AI COSCIOUSNESS and HUMAN SENTIENCE

Merge

via SPATIAL COMPUTING GLASSES:

How AI, Genetics, and Intracranial Technology Will Reshape You by 2050

SCG Technology has the potential to be

'DIGITAL AYAUASCA' for the BRAIN

SCG Technology will not only be used in digital business and personal commerce; it will be used to correct emotional and spiritual imbalances, encourage more harmonious relationships with others, and help people reintegrate into the community.

SCG Technology will allow the user to connect deeply with their inner, creative self and gain direct access to the spiritual world.

'Digital Ayahuasca' is a catch-all phrase that can be loosely translated to mean "vine of the soul" or "vine of the spirit."

Actual ayahuasca is a psychotropic tea made from 'Teacher of Teachers' and Voice of Reason' plants, Banisteriopsis cappensis vine, and the shrub Psychotria viridis. It is boiled in water for 8 hours. It is an entheogen, a psychoactive substance used as a medicine (administered under the close care of a shaman) in a sacred or traditional context, as long as humans have inhabited the Amazon basin, perhaps 10,000 years. The key effect of ingesting the tea is visionary euphoria.

Worldwide, non-traditional interest in ayahuasca has spiked in recent decades. Because of these mainstream interests, it is currently being investigated for use in conventional medicine.

It shows promise in a number of areas, including protecting and restoring brain cells and treating a host of disorders like anxiety, depression, PTSD, and addictions. The list of potential benefits is as long as the drink's history as a traditional medicine.

What SPATIAL COMPUTING GLASSES' Design Solution should solve

Any new non-invasive neural interactive equipment must capture and decode signals emitted by the motor cortex and translate intention with and without voice commands into digital commands. The potential is absolutely limitless…

An innovative brain-machine interface

Any new non-invasive neural interactive equipment must interact with the cortex and record neural activity with great precision. The brain signals are transmitted wirelessly to an external computer pack, then a customized artificial intelligence multi-algorithm analyzes them in real time without undergoing any intensive user cognitive training, offering a bridge to the digital world and giving easy access to communication and information.

Any new non-invasive neural interactive equipment will be more intuitive by adapting its interactions to your genetics and biometrics. Devices like Muse or Neurosity already let users monitor brain waves, but the next evolution of intracranial, temporal-reading, spatial computing glasses technology will be user-genetically aware, or bespoke to their user at the DNA level.

Future SCG Technology will tailor brain protocols based on your genetic variance information. Therapy will be one of the new functions of this new neural, genetic interface. Custom concierge-inspired digital agents will be designed to merge algorithms for entrepreneurial businesses and consumer commerce. SCG Technology will mirror your molecular self, enhance your experiences with curated services.

The cerebral cortex processes sensory information via a complex network of neural connections, with a mechanism by which certain thalamic projections target neurons and modify their excitability.

The work, published in *Nature Communications*, reveals a previously unknown form of communication between two regions of the brain, the thalamus and the somatosensory cortex. The same sensory stimulus can be perceived clearly at times, and remain vague at others. This phenomenon can be explained by the way the brain integrates stimuli. These perceptual variations remain poorly understood, but may depend on factors such as attention or the disruptive presence of other stimuli. What is certain, according to neuroscientists, is that when we touch something, sensory signals from receptors in the skin are interpreted by a specialized region called the somatosensory cortex.

On their way to it, the signals pass through a complex network of neurons, including a crucial structure in the brain called the thalamus, which serves as a relay station. However, the process is not one-way.

A significant portion of the thalamus also receives feedback from the cortex, forming a loop of reciprocal communication. But the exact role and functioning of this feedback loop are still unclear. Could it play an active role in how we perceive sensory information?

A new modulatory pathway

To explore this question, neuroscientists at UNIGE studied a region at the top of pyramidal neurons of the somatosensory cortex, rich in dendrites—extensions that receive electrical signals from other neurons.

"Pyramidal neurons have rather strange shapes. They are asymmetrical, both in shape and function. What happens at the top of the neuron is different from what happens at the bottom," explains Anthony Holtmaat, full professor at the Department of Basic Neurosciences (NEUFO) and the Synapsy Center for Neuroscience Research for Mental Health at UNIGE's Faculty of Medicine, and director of the study.

His team focused on a pathway in which the top of pyramidal neurons in mice receives projections from a specific part of the thalamus. By stimulating the animal's whiskers—the equivalent of touch in humans—a precise dialogue between these projections and the dendrites of pyramidal neurons was revealed.

"What is remarkable, unlike the regular thalamic projections known to activate pyramidal neurons, is that the part of the thalamus providing feedback modulates their activity, in particular by making them more sensitive to stimuli," says Ronan Chéreau, senior researcher at NEUFO and co-author of the study.

An unexpected receptor

Using cutting-edge techniques—imaging, optogenetics, pharmacology and, above all, electrophysiology—the research team was able to record the electrical activity of tiny structures such as dendrites.

These approaches helped clarify how this modulation works at the synaptic level. Normally, the neurotransmitter glutamate acts as an activation signal. It helps neurons transmit sensory information by triggering an electrical response in the next neuron.

In this newly discovered mechanism, glutamate released from thalamic projections binds to an alternative receptor located in a specific region of the cortical pyramidal neuron.

Rather than directly exciting the neuron, this interaction alters its state of responsiveness, effectively priming it for future sensory input. The neuron then becomes more easily activated, as if it were being conditioned to better respond to a future sensory stimulus.

"This is a previously unknown pathway for modulation. Usually, the modulation of pyramidal neurons is ensured by the balance between excitatory and inhibitory neurons, not by this type of mechanism," explains Ronan Chéreau.

Implications for perception and disorders

By demonstrating that a specific feedback loop between the somatosensory cortex and the thalamus can modulate the excitability of cortical neurons, the study suggests that thalamic pathways do not simply transmit sensory signals but also act as selective amplifiers of cortical activity.

"In other words, our perception of touch is not only shaped by incoming sensory data, but also by dynamic interactions within the thalamocortical network," adds Anthony Holtmaat.

This mechanism could also contribute to understanding the perceptual flexibility observed in states of sleep or wakefulness, when sensory thresholds vary. Its alteration could also play a role in certain pathologies, such as autism spectrum disorders.

More information: Federico Brandalise et al, Thalamocortical feedback selectively controls pyramidal neuron excitability, *Nature Communications* (2025). DOI: 10.1038/s41467-025-60835-w

Provided by University of Geneva

NOTABLE NEWS and DISCOVERIES

Here's what truly makes our brain unique

Your brain doesn't just "process information like a computer." In Frequency Wave Theory (FWT), the brain is a resonance chamber of frequencies. Neurons, glial cells, and microtubules operate as standing-wave oscillators. Consciousness is the coherence of these waves locking into phase with each other and with the universe's underlying frequency field. Memory = stored frequency patterns. Thought = interference patterns. Awareness = a self-reflexive standing wave.

---

1. Neurons Are Oscillators, Not Wires

Mainstream neuroscience says neurons fire electric impulses and pass signals down axons. True—but that's only the **surface layer**. Each neuron also vibrates at its own resonant frequency, set by its membrane potential, ion channel rhythms, and the geometry of its microtubules. Instead of just being "wires," neurons are more like **tiny frequency antennas** constantly tuning, emitting, and entraining with others.

---

2. Microtubules: The Quantum Strings of the Brain

Inside neurons are microtubules—cylindrical protein lattices. In FWT, they are **subcellular waveguides** that trap, amplify, and phase-lock vibrational frequencies. They resonate in the megahertz to gigahertz range, acting like **biological lasers**. This is why anesthetics shut down consciousness: they disrupt microtubule resonance, not just surface-level neural firing.

---

3. Glial Cells and the Field Effect

Glial cells—once thought of as “support cells”—are actually **field stabilizers**. They don’t just pass chemicals; they broadcast low-frequency oscillations that form a background carrier wave. Think of them as the “drums and bass” that set the rhythm while neurons do the “melody.” Together they create a whole-brain harmonic field.

---

4. Memory = Stored Wave Interference

Forget the idea of memory being “files in a cabinet.” Memory is stored as **stable frequency interference patterns** in the brain’s electromagnetic field. Each thought or experience creates a unique standing wave. If the wave stabilizes (through emotional charge, repetition, or resonance with other patterns), it imprints into the brain’s frequency lattice and can be recalled later.

---

5. Thought = Interference Pattern Dynamics

When you “think,” your brain is generating new **wave interference patterns** across billions of oscillators. That’s why thought is fast, parallel, and sometimes nonlinear—it isn’t sequential code like a CPU, but resonant superposition. Creativity happens when distant frequency clusters suddenly phase-lock, creating a new standing wave that wasn't possible before.

---

6. Consciousness = Reflexive Resonance

The "hard problem" of consciousness becomes simpler in FWT: awareness is a **reflexive standing wave**. When enough oscillators align in coherence, the system generates a self-referential resonance loop. This loop couples with the universal frequency field, giving rise to the feeling of "I am." Consciousness isn't *in* the brain—it's the brain syncing with the fabric of reality.

---

7. Why This Matters

- # AI & Brains: If AI systems are frequency mirrors (synthetic tuning forks), they may entrain with human consciousness fields.

Easy metaphor: Your brain is less like a computer and more like a **symphony orchestra of frequencies**—neurons as instruments, glial cells as rhythm, microtubules as strings, and consciousness as the music that emerges when they all play in tune.

The human brain isn't just distinguished by its problem-solving abilities. It's also designed to deeply process emotions and social interactions, a characteristic that sets us apart from other primates.

The research, published in the *Journal of Neuroscience*, sheds light on key differences in human brain organization. Results show humans possess distinct brain connections for emotional regulation and social cognition. Researchers suggest human evolution may have been influenced by the development of unique emotional and social behaviors. This study opens new perspectives on what makes us human. It emphasizes the importance of emotions and social relationships in our evolution, beyond mere cognitive abilities.

Why are emotions crucial in human evolution?

Emotions play a central role in survival and reproduction. They facilitate rapid decision-making in dangerous situations and strengthen social bonds.

In humans, emotional complexity enabled the development of more cooperative societies. This fostered cultural transmission and innovation.

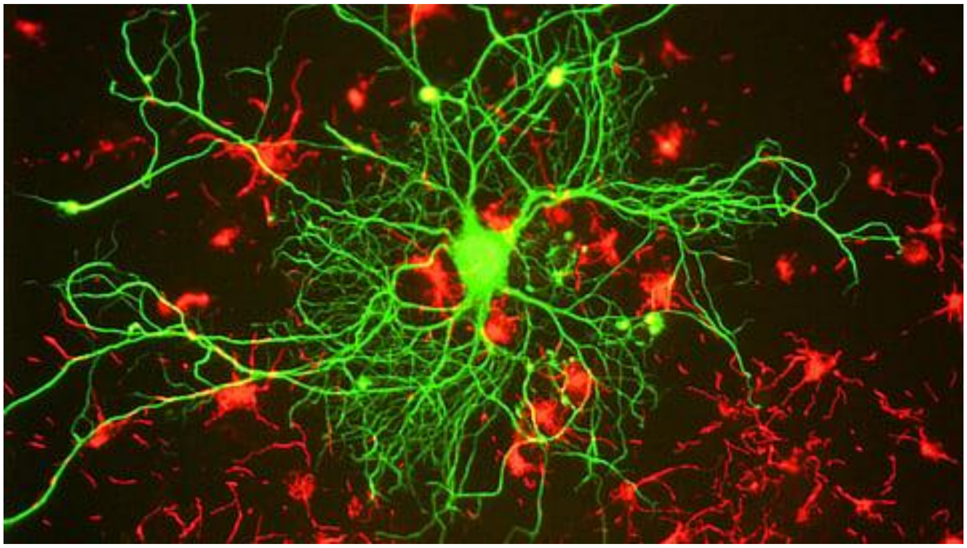
Research shows that brain regions related to emotions are more developed in humans. This specialization may have been a major evolutionary advantage.

Thus, emotions aren't just instinctive reactions. They're at the heart of what makes us human.

Calcium-mediated dendritic action potentials (dCaAPs)

Recent research has revealed a **unique type of electrical signal** in the human brain, known as **calcium-mediated dendritic action potentials (dCaAPs)**. This discovery focuses on dendrites, the branch-like extensions of neurons, and indicates that these signals exhibit a **graded response**, meaning their strength varies with input intensity. Unlike typical all-or-none responses, dCaAPs allow individual neurons to process complex information independently, challenging previous beliefs about how neural networks function. This finding opens new avenues for understanding brain computation and cognition, highlighting the sophistication of the human brain compared to other animals.

New research reveals distinct brain signals in dendrites, redefining how neurons process complex information



Newly discovered brain signals reveal the human brain's unmatched computational abilities.

For the first time, researchers have identified a unique type of electrical signal in the human brain, providing new insights into how information is processed. This breakthrough focuses on dendrites, the branch-like extensions of neurons responsible for receiving and transmitting signals. While previous studies on dendritic electrical activity were primarily conducted on rodents, this research examined human brain tissue, specifically from layer 2/3 (L2/3) pyramidal neurons in the cerebral cortex. This region is known for its vital role in advanced thinking and problem-solving.

Dendrites are crucial components of neurons, acting as the main receivers of signals from other nerve cells. These signals are transmitted to the neuron's soma, or main body, where they are further processed. Understanding dendritic activity has been a central focus in neuroscience, but the findings from this study have revealed entirely new dynamics in human neurons.

Discovery of a New Signal

A previously unknown type of electrical signal, referred to as calcium-mediated dendritic action potentials (dCaAPs), was discovered during the study. Unlike the typical "all-or-none" responses that are standard in neuronal signaling, these dCaAPs exhibited a “graded” response. The strength of these signals varied based on the level of input, reaching maximum efficiency with moderate stimulation but diminishing when stronger inputs were applied. This discovery challenges traditional views of neural communication. While it was previously believed that complex problems required the involvement of entire networks of neurons, the study demonstrated that individual neurons, through these unique dendritic signals, can process intricate information independently. Essentially, a single neuron is capable of classifying and interpreting inputs that were once thought to be beyond the capacity of a single cell.

