com/technology/2014/12/militarys-new-years-resolution-artificial-intelligence/102102/?oref=search_Roadmap%20for%20AI. (accessed March 15, 2016).

———. 2015. The Pentagon is nervous about Russian and Chinese killer robots. In *Defense One*. http://www.defenseone.com/threats/2015/12/pentagon-nervous-about-russian-and-chinese-killer-robots/124465/?oref=d-river. (accessed March 15, 2016).

Turing, Alan M. "Computing machinery and intelligence." *Mind 59*, no. 236 (October) (1950): 433-60.

Underground Economy 2013 Conference, Lyon 2-6 September 2013. In *Republic of Serbia, Administration for the Prevention of Money Laundering*. http://www.apml.gov.rs/eng961/novost/Underground-economy-2013-Conference,-Lyon-2-%E2%80%93-6-September-2013.html. (accessed September 26, 2017).

Uszok, Andrzej, Jeffrey M. Bradshaw, James Lott, Matthew Johnson, Maggie Breedy, Michael Vignati, Keith Whittaker, Kim Jakubowski, and Jeffrey Bowcock. "Toward a Flexible Ontology-Based Policy Approach for

Network Operations Using the KAoS Framework." Presented at the The 2011 Military Communications Conference (MILCOM 2011) 2011, 1108-14.
Valéry, Paul. 1937. "Our destiny and literature." In *Reflections on the World Today*. Translated by Francis Scarfe, 131-55. New York City, NY: Pantheon Books, 1948.
Washington, Glynn, Stephanie Foo, and Murray Campbell. 2014. Kasparov vs. Deep Blue (8 August 2014). In *NPR Now*. http://www.npr.org/2014/08/08/338850323/kasparov-vs-deep-blue. (accessed June 8, 2016).
Webster, George. 2011. The future of airport security: Thermal lie-detectors and cloned sniffer dogs (25 November 2011). In *CNN*. http://edition.cnn.com/2011/11/25/tech/innovation/future-airport-security/index.html#! (accessed March 15, 2016).
The week of... Sebastian, Jefferey & Thomas. In *Delta Magazine, 11 January 2016, 48:8*, TU Delft. http://delta.tudelft.nl/artikel/the-week-of-sebastian-jefferey-thomas/30983. (accessed June 1, 2016).
Weinfeld, Moshe. "The Decalogue: Its significance, uniqueness, and place in Israel's tradition." In *Religion and Law: Biblical-Judaic and Islamic Perspectives,* edited by Edwin B. Firmage, Bernard G. Weiss and John W. Welch, 3-47. Winona Lake, IN: Eisenbrauns, 1990.
Weld, Daniel, and Oren Etzioni. "The first law of robotics: A call to arms." Presented at the Proceedings of the National Conference on Artificial Intelligence (AAAI 94) 1994, 1042-47.
Wellington, Naomi. "Whole brain emulation: Invasive vs. non-invasive methods." In *Intelligence Unbound: The Future of Uploaded and Machine Minds,* edited by Russell Blackford and Damien Broderick, 178-92. Chichester, West Sussex, England: John Wiley, 2014.
Work, Robert O., and Shawn Brimley. 2014. 20YY: Preparing for War in the Roboti Age. In *Center for a New American Security Publications*, Center for a New American Security. http://www.cnas.org/20YY-Preparing-War-in-Robotic-Age#.Vuhfh8c4Vqk. (accessed March 15, 2016).
Wurtz, Robert H. "Brain mechanisms for active vision." *Daedalus: Journal of the American Academy of Arts & Sciences 144*, no. 1 (Winter 2015): 10-21.
Zivotofsky, Ari Z. 2009. What's the truth about ... the Kohen Gadol's rope? (12 August 2009). In *Jewish Action: The Magazine of the Orthodox Union*. https://www.ou.org/jewish_action/08/2009/whats_the_truth_about_-_the_kohen_gadols_rope/. (accessed April 10, 2016).

Endnotes

1 With apologies to P. Valéry, Our Destiny; A. C. Clarke, Future Isn't (article); S. Jobs, Future Isn't, and others.

2 With apologies to D. McDermott, Artificial intelligence meets natural stupidity.

3 Yogi Berra: https://upload.wikimedia.org/wikipedia/commons/f/f1/1953_Bowman_Yogi_B

erra.jpg (21 January 2016); J. Golden Kimball: https://rsc.byu.edu/sites/default/files/j%20golden%20kimball.jpg; Arthur C. Clarke: http://www.famousauthors.org/famous-authors/arthur-c-clarke.jpg; Steve Jobs: http://images.huffingtonpost.com/2015-07-17-1437155465-1424656-SteveJobs.jpg; Cassette: http://lifelibertytech.com/wp-content/uploads/2012/10/Talk-by-Steven-Jobs-Cassette.jpg

4 A version of the saying that uses the word "ain't" is often credited to Yogi Berra: "The future ain't what it used to be." According to March 9: "The baseball great Yogi Berra writing in his 1998 volume *The Yogi Book* did claim that he used this expression. A precise timeframe was not given, but the saying was accompanied with a picture from 1974. Yogi also offered an interpretation: 'I just meant that times are different. Not necessarily better or worse. Just different.'"

5 As is the case with many prominent figures, some of the stories told about Elder Kimball never happened. "Once a nephew came to J. Golden and said: 'Well, Uncle Golden, I heard another J. Golden yarn today.' 'I'll bet the damn thing isn't genuine,' wheezed the old man. 'Seems like all the stories told these days are either about me or Mae West" (A. Fife et al., *Saints*, p. 314). Cf. E. A. Eliason, *J. Golden Kimball Stories*, p. 67.

6 See the full title of Y. Berra, Yogi Book.

7 See Seven Deadly Heresies; A. C. Clarke, Future Isn't (article). Clarke was prescient in anticipating many future developments.

8 S. Jobs, Future Isn't. See Y. Heisler, In 1983 Speech for many examples of what Jobs foretold accurately.

9 For a popular summary of initial evidence of some of the complex physiological, social, and cultural changes that technology is working upon us, see, e.g., S. Greenfield, Mind Change.

