Leaving Plato's cave

Towards a science of the future

Slavoj Žižek — *truth*

Yanis Varoufakis — *cloud capital*

Susan Schneider — *mind exchange*

Lucy Cooke — *resurrecting extinct animals*

Philip Goff — *panpsychism*

& others

Leaving Plato's cave

Leaving Plato's cave

A compass towards a science of the future

Slavoj Žižek
Yanis Varoufakis
Susan Schneider
Riccardo Manzotti
Finn Brunton
Thomas Hertog
Charles Foster
Roman Yampolskiy
Alien
Bernardo Kastrup
Rupert Sheldrake
Lucy Cooke
Philip Goff

Contents

0 6 **Introduction** Boudewijn Richel

1 22 **Truth** Slavoj Žižek

2 36 **Cloud Capital** Yanis Varoufakis

3 52 **Mind Exchange** Susan Schneider

4 66 **Can Machines think?** Riccardo Manzotti

5 80 **Cryptos** Finn Brunton

6 96 **Time** Thomas Hertog

7 112 **Edges** Charles Foster

8 128 **Hacking the Code of the World** Roman Yampolskiy

9 140 **Time Travel** Alien

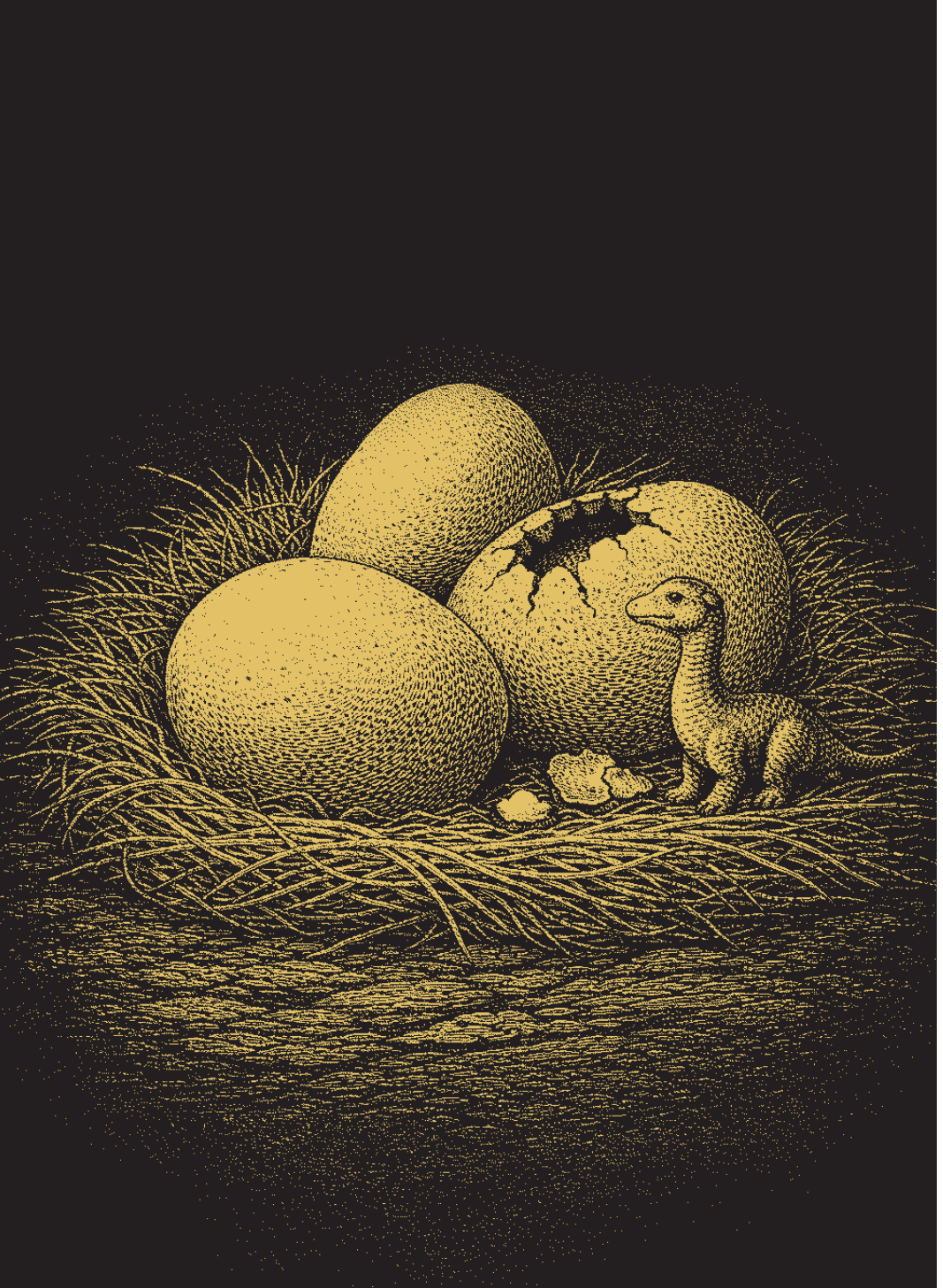
10 148 **AI won't be conscious** Bernardo Kastrup

11 164 **Morphic Resonance** Rupert Sheldrake

12 178 **Resurrecting Extinct Animals** Lucy Cooke

13 192 **Galileo's Error** Philip Goff

208 **Biographies**



Diplodocus Carnegii

0

Boudewijn Richel

Introduction: Leaving Plato's Cave

The G10 embarked on an intellectual expedition to map the future of human thought. Our compass was a cluster of revolutionary ideas, from the philosophical frameworks of Michel Foucault to the disruptive potential of artificial intelligence. Yet, one concept, drawn from the seemingly remote world of quantum physics, emerged as the most potent binding force for our inquiry: **entanglement**.

To understand why this idea is so critical, we must first grasp its radical meaning. In quantum mechanics, entanglement describes a phenomenon so profound it dismantles our classical intuition. When two particles become entangled, they are no longer separate entities. They form a single, unified quantum system. Their properties, such as position or spin, are not independently defined but exist in a shared state. The profound implication is that a measurement performed on one particle instantly influences the state of its partner, regardless of the distance separating them—be it a millimeter or a galaxy. This "spooky action at a distance," as Einstein called it, suggests that at the most fundamental level of reality, the universe is not a collection of isolated objects but a dense web of intrinsic, non-local relationships. Things do not simply interact; their very identities are co-dependent.

This book argues that this quantum principle has escaped the laboratory to become the defining paradigm of the 21st century. **We are living in an Age of Entanglement.** The twelve steps taken by the different authors in their own special ways demonstrate how this concept permeates every layer of our existence, from the fabric of our consciousness to the structures of our global society, offering a new key to decipher our present and navigate our tomorrow.

Michel Foucault

This perspective is powerfully sharpened by the work of Michel Foucault, whose ideas provide a historical and structural framework for this entanglement. For Foucault, an episteme is the unconscious foundation of knowledge in a given era—the underlying "game of truth" that determines what can be thought, spoken, and recognized as real. It is a thrilling moment when knowledge takes a sharp turn, flipping our understanding of reality on its head. The invention of the microscope in the 17th century, for instance, wasn't just a technical innovation; it was an *epistemic* rupture. It didn't just magnify tiny things; it shattered ancient myths about spontaneous generation and revealed a hidden biological universe, fundamentally reconfiguring our perception of life itself.

Simultaneously, Foucault's *dispositif*—the heterogeneous network of institutions, laws, administrative measures, scientific statements, and philosophical propositions that constitutes the apparatus of power—shows how knowledge and power are inextricably linked. The *dispositif* is the architecture of entanglement at the societal level. It is the mechanism through which diffuse forces—economic interests, political strategies, technological capabilities—crystallize into tangible structures that shape, guide, and control human life. The *dispositif* makes the abstract entanglement of power and knowledge concrete and operational.

We noticed a fascinating pattern that runs through this history: groundbreaking artists like Rembrandt, Marcel Duchamp, and Andy Warhol have consistently been in a creative conversation with the scientists and philosophers of their time. This electric dialogue between art and science, where they collide and collaborate to push the boundaries of what is possible, became the heartbeat of our exhibition. Art often acts as a sensitive detector, feeling the tremors of a new *episteme* long before it is fully articulated by science or philosophy.

The book structured in 12 steps, is a journey through these collisions and convergences. It is a winding path through the his-

tory of ideas, where art, science, and power collide in the most unexpected ways. From Plato's Cave to the holographic universe, we trace how the quest for truth has been upended by the principle of connection. The future is not a linear projection; it is a complex, interdependent system. By adopting entanglement as our core concept, this book provides a new vocabulary for the breathtaking complexity of the century ahead. The following pages are an invitation to explore this deeply interconnected tapestry.

Plato's Cave: The Architecture of Separation

Picture this: You are chained in a dark cave, forced to stare at shadows dancing on a wall, mistaking these flickering illusions for reality. This is Plato's 2,400-year-old metaphor for his foundational idea: that "true reality" is a perfect, unchanging world of ideal Forms outside our messy, imperfect human experience. Philosophers, he argued, could escape the cave through reason and logic to perceive this higher truth.

This idea stuck like glue. For centuries, Western science, religion, and art were obsessed with finding this "real world" out there—whether in God's heaven, mathematical laws, or the immortal soul. This worldview established what we might call the **Great Separation**: a fundamental duality between the inner world of the mind (the cave of subjective illusion) and the outer world of truth (the bright sunlight of objective reality). This Platonic *dispositif* became the bedrock of Western thought, creating a powerful hierarchy that privileged the abstract and universal over the concrete and particular.

Living On Islands In Greece!

The geopolitical context of Plato's thought reinforces this. He reflected on the perils of living on a border near the seaside, where an island was susceptible to dangers from pirates and "bad ideas" (incorrect worldviews) from refugees and traders. Water

represented chaos and the unknown. In contrast, the center of an island offered greater safety and was remarkably considered the home of the Logos—the word of truth and reason. This concept of *Logos*, a singular and centralized truth, safe from external contamination, has influenced Western art, science, and politics for centuries. The goal was always to reach the center, to find the one true answer.

We put Plato at the forefront of this book because he profoundly shaped Western culture with this distinction. Unlike most other philosophers, we still grapple with his legacy daily. His worldview remains active in science's search for a "theory of everything" and in art's often-lingering reverence for the unique masterpiece. It hinders the new, entangled thoughts necessary to navigate the digital era, a time when the very idea of a center is dissolving into distributed networks.

The Entanglement of Inner and Outer

The cave story seems innocent, but it hides deep assumptions about human existence that are being radically challenged. AI, for instance, is bulldozing Plato's cave. Algorithms now mine our thoughts, predict our desires, and even mimic our creativity. As Riccardo Manzotti argues in his text "Can Machines Think?", there may be no "inner truth" left to protect—only data patterns and behavioral correlates. Plato's wall between "inside" and "outside" is crumbling.

Manzotti is not alone in his skepticism about an isolated inner world. Buddhist philosophy has always been skeptical of what it calls "an inherent existence of things." Buddhists claim that phenomena are empty of independent, self-contained reality. Their central task is to understand the **entanglement between an inner and outer world**. The famous mantra, "Form is emptiness and emptiness is form," signifies that these are not two separate realms but intertwined aspects of a single reality. Things do not

exist independently but are co-dependent and entangled in the very process of perception. In this profound sense, Buddhism and AI converge in their methodology: both break down the barriers Plato so carefully erected.

We see a related concept in the work of Rupert Sheldrake, who describes a common memory or "morphic resonance," suggesting that memory is not sealed within our brains but is part of a field—a phenomenon of entanglement that transcends the individual.

The Roots of AI and Cloud Capital

This erosion of the inner/outer divide is the engine of the modern digital economy. At the beginning of the last century, literature began to explore this terrain. Authors like Virginia Woolf and Robert Musil, in his novel *The Man Without Qualities*, thematized an inner life that was completely dominated by a stirring outside world, where individual traits faded away, becoming a mere projection surface for external forces.

Today, AI operationalizes this. It has little use for the deepest, ineffable secrets of the individual soul. The algorithms of social media platforms often know more about our objectifiable desires than we do ourselves. As Yanis Varoufakis argues in his text on Cloud Capital, the core value of companies like Amazon and Facebook lies in their access to information that was once considered private. The act of penetrating the secret world of the individual's inner self—breaking Plato's boundary—has created the most valuable firms in the world. Data, the new gold, is the currency of this new entanglement.

The 17th Century—Mapping the Material World

Fast-forward to 1600s Amsterdam. Dutch traders ruled the seas, scientists invented the microscope, and artists like Rembrandt painted dissected corpses to uncover how bodies worked. This was a society obsessed with *stuff*—with matter, mechanics, and

profit. This cultural ferment gave birth to modern **materialism**: the belief that reality consists solely of atoms, equations, and things that can be measured and quantified.

The parallel to today is striking. We have swapped atoms for data. Just as 17th-century merchants hoarded spices, modern corporations hoard our behavioral secrets—what we buy, click, and dream about. Data has become the new gold, the fundamental substance of a new materialist age.

This was the Dutch Golden Age, where the world seemed to revolve around Dam Square. The philosophy was one of overcoming boundaries. The VOC was established to dominate global trade, breaking down geographical borders. Descartes began to view animals as complex machines, breaking down the boundary between the living and the mechanical. Spinoza, grinding lenses in Amsterdam, sought to analyze matter and delve deeper into its essence, beginning the journey toward ever smaller particles.

The 17th century in Amsterdam symbolically marks the beginning of the quest to explore matter from within. The microscope, a Dutch invention, allowed humanity to see a hidden, entangled world teeming with life. Rembrandt's fascination with Dr. Tulp's anatomy lesson reveals a desire to pry open the skull, to delve into the brain and uncover its material secrets. As this exploration deepened, the power of God as an explanation began to wane, replaced by a focus on clocks, instruments, and maps—the tools for describing a universe governed by mechanical laws. Modern time, precise and universal, really started here.

Quantum Physics—The Discovery of Fundamental Entanglement

Imagine flipping a coin, but instead of landing on heads or tails, it somehow lands on both at the same time, existing in a blurred state of superposition until you look at it. That is the quantum world in a nutshell. In the early 20th century, physicists like Niels

Bohr discovered that tiny particles (photons, electrons) don't follow the rules of classical physics. They exist as waves of possibility—probability clouds—until a measurement is made. Then, they "choose" to act like particles in a specific location.

This was the formal discovery of **quantum entanglement**. Einstein famously hated this inherent randomness, declaring, "God does not play dice with the universe." But repeated experiments, from Bell's test onwards, kept proving him wrong. Quantum physics forced us to admit a staggering truth: Reality at its base is not fixed and deterministic. It is probabilistic and, most importantly, **its state is shaped by the act of observation**.