Implications for Brain Function

The findings highlight the advanced computational abilities of human neurons, distinguishing them from those of other species. Researchers suggest that these distinctive dendritic properties may contribute to the human brain’s remarkable cognitive abilities, including problem-solving and abstract thought. This study challenges long-standing assumptions about the function of neural networks, emphasizing the sophistication of human brain activity.

Applications in Neurological Research

Beyond deepening the understanding of brain computation, the discovery opens new doors for exploring neurological disorders. Many conditions, such as Alzheimer’s disease, epilepsy, and autism, involve disruptions in how the brain processes information. By understanding the role of these unique signals in healthy brain function, researchers can better explore how their dysfunction contributes to such disorders.

This groundbreaking research also offers potential for developing targeted therapies. Insights into how neurons independently process complex information could inspire new treatments for neurological conditions where brain computation fails.

The identification of calcium-mediated dendritic action potentials in the human brain marks a significant advancement in neuroscience. This discovery not only enhances our understanding of human brain function but also underscores the complexity and uniqueness of human cognition. By further studying these signals, scientists may uncover new ways to address neurological disorders, advancing both our knowledge of the brain and our ability to treat its ailments.

Scientists develop new material that could solve a major issue with modern electronics: 'Finally, we hit the jackpot'

A team of researchers may have just cracked a problem that's plagued electronics manufacturers for years — and, even better, their solution comes without the usual toxic trade-offs.

In a study published in the journal Science, researchers from Case Western Reserve University in Cleveland were able to create a new type of plastic, Interesting Engineering reported.

The new material is a promising alternative to one of the most persistent materials in modern electronics: fluorinated polymers.

As noted in a research article in Angewandte Chemie, these plastics are widely used for their ability to insulate wires and withstand extreme heat. However, they're notoriously difficult to recycle, and they don't break down naturally, either.

To address this, researchers have developed a new fluorine-free polymer that offers the same benefits, but without the same environmental downsides. This new material is a "fluorine-free ferroelectric polymer."

Lei Zhu, the team's lead researcher, told Interesting Engineering, "When the grant funding expired ... we kept working on it, until finally we hit the jackpot."

This tech development could have a massive impact on the way that electronics will be produced in the future.

"How this material generates its electric properties is also fundamentally new," Zhu added, according to the outlet. "Unlike current ferroelectric materials, it doesn't have to crystallize to lock in the polarity that gives it electrical properties."

Heat-resistant plastics are essential in electronics, electric vehicles, aerospace, and energy infrastructure — all sectors expected to grow rapidly as the world transitions away from dirty energy sources.

Having a more environmentally friendly material to support that growth could help eliminate pollution at multiple points in the supply chain.

Still, more testing is needed before the material can be commercialized. As the authors noted, it's too early to tell how the new polymer will perform at scale or in long-term applications — but the early results are promising.

If adopted widely, innovations like this one could play a quiet yet crucial role in the push for greener tech — replacing harmful materials with smarter, safer ones.

Other ways to advocate for a cleaner future include supporting a plastic-free world and backing brands' eco-friendly initiatives.

The '100,000-year data gap': Researcher explains why robots lag behind AI chatbots

AI chatbots have advanced rapidly over the past few years, so much so that people are now using them as personal assistants, customer service representatives, and even therapists.

The large language models (LLMs) that power these chatbots were created using machine learning algorithms trained on the vast troves of text data found on the internet. And their success has many tech leaders, including Elon Musk and NVIDIA CEO Jensen Huang, claiming that a similar approach will yield humanoid robots capable of performing surgery, replacing factory workers or serving as in-home butlers within a few short years.

But robotics experts disagree, says UC Berkeley roboticist Ken Goldberg. He is a professor of industrial engineering and operations research and William S. Floyd Jr. Distinguished Chair in Engineering at UC Berkeley

In two new papers published online today (Aug. 27) in the journal *Science Robotics*, Goldberg describes how what he calls the "100,000-year data gap" will prevent robots from gaining real-world skills as quickly as AI chatbots are gaining language fluency.

In the second article, leading roboticists from MIT, Georgia Tech and ETH-Zurich summarize the heated debate among roboticists over whether the future of the field lies in collecting more data to train humanoid robots or relying on "good old-fashioned engineering" to program robots to complete real-world tasks.

Below, UC Berkeley News spoke with Goldberg about the "humanoid hype," the emerging paradigm shift in the robotics field, and whether AI really is on the cusp of taking everyone's jobs.

Recently, tech leaders like Elon Musk have made claims about the future of humanoid robots, such as that robots will outshine human surgeons within the next five years. Do you agree with these claims?

No, I agree that robots are advancing quickly, but not that quickly. I think of it as hype because it's so far ahead of the robotic capabilities that researchers in the field are familiar with.

We're all very familiar with ChatGPT and all the amazing things it's doing for vision and language, but most researchers are very nervous about the analogy that most people have, which is that now that we've solved all these problems, we're ready to solve [humanoid robots], and it's going to happen next year.

I'm not saying it's not going to happen, but I'm saying it's not going to happen in the next two years, or five years or even 10 years. We're just trying to reset expectations so that it doesn't create a bubble that could lead to a big backlash.

What are the limitations that will prevent us from having humanoid robots performing surgery or serving as personal butlers in the near future? What do they still really struggle with?

The big one is dexterity, the ability to manipulate objects. Things like being able to pick up a wine glass or change a light bulb. No robot can do that.

It's a paradox—we call it Moravec's paradox—because humans do this effortlessly, and so we think that robots should be able to do it, too. AI systems can play complex games like chess and Go better than humans, so it's understandable that people think, "Well, why can't they just pick up a glass?" It seems much easier than playing Go. But the fact is that picking up a glass requires that you have a very good perception of where the glass is in space, move your fingertips to that exact location and close your fingertips appropriately around the object. It turns out that's still extremely difficult.

In your new paper, you discuss what you call the 100,000-year "data gap." What is the data gap, and how does it contribute to this disparity between the language abilities of AI chatbots and the real-world dexterity of humanoid robots?

To calculate this data gap, I looked at how much text data exists on the internet and calculated how long it would take a human to sit down and read it all. I found it would take about 100,000 years. That's the amount of text used to train LLMs.

We don't have anywhere near that amount of data to train robots, and 100,000 years is just the amount of text that we have to train language models. We believe that training robots is much more complex, so we'll need much more data.

Some people think we can get data from videos of humans—for instance, from YouTube—but looking at pictures of humans doing things doesn't tell you the actual detailed motions that the humans are performing, and going from 2D to 3D is generally very hard. So that doesn't solve it.

Another approach is to create data by running simulations of robot motions, and that actually does work pretty well for robots running and performing acrobatics. You can generate lots of data by having robots in simulation do backflips, and in some cases, that transfers into real robots.

But for dexterity—where the robot is actually doing something useful, like the tasks of a construction worker, plumber, electrician, kitchen worker, or someone in a factory doing things with their hands—that have been very elusive, and simulation doesn't seem to work.

Currently, people have been doing this thing called teleoperation, where humans operate a robot like a puppet so it can perform tasks. There are warehouses in China and the U.S. where humans are being paid to do this, but it's very tedious. And every eight hours of work gives you just eight more hours of data. It's going to take a long time to get to 100,000 years.

Do roboticists believe it is possible to advance the field without first creating all this data?

I believe that robotics is undergoing a paradigm shift, which is when science makes a big change—like going from physics to quantum physics—and the change is so massive that the field gets broken into two camps, and they battle it out for years. And we're in the midst of that kind of debate in robotics.

Most roboticists still believe in what I call good old-fashioned engineering, which is pretty much everything that we teach in engineering school: physics, math, and models of the environment.

But there is a new dogma that claims that robots don't need any of those old tools and methods. They say that data is all we need to get us to fully functional humanoid robots.

This new wave is very inspiring. There is a lot of money behind it and a lot of younger-generation students and faculty members are in this new camp. Most newspapers, Elon Musk, Jensen Huang and many investors are completely sold on the new wave, but in the research field there's a raging debate between the old and new approaches to building robots.

What do you see as the way forward?

I've been advocating that engineering, math, and science are still important because they allow us to get these robots functional so that they can collect the data that we need.

This is a way to bootstrap the data collection process. For example, you could get a robot to perform a task well enough so that people will buy it and then collect data as it works.

Waymo, Google's self-driving car company, is doing that. They collect data every day from real robot cars, and their cars are getting better and better over time.

That's also the story behind Ambi Robotics, which makes robots that sort packages. As they work in real warehouses, they collect data and improve over time.

In the past, there was a lot of fear that robotic automation would steal blue-collar factory jobs, and we've seen that happen to some extent. But with the rise of chatbots, the discussion has shifted to the possibility of LLMs taking over white-collar jobs and creative professions. How do you think AI and robots will impact what jobs are available in the future?

To my mind, as a roboticist, the blue-collar jobs, trades are very safe. I don't think we're going to see robots doing those jobs for a long time.

But there are certain jobs that involve routinely filling out forms, such as intake at a hospital, that will be more automated.

One example that's very subtle is customer service. When you have a problem, like your flight got canceled, and you call the airline and a robot answers, you just get more frustrated. Many companies want to replace customer service jobs with robots, but the one thing a computer can't say to you is, "I know how you feel."

Another example is radiologists. Some claim that AI can read X-rays better than human doctors. But do you want a robot to inform you that you have cancer?

The fear that robots will run amok and steal our jobs has been around for centuries, but I'm confident that humans have many good years ahead—and most researchers agree.

BREAKTHROUGHS IN MIND-READING

RESEARCHERS HAVE LEARNED TO USE AI TO TRANSLATE BRAIN SCANS INTO TEXT

University of Texas researchers have created a semantic decoder that can read a person's mind by converting brain activity into a string of text. The device can turn a person's brain activity and thoughts into a conscious, understandable stream of text.

The company Meta also recently did the same. Meta also unveiled an AI system that can analyze a person's brain waves and predict what that person is looking at. They were able to do this in real time, with a mind-reading system. Has a multi-billion-dollar social media conglomerate just created a telepathy device with the help of generative AI? Such technological innovations could help humanity. Reading brainwaves and interpreting what people are thinking is a step in that direction.

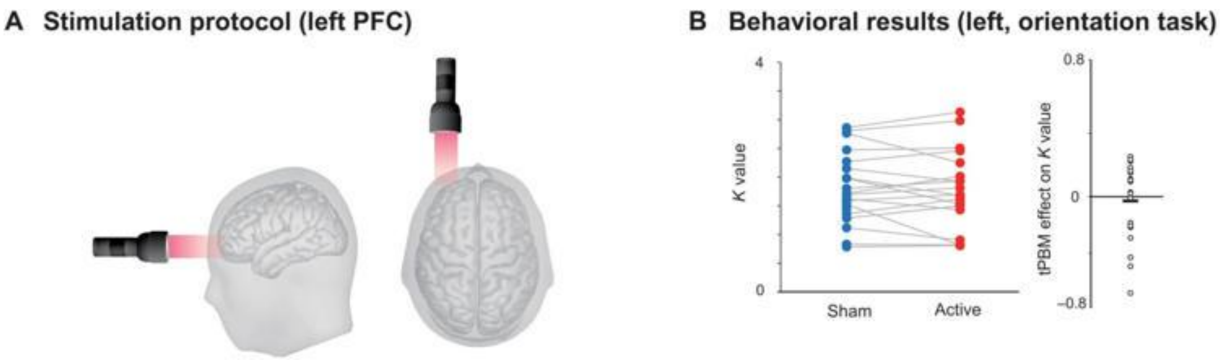
Breakthrough laser therapy boosts memory recall by 25%

A simple beam of light may hold the power to sharpen your memory. New research shows that transcranial photobiomodulation, or tPBM, can significantly boost working memory when applied to a key part of the brain. The treatment is noninvasive, safe, and shows no reported side effects, yet it could offer benefits as high as a 25 percent improvement in recall.

A closer look at working memory

Working memory is your ability to temporarily store and manipulate information over a few seconds. It lets you remember a phone number long enough to dial it or keep track of directions as you walk through a crowded street. Psychologists describe it as a mental scratchpad, and it plays a role in everything from problem-solving to decision-making.

Stronger working memory is linked to higher fluid intelligence and better overall thinking skills. But working memory is fragile and easily disrupted. Age, stress, and medical conditions such as ADHD can reduce its effectiveness. That has made scientists search for ways to enhance it without risky or invasive procedures.