10 A. C. Clarke, Future Isn't (article), p. 4.

11 Paul Valéry: http://www.katakrak.net/sites/default/files/events/paul_valery.jpg (see also https://quoadsubjectum.files.wordpress.com/2012/08/paulvalery3.jpg, http://www.aphorism4all.com/images/1355814906.jpg)

12 The biography of Valéry on the website of "The Poetry Foundation" summarizes (Paul Valéry (1871-1945), Paul Valéry (1871-1945)):

> Paul Valéry occupies a position in the history of French letters that is at once strategic and highly problematic. Critics have affixed to him various labels, all of them partially correct. He has been called the last French symbolist, the first post-symbolist, a masterful classical prosodist, an advocate of logical positivism, and a cerebral narcissist. ... [H]e is understood as having broken away from symbolism, as having rejected the cult of poetry for its own sake in favor of a cult of the mind. These views need not be contradictory. ...
>
> Some facts about Valéry might predict a less than faultless comportment on Valéry's part during World War II and France's occupation by Germany: first, he had been

quietly but strongly "anti-Dreyfusard" during the famous Dreyfus affair … Furthermore, Valéry was also friendly with Marshal Philippe Petain, one of the leaders of France's pro-German Vichy government.

However, the poet did prove sympathetic to the Free French Movement led by General Charles de Gaulle, and of the Nazis he wrote in "War Economy for the Mind": "As for our enemies, we, and the whole world, know that their politics with regard to the mind has been reduced or limited for ten years to repressing the developments of intelligence, to depreciating the value of pure research, to taking often atrocious measures against those who consecrated themselves to these things, to favoring, even as far as endowed chairs and laboratories, worshippers of the idol to the detriment of independent creators of spiritual richness, and they have imposed on the arts as on the sciences the utilitarian ends which a power founded on declamations and terror pursues."

13 **error**March 9 found one earlier occurrence of this saying than Valéry's:

> The earliest evidence of this saying located by [*Quote Investigator*] was published in 1937 in a journal called "Epilogue" within an article titled "From a Private Correspondence on Reality" by Laura Riding and Robert Graves. The authors who were both prominent literary figures asserted that the perception of the future had changed:
>
> > The human mind has reached the end of temporal progress: the future is not what it used to be, and people talk with less and less progenitive self-precipitation into the future, and behave with more and more fatally decisive immediacy. The future, that is, contains nothing but scientific development. It is an involuntary spending and manipulation of physical forces, empty of consciousness: it no longer matters.

In 1950, Mordecai M. Kaplan wrote: "Men say the future isn't what it used to be. Neither is the past. Both are in need of reconstruction, if we are to have a livable present" (M. M. Kaplan, Random Thoughts, cited in C. C. Doyle et al., *The Dictionary of Modern Proverbs*, p. 90).

14 P. Valéry, Our Destiny, pp. 135, 143-144.

15 Ibid., p. 142.

16 Luke 20:30-31.

17 L. D. d. L. Rochefoucauld, Maxims, 93, p. 29. L. D. d. L. Rochefoucauld, Maxims, 93, p. 21: *« Les vieillards aiment à donner de bons préceptes, pour se consoler de n'être plus en état de donner de mauvais exemples. »*

18 D&C 88:118; 109:7, 14.

19 http://i.huffpost.com/gen/3052140/images/o-PEELING-ORANGE-facebook.jpg

20 Image licensed from www.shutterstock.com. Image reference 67038073

21 See www.ihmc.us/groups/jbradshaw/.

22 For a summary of some of the unique aspects of IHMC's approach to the DARPA Robotic Challenge, written for the general reader, see M. Johnson *et al.*,

Seven Cardinal Virtues. For a video presentation that includes a description of the application of the principles of coactive design to the DARPA Robotic Challenge, see J. M. Bradshaw *et al.*, Lessons Learned. For a detailed description of coactive design, see M. Johnson *et al.*, Coactive Design.

23 H. Neubauer, Curious Moments, p. 634. Image licensed from Black Star / Alamy Stock Photo, Image Reference A2Y241.

24 Ibid., p. 634.

25 From A. Salamon, Why Startup Founders.

26 Many of the thoughts in this section are drawn from and paraphrased from ibid..

27 I don't know who first came up with this version of the saying, but it was not the Duc de La Rochefoucauld, to whom so many places on the Web attribute it falsely. I like this particular wording of the sentiment, which came from a talk I heard Elder Maxwell give in 1978 (N. A. Maxwell, The stern but sweet seventh commandment (Devotional, Salt Lake Institute of Religion, 8 December 1978), p. 8).

The original source from which the basic thought is derived is probably Thomas Huxley, in his Presidential Address to the British Association for the Advancement of Science, Liverpool Meeting, 14 September 1870 (T. H. Huxley *et al.*, *Scientific Memoirs* 3, p. 580; T. H. Huxley, Biogenesis and Abiogenesis, p. 244): "But the great tragedy of Science — the slaying of a beautiful hypothesis by an ugly fact — which is so constantly being enacted under the eyes of philosophers, was played, almost immediately, for the benefit of Buffon and Needham."

See also the famous statement by John Adams in his 1770 legal defense of British soldiers involved in the Boston Massacre (G. O'Toole, Facts Are Stubborn Things, emphasis added):

> I will enlarge no more on the evidence, but submit it to you, gentlemen — *Facts are stubborn things*; and whatever may be our wishes, our inclinations, or the dictates of our passions, they cannot alter the state of facts and evidence: nor is the law less stable than the fact. If an assault was made to endanger their lives, the law is clear, they had right kill in their own defence.

Several earlier versions of the phrase "Facts are stubborn things" are known, going back to 1713 (ibid.). Joseph Smith uses "Facts are stubborn things" in the title of a *Times and Seasons* article (see J. Smith, Jr., Teachings, 15 September 1842, p. 266).