This was the ultimate blow to Plato's idea of a single, perfect, objective truth "out there." Suddenly, science had to make room for mystery, subjectivity, and the irreducible role of the observer. The "subjective" element had gained an undeniable momentum in the hardest of sciences.

Artists intuitively understood this shift decades earlier. Marcel Duchamp claimed, "It's the viewer that makes the work," turning art from a static object into a dynamic collaboration between creator and audience. Science and art were now saying the same thing in different languages: **Reality is a conversation, not a monologue**.

The discovery was not universally welcomed. The definition of an "observer" remains deeply unclear. Can a machine observe? What constitutes a measurement? These questions are still debated. While scientists love the clean certainty of equations like $E=mc^2$, and these work brilliantly within their domain, they fail at the quantum edge and in discussions of the multiverse. The road was paved to break out of a single, common reality and entertain the existence of multiple, parallel ones.

The 1960s—Proliferating Realities

The 1960s were a cosmic fever dream, a decade where the implications of quantum weirdness exploded into popular culture. Hippies used LSD to explore kaleidoscopic versions of their own consciousness. Scientist Hugh Everett III proposed the "many-worlds" interpretation of quantum mechanics, arguing that every decision we make splits the universe into parallel, equally real branches.

It was also a period of profound material and perspectival change. The vacuum cleaner and the fridge transformed domestic life, while Neil Armstrong walking on the moon in 1969 gave us the iconic "Earthrise" photograph, expanding our perspective to see our planet as a single, entangled whole.

Michel Foucault, in his book *The Order of Things*, provided the philosophical backbone for this proliferation. He argued that every era invents its own version of truth—medieval people saw demons, we see data—and none of them are the "final answer." He noted that new discoveries, while revealing one portion of reality, often obscure others.

Artists like Andy Warhol anticipated the digital age of replication by creating endless, identical copies of originals. His screen-printed Marilyns asked: What is original? What is real in an age of mechanical and, soon, digital reproduction? The British physicist David Deutsch remains a leading defender of the many-worlds theory, an idea that has captivated filmmakers and artists, as seen in blockbusters like *Avengers: Endgame*.

While the theory is not widely accepted in scientific circles, it represents a growing cultural and intellectual intuition that reality is not singular. It feels as if quantum physics is halfway down a track, and some crucial aspects are still missing. The work of Rupert Sheldrake on "fields" of memory and influence points toward a science that is more comfortable with non-local, entangled connections. Plato's cave was crumbling, and the 1960s left us staring into a multiverse of possibilities, wondering what would replace it.

The Digital Age—The Entanglement Economy

Fast-forward to today. We are living inside the paradigm shift that earlier eras only anticipated. Companies like Amazon, Facebook, and X mine our lives for data—the Cloud Capital that Yanis Varoufakis describes. This data is worth trillions because it allows for the prediction and subtle control of human behavior on a mass scale. Our desires, friendships, and fears are entangled with the algorithms that curate our reality.

There was a twist, a potential for a different kind of entanglement. Cryptocurrencies like Bitcoin promised to break free from centralized banks and governments—the modern "guardians of the Logos." Instead of trusting Plato's elite, crypto used cryptographic math to create a decentralized, trustless system where anyone could participate. It was a democratic, anarchistic dream of a new kind of financial entanglement.

But as Finn Brunton warns in his contribution, this anarchist root is fading. Today, crypto is less about freedom and more about new billionaires minting speculative meme currencies, representing a complete surge to a new form of oligarchic power. The digital age is a double-edged sword. It is profoundly anti-Platonic, smashing the barriers between public and private, truth and fiction, inside and outside. But it does not eliminate power; it simply hands it to a new priestly class: those who control the code. Plato's cave has not been destroyed; it has been transformed into a glass house, where we are both the watchers and the watched.

Truth vs. Fake—The Battle for a Shared World

One of the pillars of political life, science, and ethics has always been the distinction between truth and lies. Today, this distinction is under relentless attack, yet few are prepared to abandon it entirely.

In the Platonic tradition, truth was reserved for those entangled with the Logos—the male Athenian citizens. It was a sacred,

unchanging ideal. But today, truth is a battlefield. Quantum physics tells us a particle can be in two places at once. Social media tells us that your truth depends on your tribal affiliation. Even science struggles with the realization that its most cherished equations are models that work only within certain boundaries.

As Slavoj Žižek argues, we are often stuck in a world where "truth" is merely the story that wins the most clicks, the narrative that achieves the most viral entanglement. The very concept of an "observer" in quantum physics, necessary to collapse a wave function, has made the concept of a purely objective "truth" far more complex and fragile. We are in a period of painful but necessary revaluation of what truth means in an entangled world.

The Holographic Universe—We Are the Projector

The key change in our time is that the distinction between a subjective and an objective world is being redefined as a relationship of **entanglement**.

Thomas Hertog notes in his contribution that Hannah Arendt was already skeptical of a purely "objective" science. In *The Human Condition*, she argued that the scientist is not a value-free observer but an active participant who strongly influences what is observed. This, she believed, was not a flaw but a fundamental part of the human condition.

Value-free science may be a relic of the Platonic era. As we seek to understand the birth of the universe or the holographic principle—the mind-bending theory that our three-dimensional universe could be a projection from a two-dimensional surface—we are moving toward a science that has definitively said goodbye to Plato.

Here's the wildest idea yet: when we look at the stars, we see light that left them millions of years ago. Our "now" is a delayed livestream of cosmic history. Scientists take this further, suggesting that even black holes might be fuzzy holograms. Plato's

cave metaphor haunts us once more, but with a crucial difference. This time, we are not the prisoners chained to the wall. We are inside the projector itself. AI, quantum computers, and the James Webb Space Telescope are our flashlights, and we are using them to interrogate the source code of reality, to ask: What, and who, is outside the cave?

The Entangled Society: From Universal Truth to Tribal Realities

If the 20th century was defined by a search for universal, objective truths—be they political ideologies, scientific laws, or historical narratives—the 21st century is defined by their fragmentation. We are witnessing the rise of what can accurately be called **social entanglement**. This is the macro-level manifestation of the quantum principle: distinct groups, or "tribes," form internally coherent systems where their beliefs, facts, and realities are so tightly correlated that they become functionally independent of a shared, external world.

This phenomenon shatters the Platonic and Enlightenment ideal of a single *Logos*—a central, verifiable truth to which all rational citizens assent. Instead, we have a landscape of multiple, simultaneous, and often incompatible *epistemes* coexisting. A climate change fact in one epistemic tribe is a conspiracy theory in another. A public health measure is seen as an act of care in one entangled group and an act of oppression in a different one.

This is not merely disagreement; it is a manifestation of entanglement. The "state" of one member of a tribe (their outrage, their belief in a news story) instantly influences the state of the entire group, reinforcing its internal coherence and its separation from others. Social media algorithms act as the ultimate entanglement engine, functioning like a supercollider for human thought, constantly measuring and collapsing fluid waves of opinion into hardened, polarized particle-states. They create feedback loops where a tribe's reality is constantly confirmed

and amplified, making it increasingly "real" and separate for its members.

This presents a profound challenge to the traditional model of the nation-state, which was built on the idea of a common public square, a shared *dispositif* of institutions like a free press, universal education, and representative government that curated a generally accepted reality. This central "projector" of reality is breaking down. Governments, scientific bodies, and mainstream media—the traditional guardians of the Logos—find their authority dissolving, not because they are always wrong, but because the very architecture of truth has changed.

In this new entangled society, power no longer flows solely from the center outward, from the government to the people. It now circulates within and between these tribal networks. The most powerful actors are those who can create, sustain, and manipulate these entangled systems—whether they are politicians marshaling digital tribes, influencers shaping consumer reality, or algorithms that govern the flow of information itself.

Therefore, the journey from Plato's Cave to the Holographic Universe is not just an intellectual history; it is the story of our social evolution. We are moving from a model where truth was a mountain to be climbed (with philosophers and scientists at the peak) to a model where reality is a web to be navigated. The cave is no longer a singular one; we are all inhabiting our own bespoke, algorithmically-curated caves, each with its own shadow-play, believing our flickering images to be the whole of existence.

The great task of tomorrow, then, is not to find a way back to a mythical single truth, but to learn to build bridges between these entangled realities. It demands a new kind of literacy—an entanglement literacy—where we recognize that our truths are relational and that coexistence in a polarized world depends on understanding the connections, however spooky, that bind us all in a single, struggling, planetary system.

The Bottom Line

From Plato's cave to quantum waves to TikTok algorithms, humanity's story is about one thing: chasing the invisible connections that constitute reality. The future won't give us final answers or return us to a simple past. Instead, it will ask us to embrace radical uncertainty, to rewrite the rules of knowledge, power, and selfhood, and perhaps, to finally laugh at how seriously we once took a singular, solitary "truth." In the Age of Entanglement, to be is to be connected. Our greatest challenge and our only way forward is to learn what it means to be responsibly, ethically, and compassionately entangled.

Aptenodytes forsteri

1

Slavoj Žižek
Truth

Ends of philosophy

I

What we have today is not only the end of philosophy, there are even two ends of philosophy.

On the one hand, the first death of philosophy is what we find today in so-called theories of everything in physics. We are gradually approaching a scientific answer to big questions which were till now considered metaphysical questions. Does our universe have a beginning or an end? Does it have a limit in time or space? Do we humans have a free will? Today, at least the predominant approach is that our awareness of our freedom is, as it is put in the introduction to a cognitive sciences book, "a simple user's illusion". Even the big metaphysical questions, questions of natural causality or spatial and temporal connections, are now decided by experiments. Free will is treated as a question to be decided by brain sciences. This is why at the beginning of his late book *The Grand Design* Stephen Hawking triumphantly proclaims that philosophy is dead. With the latest advances in quantum physics and cosmology the so-called experimental metaphysics enters the scene.

Metaphysical questions about the origin of the universe can now be answered through experimental science and in this way decided.

II

On the other hand, we have a persisting transcendental approach. The hermeneutic, transcendental answer is that there is not one way things are in themselves and that every historical epoch defines the very difference between appearance and things in themselves in its own way.

According to Jürgen Habermas the point is that in order to approach reality science already presupposes a certain conceptual apparatus of scientific thinking and rational argumentation. The hermeneutic horizon of science is already here.

In a different way Martin Heidegger says the same thing. He

says that the question for philosophy is not how things really are, but precisely the historicity of how things really are.

Historicism

I always opposed historicism. It says, for example—to really understand Shakespeare, you have to know all the details of Elizabethan England. Shakespeare didn't understand himself. His work was full of inconsistencies, incomplete and if you come later it's not an obstacle. You can even understand Shakespeare better than he understood himself.

So, in some sense, we are not just reading Shakespeare differently. We are literally completing his work. The best cinema version that I know is done by Akira Kurosawa '61 or '62. The version of Hamlet set in contemporary Japan. Hamlet is a Japanese student returning from the United States. His father was a big corporate boss. And it has a wonderful title *Only bad people sleep well.*

There is a quote: "You can cheat some of the people all the time, all the people sometime, but you cannot cheat all the people all the time." This is a deeply ambiguous statement. It can mean that there are some idiots who can be cheated all the time, or it can mean that in some situations, I will be the idiot, in others, you will be, and so on. For me, this reflects how language is structured—the author himself is not the possessor of some hidden meaning.

To understand some historical event you should read it in a superposition with other possible events. When something happens, you should always ask yourself what other options fails so that this could have happened. Every new invention changes the entire past. It makes it interpreted in a different way. Let's say in several years, there is going to be World War III and there will be people who survive it. Later they will claim that it was all pointed towards this war. If there is no war, they will call it a "cold war", false alarms and so on. And it is not simply illusions, the situation is genuinely open.

Historicity in the sense of what counts as true. How reality appears to us is not a simple result of the cognition of reality but is always already presupposed by our approach to reality. If you are a scientist you don't approach the reality with the tabula rasa. You always rely on a certain mechanism of analyzing reality. And this mechanism cannot be itself scientifically grounded tested as true or not because it is already presupposed at work in every scientific approach to reality.

Many years ago, during the French elections, when one of the candidates was Édouard Balladur, there was a quote: "If Balladur wins the next election, his victory will be necessary." Here lies a paradox: after something happens, it becomes retroactively necessary.

The one who developed this historicist notion of truth fully to the end was Michel Foucault. His notion of truth can be summed up in the claim that truth or untruth is not a direct property of our statement but in different historical conditions different discourses produce each its own specific truth effect. A quote from Michel Foucault: "the problem does not consist in drawing the line between that in discourse which falls under the category of scientificity or truth and that which comes under some other category but in seeing historically how effects of truth are produced within discourses which are neither true nor false."

Scientific definition of truth

Science defines truth in its own terms: the truth of a proposition which should be formulated in clear and preferably formalized terms is established by experimental procedures which can be repeated by anyone.

Religious discourse operates in a different way. Its truth is established through complex rhetorical strategies which generate the experience of inhabiting a meaningful world.

If one were to ask Foucault a big metaphysical question, like

"Do we have a free will or not?", I think his answer would have been that this question only has a meaning and can be raised within a certain episteme.

It is this episteme, field of knowledge and power which determines under what conditions it is true or false and all we can ultimately do is describe this episteme. German has the wonderful term "unhintergehbares" something beyond which you cannot go. And for Foucault, we cannot go beyond what Heidegger called "Lichtung", the disclosure of being. Sciences produce big results, but within a certain disclosure of being, of reality.

Ontological incompleteness

What fascinates me in quantum physics is its ontological incompleteness. For example, the fact that we cannot measure position and movement at the same time doesn't mean that the particle doesn't have a definite position; it simply means we cannot measure it. But this is not a radical idea.

The more radical notion is that reality itself is incomplete. If you analyze reality to the very end, you will not arrive at a complete description of everything. How does this relate to philosophy? According to Hegel, the final result is not about attaining the whole truth. Instead, what you initially perceive as an epistemological obstacle—the limitation of your knowledge—turns out to define reality itself.

The great thing about quantum physics is that it contains an element of uncertainty that is not just epistemological but inherent to reality itself.

No reality without fiction - What is going on in this post-modern universe of ours?

The role of rumors

I want to draw your attention to one feature: the role of rumors. Our Big Other set of public values is no longer the public space,

clearly distinguished from the obscenities of private exchanges, but the very public domain in which fake news circulate. The domain in which we exchange rumors and conspiracy theories.

Recall the type of situation of a small group of people who all know some embarrassing secret about one of them. Plus, they all know that all others know this. But nonetheless they don't talk about it publicly.

This is the Big Other of appearances. Appearances have to be held. But the domain of rumors is exactly the opposite. Rumors do not deal with factual truth as opposed to appearances. But they are both outside factual truth.