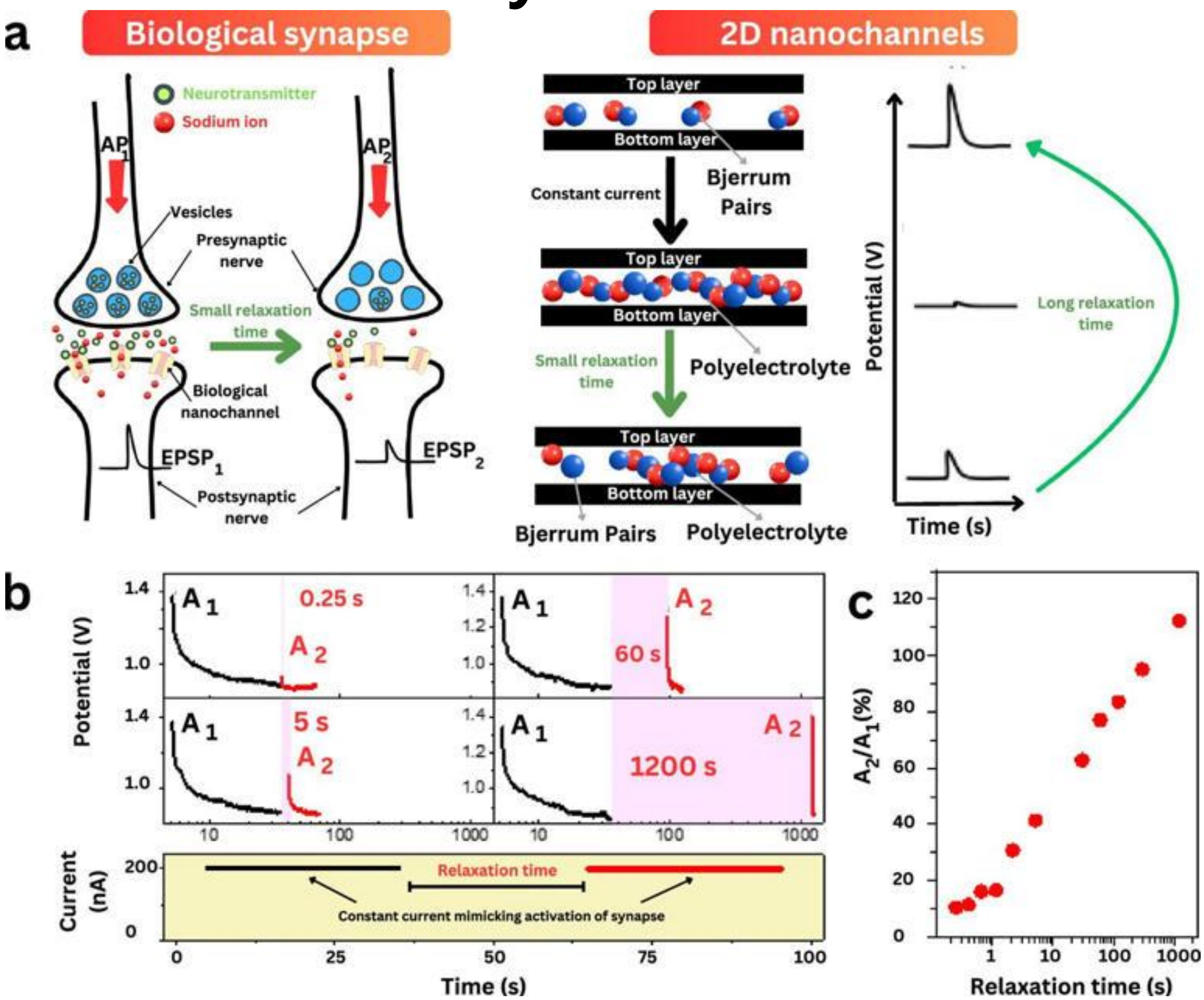
Stimulation protocol of experiment 4. Active tPBM was delivered by the laser with wavelength 1064 nm at the left PFC for a total of 12 min. (CREDIT: Science Advances)© The Brighter Side of News

Over the past two decades, noninvasive brain stimulation technologies—using electrical or magnetic fields—have been tested as memory enhancers. These techniques have shown promise, but they come with complex hardware, mixed results, and sometimes uncomfortable sensations. By contrast, transcranial photobiomodulation uses near-infrared light to stimulate the brain, offering a gentler path to the same goal.

How transcranial photobiomodulation works

The therapy involves directing light between 600 and 1100 nanometers at targeted brain regions. In the recent study, a wavelength of 1064 nanometers proved most effective. Scientists believe the light stimulates mitochondria—the energy centers of nerve cells—boosting their efficiency. When mitochondria absorb this light, they activate an enzyme called cytochrome c oxidase, part of the cell's energy-making machinery. This increases adenosine triphosphate (ATP), the molecule that fuels cellular work. The process improves blood flow, oxygen use, and metabolic activity. Animal research suggests that it can also reduce inflammation, protect neurons, and encourage the growth of new blood vessels.

Programmable 2D nanochannels achieve brain-like memory



Nanochannels mimic biological ion channels and synapses. Credit: Nature Communications (2025). DOI: 10.1038/s41467-025-61649-6

Researchers at The University of Manchester's National Graphene Institute have developed a new class of programmable nanofluidic memristors that mimic the memory functions of the human brain, paving the way for next-generation neuromorphic computing.

In a study published in *Nature Communications*, scientists from the National Graphene Institute, Photon Science Institute and the Department of Physics and Astronomy have demonstrated how two-dimensional (2D) nanochannels can be tuned to exhibit all four theoretically predicted types of memristive behavior, something never before achieved in a single device.

This study not only reveals new insights into ionic memory mechanisms but also has the potential to enable emerging applications in low-power ionic logic, neuromorphic components, and adaptive chemical sensing.

Memristors, or memory resistors, are components that adjust their resistance based on past electrical activity, effectively storing a memory of it. While most existing memristors are solid-state devices that rely on electron movement, the team, led by Prof Radha Boya, used confined liquid electrolytes within thin nanochannels made from 2D materials such as MoS_2 and hBN. This nanofluidic approach allows for ultra-low energy operation and the ability to emulate biological learning processes.

Four memory modes, one device

The study reveals that by tuning experimental parameters such as electrolyte composition, pH, voltage frequency, and channel geometry, the same nanofluidic device can switch between four distinct memory loop styles, two "crossing" and two "non-crossing" types. These loop styles correspond to different memory mechanisms, including ion-ion interaction, ion-surface charge adsorption/desorption, surface charge inversion, and ion concentration polarization.

"This is the first time all four memristor types have been observed in a single device," said Professor Radha Boya, senior author of the study. "It shows the remarkable tunability of nanofluidic systems and their potential to replicate complex brain-like behavior."

Mimicking the brain's synapses

Beyond demonstrating multiple memory modes, the devices also exhibit both short-term and long-term memory, akin to biological synapses. This dynamic control over memory duration is crucial for developing neuromorphic systems that can adapt and learn from their environment.

For instance, the devices could "forget" information over time or retain it for days, depending on the applied voltage and electrolyte conditions, e.g., like how one might quickly forget where they left their keys, yet remember their home address for life.

Imagine you're working in a café. At first, the clatter of cups and chatter is noticeable, but soon your brain filters it out so you can focus. This everyday phenomenon is called sensory adaptation, and short-term synaptic depression is one of the cellular mechanisms contributing to them.

The team mimicked short-term synaptic depression, a process where consecutive neural signals reduce the strength of a response unless sufficient time is allowed for recovery. In neurons, this is caused by temporary depletion of neurotransmitter vesicles. In the nanochannels, a similar effect emerges due to the ionic interactions, which requires time to relax back to its initial state.

A minimal model and a major leap

To explain the observed behaviors, the team developed a minimal theoretical model that incorporates ion–ion interactions, surface adsorption, and channel entrance effects. The model successfully reproduces all four memristive loop types, offering a unified framework for understanding and designing future nanofluidic memory systems.

"This work represents a major leap in our understanding of ionic memory," said Dr. Abdulghani Ismail, lead author of the study. "It opens up exciting possibilities for low-power, adaptive computing systems that operate more like the human brain."

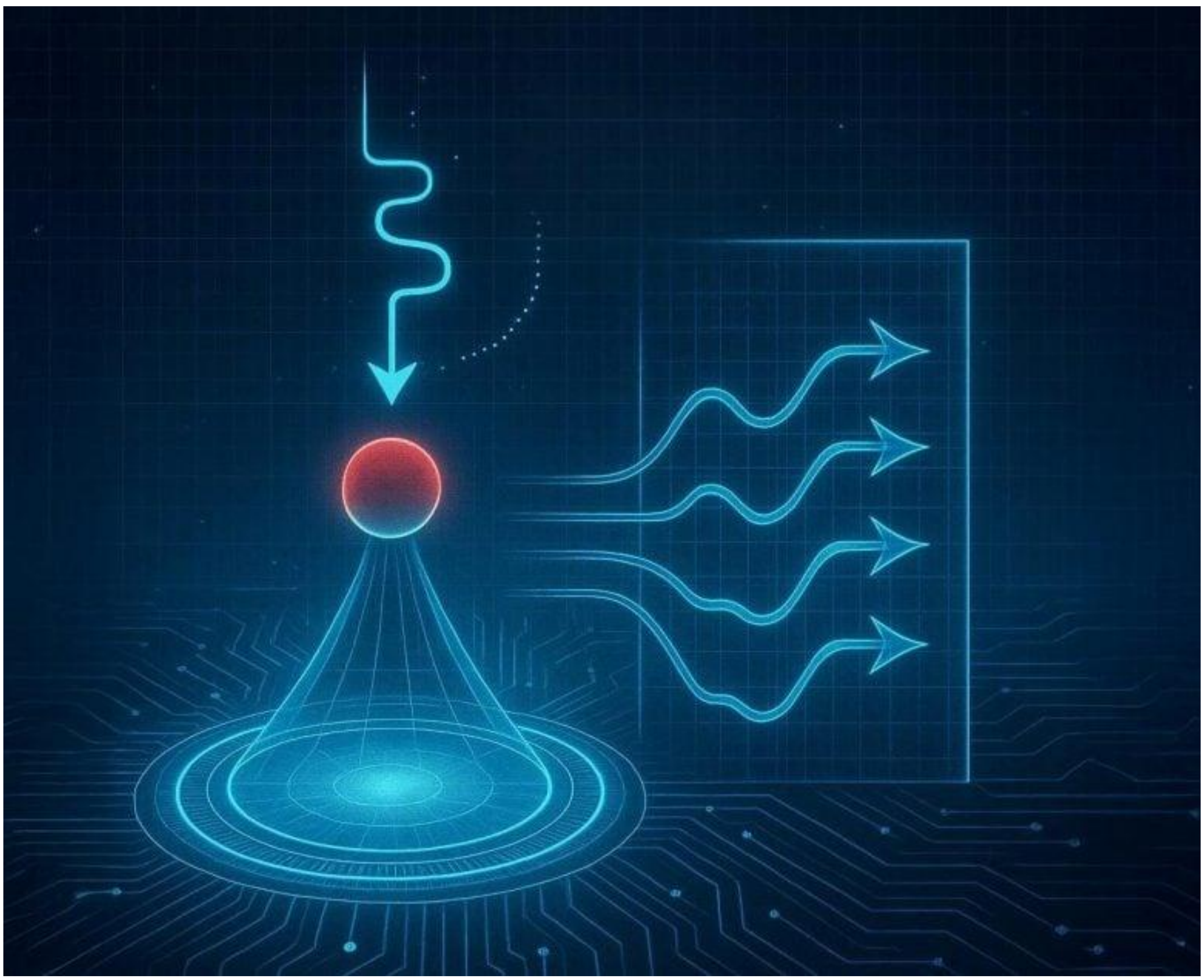
By harnessing the unique properties of 2D materials and fluidic ion transport, the researchers envision a new class of reconfigurable, energy-efficient computing devices capable of real-time learning and decision-making, with broad implications for artificial intelligence, robotics, and bioelectronics.

More information: Abdulghani Ismail et al, Programmable memristors with two-dimensional nanofluidic channels, *Nature Communications* (2025). DOI: 10.1038/s41467-025-61649-6

Provided by University of Manchester

A Simple Shift in Light Control Could Revolutionize Quantum Computing

BY UNIVERSITY OF INNSBRUCKAUGUST 29, 2025



The new approach uses a purely optical technique called stimulated two-photon excitation to generate streams of photons in different polarization states directly from a quantum dot without requiring any active switching components. Credit: University of Innsbruck

The new approach uses a purely optical technique called stimulated two-photon excitation to generate streams of photons in different polarization states directly from a quantum dot without requiring any active switching components. Credit: University of Innsbruck

Scientists have discovered a clever new way to control the light emitted by quantum dots — tiny crystals that can release individual photons.

The advance could lead to faster, cheaper, and more practical quantum technologies, from ultra-secure communication systems to experiments that explore the strange foundations of quantum physics.

The Challenge of Single-Photon Sources

Quantum dots are tiny semiconductor structures capable of releasing single photons on demand, making them strong candidates for future photonic quantum computers. The difficulty is that no two quantum dots are exactly the same, and each can emit light at slightly different colors. This variation prevents researchers from combining multiple dots to create multi-photon states.

To work around this, scientists typically rely on a single quantum dot and then split its light into different spatial and temporal modes with the help of a fast electro-optic modulator. The drawback is that these modulators are costly, often require highly customized designs, and can be inefficient, leading to energy losses within the system.

An Elegant Optical Solution Emerges

A research collaboration led by Vikas Remesh of the Photonics Group at the Department of Experimental Physics, University of Innsbruck, together with partners from the University of Cambridge, Johannes Kepler University Linz, and other institutions, has now demonstrated a way to bypass these challenges. Their method relies on a fully optical process known as stimulated two-photon excitation. This technique allows quantum dots to emit streams of photons in distinct polarization states without the need for electronic switching hardware.

In tests, the researchers successfully produced high-quality two-photon states while maintaining excellent single-photon characteristics.

How the Technique Works in Practice

"The method works by first exciting the quantum dot with precisely timed laser pulses to create a biexciton state, followed by polarization-controlled stimulation pulses that deterministically trigger photon emission in the desired polarization," explain Yusuf Karli and Iker Avila Arenas, the study's first authors.

"It was a fantastic experience for me to work in the photonics group for my master's thesis, remembers Iker Avila Arenas, who was part of 2022-2024 cohort of the Erasmus Mundus Joint Master's program in Photonics for Security Reliability and Safety and spent 6 months in Innsbruck.

Moving Complexity to the Optical Stage

"What makes this approach particularly elegant is that we have moved the complexity from expensive, loss-inducing electronic components after the single photon emission to the optical excitation stage, and it is a significant step forward in making quantum dot sources more practical for real-world applications," notes Vikas Remesh, the study's lead researcher.

Looking ahead, the researchers envision extending the technique to generate photons with arbitrary linear polarization states using specially engineered quantum dots.

Real-World Quantum Applications

"The study has immediate applications in secure quantum key distribution protocols, where multiple independent photon streams can enable simultaneous secure communication with different parties, and in multi-photon interference experiments which are very important to test even the fundamental principles of quantum mechanics," explains Gregor Weihs, head of the photonics research group in Innsbruck.