28 As Anna Solomon described it A. Salamon, Why Startup Founders:

> It's easy to point to the value in euphoria and optimism. You get lots of code written, recruit lots of funding and talent, write a perfect draft — it's the part of the cycle where you're drawn to working seventy hour weeks, checking off each and every item from your to-do list. But the "down" parts often feel like they're pointless at best, and dangerous or counterproductive at worst. … In our own pasts, we found ourselves wondering why our brains couldn't just

hang on to the momentum — why they insisted on taking us through stupid detours of despair or shame before returning us back to apparent "forward motion."

29 Ibid.

30 Ibid.

31 Decalogue.

32 Ibid.

33 What's the Truth.

34 For a link to the video of Sebastian Thrun's presentation, see www.jeffreymbradshaw.net. Prior to the ceremony, Thrun (Stanford), Tom Sheridan (MIT), Dirk Helbing (ETH Zurich), Catholijn Jonker (TU Delft) and I presented at a public seminar on "Intelligent Robots: Tools or Teammates" (Week of Sebastian, Jefferey & Thomas, Week of Sebastian, Jefferey & Thomas; Exploring, Exploring). My talk at the seminar was a critique of the current overselling of "full autonomy" and a brief overview of Human-Agent-Robot Teamwork, with examples from IHMC projects (J. M. Bradshaw *et al.*, Human-Agent-Robot Teamwork Through Coactive Design).

35 How It Works, How It Works.

36 Ibid..

37 Cf. L. D. d. L. Rochefoucauld, *Maxims*, 106, p. 31: "Knowledge is an absolute mastery of details, and since of details there is no end, ours is always an imperfect and superficial knowledge." (L. D. d. L. Rochefoucauld, *Maxims*, 106, p. 24): *« Pour bien savoir les choses, il en faut savoir le détail, et comme il est presque infini, nos connoissances sont toujours superficielles et imparfaites. »*

To their credit, Google has invested heavily in empirical research on driving in real-world environments and has openly critiqued failures of autonomous capabilities in these situations. For example, see this assessment, briefly summarized in R. Lindner, Google Warning:

> Since 2009, when Google began working on self-driving cars, its fleet has so far covered more than two million kilometers in autonomous mode. During that time, there were, according to the group's report, 272 cases where a failure of the autonomous capabilities had been determined. In 69 cases, the driver had intervened to avert accidents. Subsequent simulations had shown the probability that in 13 of these cases an accident would have occurred had there been no intervention by a human driver. In a blog post Google pointed out, however, that the number of such incidents is declining. Of the 13 dangerous situations, eight had occurred eight in 2014, while only five took place last year. But Google also warned that the number could rise again if the autonomous cars were used under complex conditions — for example in bad weather.

38 http://johnnyholland.org/wp-content/uploads/DonaldANorman.jpg.

[39] New Scientific Advisory Council, New Scientific Advisory Council; Bradshaw to Chair, Bradshaw to Chair.

[40] D. A. Norman, Design of Future Things, p. 13.

[41] Ibid., p. 15.

[42] R. Lindner, Google Warning.

[43] W. C. Fields in an unknown movie clip.

[44] D. A. Norman, Design of Future Things, pp. 113, 116.

[45] Ibid., p. 116.

[46] M. Harris, Will Nissan Beat Google and Uber.

[47] See, e.g., R. Lindner, Google Warning:

> Meanwhile, Google is looking for ways for its autonomous vehicles to close ranks with the auto industry. John Krafcik, who is in charge of the project, said on Tuesday at a press conference that took place near the auto show in Detroit that Google wants to enter into collaborations with many different companies. A number of automakers have sought contact with Google. A few weeks ago there was speculation that Google held discussions with the American automaker Ford shortly before the announcement of the alliance. Some wonder whether Ford could be contracted to manufacture Google's next generation of self-driving cars. So far, these speculations have not materialized. Ford's CEO Mark Fields distanced himself last week at the electronics show in Las Vegas from such cooperation and said he would not be limited to the role of a supplier for the technology industry.

[48] M. Harris, Will Nissan Beat Google and Uber. See also the discussion of some of IHMC's recent work with Nissan on intelligent fleet management services in Excavating.

[49] Jeffrey M. Bradshaw, 6 January 2016, Image Reference IMG_6117.jpg.

[50] Although the details of many aspects of Nissan research are confidential, I was invited to give a public presentation at a meetup of the Silicon Valley Autonomous Vehicle Enthusiasts about IHMC's research on visualization for fleet management. See J. M. Bradshaw *et al.*, Agents, Ontologies, Policies, and Visualization for a video of that presentation.

[51] AP Photo/Terry Chea, in M. Liedtke, Renault-Nissan to Introduce.

[52] J. Markoff, For Now Self-Driving Cars.

[53] J. Thune *et al.*, Hands Off. See also M. M. L. Cummings *et al.*, Who Is In Charge?

[54] N. Bostrom, Superintelligence.

[55] Doug Lenat's Cyc can be seen as a philosophical grandfather to such efforts. In 1999, Lenat wrote: "HAL was a general artificial intelligence, and Cyc is the closest thing that exists in the world to that kind of general artificial intelligence" (S. Moody, Brain). See also D. B. Lenat, From 2001.

[56] https://upload.wikimedia.org/wikipedia/commons/thumb/f/f6/HAL9000.svg/1024px-HAL9000.svg.png.

[57] A. C. Clarke, Foreword, p. xi.

[58] http://www.cyc.com/wp-content/uploads/2015/04/kbase.png.

[59] For brief summaries of some of the most common criticisms of Cyc, see **error**Dennison Lott Harris, Dennison Lott Harris; S. Stoffer, Cyc. For Lenat's own assessments of his work on Cyc, see R. V. Guha et al., Cyc: A midterm report; D. B. Lenat et al., CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks; D. B. Lenat et al., Building Large Knowledge-based Systems; D. B. Lenat, From 2001. For a recent video perspective by Lenat, see D. B. Lenat, Computers with Common Sense.

[60] See B. R. Gaines, Ounce of Knowledge for a preliminary study that suggests "the possibility of developing a quantitative science of knowledge in terms of the amount of data reduction that knowledge buys us when carrying out empirical induction."