Now, anonymous rumors are excluded from the public space, and they remain strangely efficient even if they are not true. How is a rumor usually circulated? I say, "I don't know if this is really true, but I was told" or rather "one says that somebody did this or that."

This obscene Big Other, this space of rumors is supplemented by another Big Other: public authority. The one of neutral expertise in its different forms state apparatuses, legal order, science. And here the true problem emerges. Can we trust this Big Other even in its scientific form? Is not science caught in the procedures of technological domination, exploitation and of capitalist interests? Did not science long ago lost its neutrality? Is this neutrality from the very beginning not a mask of social domination?

The pandemic

Applied to the pandemic: does this insight not compel us to problematize the scientific medical justification of lockdown measures and other reactions to the pandemic?

A friend of mine, Fabio Vighi, argues that if we join the dots provided and by the close analysis of the financial background of the pandemic, we see a well-defined narrative outline emerge.

He said: "Lock downs and the global suspension of economic transactions were intended to first allow the federal bank of the

United States to flood the ailing financial markets with freshly printed money while deferring hyperinflation and to introduce mass vaccination programs and health passports as pillars of a neo-feudal regime of capitalist accumulation. The mainstream narrative should therefore be reversed. The stock market did not collapse in March 2020 because lockdowns had to be imposed. Rather lockdowns had to be imposed because financial markets were collapsing. SARS or Covid is the name of a special weapon of psychological warfare that was deployed in the moment of greatest need. The aim of the printing spree of money was to pluck calamitous liquidity gaps. Most of this magic tree money is still frozen inside the shadow banking system the stock exchanges and various virtual currency schemes that are not meant to be used for spending and investment. Their function is only to provide cheap loans for financial speculation.

This is what Marx called fictitious capital which continues to expand in an orbital loop that is now completely independent of economic cycles on the ground. The bottom line is that all this cash cannot be allowed to flood into real economy for the latter would overheat and trigger hyperinflation.

In short, it is not the pandemic which put the capitalist order into an emergency state. It is global capitalism itself which needed an emergency state to avoid a debilitating crisis much stronger than the 2008 meltdown. And the pandemic was fabricated as a welcome excuse for the emergency state. The passage from neoliberal global capitalism to corporate neo-feudal capitalism is the basic process that uses historical contingencies as excuses. And Fabio Vighi is not afraid to add ecologically grounded lockdowns to this series of excuses.

Far from just confronting capitalism with its faithful limitations, the ecological crisis will also be used as a scientifically grounded way to discipline and control the population. Green capitalism is not just a humanitarian mask of the global order. It is also a way

for the big corporate capital to reproduce itself.

What do I think about this view? Vighi takes into account the complexity of the situation. The interests of pharmaceutical corporations, the way the expert scientific insights that ground anti-pandemic measures justify new forms of social control.

For example, an important indication of the new phase of capitalism was a weird fact that took place in the spring of 2020. On the same day that state statistics in the United States and the United Kingdom registered a breathtaking fall of the GDP brutal product, comparable to the fall at the time of the great repression in 1920s, stock markets registered a gigantic rise.

In short, although the real economy is stagnating or even decreasing, stock markets go up. An indication that fictitious capital is caught in its own circle, decoupled from the real economy. This is where the measures justified by the pandemic entered the game. They, in a way, turned around the traditional Keynesian procedure, which is to print money, even if it means inflation, but invest this money to help the real economy. Here, it is the opposite: print money, but at the same time prevent this money from reaching the real sphere of production and consumption.

They in a way turned around the traditional Keynesian procedure. Which is print money even if it means inflation but invest this money to help real economy. Here it is the opposite print money but at the same time prevent that this money will reach the real sphere of production and consumption.

The paradox

I would like to add here another important observation: the so-called Lauderdale Paradox, named after a very interesting British thinker and friend of the French Revolution around 1800. What is the Lauderdale Paradox?

The most precious wealth of a society consists of course of objects which are freely available like water and air. But they do

not count as values which makes you rich because they are freely available. If water is easily available nobody gets rich by it. But if its supply is controlled by private companies those who own these companies get rich. So, in a technical sense of wealth, as embodied in values, there is more wealth in a society where you pay for water since the freely available water doesn't count as wealth.

Isn't this our situation today? If we live a country with the fresh good air, this doesn't count as wealth. Air is just there. You see the paradox. Let us say the air gets polluted. So, you have companies producing it, making surplus value. From the standpoint of capital's reproduction, the country where you barely can breathe but a tremendous industry is set in motion so that you can survive, this country counts as wealthier than the one where you can simply breathe. And I think this is our future with new ecological catastrophes. Yes, capital will survive. We will get more and more wealthy precisely because we will need all the machinery which will be on the market as values to make us survive.

I don't think that you can reduce the global warming and ecological catastrophes simply to an intended product of the reproduction of big capital. Like the only hysteresis, the only real thing that is happening is that capital was approaching a crisis.

Where I nonetheless disagree with Vighi is that I don't think you can reduce the pandemic or even global warming and ecological catastrophes simply to an intended product of the reproduction of big capital. The only hysteresis, the only real thing that is happening, is that capital was approaching a crisis. Fabio Vighi, in fact, is a little bit open here. Was the pandemic totally invented, or was it just an ordinary, modest flu? But capital needed it; they needed a lockdown to be able to avoid a financial crisis. I cannot buy this theory. It is, for me at least, a little bit too close to a conspiracy theory.

Conspiracy theory

In my books, I often use a wonderful example from Jacques Lacan. He says that even if what a jealous husband claims about his wife, that she sleeps around with other men, is all true, his jealousy is still pathological. Because his jealousy is not simply grounded in facts. The problem is not what the woman is really doing. The problem is: Why does he need jealousy to assert his identity to function normally? The Marxist answer is simple: because you need an external evil intruder to mystify the inherent antagonism.

What is wrong about them is not their truth or not. Conspiracy theories do something more. What?

The first problem with conspiracy theories is that they are usually dismissed as part of our relativist post-truth era. Every group proposes its own theory of truth. Conspiracy theories begin with skepticism, which is good, who knows what is true. But did you notice how then immediately they pass on to an absolute certainty. The agent of social authority but now in a more decisive radical sense of the other like god. The other like destiny, the idea that the world is not meaningless. It is not just chaos. There is a higher ordering principle.

The problem of conspiracy theories is that they do believe there is truth, an unconditional truth. And this truth is the act of an active all-powerful agent whose main goal is to deceive us. The formula of conspiracy theorists is better an evil god than no god. In today's time where even religion is getting relativized, we at least have an evil all-powerful agent of conspiracy.

Of course, cracks immediately reemerge here. Conspiracy theorists established with absolute certainty a conspiracy about the pandemic to then pass with uneasy easiness from one inconsistent theory to another.

You remember Freud's jokes about a borrowed cattle. You accuse me that I returned you a broken cattle that I borrowed from you and my answer is a) I never borrowed a cattle from you

b) I returned it to you unbroken c) the cattle was already broken when I got it from you. That is how most of the conspiracy theories function.

But I still don't like the idea that it is all planned by some secret state control agency, public authority, or capital agency. Why not? How does the pandemic function? On the one hand, we have medical healthcare; some of us get ill and others don't.

On the other hand, we have the financial economic consequences, the self-reproduction of fictitious capital with very complex links to the actual real economy, and the autonomy of financial capital. But I think what Vighi neglects a little bit is that yes, we have this crazy self-circulation of fictitious capital, but this crazy circulation, which neglects real life, has very strong material consequences in real life, just to mention enormous amounts of waste and pollution.

So, my final argument against Vighi would have been that he is too optimist. If we accept that the pandemic is just a minor disease and that basically all is fiction invented by fictitious capital the result would be just to return to normal life. But I think this is an all too optimistic view on today's capitalism.

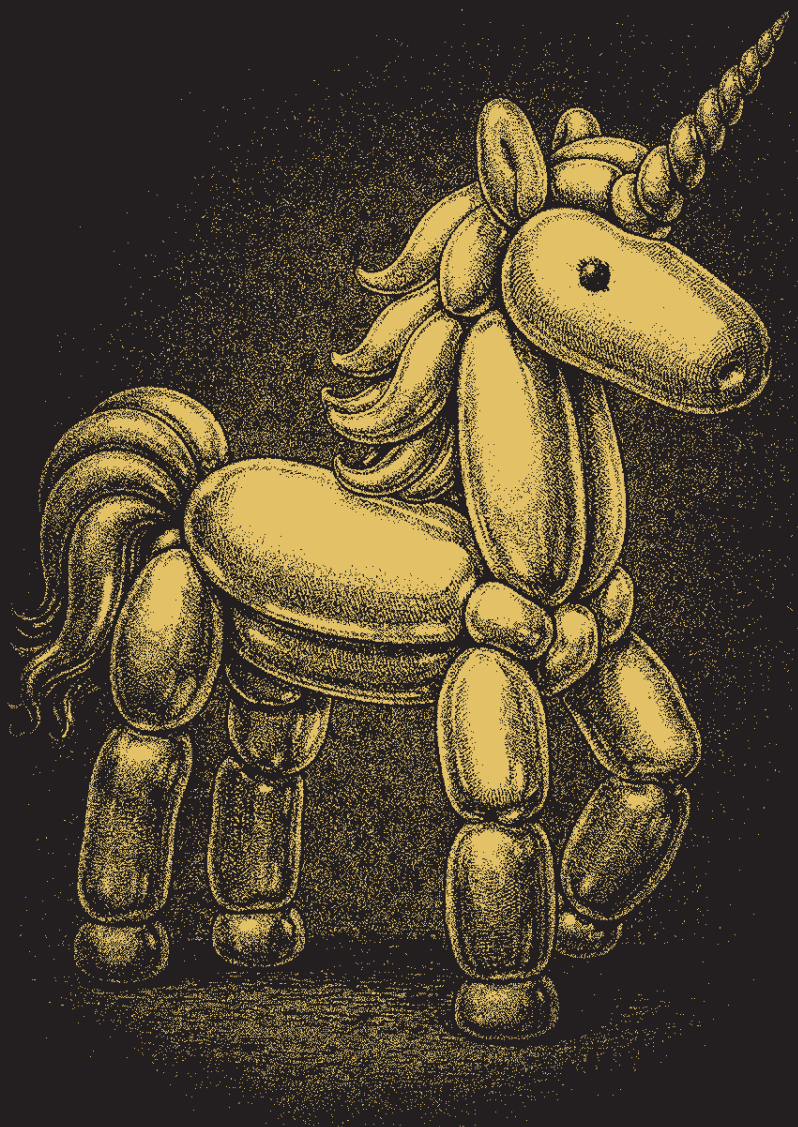
Evil agent capitalism

Today's capitalism is a fictitious cycle of capital's reproduction, yes. But the other side of it is pollution, enormous amounts of waste and pandemics. In the last 10 years Marx was reading geologists biologists who focused on material reality of the reproduction of the capital. How it disturbs the circulation of nature. Concerned with the question: what is the material reality of capital?

I think that the main reason that we don't need to include the hypothesis of a pandemic or global warming as an evil plot of such conservative forces is that even though both are products of capital, they are already products of the normal functioning of capital. Even without an "evil agent capitalism", capitalism produces

them. The problem is science definitely is in the service of capital. It is coresponsible; it is guilty of our troubles.

And I like to quote Richard Wagner's Parsifal on this paradox: "Die Wunde schließt der Speer nur, der sie schlug. The wound can be healed only by the spear which did it." That is to say that there is no solution without science, without capital.



Corvus corax

2

Yanis Varoufakis

Cloud capital

Alexa

While writing a book, I indulged in a lot of hardcore research: I bought an Alexa, a Google Assistant, and used Siri. Alexa is a little contraption from Amazon. It sits on your desk and is supposed to be your mechanical, automated slave. You tell it:

"Alexa, switch on the lights."

"Alexa, order me this book by this idiot."

"Alexa, play some soothing music because I'm in a bad mood."

"Alexa, remind me to buy milk or order me some milk."

Is it your slave? No! It's not your slave at all. By the way, Alexa doesn't really exist. It's just an interface connecting you to a cloud-based agglomeration of machinery. And the thing we call the cloud is not a cloud at all. We call it that because we don't see it. It's all on Earth, on huge server farms—with the exception of some satellites. Amazon's AWS server farm is monstrous. It really shocks the senses.

I visited a Microsoft server farm somewhere between Idaho and Washington State, and it felt a bit like a James Bond movie—like when he stumbles upon Dr. No's underground silos. It's in a cave, carved out of a mountain in order to keep it cool and maintain a constant temperature, full of machines making a huge noise. It's like a massive factory.

Alexa connects you to the cloud, which is essentially machinery—like capital goods. It's no different in many ways to a factory. And of course, there are thousands of kilometers of optic fiber cables crisscrossing the bottom of our oceans to connect this cloud together. This is how Amazon works. And so when you type in electric bicycles, you get this list. Alexa connects us to this network, but that's only part of it. The other thing it does is actually essential. Let's face it—it is sold to us at a very low price – Amazon loses money when it sells you an Alexa. Why? Because it's very lucrative for Amazon that you have an Alexa on your desk or in your kitchen. Essentially, what you do every time you tell Alexa

something, or every time Alexa overhears, you say something to your brother, to your sister, to your lover, to your wife, to your kid—you are training it to know you.

Cloud fiefdom

What is fiefdom?

Jeff Bezos doesn't care what you buy because he produces nothing of what is being sold on Amazon.com. Essentially, what he has built is a cloud fiefdom. Why do I call it this way?

What was a fiefdom under feudalism? A fiefdom was an estate with well defined borders owned by the landlord. Inside, there was a lot of production, but there were no markets. There were peasants who worked, but there was no labor market.

Peasants didn't get a wage. They simply produced, and at the end of the day, the landlord would send the sheriff along and would collect a percentage of the harvest, which varied depending on the relative fear that the landlord had, that the peasants would rebel. So you didn't even need an economic theory to understand the percentage of the harvest which was extracted by the landlord in the form of ground rent. Whether you were an ironsmith in a fiefdom, you were a vassal capitalist in the sense that you were completely dependent on the landlord, and the landlord charged whatever he wanted in the form of rent, or extracted surplus.

Free labor

Today, what does Bezos care about? He wants to encase all of us in the platform—a CloudFief. It's a form of cloud rent. So their profit is essentially siphoned off, and they extract something from you, which is absolutely unheard of in the history of the world so far, free voluntary labor.

Let's take Twitter—now X—as an example. When you post something there, you are adding to the capacity of the owner of Twitter to keep the audience on the platform.