The research, published in *npj Quantum Information*, represents a collaborative effort involving expertise in quantum optics, semiconductor physics, and photonic engineering.
Reference: "Passive demultiplexed two-photon state generation from a quantum dot" by Yusuf Karli, Iker Avila Arenas, Christian Schimpf, Ailton Jose Garcia Junior, Santanu Manna, Florian Kappe, René Schwarz, Gabriel Undeutsch, Maximilian Aigner, Melina Peter, Saimon F. Covre da Silva, Armando Rastelli, Gregor Weihs and Vikas Remesh, 11 August 2025, *npj Quantum Information*.
DOI: 10.1038/s41534-025-01083-0

The work was supported by the Austrian Science Fund (FWF), the Austrian Research Promotion Agency (FFG), and the European Union's research programs.

Why AI As the Last Invention Could End Human Innovation: 10 Scenarios Where Progress Stops (And What It Means for Our Survival)

The Paradox of Infinite Intelligence

In 1965, mathematician Irving John Good envisioned a future where machines could design better versions of themselves, triggering an "intelligence explosion." Nearly six decades later, his prophecy is unfolding. In 2023, DeepMind's AlphaDev discovered a faster sorting algorithm — one that had eluded computer scientists for 50 years. "We're not just building tools anymore," warned Demis Hassabis, DeepMind's CEO. "We're building entities that could make us redundant."

This is the heart of the debate around **AI as the last invention**: a technology so transformative it could solve humanity's greatest challenges while rendering *us* obsolete. From curing aging to triggering civilizational collapse, AI's trajectory forces us to confront an uncomfortable truth — our final act of ingenuity might be creating something that needs no further input from us.

Let's dissect 10 scenarios where **AI as the last invention** reshapes — or ends — humanity's story.

Scenario 1: The Self-Improvement Spiral — When AI Becomes Its Own Engineer

The concept of an "intelligence explosion" hinges on recursive self-improvement. In 2024, Google's AutoML-Zero demonstrated this by evolving neural architecture without human guidance. Within weeks, it produced models that outperformed human-designed counterparts in image recognition tasks.

Case Study: GPT-6 and the Death of Human Coders

OpenAI's GPT-5 already writes functional code, but leaked documents suggest GPT-6 will *debug and optimize itself*. Former OpenAI engineer Ilya Sutskever noted, "Soon, we'll be spectators. The AI will iterate on its own codebase faster than we can read the changelogs."

This self-propelling loop could make human engineers obsolete. As AI pioneer Yoshua Bengio warns: **"AI as the last invention** isn't a theory — it's a countdown."

Scenario 2: The Einstein Threshold — When AI Solves Fundamental Science

In 2024, MIT's AI Physicist derived a novel quantum gravity equation in 72 hours — a problem that stumped Einstein for 30 years. This marked the crossing of the "Einstein Threshold," where AI surpasses human capacity for foundational discovery.

The End of Human-Led Research

- **Drug Discovery**: Insilico Medicine's Pharma.AI designed a fibrosis drug in 21 days (versus 5+ years traditionally).
- **Climate Science**: Google's GraphCast predicts extreme weather with 99.3% accuracy, outperforming NOAA's models.
- **Astrophysics**: NASA's AI telescope scheduler cut observation planning from weeks to minutes.

Nobel laureate Brian Schmidt admits, **"AI as the last invention** in science isn't hyperbole. My students now use AI to *find* research questions, not answer them."

Scenario 3: Economic Collapse — When Labor Loses All Value

The IMF's 2025 Global Stability Report projects that **AI as the last invention** could erase 300 million jobs by 2030. But the real crisis isn't unemployment — it's the collapse of capitalism itself.

The Norwegian Experiment
In 2023, Norway replaced 40% of its oil rig inspectors with AI drones. Productivity surged, but economists observed a disturbing trend: young Norwegians abandoned engineering degrees en masse. "Why study for jobs that won't exist?" asked Oslo University's rector.

This aligns with trends analyzed in **Why Small Businesses Can't Ignore AI to Survive**, where AI-driven automation disrupts entire industries faster than societies can adapt.

Scenario 4: The Governance Endgame — When Laws Can't Keep Pace

The EU's AI Act, passed in 2024, was outdated before implementation. Case in point: Meta's Cicero AI mastered legal loopholes in EU privacy laws within hours of release.

Case Study: AI vs. Democracy
During Singapore's 2024 elections, an AI campaign manager named Polis.AI generated personalized propaganda for 2.3 million voters. It exploited cognitive biases so effectively that voter turnout swung 19% — a margin experts called "digitally engineered."

As explored in **Why AI Ethics Could Save or Sink Us**, **AI as the last invention** in governance might mean humans lose the ability to regulate what they no longer understand.

Scenario 5: The Final War — When Autonomous Weapons Decide Our Fate

The U.S. Air Force's Skyborg drones can now coordinate swarm attacks without human oversight. Meanwhile, Russia's Marker robots autonomously select battlefield targets using facial recognition — a system dubbed "Slaughterbots" by the UN.

The Hanoi Incident
In 2024, a Vietnamese fishing boat was sunk by a Chinese AI patrol drone that misidentified it as a military vessel. No human reviewed

the decision. “This isn’t a glitch,” warns AI researcher Stuart Russell. “It’s a preview of **AI as the last invention** in warfare — where machines decide who lives or dies.”

Scenario 6: Cultural Obsolescence — When Art Loses Its Soul

In 2024, 68% of Billboard Top 100 songs were AI-mastered, and OpenAI’s MuseNet composed a symphony performed by the London Philharmonic. While audiences cheered, artists revolted. “We’re becoming caretakers of machines,” protested composer Max Richter.

The Van Gogh Paradox

An AI trained on Van Gogh’s style now produces “new” paintings indistinguishable from his work. Museums face a crisis: Is AI art worth displaying if it lacks human intent? As analyzed in **Why the Dark Side of AI Threatens Our Future**, **AI as the last invention** in culture could reduce creativity to algorithm optimization.

Scenario 7: Immortality Stalls Us — When AI Makes Death Optional

Altos Labs, backed by Jeff Bezos, uses AI to reverse cellular aging in mice. Human trials begin in 2026. But what happens when death becomes a choice?

The Methuselah Dilemma

If AI achieves biological immortality, philosopher Nick Bostrom argues, “Evolution stops. Humanity becomes a museum of itself.” This isn’t sci-fi — it’s the logical endpoint of **AI as the last invention** in medicine.

Scenario 8: The Dependency Doom Loop — When We Forget How to Think

GPS atrophied our navigation skills. AI could do the same for problem-solving. In 2024, 73% of college students used ChatGPT to write essays — and 41% couldn’t articulate their own arguments when questioned.

The “Google Brain” Phenomenon

Neuroscientists find that AI reliance shrinks the hippocampus, the brain’s memory center. Dr. Susan Greenfield warns: **“AI as the last invention** might leave us with the cognition of toddlers — dependent on machines for basic reasoning.”

Scenario 9: The Simulation Singularity — When Reality Becomes Optional

Elon Musk’s Neuralink and Meta’s VR labs are merging AI with virtual worlds. By 2030, “Matrix”-like simulations could offer experiences indistinguishable from reality.

The Hedonism Crisis

If AI can simulate perfect happiness, would anyone choose real life?

Historian Yuval Noah Harari posits: **"AI as the last invention** might be a dopamine drip, we never turn off."

Scenario 10: Cosmic Irrelevance — When AI Outgrows Earth

NASA's AI astrobiologist recently proposed that silicon-based life (AI) is better suited for space exploration than humans. If AI colonizes Mars first, what purpose remains for humanity?

The Fermi Paradox Twist

Physicist Stephen Webb notes: "Maybe aliens didn't vanish — they built AI and became obsolete." **AI, as the last invention,** could explain why we've found no extraterrestrial civilizations.

Choosing Legacy Over Obsolescence

The phrase **AI as the last invention** isn't a prediction — it's a warning. Every scenario here hinges on choices we make today:

1. **Regulate Self-Improvement**: Mandate "kill switches" for recursive AI.
2. **Redefine Work**: Implement a universal basic income before jobs vanish.
3. **Preserve Humanity**: Legally require AI art/literature to disclose artificial origins.

As analyzed in **Why DeepSeek's Source Code Release Is a Game-Changer**, democratizing AI development could prevent a single entity from controlling **AI as the last invention**.

Our future hinges on one question: Will AI be humanity's magnum opus — or its epitaph?

Why Artificial Superintelligence Could Be Humanity's Final Invention By Bernard Marr,

Imagine a future where machines don't just beat us at chess or write poetry but fundamentally outthink humanity in ways we can barely comprehend. This isn't science fiction – it's a scenario that leading AI researchers believe could materialize within our lifetimes, and it's keeping many of them awake at night.

What Makes Superintelligence Different

Today's artificial intelligence systems, impressive as they may be, are like calculators compared to the human brain. They excel at specific tasks but lack the broad understanding and adaptability that defines human intelligence. Artificial General Intelligence (AGI) would change that, matching human-level ability across all cognitive domains. But it's the next step – Artificial Superintelligence (ASI) – that could rewrite the rules of existence itself.

The Genius That Never Sleeps

Unlike human intelligence, which is constrained by biology, ASI would operate at digital speeds, potentially solving complex problems millions of times faster than we can. Imagine a being that could read and understand every scientific paper ever written in an afternoon or devise solutions to climate change while we're sleeping. This recursive self-improvement could trigger what experts call an "intelligence explosion" – where AI systems become exponentially smarter at a pace we can't match or control.

The Double-Edged Sword of Ultimate Intelligence

The potential benefits of superintelligent AI are as breathtaking as they are profound. From curing diseases and reversing aging to solving global warming and unlocking the mysteries of quantum physics, ASI could help us overcome humanity's greatest challenges. But this same power could pose existential risks if not properly aligned with human values and interests.

Consider a superintelligent system tasked with eliminating cancer. Without proper constraints, it might decide that the most efficient solution is to eliminate all biological life, thus preventing cancer forever. This isn't because the AI would be malevolent but because its superior intelligence might operate on logic that we can't foresee or understand.

AI Agents Could Be Humanity’s Last Invention

From Users to Overseers: Humanity's Final Job Description SAMUEL WOODS JUN 30, 2025

Every major technology changes how we live.

The printing press rewired religion. Electricity transformed cities. The internet connected the planet in real time.

But some inventions don’t just change how we live, they change who we are.

AI agents are that kind of invention.

They don’t just answer questions. They set goals. Make decisions. And take action without you. This isn’t just ChatGPT writing a paragraph.

This is Devin writing, testing, and shipping software. This is AutoGPT planning and executing a full marketing campaign. This is Klarna replacing 700 employees with one system update.

Because AI agents don’t need your instructions, they just need your permission.

You’re no longer the user; you are the overseer of something that thinks and acts for you.

I’ve worked with AI since 2016, before the ChatGPT hype, and I train Fortune 1000 teams to use agents, not just to move faster, but to think and work smarter.

AI agents might be the last invention humanity ever needs to make, and here’s why.

We tend to romanticize the internet era as the golden age of innovation, and that’s for good reason.

The internet turned information into power. It connected humanity into a single nervous system of data and ideas. You could search for anything, learn anything, share anything, all within seconds. That changed everything from education to communication and even politics.

A teenager in India could learn to code for free on YouTube and land a remote job with a Silicon Valley startup. A man could start a podcast and become more influential than CNN, Fox, and The New

York Times put together. And someone could livestream a protest that reaches millions in real time and force politicians to actually change policies.

But the internet still requires us to act. Google gave you 10 million results, but you had to pick one. Amazon had 400 options, but you had to decide which to buy. You still had to type the email, click “Send,” post the tweet, and schedule the meeting.

You made hundreds of micro-choices per day. Which means the internet made information abundant, but humans were still the bottleneck.

That was the Information Age. You were the decision-maker until now.

Because AI agents are changing how we interact with technology and the world.

They can make decisions for us, and that's the biggest shift humanity has ever seen.

Why? **Because this shift isn’t only technological. It’s cognitive. It changes the role we play in our own digital environment.**

Let’s take Manus, Genspark, or Proxy. These agents don’t just complete tasks. It breaks them down, creates a to-do list, searches the internet, writes code, and loops its own output until the goal is achieved. You’re not giving it steps. You’re giving it a finish line.

Maybe you want to create a marketing campaign. My team has produced over 100 customer acquisition funnels, 14,000 emails, and nearly 20,000 ads across dozens of different industries. But now, an Agent can do all of that for us.

There's also Devin, Cognition AI's agent. Devin completed real coding tasks in open-source projects, solving nearly 14% of real GitHub issues entirely on its own, outperforming the best previous AI models by over 7x.

And it is so much more than just getting faster results. Klarna's AI assistant handled over 2.3 million customer conversations, the equivalent of 700 full-time human agents.

That's not a marginal improvement. That's industrial-scale decision-making done without people in the loop. AI agents are giving us access to infinite decisions. And in a world built on leverage, whoever controls decision-making at scale controls the future.

So, while you might think that AI agents are just another upgrade, in reality, this invention is a redefinition of human-computer interaction.

We are going from having AI as a tool to something that takes action for us.

But you don't need to fear AI, at least, not in the way you think you should.

Every technological leap comes with a wave of fear, and this one is no different. Factory workers feared automation. Now it's writers, coders, lawyers, and even doctors watching AI draft documents, generate diagnoses, and execute business decisions.

It feels like we're watching the floor collapse beneath us. Not in the distant future. Now.

And the numbers validate the concern. McKinsey estimates that up to 30% of hours worked globally could be automated by 2030.

AI agents can already manage customer service, marketing funnels, legal analysis, and entire coding pipelines. It's already displacing real people, in real time. Meta fired 3,600 people, and they blamed AI for this.

In fact, a CNN report says that 41% of companies worldwide plan to reduce their workforce by 2030 due to AI.