[61] J. E. Kelly, III, Computing, Cognition, and the Future.

[62] G. McMillan, IBM Aims to Build. In fairness to IBM, it should be noted that such media reports exaggerated the nature of their project. On the other hand, IBM's description of their work on their SyNAPSE chip tends to mislead non-specialist readers who may assume that building a "cognitive chip" is just a few small steps away from building the equivalent of a human brain. According to a description given by IBM, SyNAPSE is "a new chip with a brain-inspired computer architecture powered by an unprecedented 1 million neurons and 256 million synapses" while consuming only "70mW during real-time operation — orders of magnitude less energy than traditional chips" (Brain Power, Brain Power).

[63] https://commons.wikimedia.org/wiki/File:IBM_Watson.PNG.

[64] Anonymous, How IBM. Cf. J. E. Kelly, III, Computing, Cognition, and the Future.

[65] E. Davis *et al.*, Commonsense Reasoning, p. 92.

[66] R. Hoffman *et al.*, Rose, p. 75.

[67] See, e.g., H. Holmback *et al.*, 'Agent A, can you pass the salt?': The role of pragmatics in agent communication (expanded version accessible online) See also I. A. Smith *et al.*, Designing conversation policies using joint intention theory; M. Greaves *et al.*, What is a conversation policy?.

[68] Kasparov Playing, Kasparov Playing.

[69] Deep Blue, Deep Blue.

[70] J. Latson, Did Deep Blue Beat Kasparov?

[71] C. McGourty, Man vs Machine.

[72] G. Washington *et al.*, Kasparov Vs. Deep Blue.

[73] C. Metz, Google Is Two Billion Lines of Code.

[74] Ibid.

[75] Ibid.

[76] K. Mok, Artificial Imagination.

[77] Ibid.

[78] Anonymized, January 30 2016.

[79] Ibid.

[80] C. Sang-Hun, Google's Computer Program.

[81] http://www.goratings.org/ (accessed 21 June 2016).

[82] https://gogameguru.com/i/2016/01/DeepMind-AlphaGo.jpg.

[83] R. W. Bradshaw, January 30 2016. Robert's cogent summary of what makes AlphaGo's approach interesting is informative:

> Your best move is the one that maximizes your position. Your position is defined by your opponent's best move. So it's really a recursive definition. Of course for any reasonably complex game, you can't all the way to the win/lose state, so at some point you have to "judge" a position and/or move on its own (without looking forward to what could happen) based on its intrinsic properties. Both computers and humans do both, but computers tend to (need to) compensate for poor "leaf position" judgment abilities by searching much broader and deeper.
>
> For chess, this works well enough both because the number of sensible (or at least legal) moves is often quite limited (say tens, often less) and one can come up with simple scoring "limited intelligence required" algorithms that work pretty well (e.g. assign a number of points to each live player, and add them up--even great players have trouble coming back from a significant deficit using the standard weights). These two facts, some clever tricks, and the computing power in our reach, make Deep Blue (and its successors) a reality.
>
> On the other hand, for Go, both the number of possible moves at any point in the game is so great as to make it infeasible to search a reasonable fraction of moves deeply, and the "goodness" of a position is harder to quantify. This is where the AI comes into play — it's used to (intrinsically but opaquely) judge how "reasonable/good" a move is, and how "good" a position is, (more) similar to how a person works, drastically reducing the search space (though it still tries out *way* more possibilities a human does). This better AI is the differentiator here.
>
> Put another way, the AI problem they're trying to solve with deep neural nets is "what's the best move here" and though they make a great leap forward (agreeing with expert players over half the time now) it still needs the to be "amplified" by search to be competitive. But the fact that its good enough that when paired with a *feasible* search it can compete with expert humans makes this a measurable milestone. (And the claim is that without *any*

search it still plays at the early amateur level.) Whether that's intelligent, well, the colloquial definition of AI excludes what has been accomplished.

Of course these aren't Go experts, they're just interested in (incrementally) advancing the state of AI in general. It seems the most novel theoretical aspect of this is in "reinforcement learning." There just aren't the same vast quantities of (expert-level) Go games played to rely on volume alone (like one can with, say, machine translation) without over-fitting, so they've had to rely more heavily on these self-learning techniques which is a lot trickier to get right (similar to how we had missionary apartments that only spoke French among themselves, which helped vocabulary but was horrible in reinforcing bad accents), and this helps validate that work.

I personally find this most exciting because a lot of improvements in [machine learning] seem to have been due to being able to train on orders of magnitude more data (which, of course, is non-trivial theoretical and engineering challenge), but once you've trained on (say) all digitized bilingual text in the world there's not much further to go from there. These techniques are starting to explore the path of more effectively extracting "intelligence" out of (relatively) smaller corpora of data. It's also techniques like this that will allow it to produce results *better* than the data it trained on, which is a more fascinating proposition.

[84] D. Ormerod, Lee Sedol Defeats AlphaGo.

[85] J. Markoff, Artificial Intelligence Is Far From Matching.

[86] Ibid.

[87] https://en.wikipedia.org/wiki/File:Raymond_Kurzweil_Fantastic_Voyage.jpg.

[88] Transhumanism, Transhumanism.

[89] Ibid.

[90] Cited in S. Armstrong et al., Who Knows?, p. 55.

[91] Ibid., p. 55. For more detail, see S. Armstrong, Assessing Kurzweil.

[92] John Rennie, quoted in Ray Kurzweil, Ray Kurzweil. See also S. Armstrong, Assessing Kurzweil.

[93] S. Armstrong et al., Errors, Figure 1. Cf. S. Armstrong et al., Who Knows?, p. 51 Figure 3.1.

[94] S. Armstrong et al., Who Knows?, p. 50.

[95] Ibid., p. 51 Figure 3.1.

[96] S. Armstrong et al., Errors, Figure 2. Cf. S. Armstrong et al., Who Knows?, p. 52 Figure 3.2.

[97] S. Armstrong et al., Who Knows?, pp. 50-51.

