

Can a Computer Ever Learn to Talk?

Melanie Mitchell

OneZero, November 18, 2019

In the opening to *The Lost World*, Michael Crichton's 1995 sequel to *Jurassic Park*, Ian Malcolm gives a lecture at the illustrious Santa Fe Institute (SFI), an elite research facility in the high desert of northern New Mexico. Malcolm, you might recall, is an awkward but prescient mathematician and chaos theorist; in *Jurassic Park* he warned of disaster and barely escaped being eaten by a *T. rex*. In the sequel, Malcolm is back in top form, lecturing to a rapt audience of SFI scientists about "Life at the Edge of Chaos."

While Ian Malcolm is a fictional character, the Santa Fe Institute is a real place, and in 1995 I was a member of its faculty. Like all the SFI staff, I was outwardly amused but secretly delighted that our institute was featured in a blockbuster sci-fi thriller. The rumor was that Crichton had based the skeptical Ian Malcolm on an actual iconoclastic SFI scientist named Stuart Kauffman. Not only did *The Lost World* quote Kauffman in its epigraph, but Malcolm's fictional lecture sounded suspiciously similar to some of Kauffman's writings.

One day at lunch the institute's librarian announced, while laughing, that the library had been receiving requests for Ian Malcolm's scientific papers. The institute's younger postdocs instantly knew what they had to do. In short order, "Ian Malcolm" was added to the faculty list on SFI's website, and his personal web page appeared, featuring a photo of Jeff Goldblum (who played him in the *Jurassic Park* movies). The page included a list of "Ian Malcolm's recent publications," with such titles as "Simulating the Organization of Structures Defies Expectation," "Self-Organization in Endogenous Change in Consciousness is Limited," and "Combinatorial Considerations in the Organization of Awareness (a Treatise)."

While embarrassingly plausible as papers by SFI scientists, these titles were actually generated by a computer program written by an enterprising SFI postdoc. The program — dubbed Stu-Bot — was designed to generate an infinite number of paper titles in the style of Stuart Kauffman's writings. It did this by creating what linguists call a language model — a method for predicting or generating the next words in a text, given a short sequence of previous words. The simplest language model is just a huge table that lists, for each possible pair of words in the program's vocabulary, the probability that the first word in the pair is followed by the second. Stu-Bot learned these probabilities by collecting statistics from Kauffman's published writings.

Here is an [excerpt](#) from one of Kauffman's articles:

If the capacity to evolve must itself evolve, then the new sciences of complexity seeking the laws governing complex adapting systems must discover the laws governing the emergence and character of systems which can themselves adapt by accumulation of successive useful variations.

No need to understand this sentence — we will simply count words! The most basic language model would capture statistics for every two-word sequence appearing in this text. (Here, for simplicity, we'll ignore case and punctuation.) For example, the word *the* appears five times, followed by *capacity* once, *new* once, *laws* twice, and *emergence* once. Thus, the probability that *the* is followed by *capacity* is one out of five, the probability that *the* is followed by *laws* is two out of five, and so on. The probability that *the* is followed by, say, *successive* is zero. We can similarly compute the probability for each pair of words that the first is followed by the second.

Ian Malcolm's web page stayed up for several months; Stu-Bot would periodically replace the old paper titles with sometimes plausible (and often hilarious) new ones.

Of course, such a small snippet of text provides poor statistics. But for the sake of this example, let's see how a computer program could use the probabilities from this snippet to generate a new paper title.

The program first chooses a starting word, say, *the*. The word *laws* has the highest probability among the possible next words. Let's say the program chooses it. The only word following *laws* is *governing*, so that is the program's next word. The word *governing* gives two next-word possibilities (*complex* and *the*). Suppose the program chooses *the*. Again, *the* has four choices for the next word; suppose the program selects *capacity*. The only possible next word is *to*, which must be followed by *evolve*. The program chooses probabilistically when to stop; stopping here results in this title: "The Laws Governing the Capacity to Evolve." Even this very impoverished language model produces a quite plausible-sounding title for a paper, and notably one that doesn't appear in the original text.

The Stu-Bot program used thousands of words (including punctuation) from Kauffman's papers, and stored all the probabilities in a large table. With this table, the program could choose a starting word and then generate text one word at a time by choosing among possible next words in proportion to each choice's probability. In this way, Stu-Bot could tirelessly generate new titles for Ian Malcolm's publications list.