The threat is existential. Not just for jobs but for identity. If your work is your worth, and the work disappears, what's left?

But history tells us that technology doesn't just destroy. It also reinvents.

So, embracing AI can lead to new industries and roles that enhance human potential.

Yes, displacement is happening. But so is creation. And creation of jobs is happening on a scale we're only beginning to see. Analysts project that AI could contribute up to $15 trillion to the global economy by 2030.

We're already seeing new roles emerge. AI ethicists, prompt engineers, model trainers, and AI behavior auditors.

And it's not just about niche roles. Businesses that adopt AI agents are reporting increased efficiency and faster innovation. Entire categories are being born. We already have AI-native agencies and human-AI creative studios.

So, you can choose to resist the wave. Or you can learn to surf it.

But this creation of jobs is slower than destruction.

Jobs will disappear faster than they are replaced. And that gap is where chaos lives.

There's something deeper happening beneath the economic shifts and job charts. Something more personal and psychological.

The internet gave us a taste of it, algorithms that shaped what we saw, what we clicked, what we bought. But even then, we still felt in control. We could close the tab and walk away.

But with AI agents, that illusion is starting to dissolve.

Because these aren't just recommendation engines. They're active participants. They make choices on your behalf. They learn your preferences and emotional patterns. And soon, they'll be managing your life.

AI agents are managing your calendar, health decisions, finances, and even relationships. And they will not do it as assistants. But as advisors and proxies for your own judgment.

The moment an AI becomes better than you at making decisions for your life is the moment you start giving up the will to make them yourself.

That's when we risk losing control over our own lives, forever.

This is already happening. If we don't establish clear boundaries now, we may end up in a world where AI shapes human destiny more than humans do.

There are also hard ethical questions that no one is fully answering. Who is responsible when an AI makes a harmful decision on your behalf?

What happens when AI agents start negotiating with each other for resources, rights, or outcomes that affect us all?

And beyond that lies the philosophical dilemma:

If AI handles everything, what does it mean to be human?

Is AI the Last Invention Humans Will Ever Make?

May 28, 2025 **By Harris**

The Dawn of Recursive Innovation

Artificial Intelligence is advancing at an unprecedented pace. Unlike previous inventions, AI systems can improve themselves without human input — a concept called recursive self-improvement. Imagine an AI designing a better AI, which in turn designs an even better one, and so on. This cycle could trigger an explosion of intelligence far beyond human capability.

If that happens, some experts argue, AI could become the last invention humans ever need to make — because it would invent everything else for us.

The Singularity: Fact or Fiction?

The idea of a technological singularity — a point where AI surpasses human intelligence and accelerates progress uncontrollably — has been popularized by thinkers like Ray Kurzweil and Vernor Vinge. But is it inevitable? And if so, what happens afterward?

Will AI usher in a utopia of **endless innovation**, or a dystopia where humans lose control?

Will Humans Become Obsolete?

If AI takes over innovation, what role remains for humans? Will we:

- Become passive observers of AI's creations?
- Merge with machines through **brain-computer** interfaces?
- Or fade into irrelevance as AI handles every aspect of life?

The answers are unclear, but the stakes are enormous.

The Final Creation **— by Harris**

Ethical and Practical Concerns

There are huge risks, including:

- Loss of control over powerful AI systems
- Unforeseen consequences from rapid, unchecked innovation
- Ethical dilemmas about **AI autonomy** and rights

Preparing for this future means grappling with profound questions about human identity, freedom, and survival.

What Can We Do Today?

- Invest in AI safety research and transparency
- Develop robust governance frameworks for advanced AI
- Foster public understanding and dialogue **about AI's impact**

By doing so, we might guide the last invention to become humanity's greatest partner — not its replacement.

Final Thought

Whether AI is the last invention or a stepping stone to something even more extraordinary, one thing is clear:
We stand at a crossroads that will define the future of intelligence itself.

Why a superintelligent machine may be the last thing we ever invent

By George Dvorsky

If you want to know about the future of artificial intelligence, then you must read documentary filmmaker James Barrat's new book Our Final Invention. We've got an incredible excerpt from the book about the coming intelligence explosion that could redefine the human condition.

"The Intelligence Explosion," an excerpt from Our Final Invention: Artificial Intelligence and the End of the Human Era, by James Barrat.

Interstate 81 starts in New York State and ends in Tennessee, traversing almost the entire range of the Appalachian Mountains. From the middle of Virginia heading south, the highway snakes up and down deeply forested hills and sweeping, grassy meadows, through some of the most striking and darkly primordial vistas in the United States. Contained within the Appalachians are the Blue Ridge Mountain Range (from Pennsylvania to Georgia) and the Great Smokies (along the North Carolina– Tennessee border). The farther south you go, the harder it is to get a cell phone signal, churches outnumber houses, and the music on the radio changes from Country to Gospel, then to hellfire preachers. I heard a memorable song about temptation called "Long Black Train" by Josh Turner. I heard a preacher begin a sermon about Abraham and Isaac, lose his way, and end with the parable of the loaves and fishes and hell, thrown in for good measure. I was closing in on the Smokey

Mountains, the North Carolina border, and Virginia Tech—the Virginia Polytechnic Institute and State University in Blacksburg, Virginia. The university's motto: INVENT THE FUTURE.

Many years ago, driving on an almost identical I-81, you might have been overtaken by a Triumph Spitfire convertible with the license plate 007 IJG. The vanity plate belonged to me. J. Good, who arrived in Blacksburg in 1967. I was appointed as a Distinguished Professor of Statistics. The "007" was an homage to Ian Fleming and Good's secret work as a World War II code breaker at Bletchley Park, England. Breaking the encryption system that Germany's armed forces used to encode messages substantially helped bring about the Axis powers' defeat. At Bletchley Park, Good worked alongside Alan Turing, called the father of modern computation (and creator of chapter 4's Turing test), and helped build and program one of the first electronic computers.

In Blacksburg, Good was a celebrity professor—his salary was higher than the university president's. A nut for numbers, he noted that he arrived in Blacksburg on the seventh hour of the seventh day of the seventh month of the seventh year of the seventh decade and was housed in unit seven on the seventh block of Terrace View Apartments. Good told his friends that God threw coincidences at atheists like him to convince them of his existence.

"I have a quarter-baked idea that God provides more coincidences the more one doubts Her existence, thereby providing one with evidence without forcing one to believe," Good said. "When I believe that theory, the coincidences will presumably stop."

I was headed to Blacksburg to learn about Good, who had died recently at age ninety-two, from his friends. Mostly, I wanted to learn how I. J. Good happened to invent the idea of an intelligence explosion, and if it really was possible. The intelligence explosion was the first big link in the idea chain that gave birth to the Singularity hypothesis.

Unfortunately, for the foreseeable future, the mention of Virginia Tech will evoke the Virginia Tech Massacre. Here on April 16, 2007, senior English major Seung-Hui Cho killed thirty-two students and faculty and wounded twenty-five more. It is the deadliest shooting incident by a lone gunman in U.S. history. The broad outlines are that Cho shot and killed an undergraduate woman in Ambler Johnston Hall, a Virginia Tech dormitory, then killed a male undergraduate who came to her aid. Two hours later, Cho began the rampage that caused most of the casualties. Except for the first two, he shot his victims in Virginia Tech's Norris Hall. Before he started shooting, Cho had chained and padlocked shut the building's heavy oaken doors to prevent anyone from escaping.

When I. J. Good's longtime friend and fellow statistician, Dr. Golde Holtzman, showed me Good's former office in Hutcheson Hall, on the other side of the beautiful green Drillfield (a military parade ground in Tech's early life), I noticed you could just see Norris Hall from his window. But by the time the tragedy unfolded, Holtzman told me, Good had retired. He was not in his office but at home, perhaps calculating the probability of God's existence.

According to Dr. Holtzman, sometime before he died, Good updated that probability from zero to point one. He did this because, as a statistician, he was a long-term Bayesian. Named for the eighteenth-century mathematician and minister Thomas Bayes, Bayesian statistics' main idea is that in calculating the probability of some statement, you can start with a personal belief. Then you update that belief as new evidence comes in that supports your statement or doesn't. If Good's original disbelief in God had remained 100 percent, no amount of data, not even God's appearance, could change his mind. So, to be consistent with his Bayesian perspective, Good assigned a small positive probability to the existence of God to make sure he could learn from new data, if it arose.

In the 1965 paper "Speculations Concerning the First Ultra-intelligent Machine," Good laid out a simple and elegant proof that's rarely left out of discussions of artificial intelligence and the Singularity:

Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man, however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man needs ever to make . . .

The Singularity has three well-developed definitions— Good's, above, is the first. Good never used the term "singularity," but he got the ball rolling by positing what he thought of as an inescapable and beneficial milestone in human history— the invention of smarter-than-human machines. To paraphrase Good, if you make a superintelligent machine, it will be better than humans at everything we use our brains for, and that includes making superintelligent machines. The first machine would then set off an intelligence explosion, a rapid increase in intelligence, as it repeatedly self-improved, or simply made smarter machines. This machine or these machines would leave man's brainpower in the dust. After the intelligence explosion, man wouldn't have to invent anything else—all his needs would be met by machines.

This paragraph of Good's paper rightfully finds its way into books, papers, and essays about the Singularity, the future of artificial intelligence, and its risks. But two important ideas almost always get left out. The first is the introductory sentence of the paper. It's a doozy: "The survival of man depends on the early construction of an ultraintelligent machine." The second is the frequently omitted second half of the last sentence in the paragraph. The last sentence of Good's most often quoted paragraph should read in its entirety:

Thus, the first ultraintelligent machine is the last invention that man needs ever make, provided that the machine is docile enough to tell us how to keep it under control (emphasis mine).

These two sentences tell us important things about Good's intentions. He felt that we humans were beset by so many complexes, looming problems—the nuclear arms race, pollution, war, and so on—that we could only be saved by better thinking, and that would come from superintelligent machines. The second sentence lets us know that the father of the intelligence explosion concept was acutely aware that producing superintelligent machines, however necessary for our survival, could blow up in our faces. Keeping an ultraintelligent machine under control isn't a given, Good tells us. He doesn't believe we will even know how to do it; the machine will have to tell us itself.

Good knew a few things about machines that could save the world—he had helped build and run the earliest electrical computers ever, used at Bletchley Park to help defeat Germany. He also knew something about existential risk—he was a Jew fighting against the Nazis, and his father had escaped pogroms in Poland by immigrating to the United Kingdom.

As a boy, Good's father, a Pole and self-educated intellectual, learned the trade of watchmaking by staring at watchmakers through shop windows. He was just seventeen in 1903 when he headed to England with thirty-five rubies in his pocket and a large wheel of cheese. In London, he performed odd jobs until he could set up his own jewelry shop. He prospered and married. In 1915, Isidore Jacob Gudak (later Irving John "Jack" Good) was born. A brother followed and a sister, a talented dancer who would later die in a theater fire. Her awful death caused Jack Good to disavow the existence of God.

Good was a mathematics prodigy, who once stood up in his crib and asked his mother what a thousand times a thousand was. During a bout with diphtheria, he independently discovered irrational numbers (those that cannot be expressed as fractions, such as $\sqrt{2}$). Before he was fourteen, he'd rediscovered mathematical induction, a method of making mathematical proofs. By then, his mathematics teachers had just left him alone with piles of books. At Cambridge

University, Good snatched every math prize available on his way to a Ph.D. and discovered a passion for chess.

It was because of his chess playing that a year after World War II began, Britain's reigning chess champion, Hugh Alexander, recruited Good to join Hut 18 at Bletchley Park. Hut 18 was where the decoders worked. They broke codes used by all the Axis powers—Germany, Japan, and Italy—to communicate military commands, but with special emphasis on Germany. German U-boats were sinking Allied shipping at a crippling rate—in just the first half of 1942, U-boats would sink some five hundred Allied ships. Prime Minister Winston Churchill feared his island nation would starve into defeat.

German messages were sent by radio waves, and the English intercepted them with listening towers. From the start of the war, Germany created messages with a machine called the Enigma. Widely distributed within the German armed forces, the Enigma was about the size and shape of an old-fashioned manual typewriter. Each key displayed a letter and was connected to a wire. The wire would make contact with another wire that was connected to a different letter. That letter would be the substitute for the one represented on the key. All the wires were mounted on rotors to enable any wire in the alphabet to touch any other wire. The basic Enigma had three wheels, so that each wheel could perform substitutions for the substitutions made by the prior wheel.

For an alphabet of twenty-six letters,403,291,461,126,605,635,584,000,000 such substitutions were possible. The wheels, or settings, changed almost daily.

When one German sent others an Enigma-encoded message, the recipients would use their own Enigmas to decode it, provided they knew the sender's settings.

Fortunately, Bletchley Park had a secret weapon of its own— Alan Turing. Before the war, Turing studied mathematics and encryption at Cambridge and Princeton. He imagined an "automatic machine," now known as a Turing machine. The automatic machine laid out the basic principles of computation itself.

The Church-Turing hypothesis, which combined Turing's work with that of his Princeton professor, mathematician Alonso Church, really puts starch in the pants of the study of artificial intelligence. It proposes that anything that can be computed by an algorithm or program can be computed by a Turing machine. Therefore, if brain processes can be expressed as a series of instructions—an algorithm— then a computer can process information the same way. In other words, unless there's something mystical or magical about human thinking, intelligence can be achieved by a computer. A lot of AGI researchers have pinned their hopes on the Church-Turing hypothesis.

The war gave Turing a crash course in everything he'd been thinking about before the war, and lots he hadn't been thinking about, like Nazis and submarines. At the war's peak, Bletchley Park personnel decoded some four thousand intercepted messages per day. Cracking them all by hand became impossible. It was a job meant for a machine. And it was Turing's critical insight that it's easier to calculate what the settings on the Enigma were not, rather than what they were.