[98] R. Kurzweil, *Singularity*.

[99] J. Gray, *Immortalization Commission*, pp. 214.

[100] Ibid., p. 214.

[101] Ibid., p. 216.

[102] Ibid., p. 216. For an example of a recent title that embraces Christian transhumanism, see R. Cole-Turner, *Transhumanism and Transcendence*.

[103] S. M. Brown, Enhancing Evolution, pp. 43-44, 46.

[104] Matthew 10:39; 16:25-26; Luke 9:24; 17:33.

[105] http://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf. Brief, informal critiques of this roadmap can be found in the blogosphere include http://blog.ciphergoth.org/blog/2010/02/24/doug-clow-whole-brain-emulation-roadmap/ ; http://lproven.livejournal.com/279235.html ; http://blog.ciphergoth.org/blog/2010/02/20/david-matthewman-whole-brain-emulation-roadmap/ .

[106] S. M. Brown, Enhancing Evolution, p. 47.

[107] N. Bostrom, *Superintelligence*, p. 31. For examples of some of the preliminary thinking going on in this arena, see R. A. Koene, Feasible Mind Uploading; N. Wellington, Whole Brain Emulation; N. Bostrom, *Superintelligence*, pp. 30-36.

[108] Marc Chagall (1887-1985), *The Falling Angel (La Chute de l'Ange)*, 1947. Public Domain. http://www.wikipaintings.org/en/marc-chagall/the-falling-angel-1947 (accessed 25 November 2012).

[109] H. W. Nibley, Enoch, pp. 184-185.

[110] See G. W. E. Nickelsburg *et al.*, 1 Enoch, 8:1-3, p. 188.

[111] See D. C. Matt, Zohar 1, Be-reshit 1:56a, pp. 315-316 and n. 1545.

[112] Ibid., Be-reshit 1:56b, pp. 318-319.

[113] See ibid., Be-reshit 1:56b, p. 319; Genesis 7:23.

[114] Brian Kershisnik, *La Conversation Entre Amis*, https://www.google.com/imgres?imgurl=http%3A%2F%2Fwww.meyergallery.com%2Fthumbs%2Fcf35b0-4801-w806-0.jpg&imgrefurl=http%3A%2F%2Fwww.meyergallery.com%2Fart.php%3Fid%3D4801&docid=0BOtGWeYCBETgM&tbnid=XCSiw0jBH6GohM%3A&w=806&h=641&itg=1&client=safari&bih=941&biw=1803&ved=0ahUKEwjOnLf0t6zNAhVQxCYKHbRMBGoQMwgeKAAwAA&iact=mrc&uact=8 (accessed June 16, 2016).

[115] Alma 34:34.

[116] https://www.thegospelcoalition.org/article/the-great-divorce-on-stage (accessed June 16, 2016).

[117] C. S. Lewis, *Divorce*, p. 13.

[118] Ibid., p. 13.

[119] http://owennelsonexpression.blogspot.gr/2010/06/great-divorce.html (accessed June 16, 2016).

[120] C. S. Lewis, *Divorce*, pp. 20, 21.

[121] http://www.writeca.com/2014/05/04/digital-art-alex-andreev/ (accessed June 16, 2016).

[122] "Milton was right," says Lewis through one of his characters:

> The choice of every lost soul can be expressed in the words "Better to reign in Hell than serve in Heaven." There is always something they insist on keeping, even at the price of misery. There is always something they prefer to joy — that is, to reality. … Everyone who wishes [to come to Heaven] does. Never fear. There are only two kinds of people in the end: those who say to God, "Thy will be done," and those to whom God says, in the end, "*Thy* will be done." All that are in Hell, choose it. Without that self-choice there could be no Hell. No soul that seriously and constantly desires joy will ever miss it. Those who seek find. To those who knock it is opened" (C. S. Lewis, *Divorce*, pp. 71, 72).

[123] http://mi.byu.edu/wp-content/uploads/2013/10/Nibley-1.jpg.

[124] H. W. Nibley, But What Kind, pp. 257-260.

[125] Alma 12:18.

[126] W. Shakespeare, Hamlet, 1:2:133, p. 1145.

[127] Ibid., 4:4:33-35, p. 1172.

[128] Licensed from Shutterstock, still image from v7983493.mov.

[129] Regarding *theatromania*, see H. W. Nibley, Victoriosa Loquacitas; H. W. Nibley, Sparsiones; H. W. Nibley, Roman Games.

[130] http://www.peoplenotary.com/wp-content/uploads/2008/06/running-businessman.jpg (accessed 3 January 2016).

[131] Elder Neal A. Maxwell perceptively commented on this topic (N. A. Maxwell, Grounded, pp. 111-112, 118-119):

> It is especially helpful to remember also that the temptations and challenges we face are common to man (see 1 Corinthians 10:13), yet we must respond uncommonly. It is also useful to ponder the fact that, along with even the Savior himself, we are to experience certain things "according to the flesh" (Alma 7:12) and to learn "in process of time" (Moses 7:21). Built, therefore, into the seemingly ordinary experiences of life are opportunities for us to acquire such eternal attributes as love, mercy, meekness, patience, and submissiveness and to develop and sharpen such skills as how to communicate, motivate, delegate, and manage our time and talents and our thoughts in accordance with eternal priorities. These attributes and skills are portable; they are never obsolete and will be much needed in the next world.
>
> How often have you and I really pondered just what it is, therefore, that will rise with us in the resurrection? Our intelligence will rise with us, meaning not simply our I.Q., but our capacity to receive and to apply truth. Our talents, attributes, and skills will rise with us, certainly also our capacity to learn, our degree of self-discipline, and our capacity to work. Note that I said "our capacity to work" because the precise form of our work here may have no counterpart there, but the capacity to work will never be obsolete. To be sure, we cannot, while here, entirely avoid contact with the obsolescent and the irrelevant. It is all around us. But one can be around irrelevancy without

becoming attached to it, and certainly we should not become preoccupied with obsolete things.