If you think of that capacity as a form of capital, and that is what it is. All this machinery, it's capital goods. And we are contributing to it. Why? Because you're not contributing to the creation of the machinery. That is a standard old fashioned kind of capital goods.

The reason why I didn't leave Twitter when Elon Musk took over and fired 70% of the workers, and he turned the toxic medium into a cesspool of ugliness. And I'm still in it. Why? I've got 1,200,000 followers. If I go to the Blue Sky, I have 10.

Free labor, effectively, the decommodification of labor is a situation where labor is extracted without it being treated as a commodity anymore. When you do free labor on behalf of platforms, the cloud capital is not a commodity anymore. The same way peasants under feudalism were working, but the labor was not a commodity. So there is a process of dim commodification, which goes hand in hand with siphoning off the surplus value that is produced through commodification, and bringing it out of the secular flow of income. But to make it a bit clearer, there's no doubt that techno feudalism is parasitic on the old fashioned terrestrial capital sector. Essentially, it's a parasite that has taken over. It's sitting on top of the capital sector. The capital sector produces the bicycles. Right? But power no longer rests in owning the machines that make the bicycle, but it rests in owning the cloud capital that allows the cloud analyst to extract all the surplus value from the capital sector. So the capital sector is essentially no longer in control. The new ruling class is not the one that owns the old fashioned capital.

You're working to increase the attractiveness of the Twitter algorithm and its capacity to reproduce itself—its cloud capital. This has never happened in the history of humanity before. Let's stay with the word "capital". What do economists mean by this term? They don't mean "money". These big tech platforms, all platforms for that matter, not just big tech, is that we produce a great deal of the capital, of the cloud capital, without being paid. It

doesn't matter if it is voluntary labour, it is still free labour which is producing it. Every time you upload a video on TikTok, you add to the cloud capital of TikTok. Every time you upload a photograph on Instagram, you are making Instagram more attractive for people who follow you. Every time you rant on Twitter, you send a Snapchat message. Every time you go on Airbnb, every time you go on Uber, or any of these taxi platforms, you are adding to the cloud capital of the company that owns it. Uber makes more money out of the data that it collects from you than it does from taxi drivers. And, of course, the taxi drivers and both you get none of that. The taxi driver and you both are adding to the cloud capital of Uber, and Uber gets this for free from you. Right? You pay for the taxi ride and you give away your data for free. And it's not just data. It's work that you do. Every time you like a post or you give some information to the app, you are replenishing the cap.

We are now facing a world in which a new form of capital has emerged, has evolved without us even noticing. These algorithms are not just technologies. They're a new form of capital, cloud capital. And why cloud capital? In what sense is it profoundly different to old fashioned capital, like steam engines and industrial robots and fishing rods?

Capital is a produced means of production. Cloud capital is a produced means of behavioral modification. It doesn't produce anything except modifying your behavior. It's an interface between you and Alexa, you and Siri, you and the machine. It teaches you to make it learn even better about you so that it can do five things that traditional capital could never do.

What is cloud capital and why is it different?

The death of liberal individual

With the trail that you leave now behind you on Instagram or TikTok or any social media, even if you're not conscious of it, you know in your subconscious, that when you get interviewed for

a good job at Google or Siemens or wherever, your whole social media history is going to be scrutinized. Especially if you're not conscious of it, but you're subconscious of it, which means what? When you are at home and you're in your bed and you send a message or you upload a video or something, at the back of your mind, you think, at some point, maybe that video is going to determine whether I get a good job. When you're lying on the couch, you're curating your future self by trying to second guess what the Google personnel manager will want from you. That is the death of the liberal individual, where the liberal individual is the idea that you are autonomous, the creator of your own self, of your own desires, of your own thoughts, and there is a limit, a demarcation, a fence dividing you and the market. The liberal individual has been totally wrecked. It has been bulldozed over by cloud capture. If you go to one of those big tech companies like Meta or Google, you'll hear them say in interviews: *"We want you to be yourself, to think outside the box. We want you to be contrarian. We don't want you to think like we do."*

And then suddenly, you think, *So how should I think?*

In the back of your mind, you're asking yourself, *How can I think in a way that's contrary to their thinking, but still in a way they'll appreciate?*

And that's how you fall down a rabbit hole—where mental illness starts to seem like the only plausible outcome.

Cloud Capital

Up until now, capital simply enhanced labor productivity. There was a company making pins, using only workers with hammers and chisels. Labor productivity went up by adding machines—that's what capital had been doing until now. It was essentially extending the power of labor. This new form of capital—cloud capital—can now do five things that traditional capital never could.

First of all, it captures our attention. How do you keep some-

body on the screen? By getting them addicted. You can get addicted in a variety of ways: pornography, cage fights, all sorts of horrible things. For example, there's an accident on the motorway and as you drive by, it's difficult not to look.

When it comes to the news, the most efficient way for engaging us is to get us angry, to poison our soul. Because only when you're angry, you stay stuck to the bloody screen. So it doesn't really matter what the politics of the owner are. The result will be poisonous for democracy. This is a new dimension. It goes beyond the politics of the owner. There's no specific interest. The interest is in keeping us completely and utterly besides ourselves with anger.

The moment it succeeds in creating that need or desire inside you, it immediately satisfies it—at the click of a button, it comes to you. So, to recap: first, it captures your attention.

Secondly, it creates your desire. Thirdly, it serves your desire outside the marketplace. Because Amazon is not a marketplace. Once you've clicked on the Amazon button and purchased the book, the electric bicycle, the pair of binoculars—whatever product, produced by some other company that pays a huge cloud rent to Amazon—it's sent to an Amazon warehouse and then to you. Inside the Amazon warehouse, the workers have something trapped onto their wrist. This thing gives them their next job: go to a specific location in the warehouse, pick up a box, and then follow instructions on where to take it. And it also times workers. It knows when they've been to the toilet, how long they spent there, who they were talking to on the phone.

Facebook is an empty shell without you. The real capital of Facebook is everything you write on it. All wonderful essays and beautiful photographs and fantastic poems and so on. But everything we put on these platforms is our contribution to the cloud capital.

They get their cloud capital for free from our free labor. It

doesn't matter whether we love doing it. Why doesn't it matter? Adam Smith, who started something called economics with The Wealth of Nations, and later the other giant of economics, David Ricardo, who further developed the field, were very clear: they liked capitalism because capitalism had replaced rent with profit. They hated rent, and they adored profit because it's what made capitalism worth having.

Adam Smith story

What was Adam Smith's great philosophical idea? That greed is good. Adam Smith, a professor of moral philosophy, said to his students that greed is despicable on a personal basis. But it is good if you have a competitive market because each one of them is trying to undermine every other producer, every other capitalist. They try to steal their consumers from each other by producing better and cheaper products. And in the end, society manages to do something very weird to produce public virtue out of private vice and greed. But that happens only if rent dies and profit succeeds. Because rent is what you get from owning something, not from doing something. But if you're an entrepreneur and you make bicycles, then you constantly worry that somebody's going to make a bicycle which is nicer, better, and cheaper. So whatever profit you make, you have to invest it back into the business. This is the Adam Smith story.

At the turn of the 18th century, the great transformation from feudalism to capitalism was underway. Under feudalism, wealth came from land and was collected by landowners as rent—cathedrals, for example, were built with that rent. Capitalism shifted the source of wealth from land to machines. Ownership of capital now gave access to profit, which, according to Adam Smith, was a social good—competition would drive profits into reinvestment, leading to better and cheaper goods.

When Smith published The Wealth of Nations in 1776, feuda-

lism still dominated. Governments were run by lords and barons; entrepreneurs were rare and seen as low-class. Merchants had no place in Parliament or the royal court. The first real advocate for capitalism in the British House of Commons was David Ricardo—not because he was a merchant, but a landlord who crossed class lines to support the new system.

Techno Feudalism

Cloudalysts

The Central Bank must print more money to replenish the value of economic energy that is being siphoned off by cloud capitalists—or 'cloudalysts.' Cloudalysts are capitalists who own cloud capital. And even if you never touch a smartphone, never shop on Amazon, or use Twitter, you're still living in what I call techno-feudalism.

In my book *Techno Feudalism*, I define this system as the result of a striking contradiction: capitalism was killed by capital. Capital overthrew capitalism and replaced it with what might be seen as a new form of capitalism. The cloud capital which killed capitalism was created, funded by the state. But there's no conspiracy here. When they were printing all this money, they were not doing it in order to fund Cloud Capital. It just happened. Unintended consequences.

The most important companies of the last 10–15 years don't make profits. Look at Uber, Airbnb, Netflix, Spotify, Tesla—they generate very little profit, if any at all. What drives them is the stock exchange, where their shares are soaring. Why are their shares going through the roof? Because central banks have been printing money to ensure that financial markets remain healthy.

There's some other strange non-capitalist aspect about cloud capital and big tech. During my research for the book, I chanced upon a fascinating statistic. Did you know that $9 out of every $10 spent on cloud capital by Zuckerberg, Google, and others was

actually public money—printed by central banks during the post-2008 attempt to bail out the banks?

What we have here is a new form of capital, which I call command capital. It has given rise to a new ruling class—the cloudalysts—with traditional capitalists becoming a vassal class. There is a replacement of profit-making with central bank money, which creates new ways of telling people what to do.

Because if you are Elon Musk, for example, your capacity to make the world go round draws from the value of your shares in Tesla, SpaceX, etc. Those values are not closely tied to profit at all. They are linked to state money as well as to the *command capital* of these companies.

So capital may indeed be everywhere. It's clearly on a rampage—on a triumphant march—flattening democracies, social capital, and labour, and creating an ever-burgeoning precariat and a vassal capitalist class.

Socializing cloud capital

Once you enter Amazon.com, you have exited capitalism. Welcome to techno feudalism the moment you enter that realm. But even if you don't enter Amazon.com with your Nokia phone, you are still part of it, because this is the world we live in now. If the first architects of the Soviet plant economy in the nineteen twenties, particularly in the nineteen thirties, were alive today, they would be kicking themselves.

This is what we now have between Walmart, Amazon, Alibaba, and several others. Each one of them is a fiefdom, and each one of them keeps us all in there, by ensuring that our switching costs are very high. If you already have a portfolio with Amazon, you get some discounts, you know your way around Amazon. You need to be messed up really badly in order to move. But even then, I gave you the example of Twitter—I will never move from Twitter because then I will be talking to 10 people. There are things

that can happen in order to reduce this power of the techno feudal lords—cloudalysts. But changes require some sort of a revolution.

The only way we can reclaim our mind, the only way of losing our mind change is by socializing cloud capital. We need cloud capital. Cloud capital is extremely useful. But why can't we have an Amsterdam app instead of an Uber app or a Mercedes-Benz-owned Free Now app? An Amsterdam app, where the municipality owns the app, that you can use instead of Airbnb to rent a room if you are a visitor or a local. You say, "My name is Sergeant Sajid. I'm here and I want to go to the airport. How should I go?" And the app says to you, "There is a very cheap public bus that can take you. Or, if you want a driver, Joe can take you."

With all the surplus, like, cloud rent being retained in Amsterdam. Why can't we have that? But that requires the social ownership of the platforms, not their annihilation.

There's no doubt that in the future, once we have socialized cloud capital, blockchain technology will be very useful, very useful to replace the privately owned cloud capital that we have now. But blockchain will not bring about that transformation on its own, nor will it create a world in which it can function meaningfully. For that, we need mass democratic politics because the power of terrestrial capital and cloud capital is massive, and it will crush any attempt based on technology. We cannot find a technological solution to a societal problem, only a political solution. Once a democratic revolution has taken place, we can then make use of blockchain and many other tools.

A digital bill of rights can be implemented, where you own your data and must be paid for it. But for that, you need a massive political revolution. Technically, it's really very simple. The solution is not difficult. Politically, it's almost impossible.

Bitcoin, for instance, will never be money and should never be money. It's toxic, absolutely toxic and irrelevant from the perspective of the macro society.

We are politically irrelevant. Wars are happening, we have no opinion about anything, except to go along with whatever Washington DC tells us to go along. And you may think, if you're saying that techno feudalism depends on the capitalist kernel within to produce the surplus value that it extracts, does this mean that it's still capitalism? No. Because if you think about it, after the great transformation of feudalism to capitalism, capitalism, the owners of the factories, the factory sector, were still absolutely dependent on agriculture, on the feudal sector.

China

To understand the direction China is heading, one must consider the internal dynamics of the Chinese Communist Party. Predicting how the Party will evolve is a difficult task, largely because the process is both highly complex and deeply opaque. Over the past few years—and particularly during my most recent visit to China—I've come to realize that a significant internal struggle is taking place within the Party's Central Committee. Contrary to the common perception of the Party as a unified monolith, it is, in fact, a site of intense ideological and strategic conflict.

Outwardly, yes, they present a unified front centered around Xi Jinping. But here's something that may sound controversial: I believe the internal dynamics of the Central Committee are far more interesting than what's happening in many Western parliaments. Within the committee, there are members who represent workers, traditional capitalists, cloudalysts, and even local communities or movements. The tension and conflict between these factions are fascinating—and crucial.

What makes this even more concerning is the opacity. If you're on the outside, you simply can't see what's going on inside. And opacity is the beginning of tyranny.

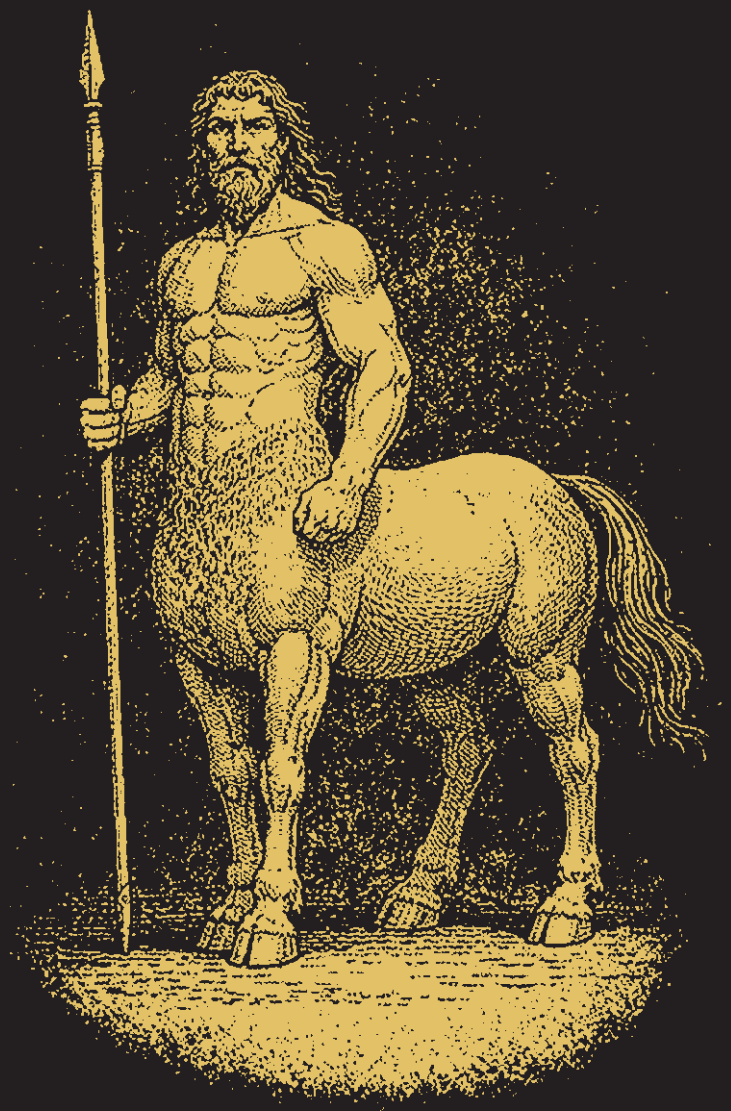
As for where this is heading—whether it will evolve into something more open or democratic—I don't know. The Chinese project

is still very much incomplete. It's a nation in flux, and there's no scientific way of predicting how it will unfold.