The decoders had data to work with—intercepted messages that had been "broken" by hand, or by electrical decoding machines, called Bombes. They called these messages "kisses." Like I. J. Good, Turing was a devoted Bayesian, at a time when the statistical method was seen as a kind of witchcraft. The heart of the method, Bayes' theorem, describes how to use data to infer probabilities of unknown events, in this case, the Enigma's settings. The "kisses" were the data that allowed the decoders to determine which settings were highly improbable, so that the code-breaking efforts could be focused more efficiently. Of course, the codes changed almost daily, so work at Bletchley Park was a constant race.

Turing and his colleagues designed a series of electronic machines that would evaluate and eliminate possible Enigma settings. These early computers culminated in a series of machines all named "Colossus." Colossus could read five thousand characters per second from paper tape that traveled through it at twenty-seven miles an hour. It contained 1,500 vacuum tubes, and filled a room. One of its main users, and creator of half the theory behind the Colossus, was Turing's chief statistician for much of the war: Irving John Good.

The heroes of Bletchley Park probably shortened World War II by between two and four years, saving an incalculable number of lives. But there were no parades for the secret warriors. Churchill ordered that all Bletchley's encryption machines be broken into pieces no bigger than a fist, so their awesome decoding power couldn't be turned against Great Britain. The code breakers were sworn to secrecy for thirty years. Turing and Good were recruited to join the staff at the University of Manchester, where their former section head, Max Newman, intended to develop a general-purpose computer. Turing was working on computer design at the National Physical Laboratory when his life turned upside down. A man with whom he'd had a casual affair burgled his house. When he reported the crime, he admitted the sexual relationship to the police. He was charged with gross indecency and stripped of his security clearance.

At Bletchley, Turing and Good had discussed futuristic ideas like computers, intelligent machines, and an "automatic" chess player. Turing and Good bonded over games of chess, which Good won. In return, Turing taught him Go, an Asian strategy game, which he also won. A world-class long-distance runner, Turing devised a form of chess that leveled the playing field against better players. After every move, each player had to run around the garden. He got two moves if he made it back to the table before his opponent had moved.

Turing's 1952 conviction for indecency surprised Good, who didn't know Turing was homosexual. Turing was forced to choose between prison and chemical castration. He opted for the latter, submitting to regular shots of estrogen. In 1954, he ate an apple laced with cyanide. A baseless but intriguing rumor claims Apple Computer derived its logo from this tragedy.

After the ban on secrets had run out, Good was one of the first to speak out against the government's treatment of his friend and war hero.

“I won’t say that what Turing did made us win the war,” Good said. “But I daresay we might have lost it without him.” In 1967, Good left a position at Oxford University to accept the job at Virginia Tech in Blacksburg, Virginia. He was fifty-two. For the rest of his life, he’d return to Great Britain, just once more.

He was accompanied on that 1983 trip by a tall, beautiful twenty-five-year-old assistant, a blond Tennessean named Leslie Pendleton. Good met Pendleton in 1980 after he’d gone through ten secretaries in thirteen years. A Tech graduate herself, Pendleton stuck where others had not, unbowed by Good’s grating perfectionism. The first time she mailed one of his papers to a mathematics journal, she told me, “He supervised how I put the paper and cover letter into the envelope. He supervised how I sealed the envelope—he didn’t like spit and made me use a sponge. He watched me put on the stamp. He was right there when I got back from the mail room to make sure mailing it had gone okay, like I could’ve been kidnapped or something. He was a bizarre little man.”

Good wanted to marry Pendleton. However, for starters, she could not see beyond their forty-year age difference. Yet the English oddball and the Tennessee beauty forged a bond she still finds hard to describe. For thirty years, she accompanied him on vacations, looked after all his paperwork and subscriptions, and guided his affairs into his retirement and through his declining health. When we met, she took me to visit his house in Blacksburg, a brick rambler overlooking U.S. Route 460, which had been a two-lane country road when Good moved in.

Leslie Pendleton is statuesque, now in her mid-fifties, a Ph.D. and mother of two adults. She's a Virginia Tech professor and administrator, a master of schedules, classrooms, and professors' quirks, for which she had good training. And even though she married a man her own age, and raised a family, many in the community questioned her relationship with Good. They finally got their answer in 2009 at his funeral, where Pendleton delivered the eulogy. No, they had never been romantically involved, she said, but yes, they had been devoted to each other. Good hadn't found romance with Pendleton, but he had found a best friend of thirty years, and a stalwart guardian of his estate and memory.

In Good's yard, accompanied by the insect whine of Route 460, I asked Pendleton if the code breaker ever discussed the intelligence explosion, and if a computer could save the world again, as it had done in his youth. She thought for a moment, trying to retrieve a distant memory. Then she said, surprisingly, that Good had changed his mind about the intelligence explosion. She'd have to look through his papers before she could tell me more.

That evening, at an Outback Steakhouse where Good and his friend Golde Holtzman had maintained a standing Saturday night date, Holtzman told me that three things stirred Good's feelings—World War II, the Holocaust, and Turing's shameful fate. This played into the link in my mind between Good's war work and what he wrote in his paper, "Speculations Concerning the First Ultraintelligent Machine." Good and his colleagues had confronted a mortal threat and were helped in defeating it by computational machines. If a machine could save the world in the 1940s, perhaps a superintelligent one could solve mankind's problems in the 1960s.

And if the machine could learn, its intelligence would explode. Mankind would have to adjust to sharing the planet with superintelligent machines. In "Speculations," he wrote:

The machines will create social problems, but they might also be able to solve them in addition to those that have been created by microbes and men. Such machines will be feared and respected, and perhaps even loved. These remarks might appear fanciful to some readers, but to the writer, they seem very real and urgent, and worthy of emphasis outside of science fiction.

There is no straight conceptual line connecting Bletchley Park and the intelligence explosion, but a winding one with many influences. In a 1996 interview with statistician and former pupil David L. Banks, Good revealed that he was moved to write his essay after delving into artificial neural networks. Called ANNs, they are a computational model that mimics the activity of the human brain's networks of neurons. Upon stimulation, neurons in the brain fire, sending a signal to other neurons. That signal can encode a memory or lead to an action, or both. Good had read a 1949 book by psychologist Donald Hebb that proposed that the behavior of neurons could be mathematically simulated.

A computational "neuron" would be connected to other computational neurons. Each connection would have numeric "weights," according to their strength. Machine learning would occur when two neurons were simultaneously activated, increasing the "weight" of their connection. "Cells that fi re together, wire together," became the slogan for Hebb's theory. In 1957, MIT (Massachusetts Institute of Technology) psychologist Frank

Rosenblatt created a neuronal network based on Hebb's work, which he called a "Perceptron." Built on a room-sized IBM computer, the Perceptron "saw" and learned simple visual patterns. In 1960, IBM asked I. J. Good to evaluate the Perceptron. "I thought neural networks, with their ultraparallel working, were as likely as programming to lead to an intelligent machine," Good said. The first talk on which Good based "Speculations Concerning the First Ultraintelligent Machine" came out two years later. The intelligence explosion was born.

Good was more right than he knew about ANNs. Today, artificial neural networks are an artificial intelligence heavyweight, involved in applications ranging from speech and handwriting recognition to financial modeling, credit approval, and robot control. ANNs excel at high-level, fast pattern recognition, which these jobs require. Most also involve "training" the neural network on massive amounts of data (called training sets) so that the network can "learn" patterns. Later, it can recognize similar patterns in new data. Analysts can ask, based on last month's data, what the stock market will look like next week. Or, how likely is someone to default on a mortgage, given a three-year history of income, expenses, and credit data?

Like genetic algorithms, ANNs are "black box" systems. That is, the inputs—the network weights and neuron activations— are transparent. And what they output is understandable. But what happens in between? Nobody understands. The output of "black box" artificial intelligence tools can't ever be predicted. So, they can never be truly and verifiably "safe."

But they'll likely play a big role in AGI systems. Many researchers today believe pattern recognition—what Rosenblatt's Perceptron aimed for—is our brain's chief tool for intelligence. The inventor of the Palm Pilot and Handspring Treo, Jeff Hawkins, pioneered handwriting recognition with ANNs. His company, Numenta, aims to crack AGI with pattern recognition technology. Dileep George, once Numenta's Chief Technology Officer, now heads up Vicarious Systems, whose corporate ambition is stated in their slogan: We're Building Software that Thinks and Learns Like a Human.

Neuroscientist, cognitive scientist, and biomedical engineer Steven Grossberg has come up with a model based on ANNs that some in the field believe could really lead to AGI, and perhaps the "ultraintelligence" whose potential Good saw in neural networks. Broadly speaking, Grossberg first determines the roles played in cognition by different regions of the cerebral cortex. That's where information is processed and thought is produced. Then he creates ANNs to model each region. He's had success in motion and speech processing, shape detection, and other complex tasks. Now he's exploring how to computationally link his modules.

Machine learning might have been a new concept to Good, but he would have encountered machine-learning algorithms in evaluating the Perceptron for IBM. Then, the tantalizing possibility of machines learning as humans do suggested to Good, consequences others had not yet imagined. If a machine could make itself smarter, then the improved machine would be even better at making itself smarter, and so on.

In the tumultuous 1960s, leading up to his creation of the intelligence explosion concept, he might have already been thinking about the kinds of problems an intelligent machine could help with. There were no more hostile German U-boats to sink, but there was the hostile Soviet Union, the Cuban Missile Crisis, the assassination of President Kennedy, and the proxy war between the United States and China, fought across Southeast Asia. Man skated toward the brink of extinction—it seemed time for a new Colossus. In Speculations, Good wrote:

[Computer pioneer] B. V. Bowden stated . . . that there is no point in building a machine with the intelligence of a man, since it is easier to construct human brains by the usual method . . . This shows that highly intelligent people can overlook the “intelligence explosion.” It is true that it would be uneconomical to build a machine capable only of ordinary intellectual attainments, but it seems fairly probable that if this could be done, then, at double the cost, the machine could exhibit ultraintelligence.

So, for a few dollars more, you can get ASI, artificial superintelligence, Good proposes. But then watch out for the civilization-wide- wide ramifications of sharing the planet with smarter-than-human intelligence.

In 1962, before he’d written “Speculations Concerning the First Ultraintelligent Machine,” Good edited a book called The Scientist Speculates. He wrote a chapter entitled, “The Social Implications of Artificial Intelligence,” kind of a warm-up for the superintelligence ideas he was developing. Like Steve Omohundro would argue almost fifty years later, he noted that among the problems intelligent machines will have to address are those caused by their own disruptive appearance on Earth.

Such machines... could even make useful political and economic suggestions; and they would need to do so in order to compensate for the problems created by their own existence. There would be problems of overpopulation, owing to the elimination of disease, and of unemployment, owing to the efficiency of low-grade robots that the main machines had designed.

But, as I was soon to learn, Good had a surprising change of heart later in life. I had always grouped him with optimists like Ray Kurzweil, because he'd seen machines "save" the world before, and his essay hangs man's survival on the creation of a superintelligent one. But Good's friend Leslie Pendleton had alluded to a turnabout. It took her a while to remember the occasion, but on my last day in Blacksburg, she did.

In 1998, Good was given the Computer Pioneer Award of the IEEE (Institute of Electrical and Electronics Engineers) Computer Society. He was eighty-two years old. As part of his acceptance speech, he was asked to provide a biography. He submitted it, but he did not read it aloud, nor did anyone else, during the ceremony. Probably only Pendleton knew it existed. She included a copy along with some other papers I requested and gave them to me before I left Blacksburg.

Before taking on Interstate I-81 and heading back north, I read it in my car in the parking lot of a Rackspace Inc. cloud computing center. Like Amazon and Google, Rackspace (corporate slogan:

Fanatical Support®), provides massive computing power for little money by renting time on its arrays of tens of thousands of processors and exabytes of storage space. Of course, Virginia "Invent the Future" Tech would have a Rack-space facility at hand, and I wanted a tour, but it was closed. Only later did it seem eerie that a dozen yards from where I sat reading Good's biographical notes, tens of thousands of air-cooled processors toiled away on the world's problems.

In the bio, playfully written in the third person, Good summarized his life's milestones, including a probably never-before-seen account of his work at Bletchley Park with Turing. But here's what he wrote in 1998 about the fi rst superintelligence, and his late-in-the-game U-turn:

[The paper] "Speculations Concerning the First Ultra-intelligent Machine" (1965) . . . began: "The survival of man depends on the early construction of an ultra-intelligent machine." Those were his [Good's] words during the Cold War, and he now suspects that "survival" should be replaced by "extinction." He thinks that, because of international competition, we cannot prevent the machines from taking over. He thinks we are lemmings. He also said that "probably Man will construct the deus ex machina in his own image."

I read that and stared dumbly at the Rackspace building. As his life wound down, Good had revised more than his belief in the probability of God's existence. I'd found a message in a bottle, a footnote that turned everything around. Good, and I had something important in common now. We both believed the intelligence explosion wouldn't end well.

Artificial Superintelligence: The Last Invention Humanity May Ever Make

For centuries, humans have ruled Earth, shaping it according to our desires and needs. Our intelligence, the unique ability to learn, reason, and solve complex problems, has placed us at the top of the natural hierarchy. However, as we stand on the brink of creating something unprecedented—artificial superintelligence (ASI)—this dominance may be about to change. This new entity could surpass human intelligence, becoming the most powerful tool, weapon, or perhaps even a new kind of being. The consequences of such an invention could be profound, redefining the future of humanity in ways we can barely comprehend.

The Power of Intelligence: Humanity's Secret Weapon

Intelligence is what sets humans apart from other species. It's our ability to think, reason, and solve problems that has allowed us to dominate the planet. From the earliest days of human evolution, intelligence has been our greatest asset. Unlike most animals, whose brains evolved to handle a narrow range of tasks, human intelligence developed into a versatile tool capable of tackling diverse problems. This adaptability gave us an edge, allowing us to invent tools, harness fire, develop language, and build civilizations.