By these remarks I do not intend to create discontent with the paraphernalia of this probationary estate, but it is a grave error to mistake the scenery and the props for the real drama which is underway. Nor do I wish to bear down too much on the fact that certain mortal vocations will be irrelevant in the next world. A mortician does useful work here, especially if it is done with excellence, compassion, and reverence for life. Whatever our vocation, we should be sweetened, not hardened. Keeping our sense of proportion *whatever* we do, keeping our precious perspective *wherever* we are, and keeping the commandments *however* we are tested reflect being settled, rooted, and grounded in our discipleship. ...

[H]ow we utilize the seemingly ordinary experiences of our life and how well we keep the commandments are true tests of our performance in this second estate. One can, while in the employ of a railroad company, learn something of patience while struggling to keep the train schedule meticulously up to date. But the patience will long outlast the printed train schedule. A discovering scientist may augment his awe and meekness before his Creator because of the breathtaking order in the universe even if his new discoveries erelong are swallowed up in even more immense discoveries.

But it is also true that routine may cause a gravedigger to become indifferent to the sorrows of the bereaved gathered about those fresh mounds of earth. The gravedigger may even become cynical about the resurrection which one day will empty all those graves. A marriage counselor can become encrusted with a protective layer of clinical indifference brought on by the routine and incessant nature of his chores. If so, his techniques will never compensate for his lack of caring. A civil servant who has forgotten how to be civil may have some sway now in the procurement division of a vast governmental direction, but he is headed in just the opposite developmental direction needed for sway in the next world.

On the other hand, one who listens more and more effectively to others with a genuine desire to understand and to help, if not always to agree, will have no regrets later on. Such an individual may occasionally run out of time here, but he is fitting himself for eternity. Love and patience are never wasted; they only appear to be. The devoted wife and mother who is a quiet but effective neighbor but whose obituary is noticed by a comparative few may well have laid up precious little here in the current coin-of-the-realm, recognition, yet rising with her in the resurrection will be relevant attributes and skills honed and refined in family and neighborhood life. Contrariwise, the civic leader whose thirst for recognition causes him to do things to be seen of men has his reward. He too will receive the gift of immortality during which expanse he can work on meekness and humility.

132 N. Bostrom, *Superintelligence*, p. 31. For examples of some of the preliminary thinking going on in this arena, see R. A. Koene, Feasible Mind

Uploading; N. Wellington, Whole Brain Emulation; N. Bostrom, *Superintelligence*, pp. 30-36.

133 Photograph by Jeffrey M. Bradshaw, DSC05961.JPG, 14 January 2014.

134 Photograph by Jeffrey M. Bradshaw, IMG_1808.JPG, 10 April 2014.

135 Photo by Evan Cantwell/Creative Services/George Mason University, 120618506E.jpg, 4 February 2015. http://gmu.smugmug.com/keyword/mental%20floss/i-XQrwdfs, accessed June 1, 2016.

136 A. M. Quatrochi et al., Mental Floss - 2012. Paper handout created to accompany the sculpture display.

137 Photo by Evan Cantwell/Creative Services/George Mason University, 120607018.jpg, 4 February 2015. http://gmu.smugmug.com/keyword/mental%20floss/i-XQrwdfs (accessed June 1, 2016).

138 From R. Parekh et al., Neuronal Morphology, as reproduced in S. Bansal, Navigating the Brain Forest.

139 C. Probst, Mental Floss.

140 S. Bansal, Navigating the Brain Forest.

141 Ibid.

142 Ibid.

143 E.g., S. Furber, To Build. In contrast to less-well-grounded researchers, Furber admits that developing a suitable computing architecture is only one of the many parts of the foundation that needs to be laid in order to "build a brain" (ibid., p. 49):

> SpiNNaker won't get us all the way to full-scale simulations of the human brain. But the machine's communications architecture could help pave the way for better-networked analog chips that could get us there. It will also help show us what information we need to make good models. Then we can really put our brains to use.

144 Significant large initiatives to understand and model the brain are currently underway. For example, the SpiNNaker computing platform described by Furber is part of a pan-European initiative called the Human Brain Project that began in October 2013 and is projected to last for ten years with a projected cost of one billion euros (Human Brain Project, Human Brain Project).

In April 2013, the US White House announced the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative, which was to be guided by a twelve-year research strategy. Participating federal agencies include the National Institutes of Health (NIH), the National Science Foundation (NSF), the Defense Advanced Research Projects Agency (DARPA), and the US Food and Drug Administration (FDA). Partners include private foundations, research institutes, universities, and industry (BRAIN Initiative, BRAIN Initiative). For

pointers to additional efforts of a similar nature, see, e.g., Artificial Brains, Artificial Brains.

145 For an entertaining TEDEd video debunking this myth, see https://www.youtube.com/watch?v=5NubJ2ThK_U.

146 For a brief overview on this topic, see https://www.youtube.com/watch?v=yE6VTvxkhFs.

147 From https://www.youtube.com/watch?v=i6VYuw52z-0 (accessed June 22, 2016).

148 See, e.g., K. M. Ford et al., Knowledge acquisition as a constructive modeling activity, pp. 11-17. For an impressive example of an automatic track-laying machine, see https://www.youtube.com/watch?v=i6VYuw52z-0, https://www.youtube.com/watch?v=_MKcTbYDP7w (accessed June 22, 2016).

149 R. H. Wurtz, Brain Mechanisms, p. 11, Figure 1.

150 Ibid., p. 11, Figure 1 caption.

151 Ibid., p. 11.

152 Ibid., p. 11.

153 Ibid., p. 12.

154 Ibid., p. 12.

155 Salvador Dali (1904-1989): *The Persistence of Memory*, 1931. In http://artearth.ru/news/painting/v-ekaterinburge-otkroetsya-vystavka-syurrealizma-posvyaschennaya-tvorchestvu-salvadora-dali (accessed June 29, 2016).