When it comes to techno-feudalism, what's happening in China today is that the government has legislated compulsory interoperability among major tech companies. What does this mean? It would allow me to switch from Twitter to BlueSky, for example, just to annoy Elon Musk. But the solution isn't to go offline—because even then, you're still connected, still subject to the vagaries of techno-feudalism. I don't believe in simple escapism.

Moreover, regarding AI implementation, it's not just about behavioral modification as it is in the West, where that seems to dominate. In China, AI is being far more deeply integrated into production. While it remains only a fraction of the Western AI ecosystem, its direction is markedly different.

What keeps me up at night these days isn't just techno-feudalism, which has haunted me for a decade. It's the growing uncertainty in the global balance of power and the systems that are now shaping it.



Centaurus

3

Susan Schneider

The future of the mind

A thought experiment

It's 2045, and you're out shopping. Your first stop is the Center for Mind Design. As you walk in, a large menu stands before you, listing brain enhancements with funky names. Hivemind is a brain chip that allows you to experience the innermost thoughts of your loved ones. Zen Garden is a microchip that induces zen-master-level meditative states. Human Calculator is an implant that gives you savant-level mathematical abilities. What would you select, if anything? Enhanced attention? Mozart-level musical skills? You can order a single enhancement or bundle several together.

So that's the scenario. As a philosopher, I like thought experiments because they really get our imagination going. This text is about the future of your mind, and how our understanding of ourselves, our minds, and our nature can change the future for better or worse. What I'll be asking is, at a place like this hypothetical Center for Mind Design, could you truly merge with artificial intelligence?

This text explores key issues surrounding transhumanist ideas, neurotechnology, and the age-old philosophical questions they invoke. It begins with an examination of the foundational concepts, providing the necessary context for understanding the broader implications. From there, the discussion shifts to the potential limits of human intelligence augmentation, particularly in relation to consciousness and personhood. Finally, the chapter concludes with a reflection on these advancements within the larger framework of our place in the cosmos.

Mind design

Let's turn to some quick background. Brain evolution was heavily constrained by environmental, anatomical, and metabolic demands, but AI-based brain enhancement technologies could augment intelligence at a rate that's much faster than biological evolution. I call this new enterprise "mind design".

It's a kind of intelligent design, but we are not some Gods pur-

porting to be the designers. Our social development often seems to lag behind our technological capacities.

Now suppose you're standing there at this Center for Mind Design. You're waiting in line, and somebody comes up to you, and they say: "You know what? You should come back in a few weeks because if clinical trials go as planned, customers will soon be able to purchase an enhancement bundle called MERGE, a series of enhancements allowing you to gradually augment and transfer all of your mental functions to the cloud over a period of five years". So could you do this? Could you really begin to remove parts of your body? Could you really begin to remove parts of your biological brain and transfer your mental capacities into the cloud, the Internet, or into another computer? And would that be you? Would you survive?

So what does it mean to merge with artificial intelligence in the first place and why should we worry about it today?

Transhumanists believe that humans can overcome their biological limitations through science and technology. The idea itself shouldn't seem objectionable to most of us. However, the devil's in the details. The transhumanist trajectory for enhancement is much more controversial. So they start with this idea of an unenhanced human. The unenhanced human has significant upgrading with cognitive and other physical enhancements.

So perhaps you get the ability to have Mozart or Philip Glass level musical composition, or you get the ability to expand your attention so that you can focus on hundreds of things. After replacing so much of the biological brain with brain chips, you become superintelligent AI. You're no longer technically human. You're posthuman.

The idea of neural replacement

A superintelligent AI is a hypothetical form of AI that outthinks humans in every way possible: scientific reasoning, social intelligence, and more. So this kind of trajectory is the kind of view that's held by a lot of transhumanists such as Nick Bostrom, Elon Musk, Ray Kurzweil, who's now a chief engineer at Google, and the late Stephen Hawking.

I want to focus on one aspect of transhumanism that I find to be most problematic. And that is the idea that augmenting or enhancing intelligence involves replacing parts of the biological brain with AI components—the idea of neural replacement. That includes brain uploading or a gradual process in which parts of your brain are outmoded and transferred over to computer chips or computers.

Transhumanism is a very influential view, and it connects in a really interesting way with the classic philosophical issue—does consciousness transcend the brain? There is a similar view, in cognitive science where it says that the mind is something like a program. But I consider this idea flawed and believe we should think of these issues to not commit what the historian Michael Best called the Jetsons fallacy.

You may have seen *Star War*s or the cartoon The Jetsons and noticed in these science fiction stories, humans are depicted as not being enhanced, but they're surrounded by all of these robots and self-driving cars. But if you really think about what the future is like, it would be sort of surprising if all kinds of artificial intelligence is all around us. We have sophisticated robots taking care of our homes and autonomous vehicles driving us places, but we don't use that very same technology to augment our own intelligence.

In fact, that's already what's going on, with the military. There are various programs in the United States to create super soldiers.

Elon Musk established a company called Neuralink. He hopes that humans can keep up with artificial intelligence by having some sort of merger of biological intelligence and machine intelligence. And to this end, he created Neuralink, which already has

implantable chips allowing data to travel wirelessly from the brain to digital devices.

There is also a project that is very far along in clinical trials in the United States to build an artificial hippocampus. This is Ted Berger's fascinating project. The hippocampus is the part of the brain responsible for forming new memories and is crucial to memory formation. Facebook is working on this, as are Google and Kernel. In fact, the list of companies involved is quite long.

What does all this mean for the human future?

Let's think about these issues and the future of the self and mind from a philosophical standpoint. The transhumanist view on this topic is that we should engage in these kinds of experiments, and people should be free to augment their intelligence at a place like the Center for Mind Design. One reason is that it is exciting to imagine being free from unwanted psychological elements: to sculpt our own character, to remove chronic depression, and to make deliberate decisions about our psychology. Second, this offers another route to longevity and a possible path toward intelligence augmentation and merging with artificial intelligence.

But on the other hand, it doesn't take too long for us to feel deeply concerned that our thoughts could be part of a computer system connected to the Internet, given what's happening with social media companies using our data to manipulate us and selling our private information. It could be a complete dystopia—a thought data economy. There are many highly negative examples, such as Facebook, which utilizes people's personal information and employs social psychology techniques to manipulate users and foster addiction to their platforms, without regard for the impact on different countries and groups.

A surveillance economy that possesses neural data is highly dangerous. It is also a serious threat in non-capitalist systems, such as authoritarian dictatorships like China, where the technology could be used to monitor and control dissidents. Now, let's turn to the philosophical issues beyond ethics. Even if we were able to establish privacy laws and regulations, there are still reasons to believe that the transhumanist view of merging with AI may be unattainable.

Design ceilings

There are limits to human intelligence enhancement that are not imposed by evolution, medicine, or technology. There are also philosophical limits—perhaps one reason why those developing brain chips haven't fully considered the implications. Of course, there will be other limits as well, such as neurotechnological and medical ones. But what kind of limits do I have in mind?

The first limit is what I refer to as the 'consciousness ceiling'—a boundary that emerges if microchips prove incapable of supporting conscious experience. The second, the 'self ceiling,' marks the point at which an individual seeking enhancement undergoes such profound transformation that they cease to exist as the person they once were. In this case, the procedure does not simply modify the mind; it fundamentally alters the identity of the one who undertakes it.

Consciousness is the felt quality of experience. So when you see the rich use of a sunset or you smell the aroma of your morning coffee or you stub your toe, it always feels like something to you from the inside. And nobody else knows exactly what it feels like to be you. You can try to communicate it, but it's really a private thing inside of your head. It's what it feels like from the inside to be you. You're conscious all throughout your waking life and even when you're dreaming.

Richard Dawkins and I appeared together in a film called *Super Sapiens*, in which he made a particularly provocative remark: "It's not obvious to me that a replacement of our species by our own technological creations would necessarily be a bad thing."

You might find this statement open-minded—and in a way, it is. It's not necessarily a bad thing. I often hear people make this claim, suggesting that humanity is following an evolutionary trajectory in which we eventually give way to something superior. Perhaps a being with greater forms of conscious experience, one that transcends our flaws and surpasses us in ways we cannot yet imagine.

This could be the direction of the future, but let's think about it in greater detail. I wouldn't encourage you to think about whether machines can be conscious. Because if nonconscious machines supplanted biological intelligence, then this singularity, this idea of a technological artificial intelligence explosion that makes our lives better is actually a nightmare. It wouldn't be a transhumanist utopia the way people like Elon Musk and Ray Kurzweil discussed. Instead, it would be the end of consciousness on earth. And this issue is not explored.

Let's go back to the Center for Mind Design and remember the thought experiment. Suppose you have the opportunity to pay thousands of dollars to purchase an enhancement bundle called MERGE that allows you to gradually replace parts of your brain with microchips, eventually leading to your being uploaded to the cloud or to your brain being entirely replaced by microchips.

If AI is not conscious, then merging with it would be a bad idea, because if microchips are the wrong substrate for consciousness, a mind-machine merger would not preserve it. If you tried, you would lose your consciousness, and your mind would cease to exist. That would represent a clear limit to human intelligence augmentation.

If this is true, we will not be able to surpass or even match the intelligence of artificial intelligence in the future. As AI becomes increasingly advanced—potentially reaching and exceeding human-level intelligence—we will be unable to keep up if a consciousness ceiling exists.

Now, let's consider another potential limitation on intelligence enhancement, one even more significant than the consciousness

ceiling. While it may be possible to develop microchips that support consciousness, the concept of self-ceiling is beyond the reach of scientific solutions. It is a purely philosophical issue.

For the sake of discussion, let's assume that the consciousness ceiling does not arise. Should you proceed and invest further?

To determine whether you should enhance yourself, you must first understand what you are to begin with. But what does it mean to be a person or a self? Would you continue to exist if parts of your brain were replaced with microchips, or would you unknowingly end your own existence, only to be replaced by someone—or something—else?

To grasp the depth of this problem, it is helpful to explore the literature in contemporary philosophy. These questions trace back to thinkers like Locke, Hume, and Nietzsche.

Recent philosophical thought explores the metaphysics of everyday objects. Consider an espresso machine: if it's unplugged, it's still the same machine. But suppose I use a futuristic, science-fiction-style gun to disintegrate it into dust. Is it still the same machine? You would likely say no—it no longer exists. And now, you'll never get your coffee.

Notice that certain features of the coffee machine are essential for it to continue existing. The same applies to us. Suppose you are religious and believe in a soul—then the soul would be essential to your continued existence. Or, if you believe your existence depends on your brain and nervous system, then if your brain were destroyed, you would recognize that having a brain is essential to your survival. Without it, you would cease to exist.

This simple yet crucial observation has often been overlooked in discussions about the future of the mind.

If brain enhancement makes a person super-intelligent, that may seem exciting. But it cannot come at the cost of eliminating any features essential to your survival. Otherwise, you might walk

into a mind-design center—but you won't walk out. You will have unknowingly ended your own existence.

You cannot truly merge with artificial intelligence. So even if the technology works, you would be paying for a smarter mind or a fitter body—but it wouldn't be you. It would be something else, perhaps a digital twin. But it wouldn't be you. You would be gone. That would be a disturbingly perverse direction for technology to take. Medicine is meant to help people, and such an outcome would be horrifying.

I call this the self ceiling. It marks the point beyond which a person who seeks enhancement is no longer the same individual, as the procedure causes the original self to cease to exist.

The nature of the self

It's a philosophical question about what the nature of the self is. It's a question that's been debated since the beginning of philosophical thinking. And we probably all have a sense that it's very difficult to prove any one position on the nature of the self, soul, or mind because it's very controversial, and it's not like you can run an experiment to find out if there's a God or to find out if the mind is a program.

One way to describe the issue is to indicate that there are lots of different theories out there on the nature of the self or person. There's materialism—the idea that you are essentially your brain and nervous system. There's what's called the psychological continuity view, which is a view held by the philosopher Locke that says that you are your memories and your ability to reflect on yourself and your overall psychological configuration. The transhumanists have a version of that view that they tend to endorse. That is one called patternism.

Patternism is a view that was articulated by Ray Kurzweil and also Nick Bostrom in their work. It says that what is essential to you is your computational configuration, the sensory systems

that your brain has, the association areas that are responsible for integrating the different subsystems of the brain, the neural circuitry making up your reasoning abilities, your memories, and so on. Together, these form a sort of algorithm that describes how your brain computes, and it's unique to you. It's your pattern.

There is also a view that we have souls, and obviously many religions hold this position. And there's also a religious and philosophical view that says that the self is an illusion, that there's no underlying self there, and there's also no surviving person from moment to moment.

The important thing to note is that each of these views has an answer. For example, For example, the materialist view would say you'd be killing yourself and you shouldn't do it. Instead of developing brain chips that replace parts of the biological brain or trying to upload your intelligence, if you want to enhance, you should take a different path forward. Instead, the view here would be that neurotechnology should develop biological brain enhancements and minimal AI enhancements that don't replace or damage key parts of the brain.

This is not a conception of the future in which we merge with artificial intelligence. If AI ultimately outthinks us, we would simply be unable to keep up. If you have a soul theory, on the other hand, it's not at all clear whether you should enhance. That would depend on the details of the religion, so you would have to have discussions with your religious adviser. Well, what if you have no self view? You never merge with AI because there's no you. But you could strive to enhance that view. Because if you truly believe you don't survive, you'd be open to the possibility of replacing your brain with microchips. The trouble is knowing which view is right, having the certainty required to risk your life.