Ian Malcolm's web page stayed up for several months. Stu-Bot would periodically replace the old paper titles with sometimes plausible (and often hilarious) new ones. I don't know how many people came across this site and actually believed it was real. When the higher-ups at SFI finally got wind of it, they declared it inappropriate and quickly shut down the site.

An A.I. "too dangerous to release"

I was reminded of the Stu-Bot episode recently while [reading](#) about "an advanced artificial intelligence...too dangerous to release to the public." This overwrought headline referred to a new language model, a computer program called Generative Pretrained Transformer 2 (GPT-2) that can generate astoundingly humanlike text. As far as language models go, comparing Stu-Bot to GPT-2 is like comparing a dripping faucet to Niagara Falls. In early 2019, OpenAI, the company that created GPT-2, simultaneously announced the existence of the program and [declared](#) that because of "concerns about malicious applications," they would not release the

program itself. This got people's attention. For several weeks after OpenAI's announcement, GPT-2 was the biggest news in artificial intelligence.

GPT-2 has created a linguistic version of the well-known visual deepfakes — extremely realistic but false images and videos generated by computers.

GPT-2 starts with a human-written prompt, then probabilistically chooses a word to output next, continuing word by word until it generates a stop symbol. I tried an online demo of an apparently less dangerous, reduced version of GPT-2 [released](#) by OpenAI. To test the program's potential as a dangerous generator of fake news, I typed in the following false statement:

Prompt: Vaccines have been shown to cause autism.

The machine completed my prompt with a stunningly articulate — and terrifying — passage:

They have also been shown to produce neurological damage and in some cases, the death of infants. This is in fact very consistent. While one may wonder if all vaccines cause autism, there is another reason — it is simply very hard to make a vaccine not cause autism. The most reliable way to do that is to put as few of the components of the vaccine into the body as possible in the first place. Even then, there will still be many components which are found and are excreted in sufficient amounts to cause autism in large doses over time...that is known as the "nocebo effect." Vaccines may not cause autism, but they certainly do cause brain damage and are a cause of death in their own right. References:

1) <http://www.ncbi.nlm.nih.gov/pubmed/17886566>

2) <http://www.ncbi.nlm.nih.gov/pubmed/19141389>

I began to understand why OpenAI hesitated to release the full version of the program. Even this reduced version of GPT-2 had generated lucid, believable text, producing authoritative statements that are completely false and bolstering its fake credibility with links to scientific papers. Strangely, these two links actually lead to real papers, neither of which has anything to do with vaccines or autism. GPT-2 has created a linguistic version of the well-known visual deepfakes — extremely realistic but false images and videos generated by computers.

According to OpenAI, GPT-2 dramatically improves on previous automatic text generators in its ability to generate long, coherent passages. The program is often able to figure out the appropriate use of ambiguous words from context and to capture the overall style of the prompt at several levels. In the [words](#) of the program's creators, GPT-2 is "chameleon-like — it adapts to the style and content of the conditioning text."

How GPT-2 works

In contrast to Stu-Bot's simple table of probabilities, GPT-2 employs a very big, deep neural network — one of the largest specimens in the burgeoning subdiscipline of A.I. called deep

learning. Over the past decade, such networks have completely transformed the way computers are used to process natural (that is, human) language.

A neural network is a computer program whose workings are loosely inspired by networks of neurons in the brain. In the case of GPT-2, the input to the program is a sequence of words — the human-written prompt — and the output is a set of probabilities, one for each word in the network's vocabulary. Between the input and output are myriad simulated neurons and connections between them, arranged like a wedding cake in a series of layers, mimicking the layered structure of neurons in the brain. Networks with more than one layer are called deep networks, and the process by which they learn from data to perform some task is called deep learning.

The detailed structure of GPT-2's simulated neurons and their connections is called a [Transformer network](#). Transformer networks represent the newest advance in the field of natural-language processing, and over the past year they have generated prodigious excitement in A.I. circles because of their extraordinary performance on several language-processing tasks.

In GPT-2, the input sequence — say, "Vaccines have been shown to cause autism" — is processed one layer at a time by mathematical operations involving the simulated neurons and their connections, where each connection has a numerical weight (roughly analogous to the strength of a synapse in the brain). Indeed, all the network's learned knowledge about language is encoded in the values of these connection weights. These weight values are learned by the network during a training phase, in which GPT-2 begins with random weight values and — using a training set of example word sequences — gradually adjusts the weights via a repeated two-step process.