156 Memory Distortions, Memory Distortions, after F. C. Bartlett, *Remembering*.

157 See, e.g., F. C. Bartlett, *Remembering*.

158 Memory Distortions, Memory Distortions.

159 F. C. Bartlett, *Remembering*.

160 C. S. Lewis, *Divorce*, pp. 67-68 brilliantly imagines how this phenomenon might color our perceptions of events if we were allowed to witness the scenes of our mortal lives being played out retrospectively in heaven:

> [We] cannot in [our] present state understand eternity … [but we] can get some likeness of it if [we] say that both good and evil, when fully grown, become retrospective. Not only [the light of the spirit world] but all this earthly past will have been Heaven to those who are saved. Not only the twilight in [the spirit prison], but all their life on earth too, will be seen by the damned to have been Hell. That is what mortals misunderstand. They say of some temporal suffering, "No future bliss can make up for it," not knowing that Heaven, once attained, will work backwards and turn even that agony into a glory. And of some sinful pleasure they say, "Let me but have *this* and I'll take the consequences;" little dreaming how damnation will spread back and back into their past and contaminate the pleasure of the sin. Both processes begin even before death. The good man's past begins to change so

that his forgiven sins and remembered sorrows take on the quality of Heaven; the bad man's past already conforms to his badness and is filled only with dreariness. And that is why, at the end of all things … the Blessed will say "We have never lived anywhere except in Heaven," and the Lost "We were always in Hell." And both will speak truly.

161 E. Loftus, How Reliable.

162 Ibid..

163 E. Loftus, *Eyewitness Testimony*.

164 E. Loftus, Creating False Memories.

165 D. H. Dodd et al., Leading Questions

166 C. C. Green et al., *Emerging Cognitive Neuroscience*.

167 Ibid., p. 80.

168 For a readable summary of the state of the art in the study of the nervous system and its functions, see F. H. Gage, What Is the Brain Good For?.

169 M. Changizi, Artificial Brains.

170 P. Tucker, Pentagon Is Nervous.

171 Summer Study on Autonomy, Summer Study on Autonomy; F. Kendall, Terms of Reference. For additional background on this study, see P. Tucker, Military's New Year's Resolution. For remarks by the Deputy Defense Secretary, Robert O. Work, that quote from a draft of the study, see P. Tucker, Pentagon Is Nervous.

For a brief overview of some of the "myths" of autonomy for the general reader, see J. M. Bradshaw *et al.*, Seven Deadly Myths. For a video presentation for a general academic audience describing and illustrating these myths, see J. M. Bradshaw, Human-Agent-Robot Teamwork.

172 For early efforts to explore computational approaches for these laws, see, e.g., R. Clarke, Asimov's laws of robotics: Implications for information technology, Parts 1 and 2; Y. Shoham et al., On the synthesis of useful social laws for artificial agent societies; D. Pynadath *et al.*, Revisiting Asimov's first law: A response to the call to arms; D. Weld et al., The first law of robotics: A call to arms.

173 See, e.g., J. M. Bradshaw et al., Policy-Based Governance; A. Uszok *et al.*, Toward a Flexible Ontology-Based Policy Approach for Network Operations Using the KAoS Framework; J. M. Bradshaw et al., KAoS; S. Joseph et al., Digital Policy Management.

174 J. Markoff, How Tech Giants.

175 R. O. Work et al., 20YY.

176 What's Real, What's Not, What's Real, What's Not.

177 C. Pellerin, Deputy Secretary: Third Offset.

178 For readable introductions to the "third offset stragegy," see, e.g., K. H. Hicks et al., Assessing; M. Eaglen, What Is the Third Offset Strategy; C. Pellerin, Deputy

Secretary: Third Offset. For a brief sketch of what future technology may bring, see M. Rosenberg et al., Pentagon's 'Terminator Conundrum'.

[179] Offset Strategy, Offset Strategy.

[180] A. Tsang, Amazon 'Reviewing'.

[181] Software as Weaponry, Software as Weaponry.

[182] See C. C. Green *et al.*, Emerging Cognitive Neuroscience, p. 95: "While modeling the whole brain is highly unlikely in the next two decades, it is not unreasonable to imagine that significant subsystems could be modeled. Moreover, it seems likely that increasingly sophisticated cognitive systems will be constructed in those two decades that, while not aiming to mimic processes in the brain, could nonetheless perform similar tasks well enough to be useful, especially in constrained situations."

[183] A. Tsang, Amazon 'Reviewing'.

[184] S. Lohr, Stepping Up Security.

[185] Summer Study on Autonomy, Summer Study on Autonomy, p 88.

[186] G. Webster, Future of Airport Security. According to the CNN article in which this image appeared, the thermal imaging system behind this image portends a new approach to detecting deception visually:

> Feeling guilty? Got something to hide? A team of UK-based researchers claim to have developed a thermal lie-detection camera that can automatically spot a burning conscience.
>
> The system could be used during customs interviews and at passport control to check whether people entering the country are giving a true account of themselves.
>
> The thermal-imaging camera captures variations in facial temperature in response to questioning. "When someone is making something up on the spot, brain activity usually changes and you can detect this through the thermal camera," said professor Hassan Ugail, who leads the research.
>
> At present, the UK's Home Office and HM Revenue & Customs are sponsoring the system's development, but will not reveal the name of the airport where it's being tested.

[187] C. C. Green et al., Emerging Cognitive Neuroscience, pp. 18-41. The study, which was published in 2008, was specifically looking out two decades, i.e., to the period ending in 2028.

[188] For an early snapshot of the DoD's view of the establishment of the US Cyber Command, see W. E. Leigher, Learning to Operate, p. 32.

[189] Underground Economy, Underground Economy.

[190] Software as Weaponry, Software as Weaponry.

[191] S. Malcomson, Splinternet; D. Alba, World May Be Headed.

[192] T. S. Bernard et al., Equifax Says Cyberattack.

[193] F. Manjoo, Seriously, Equifax?.

[194] E.g., J. Markoff, As Artificial Intelligence Evolves.

[195] A. Chen, The Agency.

[196] E.g., G. Greenberg, Hackers Cut; Hacking Power Networks, Hacking Power Networks.