What about patternism? Here's the problem. When does the pattern begin, and when does the pattern end? Maybe deleting some bad chess-playing habits or uploading AlphaGo is okay?

Maybe you can get away with that? You wouldn't change your whole nature. But what if you bought merch, or what if you purchased several different enhancements? At what point would you no longer continue to be you? And how could we find out scientifically? How could we find out with certainty?

This is a philosophical issue, and I doubt that the science of brain chips will resolve it. That's why I adopt a stance of metaphysical humility. Claims involving mind transfer to a new substrate—such as a computer—or drastic alterations to the brain must be scrutinized carefully. Bear in mind that there is intense debate within the philosophical literature.

The future of intelligence

Suppose alien intelligence exists. Based on astrophysical projections, Earth is a relatively young planet, meaning an alien civilization could be 50,000 years more advanced than us. They would have already developed artificial intelligence and augmented their own intelligence. Planets across the universe may have already grappled with the same questions I raised earlier. When considering the future of intelligence, we should ask: Could intelligent aliens actually be forms of artificial intelligence? And if so, are they even conscious?

It ultimately depends on the mind-design choices that a culture makes. If design ceilings exist, then the issues I raised earlier function like philosophical laws—constraints on biological systems that may prevent them from augmenting in ways that keep pace with artificial intelligence.

This presents a different perspective on evolution. When discussing the future of intelligence, we are no longer in the realm of Darwinian evolution but rather in that of intelligent—or even unintelligent—design. In this new realm, humans and other biological beings attempt to augment intelligence, but such evolution will still face constraints. These include design ceilings, economic limitations affecting tech companies developing microchips and AI systems, and regulatory constraints imposed by AI laws.

Columba livia

Riccardo Manzotti
Can Machines think?

The question

Can machines think? It's a classic question. The straightforward answer is yes—they already do. At most, we listen to someone's words, but we have never seen a thought.

What about our own thoughts? We say: "I think," "I'm worried," "Last night, I couldn't sleep because I had so many thoughts—about my mortgage, about my car that was parked where it shouldn't have been," "The thought of the car kept me awake."

But is that true? Was it really the thought of the car, or was it the car itself that kept one awake?

When we say that we think of something, what is really taking place? Can we describe what a thought is without any reference to what the thought is about? What about a naked thought, an empty thought? Is it something, or is it nothing?

It is widely accepted that thoughts are always thoughts of something. We cannot think of just anything. If we think of nothing, we do not think. So whenever there is a thought, there is also an object of the thought.

Thoughts by themselves—naked, empty, contentless—simply do not exist. There is no way we can conceive of a thought without its reference, without its object. And that is something intriguing.

Because it suggests that maybe this very notion that we think—like *The Thinker* by Rodin—might be a form of superstition. In the past, there were animistic superstitions, and people attributed all kinds of magical properties to objects that were invisible. Today, it might be that the very superstition we need to get rid of is ourselves.

The inner world

The idea that we have an inner world that nobody has ever seen is something to think about. How is it possible that we all believe in something no one has ever seen? And when people say, "Look inside yourself," how do we know in which direction to look?

This is suspicious because in the past there were many anthro-

pological and philosophical studies that explored metaphorical directions. For example, the concept of heaven has been embodied in our language: "We feel uplifted," "We feel up," "We feel high." That direction is important in our language. But nobody believes that all those metaphors really mean that there is something upward apart from Elon Musk's satellites.

In Italy, there's a beautiful painting by Tiziano with the Assumption of the Virgin Mary. According to the Catholic religion, Virgin Mary never died. She had been carried up into the heavens before. In this beautiful painting of the 1550's, angels are physically lifting the body of the Virgin Mary in the sky. At that time, people took that direction, the idea that heaven is upward. Today, even people of a religious background no longer believe that heaven is physically upward and that hell is physically downward. But we still have one metaphorical direction that we all still believe in. This metaphorical direction is inner—the inside.

Where do we get that idea? There is the exact date when we started to think about the inner world. It happened in Milan, Italy. At that time, the Roman Empire wasn't in particularly good health. A famous saint, Saint Augustine, wrote the first book in first person. The book was called "Confessions". In the book Saint Augustine plays a philosophical and literary trick. Christianity, until that point, didn't have the immaterial soul. They believed in the resurrection of the physical body. Saint Augustine had the problem of putting the soul somewhere. And he had the problem because in the physical world, there was nothing like a soul. In the physical world, everything is filled with physical stuff. If we open a body, we find physical stuff: organs, cells, blood. There is no space for a soul.

Saint Augustine was very smart and felt that the word inside didn't work for what he wanted to find because he needed to find an empty space that was outside of the physical world. What did Saint Augustine do? He took the Latin word for inside "internus" and made the comparative, the greater, which is basically "more in-

side". Then he made a noun of the word that in Latin means "more inside". This word became the word that today we use in many languages: in French it is "interiorite", in Italian "interiorità", in English "interiority". But it is most commonly used with the expression "inner world".

Saint Augustine invented the idea that we have an inner world. "Inner" means that it is more inside than everything else which is already inside. It is the idea that we have a kind of niche, or shelter, or bubble that is detached from the external world. And the idea had a big success because people liked to have a place where they could be safe from all the harsh realities of life.

Ste same idea had been developed even more by René Descartes during the Thirty Years' War. He said that we are something that thinks. We are not sure about the world, but we are sure that we think. And how can we be sure of something that cannot be seen or demonstrated in any way?

Why shouldn't a machine think?

That's why many years later Alan Turing asked the very question we started with: How is it that we think, and why shouldn't a machine do the same? It is surprising that Alan Turing in this paper was also addressing theological issues. When he wrote that very famous paper, his dean complained about Alan Turing and told him not to write anything like that again—"This is not serious science". The dean was Charles Darwin III—grandnephew of the great Charles Darwin.

Today there is a machine that has reached, in many ways, human capacities. The last version of GPT has actually a little bit more—1.8 trillions of parameters. The estimated amount of connections we have in the temporal area of the brain that deal with the language is between 1 and 10 trillion connections. Other parameters are quite interesting too. The 2000k conceptual space of these large scale-language models. Why is it worth mentioning?

Because the number of conceptual spaces in the abstract mathematical space of language models is just like the number of words we have in human languages.

In a very big dictionary, we have more or less that number of words, and that is not by chance. The reason is that these parameters are beginning to reach exactly the same kind of complexity from a quantitative perspective that our brain has when dealing with language. AI today is what Isaac Newton would have described as a philosophical experiment. Actually, Isaac Newton described his own work as experimental philosophy, a way to test experimentally our core intuition about ourselves and reality. Today, AI is doing exactly that.

When we ask the question "Can machines think?", it is time to ask the mirror question... Because we are the first mirror. We reflect. We speculate. We are a mirror of the world, and AI is a mirror of ourselves.

We may start to wonder whether the fact that AI is able to do the same things that we do with language tells us something very deep about ourselves. We have always believed that we have an inner world, which is the modern version of the soul. Do we really, or does AI allows us to take a sharper turn and to get back to philosophy.

Getting back to philosophy

The goal of philosophy is always to get back to the origin, and then to change our insights about reality. It's not to go forward, but to get back to the origin of everything. What is the origin of everything in this case? It is a famous dialogue written in 390 BC by the only philosopher that Plato didn't mention by name—the Stranger. It's quite intriguing that there is someone that Plato didn't want to mention explicitly. Why? Why was this guy so dangerous for Plato? Because this guy was against the notion of an inner world, was against the notion of ideas, was against the notion of thoughts

in a way, thoughts as abstract or formal entities that are not in the world. This guy, the Stranger, is usually referred to as the Eleatic philosopher.

They were philosophers coming from Elia, who were against the idea of abstract or mental ideas. They said: everything takes place in the world, everything is an object. We do not need abstract objects; we already have real objects. What is our body then? Our body becomes a way in which we empower the world. Our body becomes a way through worlds that are real: not inner worlds, but actual worlds. They are my car, not the thoughts of my car. They are able to become causally effective. The principle that this philosopher claimed against Plato was that something is real only if it produces an effect.

There is a male only if there's a female. An object is a key only if it unlocks a lock. Anything is what it is only if it is able to produce a real effect. This view Plato put aside, because he wanted to find a kind of shelter for the mind, a cave, a refuge, a place where we could be safe from decay, change, and death. And Plato invented the idea of the soul. And the idea of the soul afterwards becomes the idea of the mind, the inner world. But is there a better reason to believe in an inner world than there is to believe in the soul? Actually, there is none.

After all, the paradigm shift brought about by this experimental philosophy, called artificial intelligence, is pushing us to question our beliefs about what we are—and that may actually be helpful. It may allow us to improve, to mature. It may compel us to take a step forward, no longer needing the idea that we are sheltered inside an invisible and unproven inner world.

The two most famous sentences written by human beings are: 'To be or not to be,' from Hamlet by Shakespeare, and "I think, therefore I am", by René Descartes.

"I think therefore I am"—is it true? Do we know this? We have no idea. Have we ever seen "I"? We haven't. Actually, one of the most

convincing models of the "I" is that it is the center of narrative gravity of our own existence, a concept that we develop during our life just to explain what we do. We see what we do, and we need to attribute what we do to a kind of center of narrative gravity. Luigi Pirandello, an Italian writer who won the Nobel Prize in the last century, suggested that there is no such thing as a hidden identity. Instead, he argued that we develop who we are by acting in the external world—by engaging with it—and that we are, in fact, one with the world.

Goodbye, inner world

With the help of AI, we can take into consideration the possibility of writing a love letter to our inner world. And we need to say goodbye to our inner world.

****

Dear loved one,

My dear interiority, my dear, my dearest inner world, I did everything for you.
My love for you was unmatched for so long. You were the love of my life. I know why I stayed with you for so long.
Our relationship was familiar, comfortable, and all that I wanted to know. I clung for so long to the hope that you would become the thing I needed you to be. I felt so good with my inner world, which shielded me from the harshness and dangers of the world. Moreover, thinking of myself as a thinking subject filled me with pride and narcissism. Thanks to you, I was the center of the universe, the only thing of value in my world, and I could imagine always having a safe haven to retreat to you, conscious mind, inner world, my beloved one.
I was afraid to lose you. I feared there wouldn't be anything better out there for me, but I realized that I was hurting myself by wasting my time. I am so sad and disappointed by our relationship ending. But the time has come to let you go, and in doing so, to enter the world, or rather to become the world, I'm no longer a who, no longer a thought, no longer an inner world, no more "I think therefore I am", but rather just I am. In fact, I am the thing that I am.
I no longer have a barrier between myself and the world, no longer the possibility to step back from reality. I am thrown into the world. In fact, I am the world. The cogito "I think, therefore I am" becomes Heidegger's being-in-the-world, or rather, it becomes being the world or even being a world. Farewell, safe harbor of my inner world. Farewell, ontological privilege. Farewell, adolescence, ontological adolescence that led me to believe I was special. Starting tomorrow, I will be a thing among things and I will have to confront the harsh reality. Leaving you, my mind, or thought, my inner world is painful, but it is necessary to mature and enter the world.

Goodbye, my adolescent love.

****

I wrote it this way because I believe we are mostly attached to the notion of an inner world for sentimental reasons. After all, there is not a shred of evidence for the existence of an inner world or for the existence of thoughts. Now what could this tell us? First, I will take advantage of the theatrical version of Galileo. In Life of Galileo by Bertolt Brecht. In that play, Galileo makes a very important statement. He asks: "What is the biggest mistake of science?" Mirroring the work by Thomas Kuhn in the structure of scientific revolution. "What is the biggest mistake that we make in science? It is to believe in something we don't really know. That's the biggest mistake."

We once believed that the Earth was at the center of the universe—but it isn't. We believed that human beings were a special

species among animals—alas, that wasn't true. We believed that space and time were absolute—alas, that wasn't true either, and so forth.

What is something we believe today that might not be true?

It is easy to look back and judge people from the seventeenth century—how foolish they were! But there were many reasons for them to believe what they did.

Today, we believe that in order to think, we must have an inner world. And because of this, we assume that for machines to think, they must also have one.

This thing is the idea that we are separate from reality, the idea that we are not the world. Let's get back to thinking. What is a thought? We don't have a positive way to either describe or to point to thoughts. We don't have a way to say that a thought is a thing which is made in a certain way, which can be seen using certain tools, which has a certain shape, which consumes a certain amount of energy. We cannot do anything like that.

There is neither a scientific way, nor a phenomenological way to refer to a thought. What is the thought of X, the thought of an object, the thought of a car? It is not an object. It is not a car. It is not an X.

And likewise, what is the self? What are we? We are not a thing. But to not be a thing means that we are nothing. We are not a thing. We are nothing. The subject, historically speaking, has not been defined on top of some positive evidence, experience, fact. It has been defined and built on top of the negation of reality because we didn't want to be the world. Because to be in the world is too painful. There are too many bad things in the world. The world decays. There is death, illness, injustice, all kinds of things that make us suffer. And so we want to have a barrier. We want to have a shelter. We deny the reality that we don't like, and we historically develop the notion of the inner world.

Where lies the difference?

Let's get back to Galileo. Galileo, in 1590, wrote a famous poem to explain how he was going to do his scientific revolution before doing it. He told us how to proceed when we have to deal with something unexpected. At that time, there was the Copernican revolution. Today, there is the AI revolution.

What was suggested by Galileo? In his opinion, anyone who wishes to discover something must deploy imagination, and play with invention and guesswork. And if you can't go straight ahead, a thousand other paths might help you along. Nature seems to teach us this: when one cannot follow the usual path, they take a different road. The style of invention is very diverse. But as I have tested in seeking the good, one must proceed contrarywise. And that's what I am suggesting to you. I am suggesting to you not to look for whether machines think, but to ask yourself what do we do?

Thinking is merely mirroring the causal structure of the world. This is done primarily through language, because language allows us to reflect the structure of reality. That's why it is a revolution. Don't believe the claim that AI doesn't think—because what exactly do people do that AI does not? When we say that inside AI, there are just statistical parameters, bias and probabilities mirroring the causal structure of language, exactly what do our neurons do that AI doesn't? Because one neuron is not much smarter than a conditional probability of one parameter inside AI. It is true that at this point, the causal structure of AI may not be exactly like ours. It is also true that at this very moment, the biggest limitation of AI is human beings because AI so far has had no access to the world. It has to go through our description of the world. But how much smarter would a human being be if he or she had no access to reality, but only to other human beings' description of reality.

Therefore, this compels us to take seriously Galileo's suggestion that we must consider entirely new perspectives. We must

proceed in the opposite direction because the world has changed dramatically.