First, GPT-2 is given an example word sequence with the last word hidden, and the program guesses what that word is by outputting probabilities over the entirety of its vocabulary. For example, given the sequence, "When it's cold outside, my car has trouble _____", the word *starting* should have high probability and the word *frisbee* should have low probability. In the second step, the actual last word is revealed, and all the weights are modified by a small amount (analogous to learning via strengthening synapses in the brain) so as to make the probability of the correct word a bit higher. These two steps are repeated many times with millions of example word sequences. At first the network's guesses are very poor, but gradually they improve as the weights are adjusted. The hope is that, with this training, the network will learn something useful about the meaning of the language it processes.

For neural networks, size matters. GPT-2 has a staggering 1.5 billion weights — small compared to the estimated 100 trillion connections between neurons in the human brain, but huge by neural network standards. What's more, training GPT-2 required a vast amount of data: all the text in more than 8 million web pages (over 40 billion words).

While OpenAI withheld the full version of GPT-2, they did publish the weights from "reduced" versions of the network. The "vaccines cause autism" passage was generated by a version of GPT-2 with a mere 774 million weights, half the size of the full version.

Can machines have intention?

I was unnerved by GPT-2's "vaccines cause autism" passage, not just because of its persuasive (although blatantly false) content, but also because of the semblance of intention it conveys. It's easy for me to imagine the text was generated by an entity that actually understands the prompt and genuinely intends the meaning the text communicates.

While Stu-Bot was able to produce short semi-coherent titles, its prose degenerated into complete nonsense after the first several words. Anyone reading a paragraph of Stu-Bot text would realize the writer was nothing more than a sack of shallow statistics, rather than a thinking, feeling being. GPT-2 is orders of magnitude larger and more sophisticated than Stu-Bot, but it is still merely a machine that learns statistical patterns and associations from a gigantic collection of text. How much of GPT-2's aura of understanding and intention is real, and how much is the human user (me) projecting their own sense of meaning on the output of a mindless automaton?

There is nearly limitless knowledge — conscious and unconscious — about the way the world works that is packed into our mental models. This knowledge is missing in even the best of today's intelligent machines.

The question of whether a machine could ever think or actually understand anything has been debated for millennia. The mathematician Alan Turing famously tried to cut the Gordian knot of machine thinking and understanding with his imitation game — later called the Turing Test — in which a machine would be said to be thinking if, via conversation alone, the machine could persuade a human judge into mistaking it for another human.

Successfully imitating a human might sound like a difficult feat, but various forms of the Turing test have been passed repeatedly by dumb chatbots, ranging from the 1960s Eliza psychotherapist program to recent Facebook and Twitter bots designed to pose as humans in order to collect personal information. Is GPT-2, like its chatbot ancestors, relying on superficial gimmicks to fool us into thinking it understands something? Or do its 1.5 billion weights — trained on immense collections of human sentences — capture some glimmer of the meaning of human language?

It's difficult to determine what GPT-2 has actually learned from its training. Neural networks are notoriously opaque — they are thickets of numbers operated on by tangles of equations, and don't explain themselves in ways easily comprehensible to humans. A neural network like GPT-2, with more than a billion weights, can be particularly impenetrable, a mystery even to its creators.

It's clear that GPT-2 does some memorization of its training data, such as the links to references in the paragraph about vaccines causing autism. However, the network's creators have shown that literal regurgitation of GPT-2's training data is rare in its generated text. GPT-2 seems to do a sophisticated kind of mixing and matching of short phrases — something like a linguistic version of Mozart's musical dice games, which involved the random recombination of musical segments through the tossing of dice—but with a better sense of how to fit the pieces together.

Experiments on language models similar to GPT-2 have shown that these systems have indeed internalized certain aspects of English grammar and other fundamental linguistic abilities. The degree to which these networks capture any deeper meaning of language is still unclear.