[197] See, e.g., J. M. Bradshaw et al., Sol; L. Bunch et al., Human-Agent Teamwork; L. Bunch et al., Principles for HCI Interaction Design 2; R. M. Jones et al., Modeling and Integrating. For a readable summary of early efforts for the general reader, see New Tools, New Tools.

[198] L. Bunch et al., Principles for HCI Interaction Design 2.

[199] A. M. Turing, Computing machinery and intelligence. See also E. A. Feigenbaum et al., Alan Turing Top Ten Panel.

[200] L. Bunch et al., Principles for HCI Interaction Design 2.

[201] L. Bunch et al., Human-Agent Teamwork; J. M. Bradshaw et al., Multi-Agent Systems.

For general overviews of software agent technology, see J. M. Bradshaw et al., Human-Agent Interaction; J. M. Bradshaw, Software Agents.

For a video introduction to software agents and design principles, see J. M. Bradshaw, Designing Software Agents.

For easy-to-read introductions to the topic, see C. Noessel, Designing Agentive Technology: AI That Works for People; D. A. Norman, The Invisible Computer: Why Good Products Can Fail, the Personal Computer is So Complex, and Information Appliances Are the Solution; D. A. Norman, Things That Make Us Smart: Defending Human Attributes in the Age of the Machine.

[202] See, e.g., a summary of the view of most mainstream AI researchers in T. G. Dietterich et al., Viewpoint.

[203] E.g.:

- Allen Institute for Artificial Intelligence (https://en.wikipedia.org/wiki/Allen_Institute_for_Artificial_Intelligence, http://allenai.org)
- Centre for the Study of Existential Risk (https://en.wikipedia.org/wiki/Centre_for_the_Study_of_Existential_Risk, http://cser.org)
- Future of Humanity Institute (https://en.wikipedia.org/wiki/Future_of_Humanity_Institute, http://www.fhi.ox.ac.uk)
- Future of Life Institute (https://en.wikipedia.org/wiki/Future_of_Life_Institute, http://thefutureoflife.org)

- Global Catastrophic Risk Institute (https://en.wikipedia.org/wiki/Global_Catastrophic_Risk_Institute, http://gcrinstitute.org)
- Institute for Ethics and Emerging Technologies (https://en.wikipedia.org/wiki/Institute_for_Ethics_and_Emerging_Technologies, http://ieet.org)
- Machine Intelligence Research Institute (MIRI), formerly the Singularity Institute for Artificial Intelligence (SIAI) (https://en.wikipedia.org/wiki/Machine_Intelligence_Research_Institute, https://intelligence.org)
- OpenAI (https://en.wikipedia.org/wiki/OpenAI, https://www.openai.com/blog/introducing-openai/)

[204] E.g., AAAI Presidential Panel, AAAI Presidential Panel; AI100, AI100.

[205] E.g., T. G. Dietterich *et al.*, Viewpoint.

[206] E.g., N. Bostrom *et al.*, Global Catastrophic Risks; N. Bostrom, Superintelligence.

[207] E.g., LessWrong (https://en.wikipedia.org/wiki/LessWrong, http://lesswrong.com).

[208] Open Letter: Research Priorities, Open Letter: Research Priorities; Autonomous Weapons, Autonomous Weapons.

[209] R. D. Atkinson, 2015 ITIF Luddite Award Nominees.

[210] S. Russell *et al.*, Think-Tank Dismisses.

[211] A few examples starting at the most basic level: object-oriented programming, semantic technologies (e.g., OWL and other related W3C standards), speech recognition, industrial robot motion planning and localization, facial recognition and a host of other vision processing algorithms used in photography, industrial robotics, security, and cinema.

The result of this assimilation has sometimes been called the "AI effect," which "occurs when onlookers discount the behavior of an artificial intelligence program by arguing that it is not *real* intelligence" (AI Effect, AI Effect). This phenomenon was famously lamented by Douglas Hofstadter (D. R. Hofstadter, Gödel, Escher, Bach, p. 601):

> It is interesting that nowadays, practically no one feels that sense of awe any longer — even when computers perform operations that are incredibly more sophisticated than those which sent thrills down spines in the early days. The once-exciting phrase "Giant Electronic Brain" remains only as a sort of "camp" cliché, a ridiculous vestige of the era of Flash Gordon and Buck Rogers. It is a bit sad that we become blasé so quickly.
>
> There is a related "Theorem" about progress in AI: once some mental function is programmed, people soon cease to consider it as an essential ingredient of "real thinking." The ineluctable core of intelligence is always in that next thing which hasn't yet been programmed. This "Theorem" was first proposed to me

by Larry Tesler, so I call it Tesler's Theorem: *"AI is whatever hasn't been done yet."*

The problem was characterized by the well-known AI pioneer Marvin Minsky as a sort of argument by "redefinition" against AI: i.e., an effort to minimize any appearance of progress in the field of AI "by continually modifying the definition of intelligence in order to exclude all artificially reproduced phenomena" (R. Cordeschi, Discovery of the Artificial, p. 233. Cordeschi cites M. Minsky, Steps, p. 396, but I have been uable to track down anything in the referenced paper or in other writings by Minsky that corresponds to this idea.)

Part of the problem in properly characterizing AI is the lamentable tendency of some popular definitions to make humans the measure of AI research progress, e.g.:

> Artificial intelligence is the science of making machines do things that would require intelligence if done by men" (M. Minsky, Semantic Information Processing, p. v).
>
> Artificial Intelligences (AI) is the study of how to make computers do things which, at the moment, people do better (E. Rich *et al.*, Artificial Intelligence, p. 3).

[212] Of course, the tendency toward apocalyptic panic is not confined to a small segment of Church members. See, e.g., H. Fountain, Apocalyptic Thoughts.

[213] G. B. Hinckley, Thanks, p. 88.

[214] http://mi.byu.edu/wp-content/uploads/2013/10/Nibley-1.jpg.

[215] B. J. Petersen, Nibley, pp. 97-98**.

[216] B. N. Madsen, Truman G. Madsen, p. 107.

[217] J. R. Holland, Tomorrow, p. 127.

[218] D&C 89:4.