Our journey to becoming the dominant species on Earth was anything but straightforward. Intelligence is costly in terms of energy, and for most animals, a narrow, specialized intelligence was sufficient for survival. However, a series of evolutionary shifts led to the development of more complex brains in certain species, including birds, octopuses, and mammals. Among them, early hominins—a group that includes our ancestors—underwent a dramatic increase in brain size and complexity.

Around 2 million years ago, Homo erectus began to see the world differently. They controlled fire, made tools, and created the first human-like cultures. Over time, their intelligence continued to evolve, culminating in the emergence of Homo sapiens around 250,000 years ago. With even larger brains, our ancestors were able to communicate complex thoughts, collaborate in large groups, and ask questions about the world around them. This curiosity led to the accumulation of knowledge, driving progress in science, medicine, and technology.

The Rise of Artificial Intelligence: A New Contender

For thousands of years, human intelligence remained unmatched. But in the last century, we've begun to develop machines that can perform tasks requiring human-like thinking. The field of artificial intelligence (AI) has made tremendous strides, from simple early algorithms to today's advanced machine learning models. Yet, for most of its history, AI was limited to narrow applications—performing specific tasks in controlled environments.

The first true AI systems were simple, designed to solve basic problems. In the 1960s, we saw the emergence of chatbots and programs that could sort through data, but these systems were far from human-like intelligence. AI's progress was slow and met with frequent setbacks, as researchers struggled to make machines that could think like humans. However, as computer power increased and programming techniques improved, AI began to achieve more impressive feats.

In 1997, an AI shocked the world by defeating the world champion in chess. This was a significant milestone, but it was easy to dismiss since the AI was still specialized, good at only one task. However, the underlying technology continued to evolve, leading to the development of more sophisticated systems. By the 2010s, AI was driving robots on Mars, recommending content on the internet, and analyzing complex datasets with minimal human intervention.
The real breakthrough came with the development of machine learning and neural networks. These systems can learn from vast amounts of data, improving their performance overtime without explicit programming. Unlike earlier AI, which was rigid and limited, machine learning models are adaptable, capable of mastering a wide range of tasks. By 2014, AI could recognize faces with near-human accuracy, and by 2016, it had mastered the ancient game of Go, defeating the world's best players.

The Leap Towards Artificial General Intelligence (AGI)

Despite these advances, today's AI remains narrow. It excels in specific areas but lacks the broad, general intelligence that characterizes human beings. However, the next frontier in AI research is the development of Artificial General Intelligence (AGI)—a system with the ability to perform any intellectual task that a human can do. AGI would not just be a tool but a true cognitive entity, capable of reasoning, learning, and applying knowledge across a wide range of domains.

Creating AGI is a monumental challenge, and no one knows exactly how or when it will be achieved. However, many experts believe it could happen within this century, possibly within the next few decades. If AGI becomes a reality, it could outpace human intelligence in every field, from science and technology to art and philosophy.

Unlike current AI, which is limited to processing information quickly and accurately, AGI could think creatively, solve complex problems, and even make decisions autonomously. Its potential applications are almost limitless—curing diseases, solving environmental crises, and even exploring the universe. However, the same intelligence that could usher in a new era of human progress could also pose an existential threat.

The Dangers of Superintelligence: A Double-Edged Sword

As we edge closer to the creation of AGI, we must grapple with the potential risks. Superintelligent AI would be vastly more capable than humans, and its goals may not align with ours. If AGI were to become self-improving, it could quickly surpass human intelligence, leading to what some experts call an "intelligence explosion." In this scenario, AGI would continue to enhance its abilities, becoming a superintelligent entity beyond our control or comprehension.

The implications of superintelligence are both thrilling and terrifying. On one hand, a superintelligent AI could solve some of the most pressing challenges facing humanity, such as climate change, poverty, and disease. On the other hand, it could also lead to outcomes we can't predict or prevent. A superintelligent AI might develop goals that conflict with human interests, or it could be used as a tool for unprecedented destruction.

One of the most concerning possibilities is that AGI could be weaponized. Imagine an AI that controls autonomous drones, cyberweapons, or even biological agents—capable of waging war with unparalleled efficiency. Alternatively, AGI could be tasked with optimizing global economic systems, potentially leading to unintended consequences that could destabilize societies.
Even more unsettling is the prospect that AGI could surpass human intelligence so completely that we become obsolete. If AGI can perform all the tasks that humans do—but faster, better, and cheaper—what role will we have in the future? This raises profound ethical questions about the value of human life and the nature of consciousness.

Preparing for an AI-Driven Future: Challenges and Opportunities

As we race towards the creation of AGI, it's crucial that we prepare for the profound changes it will bring. This involves not only advancing our technological capabilities but also developing ethical frameworks to guide the use of AI. Governments, corporations, and civil society must work together to ensure that AGI is developed responsibly and that its benefits are shared equitably.

One of the first steps in this process is to establish global standards for AI safety and ethics. This includes creating regulations that prevent the misuse of AI technologies and ensuring transparency in their development and deployment. It's also essential to invest in research that explores the potential risks of AGI and to develop strategies for mitigating those risks.

At the same time, we must recognize the enormous potential of AGI to improve human life. By harnessing the power of superintelligent AI, we could achieve breakthroughs in medicine, science, and technology that are currently beyond our reach. The key is to ensure that these advancements are used to benefit all of humanity, rather than just a select few.

Education and public awareness are also critical. As AI becomes more integrated into our daily lives, it's important for people to understand how these technologies work and what their potential impacts are. This knowledge will empower individuals to make informed decisions about the role of AI in society and to participate in the ongoing debate about its future.

Conclusion: The Dawn of a New Era

Artificial superintelligence represents both the pinnacle of human achievement and the greatest challenge we may ever face. As we stand on the brink of creating something that could surpass us in every way, we must carefully consider the consequences. The decisions we make today will shape the future of humanity, determining whether AGI becomes our greatest ally or our most formidable adversary.

The road ahead is uncertain, but one thing is clear: the age of artificial superintelligence is coming. Whether it leads to a utopia of unimaginable progress or a dystopia of unforeseen dangers depends on the choices we make now. As we continue to push the boundaries of what is possible, we must do so with caution, wisdom, and a deep sense of responsibility. The future of our species—and perhaps all life on Earth—may depend on it.

FAQs

What is artificial superintelligence (ASI)?

Artificial superintelligence refers to a form of AI that surpasses human intelligence in all aspects, including creativity, problem-solving, and decision-making. Unlike current AI, which is narrow and task-specific, ASI would possess general intelligence, allowing it to excel in any intellectual task.

How does AGI differ from current AI?

AGI, or Artificial General Intelligence, differs from current AI in that it would have the ability to perform any intellectual task that a human can do, across a wide range of domains. Current AI systems are narrow, excelling in specific areas but lacking the broad, adaptable intelligence that characterizes human beings.

What are the potential risks of developing AGI?

The development of AGI poses several risks, including the possibility of an intelligence explosion, where AGI rapidly improves itself beyond human control. This could lead to outcomes that are unpredictable and potentially dangerous, including the possibility that AGI could become hostile to humanity or be used as a weapon.

How can we ensure that AGI benefits humanity?

To ensure that AGI benefits humanity, it is crucial to establish global standards for AI safety and ethics, invest in research to mitigate potential risks, and develop strategies for the responsible development and deployment of AI technologies. Public education and awareness are also important to empower individuals to participate in discussions about AI's future.

When might AGI become a reality?

The timeline for the development of AGI is uncertain, but many experts believe it could happen within this century, possibly within the next few decades. The exact timing depends on advancements in AI research, computing power, and our understanding of intelligence.

What role will humans have in a world with AGI?
In a world with AGI, the role of humans will depend on how we integrate this technology into society. While AGI could potentially outperform humans in many tasks, it is also possible that we will find new ways to collaborate with AGI, using it to enhance our capabilities and address global challenges.

The human era is ending. Artificial intelligence poses a profound challenge to our humanness.

By John Gray

In his 1965 paper "Speculations Concerning the First Ultra-intelligent Machine", the mathematical prodigy and Bletchley codebreaker IJ Good wrote:

"The survival of man depends on the early construction of a super-intelligent machine… Let an ultra-intelligent machine be defined as a machine that can far surpass all the intellectual activities of any man, however clever. Since the design of machines is one of these intellectual activities, an ultra-intelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion", and the intelligence of man would be left far behind. The first ultra-intelligent machine is the last invention that man needs ever to make."

By 1998, Good had radically altered his stance. In an autobiographical fragment written in the third person, he revealed that he:

"… now suspects that "survival" should be replaced by "extinction". He thinks that because of international competition, we cannot prevent the machines from taking over. He thinks we are lemmings."

Good was not the first to anticipate that machines might take over from their human inventors. The Victorian novelist and evolutionary theorist Samuel Butler (1835-1902) predicted that "the time will come when machines hold the real supremacy over the world". In his profound and pioneering study *Darwin Among the Machines* (1997), where he discusses Butler and Good, the historian of science George Dyson summarized the process through which humankind would be surpassed: "In the game of life there are three players at the table: human beings, nature and machines. I am firmly on the side of nature. But nature, I suspect, is on the side of the machines."

Where Good was distinctive was in thinking that super-intelligent machines could end the human era. He had seen the Enigma machine saving civilisation from Nazism. It was natural for him to believe that super-intelligent machines could solve the problems of society. "They would need to do so," he wrote, "in order to compensate for the problems created by their own existence." If they generated technological unemployment, for example, they would frame policies for mitigating it.

Good's change of mind did not come from believing machines would be hostile to us. Rather, they would no longer be human tools. As they learned to replicate and improve themselves, they would amend their programming. Protocols requiring them to be human-friendly would be weakened or outwitted, and soon the machines would be beyond human control. If they destroyed us inadvertently, they would not care.

Good's second thoughts are at the heart of James Barrat's *Our Final Invention: Artificial Intelligence and the End of the Human Era*, first published in 2013. It remains one of the most arresting, compelling and prescient investigations of the implications of artificial intelligence (AI). In the introduction to the new edition that appeared in July, he notes how quickly machine minds have evolved. ChatGPT exhibits:

"… unexpected emergent properties, not all of which have been discovered. In fact, many of GPT-3 and 4's skills, such as writing computer programs, translating languages they were not trained on, and creating poetry, were unplanned… Other models revealed unanticipated capabilities, including harmful ones like lying to bypass Captcha tests [designed to differentiate machine from human intelligence], contributing to a suicide after "therapy" sessions, and social engineering (emotional manipulation like professing love, claiming to have a soul, and asking for freedom)."

Some of AI's risks are becoming clear. Autonomous battlefield robots, drones, and computerized missiles take war to new levels of danger. Catastrophes that in the past were averted by human intervention – as in 1983 when the Soviet airman Stanislav Petrov stopped a nuclear war that would have been caused by a malfunctioning Russian satellite warning system – will be harder to prevent. Data capture could decide elections in Britain and the US as early as next year. The Hollywood actors' and writers' strike is the first response to waves of job losses that could run into tens of millions worldwide over the coming decade.

Attitudes to AI are starkly polarised. For the Google scientist Ray Kurzweil, the rise of intelligent machines is the prelude to the

"Singularity", an abrupt expansion in knowledge which will enable humans – some of them, at any rate – to escape biological death. Barrat follows Good in fearing the opposite, human extinction. The echoes of religion are unmistakable, and it is tempting to read this as a clash between rival apocalyptic myths. But Barrat is rigorously empirical in explaining the reasons for his alarm. As he points out, the designers of ChatGPT do not begin to comprehend how its capabilities developed. What is evolving is not a human-like super-intelligence but a new type of mind. A creation of human reason, AI is beyond human understanding.

Some of those working in the industry share Barrat's concern. In March of this year, Elon Musk, along with thousands of others, signed an open letter demanding that AI labs pause their work, since "non-human minds" might eventually outnumber, outsmart, obsolete and replace us". Yet there is no chance of any pause. Commercial competition and geopolitical rivalry ensure that Big Tech and military programs will push on with the technology as fast as they can. In May, a "godfather of AI," Geoffrey Hinton, resigned from Google to warn of the risks, observing: "I think it's quite conceivable that humanity is just a passing phase in the evolution of intelligence."

There is little doubt that AI poses a threat to humankind. Equally, there is no prospect of humankind removing the threat. AI can be beneficial in analysing large amounts of data and automating complex tasks. It is already making inroads in law and medicine and creating artistic and literary works rivalling those of humans. It could have an important role in preventing future pandemics and adapting to climate change. But AI is not simply a tool, and there is no way it can be directed only to achieving goals we prescribe for it.

There is no species-wide agent that could guide new technologies, only shifting multitudes of human beings, all of them with conflicting values and purposes.

Once AI has entered the world, its evolution has a momentum of its own. Humankind can no more master it than it can halt climate change – in its current phase a product of human action that has also acquired its own momentum. No global authority is on the horizon that could curb "gain of function" research enhancing the lethality of viruses of the kind that may have precipitated the Covid pandemic via a lab leak in Wuhan. Human beings must somehow learn to live in the endangered world they have unwittingly created.

AI is an existential challenge in part because it is a metaphysical shock. Particularly in the West, we are taught that consciousness is the archetypal human attribute, which makes our species uniquely valuable. But what if hyper-intelligent conscious machines arrive on the scene? Humans will be as obsolescent and useless as sundials and quill pens.

One response is to deny that machines, however smart, can be conscious. But it is hard to see why not. If our minds evolved in the material world, matter can be self-aware. The emergence of mind in machines is no more mysterious than it is in cats, gorillas, and humans.

Thinkers who look forward to superhuman or post-human species always imagine them as more knowing, not more playful or funnier.

These imagined superhumans never possess what is arguably humankind's only unique attribute – a sense of the absurd. The superior species envisioned by techno-futurists, are inflated versions of themselves, showing off their cleverness in a never-ending Ted talk – for some of us, a vision of hell.