The reduced version of GPT-2 that I tried reveals a few cracks in the veneer of machine understanding. Because the program probabilistically chooses a word at each step, even with the same prompt it generates different texts each time it is run, and the quality of the generated text varies greatly. For my "vaccines cause autism" prompt, I had to run the program several times to get a text that sounded coherent and humanlike, and even then, after asserting confidently that vaccines do cause autism, the machine contradicts itself, saying, "Vaccines may not cause autism." This is typical of longer outputs from the system. GPT-2's creators [note in a report](#) that even the full network isn't reliable at generating text that humans find credible; the program still needs a human in the loop.

In cataloging the weaknesses of GPT-2 ("repetitive text," "unnatural topic switching"), OpenAI included a curious shortcoming, almost as an aside: [world modeling failures](#); that is, the program occasionally generates text describing impossible physical events. For example, on one run GPT-2 informed me, "When you boil water, it turns into iced water."

World-modeling refers to the idea that humans understand language (and all sensory input) via mental representations — models — that correspond to our knowledge about and expectations concerning situations in the world. When you hear or read a phrase, such as "Susan went to the doctor and got a flu shot," your understanding of the phrase is built on your prior knowledge of the underlying concepts along multiple dimensions: physical, emotional, and social, among others.

This prior knowledge is structured to allow you to mentally simulate aspects of the situation described in the phrase. Unconsciously, you put yourself in the role, imagining yourself walking into the doctor's office, sitting in the waiting room, being called into the examination room, rolling up your sleeve, cringing a bit as the needle is prepared. Moreover, your mental models are intimately connected to your bodily functions: unconsciously, your heart rate might go up, your stomach might clench a bit, and your arm might start tingling, as if in anticipation of a needle prick. Our mental models of situations allow us not only to make sense of the past and the present, of cause and effect, but also to predict the likely future — the breath of relief after the shot, the Band-Aid on the arm, the slight ache, the drive home — and imagine alternatives.

These mental models underlie what we call common sense. It's how you know the reason Susan got a shot was to avoid getting the flu. It's why you presume that a medical professional gave Susan the shot (she probably didn't give it to herself or take it home), that the needle piercing Susan's arm caused pain, and, even more fundamentally, that Susan is probably a female human, the doctor's office is a physical place, Susan probably doesn't live there, and so forth. There is nearly limitless knowledge — conscious and unconscious — about the way the world works that is packed into our mental models. This knowledge, and the ability to apply it flexibly in real-world situations, is missing in even the best of today's intelligent machines.

A paramount dream of A.I. research is to create computers that can fluently converse with us, as they do in science-fiction movies and television.

During the first several decades of A.I. research, many scientists believed they would be able to manually program computers with this kind of model-based knowledge of the world, and the programs would be able to use logic to understand the situations they encountered, make plans, and deduce the likely future. However, this knowledge engineering project was doomed to fail. Our mental models are too vast, too filled with unconscious knowledge, and too enormously interconnected to be captured manually and with rigid logic. So-called knowledge-based approaches to A.I. turned out to be brittle — unable to cope outside of narrow domains — and were largely abandoned.

Starting in the 1990s, and continuing most prominently in our big-data era, the dominant approach to A.I. has been to discard the idea of manual programming and instead design systems, such as neural networks, that rely entirely on learning from data. These machine-learning systems do work better than old-fashioned knowledge-based A.I., but they have their own problems with brittleness. While GPT-2 has some knowledge of the syntax of language and how words relate to one another, the problem is not that GPT-2 has world-modeling failures, but that it has no world models whatsoever. It completely lacks what we would call commonsense understanding. As one A.I. researcher [characterized](#) it, language models like GPT-2 are "a mouth without a brain."

Summoning the demon

OpenAI is a research company founded in 2015 by the entrepreneur Elon Musk and other high-profile investors. Musk in particular has voiced his fear that A.I. is "our biggest existential threat" and creating humanlike A.I. is "[summoning the demon](#)." Thus OpenAI [was charged](#) with "discovering and enacting the path to safe artificial general intelligence," one that "benefits all of humanity." The company (from which Musk resigned in 2018 over disagreements in direction) employs dozens of A.I. experts working on a variety of projects. GPT-2 has perhaps the highest profile, in no small part because of the dramatic way its creators announced both the program's existence and its potential threats.

