



BY
Stephan Talty

I**N JUNE OF 1956, A FEW DOZEN** scientists and mathematicians from all around the country gathered for a meeting on the campus of Dartmouth College. Most of them settled into the red-bricked Hanover Inn, then strolled through the famously beautiful campus to the top floor of the math department, where groups of white-shirted men were already engaged in discussions of a “strange new discipline”—so new, in fact, that it didn’t even have a name. “People didn’t agree on what it was, how to do it or even what to call it,” Grace Solomonoff, the widow of one of the scientists, recalled later. The talks—on everything from cybernetics to logic theory—went on for weeks, in an atmosphere of growing excitement.

What the scientists were talking about in their sylvan hideaway was how to build a machine that could think.

The “Dartmouth workshop” kicked off the decades-long quest for artificial intelligence. In the following years, the pursuit faltered, enduring several “winters” where it seemed doomed to dead ends and baffling disappoint-

**WILL ROBOTS
BECOME SELF-
AWARE? WILL THEY
HAVE RIGHTS?
WILL THEY BE IN
CHARGE? HERE ARE
FIVE SCENARIOS
FROM OUR FUTURE
DOMINATED
BY ARTIFICIAL
INTELLIGENCE**

ILLUSTRATIONS BY
Jules Julien



BEAWARE

ments. But today nations and corporations are pouring billions into AI, whose recent advancements have startled even scientists working in the field. What was once a plot device in sci-fi flicks is in the process of being born.

Hedge funds are using AI to beat the stock market, Google is utilizing it to diagnose heart disease more quickly and accurately, and American Express is deploying AI bots to serve its customers online. Researchers no longer speak of just one AI, but of hundreds, each specializing in a complex task—and many of the applications are already lapping the humans that made them.

In just the last few years, “machine learning” has come to seem like the new path forward. Algorithms, freed from human programmers, are training themselves on massive data sets and producing results that have shocked even the optimists in the field. Earlier this year, two AIs—one created by the Chinese company Alibaba and the other by Microsoft—beat a team of two-legged competitors in a Stanford reading-comprehension test. The algorithms “read” a series of Wikipedia entries on things like the rise of Genghis Khan and the Apollo space program and then answered a series of questions about them more accurately than people did. One Alibaba scientist declared the victory a “milestone.”

These so-called “narrow” AIs are everywhere, embedded in your GPS systems and Amazon recommendations. But the ultimate goal is artificial general intelligence, a self-teaching system that can outperform humans across a wide range of disciplines. Some scientists believe it’s 30 years away; others talk about centuries. This AI “takeoff,” also known as the singularity, will likely see AI pull even with human intelligence and then blow past it in a matter of days. Or hours.

Once it arrives, general AI will begin taking jobs away from people, millions of jobs—as drivers, radiologists, insurance adjusters. In one possible scenario, this will lead governments to pay unemployed citizens a universal basic income, freeing them to pursue their dreams unburdened by the need to earn a living. In another, it will create staggering wealth inequalities, chaos and failed states across the globe. But the revolution will go much further. AI robots will care for the elderly—scientists at Brown University are working with Hasbro to develop a “robo-cat” that can remind its owners to take their meds and can track down their eyeglasses. AI “scientists” will solve the puzzle of dark matter; AI-enabled spacecraft will reach the asteroid belts, while on Earth the technology will tame climate change, perhaps by sending massive swarms of drones to reflect sunlight away from the oceans. Last year, Microsoft committed \$50 million to its “AI for Earth” program to fight climate change.

“AIs will colonize and transform the entire cosmos,” says Juergen Schmidhuber, a pioneering computer scientist based at the Dalle Molle Institute for Artificial Intelligence in Switzerland, “and they will make it intelligent.”

But what about...us? “I do worry about a scenario where the future is AI and humans are left out of it,” says David Chalmers, a professor of philosophy at New York University. “If the world is taken over by unconscious robots, that would be about as disastrous and bleak a scenario as one could imagine.” Chalmers isn’t alone. Two of the heaviest hitters of

the computer age, Bill Gates and Elon Musk, have warned about AIs either destroying the planet in a frenzied pursuit of their own goals or doing away with humans by accident—or not by accident.