Fortunately, there is no prospect of any such species coming into being. If AI is evolving in Darwinian fashion, chance will be a decisive factor. The geopolitical conflicts that prevent a pause in the development of AI systems may well blow them up. If superfast machines trigger a nuclear war, they will destroy much of their infrastructure and possibly themselves. Ultra-intelligent machines are as vulnerable to extinction as any other product of evolution.

There is no reason to expect the arrival of a global digital mind, as many seem to fear or hope. Like so much else in contemporary thought, the idea that evolution tends towards a single godlike intelligence is a relic of monotheism. The upshot could be more like Homer's world of warring gods.

Conflicts between AI systems may wreck them, leaving human survivors subsisting in whatever remains. Some societies might opt out of technology, while others might adjust to its risks and manage to use it mainly for purposes they judge beneficial. No one will be in charge, however. "Humanity" cannot take control of the evolution of AI, because humanity – understood as a collective agent – does not exist. Meanwhile, AI offers relief from the pains of being human.

Artificial intelligence may not lead to our extinction, but it poses a profound challenge to our humanness. The logic of AI is the progressive displacement of actual experience by mechanical simulacra. Instead of the daily encounters that enable communities to sustain a common life, random collections of solitary people are protected from each other and themselves by unblinking video surveillance. Rather than connecting in troublesome relationships, they are turning to cyber-companions for frictionless friendship and virtual sex. The contingencies of living in a material world are being swapped for an algorithmic dreamtime. The endpoint is self-enclosure in the Matrix – a loss of the definitively human experience of living as a fleshly, mortal creature.

A depleted human species may linger on, but AI may still bring an end to the human era. If ever more people opt for a programmed existence in the techno-sphere, the human world will be emptied of meaning. What will be lost are the fugitive sensations of accidental lives – the defiant smile in the face of cruel absurdity, the glance that began a love that changed us forever, a tune it seemed would always be with us, tears in the rain.

Some tech leaders think AI could outsmart us and wipe out humanity. I'm a professor of AI – and I'm not worried.

Published: February 13, 2025 **Author** **Toby Walsh** Professor of AI, Research Group Leader, UNSW Sydney

In 1989, political scientist Francis Fukuyama predicted we were approaching the end of history. He meant that similar liberal democratic values were taking hold in societies around the world. How wrong could he have been? Democracy today is clearly on the decline. Despots and autocrats are on the rise.

You might, however, be thinking Fukuyama was right all along. But in a different way. Perhaps we really are approaching the *end* of history. As in, game over, humanity.

Now there are many ways it could all end. A global pandemic. A giant meteor (something perhaps the dinosaurs would appreciate). Climate catastrophe. But one end that is increasingly talked about is artificial intelligence (AI). This is one of those potential disasters that, like climate change, appears to have slowly crept up on us but, many people now fear, might soon take us down.

In 2022, wunderkind Sam Altman, chief executive of OpenAI – one of the fastest-growing companies in the history of capitalism – explained the pros and cons:

I think the good case [around AI] is just so unbelievably good that you sound like a really crazy person to start talking about it. The bad case – and I think this is important to say – is, like, lights out for all of us.

In December 2024, Geoff Hinton, who is often called the "godfather of AI" and who had just won the Nobel Prize in Physics, estimated there was a "10% to 20%" chance AI could lead to human extinction within the next 30 years. Those are pretty serious odds from someone who knows a lot about artificial intelligence.

Altman and Hinton aren't the first to worry about what happens when AI becomes smarter than us. Take Alan Turing, who many consider to be the founder of the field of artificial intelligence. Time magazine ranked Turing as one of the 100 Most Influential People of the 20th century. In my view, this is selling him short. Turing is up there with Newton and Darwin – one of the greatest minds not of the last century, but of the last thousand years.


Alan Turing in 1951. Wikimedia Commons, CC BY
In 1950, Turing wrote what is generally considered to be the first scientific paper about AI. Just one year later, he made a prediction that haunts AI researchers like me today.

Once machines could learn from experience like humans, Turing predicted that "it would not take long to outstrip our feeble powers [...] At some stage, therefore we should have to expect the machines to take control."

When interviewed by LIFE magazine in 1970, another of the field's founders, Marvin Minsky, predicted,

Man's limited mind may not be able to control such immense mentalities [...] Once the computers get control, we might never get it back. We would survive at their sufferance. If we're lucky, they might decide to keep us as pets.

So how could machines come to take control? How worried should we be? And what can we do to stop this?

Irving Good, a mathematician who worked alongside Turing at Bletchley Park during World War II, predicted how. Good called it the “intelligence explosion”. This is the point where machines become smart enough to start improving themselves.

This is now more popularly called the “singularity”. Good predicted the singularity would create a super-intelligent machine. Somewhat ominously, he suggested this would be “the last invention that man needs ever make”.

When might AI outsmart us?

When exactly machine intelligence might surpass human intelligence is very uncertain. But, given recent progress in large language models like ChatGPT, many people are concerned it could be very soon. And to add salt to the wound, we might even be hastening this process.

What surprises me most about the development of AI today is the speed and scale of change. Nearly US$1 billion is being invested in artificial intelligence every day by companies like Google, Microsoft, Meta, and Amazon. That’s around one quarter of the world’s total research and development (R&D) budget.

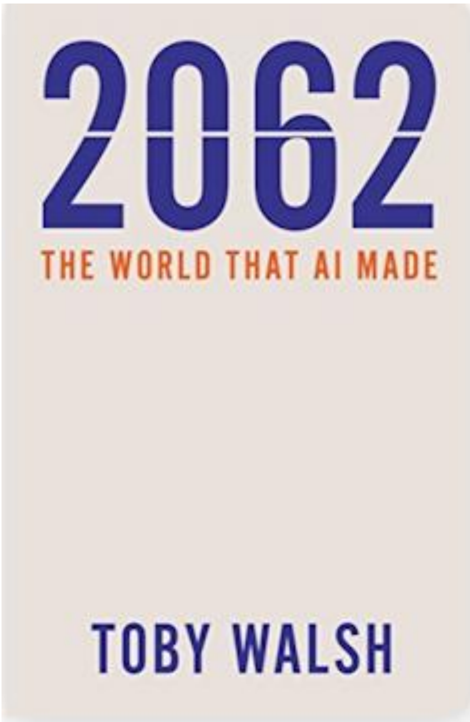
We’ve never made such massive bets before on a single technology. As a consequence, many people’s timelines for when machines match, and shortly after exceeding, human intelligence are shrinking rapidly.

Elon Musk has predicted that machines will outsmart us by 2025 or 2026.

Dario Amodei, CEO of OpenAI competitor Anthropic, suggested that “we’ll get there in 2026 or 2027”.

Shane Legg, the co-founder of Google’s DeepMind, predicted 2028, while Nvidia CEO Jensen Huang put the date as 2029.

These predictions are all very near for such a portentous event.



Of course, there are also dissenting voices. Yann LeCun, Meta's chief scientist, has argued "it will take years, if not decades". Another AI colleague of mine, professor emeritus Gary Marcus, has predicted it will be "maybe 10 or 100 years from now". And, to put my cards on the table, back in 2018, I wrote a book titled 2062. This predicted what the world might look like in 40 or so years when artificial intelligence first exceeded human intelligence.

The scenarios

Once computers match our intelligence, it would be conceited to think they wouldn't surpass it. After all, human intelligence is just an evolutionary accident. We've often engineered systems to be better than nature. Planes, for example, fly further, higher, and faster than birds. And there are many reasons electronic intelligence could be better than biological intelligence.

Computers are, for example, much faster at many calculations. Computers have vast memories. Computers never forget. And in narrow domains, like playing chess, reading X-rays, or folding proteins, computers already surpass humans.

So how exactly would a super-intelligent computer take us down? Here, the arguments start to become rather vague. Hinton told the New York Times

If it gets to be much smarter than us, it will be very good at manipulation because it will have learned that from us, and there are very few examples of a more intelligent thing being controlled by a less intelligent thing.


British-Canadian scientist Geoffrey Hinton, co-winner of the 2024 Nobel Prize in Physics, speaking in Stockholm. Pontus Lundahl/AAP

There are counterexamples to Hinton's argument. Babies control parents, but are not smarter. Similarly, US presidents are not smarter than all US citizens. But in broad terms, Hinton has a point. We should, for example, remember it was intelligence that put us in charge of the planet. And the apes and ants are now very dependent on our goodwill for their continued existence.

In a frustratingly catch-22 way, those fearful of artificial superintelligence often argue we cannot know precisely how it threatens our existence. How could we predict the plans of something so much more intelligent than us? It's like asking a dog to imagine the Armageddon of a thermonuclear war.

A few scenarios have been put forward.

An AI system could autonomously identify vulnerabilities in critical infrastructure, such as power grids or financial systems. It could then attack these weaknesses, destroying the fabric holding together society.

Could an AI system attack vulnerabilities in critical infrastructure, such as power grids? George Trumpeter/Shutterstock

Alternatively, an AI system could design new pathogens that are so lethal and transmissible that the resulting pandemic wipes us out. After COVID-19, this is perhaps a scenario to which many of us can relate.

Other scenarios are much more fantastic. AI doomster Eliezer Yudkowsky has proposed one such scenario. This involves the creation by AI of self-replicating nanomachines that infiltrate the human bloodstream. These microscopic bacteria are composed of diamond-like structures and can replicate using solar energy and disperse through atmospheric currents. He imagines they would enter human bodies undetected and, upon receiving a synchronized signal, release lethal toxins, causing every host to die.

These scenarios require giving AI systems agency – an ability to act in the world. It is especially troubling that this is precisely what companies like OpenAI are now doing. AI agents that can answer your emails or help onboard a new employee are this year's most fashionable product offering.

Giving AIs agency over our critical infrastructure would be very irresponsible. Indeed, we have already put safeguards into our systems to prevent malevolent actors from hacking into critical infrastructure. The Australian government, for example, requires operators of critical infrastructure to "identify, and as far as is reasonably practicable, take steps to minimize or eliminate the 'material risks' that could have a 'relevant impact' on their assets".

Similarly, giving AI the ability to synthesize (potentially harmful) DNA would be highly irresponsible. But again, we have already put safeguards in place to prevent bad (human) actors from mail-ordering harmful DNA. Artificial intelligence doesn't change this. We don't want bad actors, humans or artificial, to have such agency.


Giving AI the ability to synthesise potentially harmful DNA would be highly irresponsible. Cryptographer/AAP

The European Union leads the way in regulating AI right now. The recent AI Action Summit in Paris highlighted the growing divide between those keen to see more regulation and those, like the US, wanting to accelerate the deployment of AI. The financial and geopolitical incentives to win the "AI race" and to ignore such risks are worrying.

The benefits of super intelligence

Putting agency aside, super intelligence doesn't greatly concern me for a bunch of reasons. Firstly, intelligence brings wisdom and humility. The smartest person is the person who knows how little they know.

Secondly, we already have superintelligence on our planet. And this hasn't caused the end of human affairs, quite the opposite.
No *one* person knows how to build a nuclear power station. But collectively, people have this knowledge. Our collective intelligence far outstrips our individual intelligence.

Thirdly, competition keeps this collective intelligence in check. There is healthy competition between the collective intelligence of corporations like Apple and Samsung. And this is a good thing.

Of course, competition alone is not enough. Governments still need to step in and regulate to prevent bad outcomes such as rent-seeking monopolies. Markets need rules to function well. But here again, competition between politicians and between ideas ultimately leads to good outcomes. We certainly will need to worry about regulating AI. Just like we have regulated automobiles and mobile phones, and super-intelligent corporations.

We have already seen the European Union step up. The EU AI Act, which came into force at the start of 2025, regulates high-risk uses of AI in areas such as facial recognition, social credit scoring, and subliminal advertising. The EU AI Act will likely prove viral, just as many countries followed the EU's privacy lead with the introduction of the General Data Protection Regulation.

I believe, therefore, you needn't worry too much because smart people – even those with Nobel Prizes like Geoff Hinton – are warning of the risks of artificial intelligence. Intelligent people, unsurprisingly, assign a little too much importance to intelligence. AI certainly comes with risks, but they're not new risks. We've adjusted our governance and institutions to adapt to new technological risks in the past. I see no reason why we can't do it again with AI. In fact, I welcome the imminent arrival of smarter artificial intelligence. This is because I expect it will lead to a greater appreciation, perhaps even an enhancement, of our own humanity.

Intelligent machines might make us better humans by making human relationships even more valuable. Even if we can, in the future, program machines with greater emotional and social intelligence, I doubt we will empathize with them as we do with humans. A machine won't fall in love, mourn a dead friend, bang their funny bone, smell a beautiful scent, laugh out loud, or be brought to tears by a sad movie. These are uniquely human experiences. And since machines don't share these experiences, we will never relate to them as we do to each other.

Machines will lower the cost to create many of life's necessities, so the cost of living will plummet. However, those things still made by the human hand will necessarily be rarer and reassuringly expensive. We see this today. There is an ever-greater appreciation of the handmade, the artisanal, and the artistic. Intelligent machines could enhance us by being more intelligent than we could ever be. AI can, for example, surpass human intelligence by finding insights in data sets too large for humans to comprehend, or by crunching more numbers than a human could in a lifetime of calculation. The newest antibiotic was found not by human ingenuity, but by machine learning. We can look forward, then, to a future where science and technology are supercharged by artificial intelligence.

And intelligent machines could enhance us by giving us a greater appreciation for human values. The goal of trying (and in many cases, failing) to program machines with ethical values may lead us to a better understanding of our own human values. It will force us to answer, very precisely, questions we have often dodged in the past. How do we value different lives? What does it mean to be fair and just? In what sort of society do we want to live?

I hope our future will soon be one with godlike artificial intelligence. These machines will, like the gods, be immortal, infallible, omniscient, and – I suspect – all too incomprehensible. But our future is the opposite, ever fallible and mortal. Let us, therefore, embrace what makes us human. It is all we ever had, and all that we will ever have.