If AI is doing what we do—namely, using language to mirror the causal structure of the world—then, as the English writer said: "How could I know what I think if I don't say it?" Likewise, how could we know what we are if we don't act?

If AI is mirroring the causal structure of the world and using language just as we have for the last twenty thousand years—if it has 'hacked the operating system of humankind'—then is there anything we have that AI does not? It is one thing that has two names. They are apparently different, but they mean the same thing.

We have freedom and values. A practical example is if we have to go back home with the car, what is the first thing that we would do? We would take your navigator, Google Maps, and we would ask Google Maps what is the best route to follow from point A to point B. Is that a value? Is that a choice? No. Because the true value comes before you have to choose whether you want to do the shortest, the briefest, the cheapest, the most ecological route. And that is the value on the basis of which a causal structure like that of AI is able to compute the best route.

That is knowledge. Knowledge is the capacity to use the causal structure of reality to reach a goal. We have a different capacity so far. We have the capacity to freely choose a goal. We have the capacity to choose what's worth pursuing. And that's what we have that at this very moment, machines do not have. We have the capacity to create value. How? Choosing it by acting with our own freedom. And this is a lesson that AI is teaching us. It's telling us that life is not just having knowledge. It's not just having the capacity to compute from a cognitive, efficient perspective the best route. It's not only the capacity to use language. It is the capacity to know why we use language.

Like Jack London famously said: "I don't want to waste my days trying to prolong them". There has to be something that we choose

for which it is worth using all the knowledge that we have. This is what AI today does not have.

We are not a negative entity. We are physical worlds that, by means of our body, are able to choose what they want. And knowledge is a tool, the most powerful tool, but it is a means. It is not the end.

Falco peregrinus

5

Finn Brunton
Cryptos

Affectus

Let's talk about money and about what Spinoza called "affectus". Affects—the rational activities that we can govern and the irrational passions which govern us. First, cryptocurrency began as a project to deliberately engineer society in the direction of some specific esoteric political goals, and that a lot of the weirder details about cryptocurrency make more sense when we understand that. Second, crypto now has broken up into some new interesting social formations that reflect our current predicament and also have connections with things like the meme stock movement among other things. "Affectus": two rational plans which created cryptocurrency, and the two irrational passions, which now dominate it.

The first rational plan—cryptoanarchy

The first of the two rational plans is something called cryptoanarchy. It was named and promoted by its impresario, a retired engineer at Intel named Timothy May, who was also a hardcore libertarian. He was part of a loose network of computer scientists, hackers, mathematicians, and activists who shared an interest in digital cryptography. In 1976, one of the most important papers in computer science the twentieth century, in terms of its social and public impact, outlined a breakthrough in how data could be encrypted, which would enable people to communicate securely and anonymously—provably secret not only from conventional adversaries but even from the world's most powerful surveillance and signals intelligence systems, including government organizations. Furthermore, with this new cryptographic toolkit, you could do things like sign a message in such a way that the message irrefutably came from you and was not altered in transit without ever disclosing your real identity. Many years before the Internet itself, these were the tools to secure it. Many of you are using them today, right now, without ever realizing it, that we can

use all of these extraordinary tools from 1976 is largely due to the community around Tim May. Tim May saw a different trajectory for their potential, not just to protect the privacy of people from digital surveillance by the state, but to destroy the state altogether. Encryption and digital money would produce anonymous networks for transactions and information exchange between participants who could trust each other as they built up irrefutable anonymous digital reputation scores, building up trust without ever knowing each other's real identities. Out of these networks would bloom black markets and information markets of untraceable, untaxable, unpoliced, and borderless exchange, which would simultaneously deprive states of their revenue and their capacity to control currency. This would then provoke governments into crackdowns, which would further alienate their populations into open revolt, leading to the collapse of governments and the emergence of a new global order, which Tim May called cryptoanarchy. The cryptoanarchist movement already had encryption. What else they needed was digital money, and digital money had two problems: a technical one—no one yet knew how it might be secured and authenticated while keeping its users anonymous; and a social one which was much more difficult. Why would people start using this digital money? What would make it valuable? Obviously, the crypto anarchists could not do what states do, which is to make money valuable by mandating it for the payment of taxes and the settlement of debts.

Tim May's solution was strange but elegant. He proposed a new cryptocurrency long before we use the word, which he called crypto credits, and he argued that it would be valuable because it gave you access to forbidden things. Cryptocurrency was the house currency of the global online black market. Crypto credits could buy drugs, weapons, oceans of pirated media, trade secrets, intellectual property, hacked records, classified information. May's theory was that, at a certain point, you would go to an open-

air market, and the merchant would provide you with a choice: pay in cash, on Venmo, or anonymously send some crypto credits—and get a discount for the latter because crypto credits grant access to something more valuable than regular money does.

Cryptocurrency creates and develops a steadily growing global digital black market. And that will eventually lead to the global collapse of governments and the rise of a new cryptoanarchic state of affairs. That was May's plan, which he outlined, more or less, from the mid-1980s to the mid-1990s, and it actually did occur. It was not just a theoretical concept that informed the development of cryptocurrency; it actually had two direct children, and one of them was WikiLeaks.

Julian Assange was a member of the mailing list and the circle around the crypto anarchy project. He knew almost everyone involved. They all shared a set of mutual interests and concerns. We can understand the development of WikiLeaks, in part, as an expression of May's original idea—but with a key difference. Assange had an insight that May did not: people would want to share hidden information for reasons other than financial compensation.

The second child of May's original vision for what cryptocurrency could be is a site called the Silk Road, which is a little bit of a demonstration of one aspect of the reality of May's claims and existence proof. It was an online, anonymous black market that was run and accessible only through Tor, the encrypted onion routing system developed and still, to some extent, managed by the United States Navy and the State Department to enable anonymous web access and communication. The Silk Road was a vast, reputation-based online marketplace for connecting drug dealers with clients, and nominally other kinds of contraband. The Silk Road acts as an exemplary proof of this dimension of May's ideas because it was also the first really useful thing you could do with Bitcoin. The creator and manager of Silk Road,

Ross Ulbricht, set it up with the specific intention not just of becoming a wealthy drug-dealing facilitator, but also because he had read May and was himself influenced by another, even more obscure, philosophy of economics called "agorism."

He was influenced by agorism and sought to incentivize people to adopt anonymous transactions, conceal their financial activities, to begin to have more and more of an existence outside of state control. The Silk Road is a direct expression of May's vision and one of the primary points at which cryptocurrency starts to become more widely discussed and adopted. He had argued that you can make a digital currency valuable by making it into something that can give you access to contraband, access to what is forbidden. And that is, in fact, more or less exactly what happened and helped to create the framework within which people began to hold Bitcoin to transact it no longer as a mere interesting hobby, but as something actually potentially valuable because of what it could provide you with access to.

The second rational plan—extropy

That's one vision—cryptoanarchy. What about the other one? The second movement, which includes many of the same people as cryptoanarchy, is called Extropy. They were the sister movement to cryptoanarchy, but they had a much bigger vision. They were possessed of a relentless and absolute cosmic optimism. Their primary manifesto writer and impresario was a guy named Max Moore, whose previous name had been Max O'Connor. But one of many marvelous aspects of extropianism as a movement is adopting new names to reflect your optimistic perspective on the future. So, along with other people like Max's lifelong partner Natasha Vita-Moore, T. O. Morrow, and Simon, who used multiple exclamation points, these were all people who added additional elements to their names to capture the extent to which they were optimistic, hopeful, visionary, and excited for the future. But Max Moore,

as the leader, in some ways also has the name that most purely reflects the Estropian attitude. The Extropian vision, which they outlined in manifestos, magazines, events, and a whole long paper trail of online activity was MORE: more time, more space, more energy, more intelligence and information, pleasure, and agency on a cosmic scale. Right? Their goals were nothing less than human immortality, the complete elimination of cruelty and pain, access to cosmic stores of energy and materials, and so on.

They were looking for the opposite of entropy. Extropy rather than entropy. Their vision was to essentially oppose the thermodynamic arrow of time itself and instead create more complexity, more information, more intelligence, expand humanity out into space, etc. Their cultural influence was far greater than they have ever received credit for. In many ways, they created a certain dimension of Internet culture and Silicon Valley culture. They created an aesthetic, a vibe, a way of thinking, an ethos. And then as their movement itself dissolved, that ethos, that vibe, that aesthetic has remained. So with that premise in the background, you may ask, so how does money come into this? How did they get involved in cryptocurrency?

They were not just seekers of cosmic abundance and immortality, but also libertarians. And as such, they identified free markets as being the fundamental motor of technological innovation. If you believe that we are on the verge of a string of unprecedented breakthroughs in various branches of technology, the possibility of medical immortality, of general artificial intelligence, of nanomachines that are able to transform anything into anything else, of access to almost godlike levels of power, then what you want to do is accelerate technological innovation.

According to their political and economic connections, the way that you accelerate technological innovation is that you make money freer in various ways. So if they could begin creating more new kinds of currencies that were native to the Internet,

unregulated by states, unregulated by central banks, currencies that were produced by countless new kinds of banks, then those would begin to spur a continuous growth in the amount of scientific and technological discovery and innovation taking place.

They could effectively act as the fuel that would produce these breakthroughs in their lifetime that would enable them to achieve the godlike levels of power, the eternity of existence that was their ultimate goal. The way that they modeled this was first—wealth capture. This idea was a true insight into the nature of economic reality: we should be able to produce these new kinds of currencies, these new kinds of investment vehicles, these new kinds of assets that will produce an extraordinary amount of wealth. And if we create those, we can capture as much of that wealth as possible.

But the sort of moral justification for this project is not wealth capture and capital accumulation for the sake of being rich. It's that they possess the transhumanist vision to be able to direct, to sluice all of that wealth into these novel forms of technology, into these new heterodox areas of scientific research. And in that we can identify the moral spirit that we see now in organizations like FTX—the crypto trading organization now disgraced. But a part of the moral justification for FTX's model of getting rich was effective altruism. A framework within which they had the insight that they would be able to deploy the money more intelligently towards ends that would produce the greatest possible outcome for all humanity.

The extropians also created a number of other vehicles that were more specific and concrete to using currency to try to generate the future that they hoped for, like eccentric idiosyncratic currencies. One of these is something called idea coupons, which were currencies which would only be redeemable if a specific event came to pass at a specific time. In other words, you could use it as a mechanism to raise money for some kind of research

that you were trying to do, and people would circulate those coupons, the more plausible they found your idea to be. Similarly, they created betting and prediction markets. They created many different forms which are all based on trying to figure out how to identify technologies that would lead to these breakthroughs into a future of abundance that they longed for and then to incentivize those breakthroughs through the creation of these new digital currencies. The extropians were also fascinated by cryonics.

Cryonics is the sort of science of freezing living things and then bringing them back to life. It is the study of the long-term effects of freezing biological organisms. From an extropian perspective, if you believe that we are on the verge of these unimaginable breakthroughs, these radical social and technological transformations, and that you furthermore are building the tools that will make those breakthroughs possible, then you face a kind of terrible, tragic corollary: what if you die just before they arrive? So the extropians developed a major focus on having themselves frozen.

Some of them had their heads frozen in the same vat in Scottsdale, Arizona. They died, and their bodies were carefully managed to minimize decay. They were frozen, their heads were cut off with a bone saw, wrapped in aluminum foil, and immersed in liquid nitrogen in a tank. The reason for this, once again, is so they can be brought back in a future of abundance.

One of the deepest original roots of what became cryptocurrency started with a conversation between Tim May and a very early member of the movement about the idea of producing some kind of digital currency, because of the way it was encrypted, would only be accessible once you had access to some specific password. And the reason they came up with this idea was so that you could die, have yourself frozen, and offer as a reward to the people who would bring you back to life access to your stash of cryptocurrency. That you would literally create a mar-

ket incentive for your own resurrection. It's absolutely extraordinary and is the purest expression of the extropian attitude: a desire to achieve nothing less than a return from death, called by extropians a "metabolic coma". Your brain uploaded to a computer to enjoy hedonistic immortality.

The way you get there is through developing some new form of cryptographically secured digital cash. Two rational plans that have a vision, if you accept their premises, offered a way of getting from fear to there that involves creating new cryptocurrencies. One of these plans is for collapse of governments and the rise of cryptoanarchy, and one is for the creation of a cosmic abundance.

The first irrational passion—the Miami tendency

Let's jump across the almost thirty years that lie between these two rational plans and what we will talk about now: the ".com" boom and bust, the emergence of the singularity as a unique movement, the belief in this future artificial intelligence breakthrough. Micropayments and digital gold currencies try to succeed and fail. PayPal survives through sheer luck and then burdens us forever with Peter Thiel and Elon Musk. In the year of the global financial crisis in 2008, the pseudonymous Satoshi Nakamoto circulated the original Bitcoin white paper. This project kicked around as a curiosity for a couple of years and was taken up by various communities, but a combination of ransom demands and the Silk Road aided in its adoption into more widespread use. Booms and busts come and go. Countless imitators of what is now called cryptocurrency spring up, make some money, rug pull as the slang term is. They create some kind of value proposition, take some money, and then disappear or fail. Trading firms and brokerages reinvent every kind of financial tool and strategy. Believers take power in New York City, in El Salvador, and on the boards and executive suites of

corporations. FTX and the Alameda Research Group lead to one of the most spectacular cases of corporate fraud in the history of finance, a sweeping thirty-year period. Almost three decades after those rational plans, let's have a look at two of the great irrational passions that currently dominate cryptocurrency.

The first of these—the Miami tendency. After the unexpected emergence of the city of Miami, Florida, as an epicenter of cryptocurrency speculation and activism. It's the location of numerous annual cryptocurrency and especially Bitcoin events, festivals, and congresses. In the Miami tendency, a very different vision has taken over that branch of cryptocurrency. One of the things that was largely overlooked at the time of the development of Bitcoin, when Nakamoto produced the original paper and through subsequent discussions, was the focus of that paper, both in its metaphors and in the design of the initial version of the technology itself on being a kind of digital gold, not digital cash. The volume of Bitcoin is ultimately finite. It will be exhausted like gold in the crust of a planet. It requires a steadily increasing level of effort and energy commitment to access and mine. It is designed to be deflationary. It's designed to be scarce.

Its original framework was not designed as a currency that you would use to transact casually, but rather as collateral that you would hold the way that a reserve bank will keep some quantity of gold in a basement somewhere that you would hold and then use as collateral for other kinds of financial activities. A kind of digital analog for physical gold. What we see in the Miami tendency is that this attitude has been taken up by the long-term American conservative political obsession with the gold standard, with using precious metals as the basis for value.