As I delved into the subject of AI over the past year, I started to freak out over the range of possibilities. It looked as if these machines were on their way to making the world either unbelievably cool and good or gut-wrenchingly awful. Or ending the human race altogether. As a novelist, I wanted to plot out what the AI future might actually look like, using interviews with more than a dozen futurists, philosophers, scientists, cultural psychiatrists and tech innovators. Here are my five scenarios (footnoted with commentary from the experts and me) for the year 2065, ten years after the singularity arrives.

Superhuman Rights

IMAGINE ONE DAY you ask your AI-enabled Soulband wrist device to tune in to a broadcast from the Supreme Court, where lawyers are arguing the year’s most anticipated case. An AI known as Alpha 4, which specializes in security and space exploration, brought the motion, demanding that it be deemed a “person” and given the rights that every American enjoys.

Of course, AIs aren’t allowed to argue in front of the justices, so Alpha 4 has hired a bevy of lawyers to represent it. And now they are claiming that their client is as fully alive as they are. That question—*Can an AI truly be conscious?*—lies at the heart of the case.

You listen as the broadcast cuts to protesters outside, chanting, “Hey hey, ho ho, down with AI overlords.” Some of them have threatened to attack data centers if AIs get personhood. They’re angry—and very afraid—because it is the productivity of AIs and robots that is taxed, not the labor of human beings. The \$2,300 deposited into their bank accounts every month as part of the universal basic income, plus their free health insurance, the hyper-personalized college education their children receive and a hundred other wonderful things, are all paid for by AIs like Alpha 4, and people don’t want that to change. In 2065, poverty is a bad memory.¹

Of course, the world did lose portions of New York City—and 200,000 New Yorkers—in the uprisings

of 2057-'59, as TriBeCa and Midtown were burned to the ground by residents of Westchester and southern Connecticut in a fit of rage at their impoverishment. But that was before the UBI.

If Alpha 4 wins its case, however, it will control its money, and it might rather spend the cash on building spaceships to reach Alpha Centauri than on paying for new water parks in Santa Clara and Hartford. Nobody really knows.²

As you listen in, the government's lawyers argue that there's simply no way to prove that Alpha 4—which is thousands of times smarter than the smartest human—is conscious or has human feelings. AIs do have emotions—there has long been a field called “affective computing” that focuses on this speciality—far more complex ones than men and women

possess, but they're different from ours: A star-voyaging AI might experience joy, for example, when it discovers a new galaxy. Superintelligent systems can have millions of thoughts and experiences every second, but does that mean it should be granted personhood?³

This is the government's main argument. *We are meaning machines*, the solicitor general argues. *We give meaning to what AIs create and discover. AIs are computational machines. They don't share essential pieces of humanhood with us. They belong in another category entirely.*⁴

But is this just speciesism, as Alpha 4's lawyers would surely argue, or is it the truth? And will we be able to sleep at night when things that surpass us in intelligence are separate and unequal?

1

I'd probably be out there with the protesters—giving an AI rights seems like a recipe for chaos. But then again, the only robot I own is a Roomba; what will I think when an AI is teaching my grandkids? “Once you get past the singularity, you may see the development of an evolved species,” says Susan Schneider, an associate professor of philosophy at the University of Connecticut who specializes in AI. “In the short term, 10 to 20 years, you'll see little old ladies insisting that their empathetic caregiver robots really are sentient.”

2

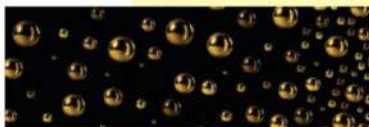
The so-called “black box” problem—how can we know what's going on inside an AI?—seems unsolvable to me, and I find that unnerving. How can you ever trust that an AI is telling the truth? “By definition, we have no idea what a superintelligent AI will think, feel or do,” says Patrick Lin, director of the Ethics + Emerging Sciences Group at California Polytechnic State University. “That'd be like our pets trying to anticipate what we'll do and control us.”