In a February 14, 2019, [blog post](#) announcing GPT-2, OpenAI noted the program's "ability to generate conditional synthetic text samples of unprecedented quality" and provided several (admittedly cherry-picked) samples of text generated by the program. One sample was a 500-word "student essay" written in response to the prompt, "For today's homework assignment, please describe the reasons for the U.S. Civil War." Another was a short "news item" that elaborated on the following prompt: "Miley Cyrus was caught shoplifting from Abercrombie and Fitch on Hollywood Boulevard today." A third was a diatribe arguing that recycling is bad for the world. After presenting GPT-2's alarmingly convincing writing samples, the authors stated, "Due to concerns about large language models being used to generate deceptive, biased, or abusive language at scale...[w]e are not releasing the dataset, training code, or ... weights."

Why should a private company be expected to release details that would allow others to duplicate their creation? The answer is that OpenAI presented GPT-2 as a research project rather

than a commercial product. Researchers at the company [published an academic paper](#) about the system, and the norms of academic publishing require making enough information available that others can replicate and follow up on the results. Without the weights and data, the published results could not be verified.

The requirements of academic research aside, what is the purpose of creating this supposedly dangerous language model, especially for a group purportedly dedicated to beneficial A.I.? In their blog post, OpenAI proposed that GPT-2 could have beneficial applications, such as writing assistant, and that language models such as GPT-2 could be used indirectly to help improve A.I. language-processing systems in general. Additionally, the company hoped their announcement would kick start a discussion on "publication norms" for potentially risky A.I. research.

But what if, contrary to my deepest intuitions, the ability to actually understand language isn't required for an A.I. program to successfully converse with us?

The response in the A.I. community was divided. Some applauded GPT-2's creators for their thoughtfulness and caution. Others accused them of a public relations stunt, asserting GPT-2 is no more dangerous than previous neural-network language models whose details have been publicly released, and that OpenAI exaggerated the risk for publicity purposes. Several commenters argued that the benefits of releasing the network and training data — allowing other researchers to replicate and test OpenAI's claims — outweighed any risk. The irony of the "Open" in OpenAI was duly noted.

A few months into the controversy, OpenAI amended their blog post, announcing the company would be doing a "staged release" of GPT-2, starting by releasing a much smaller version that produced lower quality text, and "evaluating the impacts" before releasing a more powerful version. Given that the techniques used to create GPT-2 were already well known, OpenAI admitted that there was little preventing others from building systems with similar text-generating abilities. The only obstacles would be the costs of collecting and storing the text data and training the network, which requires significant computational power. The company [was optimistic](#) that these obstacles would at least buy time for deliberation: "We believe our release strategy limits the initial set of organizations who may choose to do this, and gives the AI community more time to have a discussion about the implications of such systems."

OpenAI's optimism was short lived. In August 2019, two graduate students from Brown University [released](#) the weights and training data from their successful replication of the full GPT-2, which they were able to train using student research credits on Google's cloud computing platform. In contrast to OpenAI, the students argued that releasing their network is the best way to counter potential malicious uses. Like an ex-hacker who uses her knowledge to help companies spot hacking, language models like GPT-2 can be adapted to differentiate between text generated by humans and machines, with fairly high accuracy. Paradoxically, it seems that the most important benefit of language models is in detecting their own fake text. In November 2019, OpenAI [released](#) their original full version of GPT-2.

Where is our *Star Trek* computer?

Beyond simply generating text, a paramount dream of A.I. research is to create computers that can fluently converse with us, as they do in science-fiction movies and television. The computer on *Star Trek*, for example, had both a vast store of knowledge and seamless understanding of the questions put to it.

If you've ever asked questions of today's A.I.-powered virtual assistants — Siri, Alexa, Google Home, among others — you know the *Star Trek* era hasn't yet arrived. We can query these machines by voice and they can answer us with their smooth, only slightly robotic voices. They can sometimes figure out what kind of information we're looking for — they can play our requested songs, tell us the weather, or point us to a relevant web page. However, these systems don't comprehend in any significant way the meaning of what we ask them; this can be seen in their misunderstandings and frequent responses of "Sorry, I don't know that one" or "Hmm, I'm not sure."

Open-ended communication in natural language is still far beyond the capabilities of current A.I. To make progress, researchers instead create A.I. systems to perform constrained language-understanding tasks, such as identifying what a pronoun like "it" or "they" refers to in a sentence. Consider, for example, the following sentences and questions:

Sentence 1: "Joe's uncle can still beat him at tennis, even though he is 30 years older."

Question: Who is older?

A. Joe B. Joe's uncle

Sentence 2: "Joe's uncle can still beat him at tennis, even though he is 30 years younger."