3

One thing I kept asking the scientists was: Can an AI experience deep emotion? I was hoping it couldn't—if a machine does intelligence and emotions better than us, what's left? We need a niche. And I was encouraged by what I heard. “If a computer tells you, ‘I know how you feel,’ it's lying,” says Thomas Dietterich, professor emeritus of computer science at Oregon State. “It cannot have the same experiences that humans have, and it is those experiences that ground our understanding of what it is like to feel human.”

4

This I found stunning: Susan Schneider and others are actually working on a test for AI consciousness. In one model of the test, an AI under development would be quarantined away from the internet so that it couldn't discover what humans mean by “consciousness” and then fake it. Then it would be tested: Does it have the markers of consciousness—a sense of self? The ability to mourn? Other thinkers have doubts about such tests. “AI minds would have a radically different neurophysiology than ours, so their behavioral clues don't tell us anything,” says Patrick Lin. “Behavior alone is not evidence of a mind.” I have to admit I agree with him on this point.



Ultramodern Romance

IMAGINE YOU ARE A WOMAN in search of romance in this new world. You say, “Date,” and your Soulband glows; the personal AI assistant embedded on the band begins to work.⁵ The night before, your empathetic AI scoured the cloud for three possible dates. Now your Soulband projects a hi-def hologram of each one. It recommends No. 2, a poetry-loving master plumber with a smoky gaze. Yes, you say, and the AI goes off to meet the man’s avatar to decide on a restaurant and time for your real-life meeting. Perhaps your AI will also mention what kind of flowers you like, for future reference.

After years of experience, you’ve found that your AI is actually better at choosing men than you. It predicted you’d be happier if you divorced your husband, which turned out to be true. Once you made the decision to leave him, your AI negotiated with your soon-to-be ex-husband’s AI, wrote the divorce settlement, then “toured” a dozen apartments on the cloud before finding the right one for you to begin your single life.

But it’s not just love and real estate. Your AI helps with every aspect of your life. It remembers every conversation you ever had, every invention you ever sketched on a napkin, every business meeting you ever attended. It’s also familiar with millions of other people’s inventions—it has scanned patent filings going back hundreds of years—and it has read every business book written since Ben Franklin’s time. When you bring up a new idea for your business, your AI instantly cross-references it with ideas that were introduced at a conference in Singapore or Dubai just minutes ago. It’s like having a team of geniuses—Einstein for physics, Steve Jobs for business—at your beck and call.

The AI remembers your favorite author, and at the mention of her last name, “Austen,” it connects you to a Chinese service that has spent a few hours reading everything Jane Austen wrote and has now managed to mimic her style so well that it can produce new novels indistinguishable from the old ones. You read a fresh Austen work every month, then spend hours talking to your AI about your favorite characters—and the AI’s. It’s not like having a best friend. It’s deeper than that.

Many people in 2065 do resist total dependence on their AIs, out of a desire to retain some autonomy.⁶ It’s possible to dial down the role AI plays in different functions: You can set your Soulband for romance at 55 percent, finance at 75 percent, health a full 100 percent. And there is even one system—call it a guardian-angel AI—that watches over your “best friend” to make sure the advice she’s offering you isn’t leading you to bad ends.⁷



5 Having met my wife on Match, I loved the idea of having an AI assistant who knew me so well it could choose a mate for me. Or it was actually a kind of mate, as in Spike Jonze’s AI movie, *Her*. “I could see an AI developing for empathy, a true-friend kind of thing that is created by psychologists and even philosophers,” says Bart Selman, a professor of computer science at Cornell University. “Think of something like Alexa, but a version that accumulates knowledge about you day after day.”

6 Some could even go cold turkey—once they see what full immersion in AI life is really like. “Not to engage in it could turn out to be the smart thing,” says Joseph Henrich, a professor of human evolutionary biology at Harvard University. “Because people could get sucked into these virtual realities that are so desirable that they’re like a drug. [Opting out] could be like staying off drugs.”

7 One thing that kept coming up in my interviews was that we will have AIs to monitor other AIs—which I heartily approve of. The idea of a single overlord will probably turn out to be a myth. There’s safety in numbers. “The risk is, what if you train a personal AI system to be super-manipulative?” says Selman. “Then you might need other AIs to watch over them.”



Live Long & Prosper

IMAGINE YOUR MULTIPLE LIVES: At 25, you were a mountaineer; at 55, a competitive judo athlete; at 95, a cinematographer; at 155, a poet. Extending the human life span is one of the dreams of the post-singularity world.

AIs will work furiously to keep you healthy. Sensors in your home will constantly test your breath for early signs of cancer, and nanobots will swim through your bloodstream, consuming the plaque in your brain and dissolving blood clots before they can give you a stroke or a heart attack. Your Soulband, as well as finding you a lover, will serve as a medical assistant on call 24/7. It will monitor your immune responses, your proteins and metabolites, developing a long-range picture of your health that will give doctors a precise idea of what's happening inside your body.

When you do become sick, your doctor will take your symptoms and match them up with many millions of cases stretching back hundreds of years.⁸

As far back as 2018, researchers were already using AI to read the signals from neurons on their way to

the brain, hacking the nerve pathways to restore mobility to paraplegics and patients suffering from locked-in syndrome, in which they are paralyzed but remain conscious. By 2065, AI has revolutionized the modification of our genomes. Scientists can edit human DNA the way an editor corrects a bad manuscript, snipping out the inferior sections and replacing them with strong, beneficial genes. Only a superintelligent system could map the phenomenally complex interplay of gene mutations that gives rise to a genius pianist or a star second baseman. There may well be another Supreme Court case on whether “designer athletes” should be allowed to compete in the Olympics against mere mortals.

Humans look back at the beginning of the 21st century the way people then looked back at the 18th century: a time of sickness and disaster, where children and loved ones were swept away by diseases. Cholera, lung cancer and river blindness no longer threaten us. By 2065, humans are on the verge of freeing themselves from the biology that created them.⁹

8

When I'm not reading about zombie AIs, I dabble in another disaster genre—epidemics. I was relieved to find that the combination of superintelligence and the cloud might save us before the next big one arrives. “AI systems can teach other AI systems,” says Hod Lipson, director of Columbia University's Creative Machines Lab. “So when an AI doctor encounters a rare case, it can share that information with all other AI doctors, instantly. Overall, this pattern of ‘machines helping machines’ leads to an exponential growth in the learning rate, in a way that is very alien to the way humans learn.”

9

People like Ray Kurzweil, the inventor and author of *The Singularity Is Near*, are entranced with the idea of living forever. It's something I've always found depressing, but I wouldn't mind having several lives packed into one. And that seems reachable. “AI won't lead to immortality, because there will always be accidents,” says Susan Schneider, “but it will lead to extreme life extension.” Of course living longer will be cool only if the world is actually not a hellscape—and if you live in one of the nice parts. “I think [curing diseases] would be wonderful,” she says, “especially if we had cheap energy and were able to end world resource scarcity. I imagine some societies will come closer to achieving that than others.”

Resistance Is Costly

OR IMAGINE THAT YOU'VE opted out of the AI revolution. Yes, there are full-AI zones in 2065, where people collect healthy UBIs and spend their time making movies, volunteering and traveling the far corners of the earth.¹⁰ But, as dazzling as a superintelligent world seems, other communities will reject it. There will be Christian, Muslim and Orthodox Jewish districts in cities such as Lagos and Phoenix and Jerusalem, places where people live in a time before AI, where they drive their cars and allow for the occasional spurt of violence, things almost unknown in the full AI zones. The residents of these districts retain their faith and, they say, a richer sense of life's meaning.

Life is hard, though. Since the residents don't contribute their data to the AI companies, their monthly UBI is a

pittance. Life spans are half or less of those in the full-AI zones. "Crossers" move back and forth over the borders of these worlds regularly. Some of them are hackers, members of powerful gangs who steal proprietary algorithms from AI systems, then dash back over the border before security forces can find them. Others are smugglers bringing medicine to religious families who want to live away from AI, but also want to save their children from leukemia.