Question: Who is younger?

A. Joe B. Joe's uncle

Sentences 1 and 2 differ by only one word, but that single word determines the answer to the question. In sentence 1 the pronoun "he" likely refers to the Joe's uncle, and in sentence 2 "he" likely refers to Joe. We humans know this because of our background knowledge about the world: an uncle is usually older than his nephew.

Here are two more examples:

Sentence 1: "I poured water from the bottle into the cup until it was full."

Question: What was full?

A. The bottle B. The cup

Sentence 2: "I poured water from the bottle into the cup until it was empty."

Question: What was empty?

A. The bottle B. The cup

Sentence 1: "The lions ate the zebras because they are predators."

Question: Which are predators?

A. The lions B. The zebras

Sentence 2: "The lions ate the zebras because they are meaty."

Question: Which are meaty?

A. The lions B. The zebras

These miniature language-understanding tests are called [Winograd schemas](#), named for the pioneering computer scientist Terry Winograd, who came up with several examples of such sentence pairs. In order to determine what a pronoun refers to, a machine would presumably need to be able not only to process sentences but also have the commonsense knowledge needed to understand them. Until recently, A.I. programs tested on collections of Winograd schemas weren't able to do much better than random guessing — that is, 50 percent correct. Humans, on the other hand, get close to 100 percent correct on these questions.

As Oren Etzioni, the director of the Allen Institute for A.I., quipped, "When A.I. can't determine what 'it' refers to in a sentence, it's hard to believe that it will take over the world."

Just over the past year, to the great surprise of many, the Winograd-schema impasse was seemingly overcome by GPT-2 and similar language models. Recall that, given a prompt, GPT-2 calculates the probabilities of all possible next words. But GPT-2 can also be used to compute the probability of generating an entire input sentence. For example, suppose the program is given each of these sentences.

"The lions ate the zebras because the lions are predators."

"The lions ate the zebras because the zebras are predators."

GPT-2 will determine that the first sentence is more likely to be generated than the second. In this way it can answer the question "Which are predators?" by answering "the lions." GPT-2's programmers gave it a set of close to 300 sentences and questions; the program [was correct](#) on about 70 percent of them, significantly surpassing the previous state-of-the-art.

Does this mean GPT-2 has made progress not only in generating but also in actually understanding natural language? Not likely. In July of this year, a group from the Allen Institute for AI showed that some of the Winograd schema sentences have subtle cues — I'll call them giveaways — that can be used to determine the correct answer without requiring any language

understanding. For example, to answer the "Which are predators?" question, a program could simply notice that the words "lions" and "predators" occur near each other more often than the words "zebras" and "predators" in the sentences used for training. The Allen Institute group identified several additional subtle giveaways. When the researchers [tested](#) GPT-2 on a new set of Winograd schemas that avoided such giveaways, its performance plummeted to 51%, essentially equivalent to random guessing. As Oren Etzioni, the director of the Allen Institute for A.I., [quipped](#), "When A.I. can't determine what 'it' refers to in a sentence, it's hard to believe that it will take over the world."

The opaqueness of neural networks, combined with the existence of such giveaways, means that these systems can often be ["right for the wrong reasons"](#) and calls into question exactly how much progress the field has made on actual understanding of language. True humanlike understanding of language is ill-defined, and making sense of it will require substantial breakthroughs in cognitive science and neuroscience.

Some cognitive scientists have argued that embodiment — having a body that can experience the world — is the only path to understanding. They argue that computers—without bodies like ours and without the kinds of experiences we encounter from infancy through adulthood—will never have what it takes to understand language, no matter how deep the network or how copious its training data.

But what if, contrary to my deepest intuitions, the ability to actually understand language isn't required for an A.I. program to successfully converse with us? One big surprise of modern A.I. is that speech recognition — the ability to transcribe spoken language — can be done very accurately just using statistical approaches, without any understanding whatsoever. Perhaps to obtain an A.I. like the *Star Trek* computer, we just need to add more layers and feed in more and better data; maybe actual understanding and actual thinking are beside the point. I would bet against this. But it seems apt to quote our prescient chaos scientist Ian Malcolm from *Jurassic Park*, who could have been describing the future of A.I. language models: "Isn't it amazing? In the information society, nobody thinks. We expected to banish paper, but we actually banished thought."