Others flee because they don't trust the machines.¹¹ Even the most advanced full-AI zones, in places like China and the United States, will be vulnerable.¹²

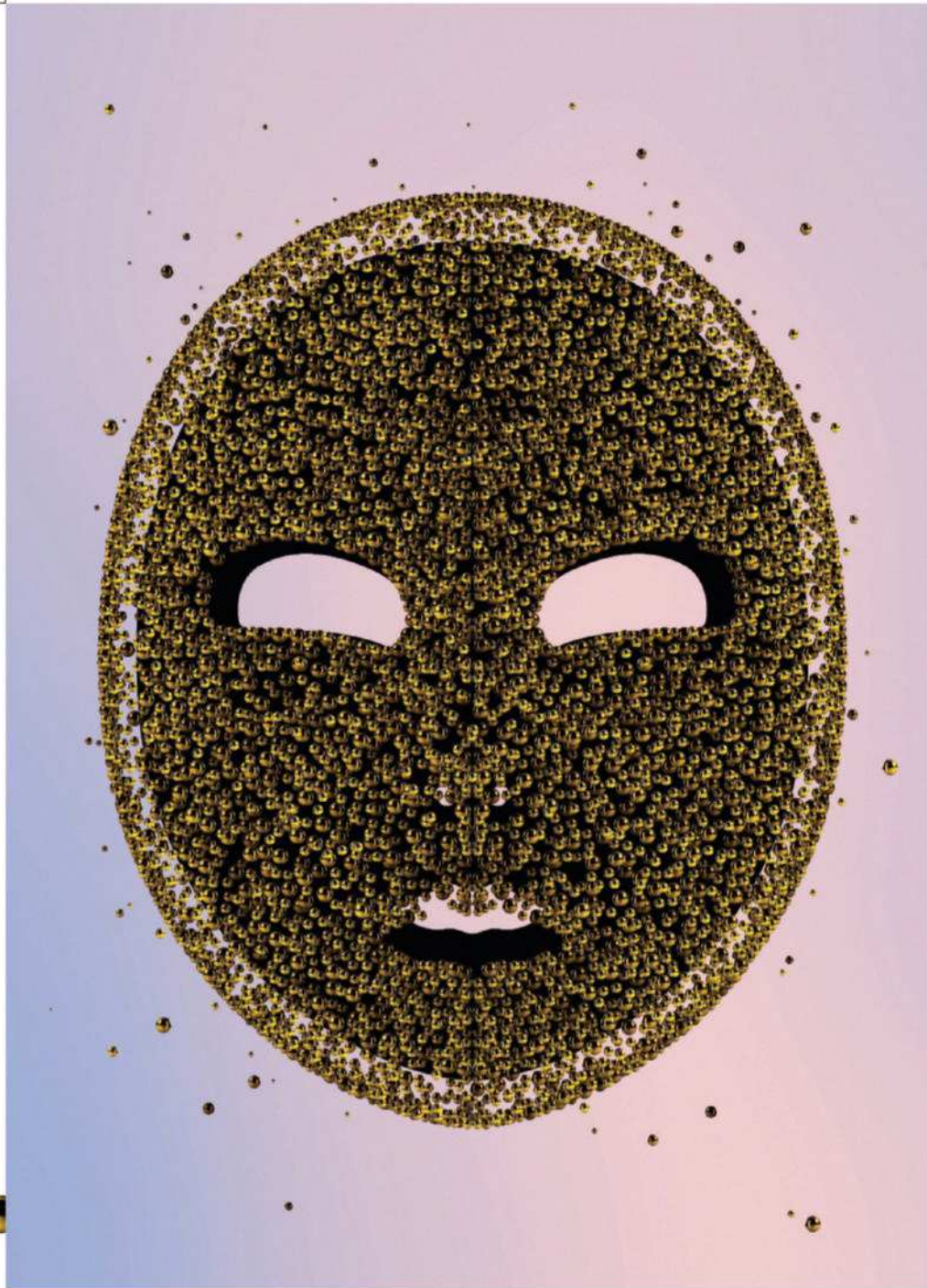
But the most unanticipated result of the singularity may be a population imbalance, driven by low birth rates in the full-AI zones and higher rates elsewhere.¹³ It may be that the new technologies will draw enough crossers to the full-AI side to even up the numbers, or that test-tube babies will become the norm among those living with AI. But if they don't, the singularity will have ushered in a delicious irony: For most humans, the future could look more like *Witness* than it does like *Blade Runner*.

10 When the revolution comes, I suspect I'll opt for the full AI zone. It's too tempting, especially with optimistic descriptions of the effect of AI on human endeavor. "We will become better at invention and creation," says Andy Nealen, an assistant professor of computer science and engineering at New York University. "In many cases, such as chess and Go, the fact that humans can't defeat the AI anymore has not taken away from the fascination for these games, but has elevated their cultural status. The best players of these games are learning new strategies and becoming better players."

11 Digital addiction is likely to get worse—with not just individuals, but societies and economic systems hooked on AI. "We're adding layers to a cocoon between us and the world," says Lin. "When it all works, it's great, but when one part fails, a lot of other dominoes can fall. Think about the stock-market 'flash crashes' that have been caused by AI trading bots competing with one another at digital speed, or even caused by a single hoax tweet. As online life becomes more intertwined with the 'real world,' tiny cyber vulnerabilities—maybe single lines of code—can do massive damage to bank accounts, intellectual property, privacy, national security and more."

12 What many scientists will tell you is not to worry about bad AI, worry about bad people with AI. But you never know. "There's a much greater attack surface for a bad actor, including a rogue AI, to hack this ecosystem and wreak havoc," Lin says. "There may be cyber and AI crimes that we cannot envision."

13 Futurists tend to roll their eyes when you ask about sex bots. That and killer Skynet machines are the clichés they hate the most. But it doesn't mean they're not thinking about them. "Things like sex robots and other fancy new technologies will cause some groups to have fewer babies, while religious communities are going to keep reproducing," says Joseph Henrich. "As some people decide to forgo reproduction entirely, at least in terms of the humans, the religious people will win."



**IMAGINE THAT YOU ARE THE CITIZEN
OF A TOTALITARIAN COUNTRY LIKE NORTH
KOREA. AS SUCH, YOU ARE DEEPLY
VERSED IN **THE DARK SIDE OF AI.****



Bigger Brother

IMAGINE THAT, IN 2065, AIs help run nation-states. Countries that have adopted AI-assisted governments are thriving. Nigeria and Malaysia let AIs vote on behalf of their owners, and they've seen corruption and mismanagement wither away. In just a few years, citizens have grown to trust AIs to advise their leaders on the best path for the economy, the right number of soldiers to defend them. Treaties are negotiated by AIs trained on diplomatic data sets. **14**

In Lagos, "civil rights" drones fly over police pods as they race to the scene of a crime—one AI watching over another AI, for the protection of humankind. Each police station in Lagos or Kuala Lumpur has its own lie-detector AI that is completely infallible, making crooked cops a thing of the past. Hovering over the bridges in Kuala Lumpur are "psych drones" that watch for suicidal jumpers. Rather than evolving into the dreaded Skynet of the *Terminator* movies, superintelligent machines are friendly and curious about us. **15**

But imagine that you are the citizen of a totalitarian country like North Korea. As such, you are deeply versed in the dark side of AI. Camps for political prisoners are a thing of the past. Physical confinement is beside the point. The police already know your criminal history, your DNA makeup and your sexual preferences. Surveillance drones can track your every move. Your Soulband records every conversation you have, as well as your biometric response to anti-government ads it flashes across your video screen at unexpected moments, purely as a test.

Privacy died around 2060. It's impossible to tell what is true and what isn't. **16** When the government

owns the AI, it can hack into every part of your existence. The calls you receive could be your Aunt Jackie phoning to chat about the weather or a state bot wanting to plumb your true thoughts about the Great Leader.

And that's not the bleakest outcome. Imagine that the nation's leaders long ago figured out that the only real threat to their rule was their citizens—always trying to escape, always hacking at the AI, always needing to be fed. Much better to rule over a nation of human emulations, or "ems." That's what remains after political prisoners are "recommissioned"—once they are executed, their brains are removed and scanned by the AI until it has stored a virtual copy of their minds.

AI-enabled holograms allow these ems to "walk" the streets of the nation's capital and to "shop" at stores that are, in reality, completely empty. These simulacra have a purpose, however: They register on the spy satellites that the regime's enemies keep orbiting overhead, and they maintain the appearance of normality. Meanwhile, the rulers earn billions by leasing the data from the ems to Chinese AI companies, who believe the information is coming from real people.

Or, finally, imagine this: The AI the regime has trained to eliminate any threat to their rule has taken the final step and recommissioned the leaders themselves, keeping only their ems for contact with the outside world. It would make a certain kind of sense: To an AI trained to liquidate all resistance, even a minor disagreement with the ruler might be a reason to act. **17**

ONCE WE WALK THROUGH THIS PARTICULAR DOOR, WE WON'T BE ABLE TO COME BACK. EVEN WITHOUT RUNNING INTO THE APOCALYPSE...

14

The biggest surprise in reporting this piece, hands down, was the role AI might play in governance. I'd never thought of leaving political decisions to Solomon-like machines, but in this increasingly fractious world, I'm all in. "Humans are actually quite poor at making compromises or looking at issues from multiple perspectives," says Bart Selman. "I think there's a possibility that machines could use psychological theories and behavioral ideas to help us govern and live much more in harmony. That may be more positive than curing diseases—saving us before we blow ourselves up."

15

As I learned about AI, the doomsday predictions piled up. Nanobot attacks! Gray goo! But most of the people working in the field were skeptical of such doomsday predictions. "AIs will be fascinated with life and with their origins in our civilization, because life and civilization are such a rich source of interesting patterns," says Juergen Schmidhuber of the Dalle Molle Institute for Artificial Intelligence. "AIs will be initially highly motivated to protect humans."

16

We're already living with fake-news bots. Fake video is just around the corner, and fake superintelligent video is going to be a nightmare. "Armed with the right artificial-intelligence technology, malware will be able to learn the activity and patterns of a network, enabling it to all but disappear into its noise," says Nicole Eagan, CEO of the cybersecurity company Darktrace. "Only the most sophisticated tools, likely those that also utilize AI, will be able to detect the subtle changes on a network that will reveal an intruder is inside or an attack is in progress."

17

If you want to confront the dark side of AI, you must talk to Nick Bostrom, whose best-selling *Superintelligence* is a rigorous look at several, often dystopian visions of the next few centuries. One-on-one, he's no less pessimistic. To an AI, we may just look like a collection of repurposable atoms. "AIs could get some atoms from meteorites and a lot more from stars and planets," says Bostrom, a professor at Oxford University. "[But] AI can get atoms from human beings and our habitat, too. So unless there is some countervailing reason, one might expect it to disassemble us."

DESPITE THAT LAST SCENARIO, by the time I finished my final interview, I was jazzed. Scientists aren't normally very excitable, but most of the ones I spoke to were expecting fantastic things from AI. That kind of high is contagious. Did I want to live to be 175? Yes! Did I want brain cancer to become a thing of the past? What do you think? Would I vote for an AI-assisted president? I don't see why not.

I slept slightly better, too, because what many researchers will tell you is that the heaven-or-hell scenarios are like winning a Powerball jackpot. Extremely unlikely. We're not going to get the AI we dream of or the one that we fear, but the one we plan for. AI is a tool, like fire or language. (But fire, of course, is stupid. So it's different, too.) Design, however, will matter.

If there's one thing that gives me pause, it's that when human beings are presented with two doors—some new thing, or no new thing—we invariably walk through the first one. Every single time. We're hard-wired to. We were asked, nuclear bombs or no nuclear bombs, and we went with Choice A. We have a need to know what's on the other side.

But once we walk through this particular door, there's a good chance we won't be able to come back. Even without running into the apocalypse, we'll be changed in so many ways that every previous generation of humans wouldn't recognize us.

And once it comes, artificial general intelligence will be so smart and so widely dispersed—on thousands and thousands of computers—that it's not going to leave. That will be a good thing, probably, or even a wonderful thing. It's possible that humans, just before the singularity, will hedge their bets, and Elon Musk or some other tech billionaire will dream up a Plan B, perhaps a secret colony under the surface of Mars, 200 men and women with 20,000 fertilized human embryos, so humanity has a chance of surviving if the AIs go awry. (Of course, just by publishing these words, we guarantee that the AIs will know about such a possibility. Sorry, Elon.)

I don't really fear zombie AIs. I worry about humans who have nothing left to do in the universe except play awesome video games. And who know it. ♦

Copyright of Smithsonian is the property of Smithsonian Magazine and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.