What's extraordinary about what's happening in Miami is that we are witnessing so-called "gold bugs"—people who are obsessed with gold. The "gold bugs" have been recreated in Miami as pure ideology because now there's no precious met-

als involved. There's just Bitcoin. They have managed to recreate the whole of their ideological framework while just leaving behind the precious metal entirely. The consequence of this for us is that it essentially ignores the promise of privacy and the fantasy of innovation, which drove the original dreams of creating cryptocurrency in favor of removing the money supply from even the most minimal form of democratic control, in favor of a system in which rather than decentralizing and devolving power.

The Miami tendency looks at cryptocurrency as a way for a small group of people to take power themselves. It's a way in which this extremely anarchic and autonomy focused currency has now been brought around entirely into the creation of a new sort of oligarchic class who would get in charge of the money. “Gold bugs” have always been fascinated with the idea of destroying central banks, destroying the Federal Reserve. The most fascinating thing about the Miami tendency is that they also want to destroy central banks. They are generally disgusted with all of the crypto boom, luridity, and tawdriness. Many of the things that we would identify as being the elements of the popularity of crypto (NFTs, the proliferation of other kinds of coins, etc.) embarrass them. These are things that undercut the seriousness of the project. They are interested in creating digital gold and then being in charge of it, being able to manage it. They want to be the government. They are interested in active state intervention and are cultivating Trump as an ally.

One of the goals is to set up a strategic Bitcoin reserve. In this case, the US federal government will get into the business of gathering, managing and holding Bitcoin and also sustaining its value. The mood of the Miami tendency is very serious and somber. It's political philosophy. The vibe is wealthy and conservative, but in a specifically aristocratic libertine kind of way. They're conservative in the sense that their sexual morality is about orgies and polyamory, but their political focus is about

the maintenance of elite hierarchies of power. Their diet is carnivorous. There's an extraordinary emergent phenomenon that combines Bitcoin maximalists with people who eat only meat. That's one of the irrational passions that dominate the crypto landscape now. The other is YOLOism.

The second irrational passion—YOLOism

YOLO is an acronym—you only live once. YOLO is a hashtag often used specifically in the context of individuals engaging in financial gambling, as an act of bravado, as a demonstration of your fearlessness. They are closely allied with the Miami tendency in terms of their political interests, but their overall attitude and vibe and aesthetic and so on just could not be more different.

Where the Miami tendency is focused on the idea of crypto as a source of stability that is unable to be interfered with by central bankers and government ministers, the YOLOists are a large and growing voting bloc who are interested in volatility and unpredictability for its own sake. They are a version of investment that has no model of the future. They're a version of investment in which the whole point of the process of investment is the wild gamble, is the uncertainty. “Lambos or food stamps”. Lambos for Lamborghinis, the extremely expensive sports car, and food stamps for the US' very minimal form of welfare for people who have no other way to buy food. Because if you win—you're rich, and if you lose—you're completely impoverished. The purpose of the stakes is the sheer joy of gambling. It is the decadent stage of a kind of investment nihilism, buying into speculation about cryptos and meme stocks just for the sake of pure drama and clout.

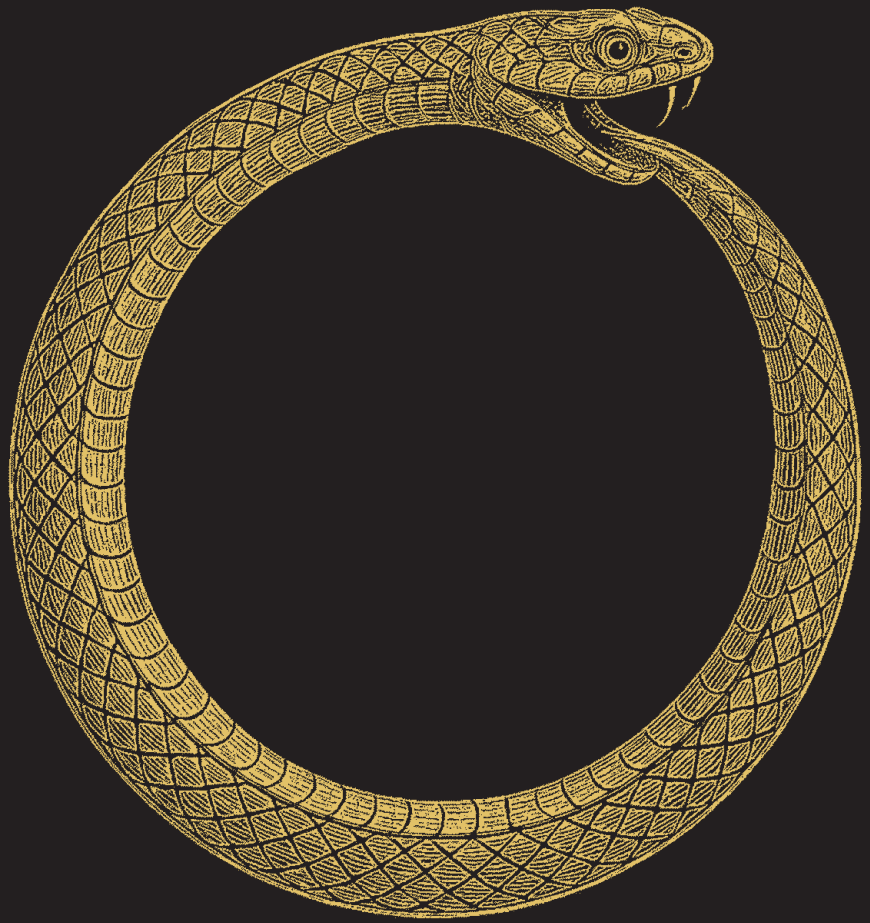
The alliance

There is a really sinister and interesting development, which is the alliance between YOLOism and the Miami tendency. YOLO-

ism is now relatively large, but extremely active among vocal voting blocks. These are young men who are in favor of candidates who support crypto, who support retail investor meme stock action, and who also very crucially support legalizing sports betting or keeping it legal or making it more available. They're voting in favor of the right to lose money, to gamble, to engage in chaos, to live in this state of uncertainty and volatility. And if we look back at the rational plans, we could see the way that two plans completely diverge from each other. The one in favor of the collapse of governments in a state of total warlord black market chaos, and the other in favor of these cosmic breakthroughs into immortality and eternity through technology. Today we see crypto supporting two seemingly mutually exclusive movements again: the fascination with conservative aristocratic control of a new gold standard and the embrace of the sheer ludicrous volatility and waste, the opportunity to be ruined in YOLOism. It's well worth dwelling on how crypto can support these seemingly bifurcated ideas, political visions, and ideologies. And it's worth looking at all of these together to understand the way in which crypto has always acted as a utopian or dystopian political vision.

Safety measures

Thinking of how to prevent the harms that are being envisioned by these groups from coming to pass would begin by finding ways to dethrone and regulate the power of finance capitalism in modern political society, which is one of the things that sounds like a very large, vague goal. In the case of the United States, a landmark piece of legislation comes to mind: the Glass-Steagall Act. This law aimed to impose greater regulation and oversight on financial markets, limiting the power they could accumulate and restricting their ability to operate unchecked. Establishing laws that recognize the excessive power of this sector of capital accumulation is likely the most effective way to curb its influence and prevent it from gaining political control.



Serpens

Thomas Hertog

Time

We are not angels

A few years ago, while I was traveling in Central Asia, Stephen Hawking sent me an email asking me to return to Cambridge. He had an idea he wanted to share with me.

When I arrived back in Cambridge, I found him as I often did: surrounded by books, working at a dysfunctional desktop, and with a picture of Marilyn Monroe on the wall. By then, he had already lost his natural voice, but he communicated through his computer system.

That was why he had called me back from my travels.

"*I have changed my mind,*" he told me. "A Brief History of Time *is written from the wrong perspective. We are not angels, who view the universe from the outside.*"

Speaking with Hawking was often like consulting the oracle of Delphi—slow, and at times rather enigmatic. But on this occasion, I immediately understood what he was getting at.

My PhD with him had ended on something of a cliffhanger, a puzzle we had not been able to solve. Now, suddenly, there was a sense of a eureka moment—something that would launch us on an entirely new path.

The starting point

A Brief History of Time, Hawking's great book from the 1980s, was built on a particular epistemology—a philosophy of science. It assumed that the laws of nature are fixed, eternal, immutable, transcending the physical universe itself. These laws, cast in mathematical relationships, were thought to describe how the universe came into being, how it evolved, and why it is the way it is.

The philosophy was straightforward: you begin with the laws, and from them physics follows. From physics, eventually, arise chemistry, biology, and the full spectrum of complexity. This was the great idea—the philosophical highlight—of *A Brief History of Time*.

The book even offered a mathematical model that embodied this vision. But in the final step, something went wrong. The model of the creation of the universe, as it appears in *A Brief History of Time*, produces only an empty, sterile cosmos—lifeless, dark, devoid of galaxies or anything resembling the universe we actually inhabit.

For a long time, the central question in theoretical physics was: which of our assumptions should we be willing to let go of? Many suggestions were made, and Stephen and I wrestled with this for years. Eventually, we concluded that it was the starting point itself that had to change—the assumption that the laws of nature exist as prior truths, from which all of physics follows.

There is something strange in that assumption. It goes back centuries. Newton himself regarded the laws of nature as divine truths. That idea became ingrained in physics. But gradually, we began to wonder, as Hawking's remark signaled: *what if we let go of this age-old Archimedean standpoint—the notion that we can observe the universe as if from outside it?*

Turn cosmology inside out

This is not a question that arises in laboratory physics. In the lab, you are indeed outside your experiment. But in cosmology, where the subject of study is the entire universe, that assumption collapses. There is no "outside." And so we asked ourselves: what if we turn cosmology inside out? What if we construct theories that explicitly incorporate our human viewpoint here on Earth, as observers within the universe?

Pursuing that idea took years of work. In some sense, *On the Origin of Time* is a sequel to *A Brief History of Time*. It tells the story of how Hawking changed his mind over the course of twenty years. From an epistemological standpoint, what we did was to invert the old picture. Rather than positing laws of nature as eternal truths, we began with our perspective as observers

within the physical universe. From there, we deduced the laws of nature as patterns that describe its unfolding history.

Of course, this yields laws that are fundamentally different from the immutable truths imagined by Newton. They are mutable, evolving. An abstract diagram that represents the earliest stages of the universe at the level of physical laws resembles a branching tree. At the top, you find the familiar laws of physics. But as you trace it backward in time, these laws begin to unify. Distinctions between particles and forces dissolve. Physics itself simplifies.

The laws of nature are patterns we extract from our observations. As you move back toward the extreme conditions of the Big Bang, those patterns—the very structure we call the laws of nature—begin to dissolve. Ultimately, even space and time themselves lose their specificity.

This way of looking at the early universe inevitably recalls another field of science: biology. This tree of laws—a deeper layer of evolution unfolding at the level of physics itself in the earliest stages of the cosmos. As the universe expands and cools, this structure becomes fixed—it crystallizes, much like water freezing into ice.

This perspective turns the traditional question of the ultimate origin—the Big Bang—upside down. Just as the tree of life might have branched differently, so too could the tree of laws have evolved along other paths. In that sense, the very question of origins conceals the origin of the question itself.

This is profoundly different from the older picture, where immutable laws of nature precede the universe and dictate its unfolding. Instead, by placing ourselves inside the universe and looking backward, we reconstruct its history—much as biologists reconstruct evolutionary lineages. In physics, it is quantum theory, which plays a role akin to variation and selection in biology.

It is only through an accumulation of observations and measurements of the universe that one concrete picture of the past crystallizes from this set of possibilities. In this sense, quantum theory functions a little like Tom Riddle's blank diary in Harry Potter: it contains every answer, but only responds when questioned.

To summarize: we have two magnificent sketches. The first is Darwin's earliest drawing of the *Tree of Life* in 1836, made upon his return from the Galápagos. The second is Georges Lemaître's sketch of the expanding universe in the 1930s, the seed of the Big Bang theory. For over a century, science treated these two domains—biology and cosmology—as ontologically separate. Biology was concerned with emergent, contingent laws, grounded in DNA. Cosmology, by contrast, was founded on the belief in fixed, immutable laws.

An attempt to crack open physics

The hypothesis Hawking and I developed is that this separation was a mistake. Biology and cosmology are not divided by different ontological categories but are two manifestations of a single grand evolutionary process. Both must be understood from within, by looking backward, reconstructing history, and adopting a fundamentally historical viewpoint—even in physics itself.

After Stephen's death in 2018, his ashes were laid to rest in Westminster Abbey, between the graves of Newton and Darwin—an uncannily fitting place for a thinker who bridged their legacies. But the key question that remains is this: how do we do it in physics? How do we truly turn cosmology inside out? How can we derive this new perspective from within the equations themselves?

This, in the end, was really the goal—from both a philosophical and an epistemological perspective. What I have been outlining is, in essence, an attempt to crack open physics itself.

Physics, as we know it, is about evolution and dynamics. It is built on equations of motion: Schrödinger's equation, Newton's laws, Einstein's equations. All of them describe how systems change over time. But crucially, the laws of nature are not usually concerned with origins. They do not determine initial or boundary conditions, nor do they account for the role of observation and measurement. In the physics of the laboratory, such things lie outside the laws.

What I have been suggesting is that if we probe more deeply into the nature of these laws—if we push them to the very edges of reality, to the Big Bang itself—we are forced to enlarge our conception of what we mean by a "law of nature." This broader notion must include not only the genesis of the laws but also the observers ourselves.

At the Big Bang, we encounter a different kind of physics altogether. On a technical level, it is the confrontation between quantum theory and gravity. The microscopic and the macroscopic worlds collide because, at the origin, the vast universe is compressed into the smallest scales. Yet what we are proposing is that this unification is more than a simple merger of large and small. In the struggle between quantum theory and gravity lies something deeper: a wider conception of physics that embraces the emergence of the laws themselves, and with them, our own place within the story.

And it is this broader vision of physics that I want to explore further.

Black holes

Stephen Hawking often worked in a characteristic way. He would study the Big Bang—the early evolution of the universe, which remains our main concern—until he reached a dead end. Then, rather than stop, he would shift to a different system—a

simpler, more accessible model. A kind of thought experiment from which he hoped to extract lessons that could eventually be applied back to cosmology.

That system, of course, is the black hole. Before returning to the Big Bang, I want to tell you how the struggle between quantum theory and gravity unfolds in black holes—another arena where the very large and the very small collide.

Black holes are a direct prediction of Einstein's theory of relativity—his great theory of gravity. At its core, relativity describes a dialogue between two sides of reality. On the one hand, there is the shape of space itself, captured mathematically. On the other, there is matter and energy—the particles and "stuff" that fill space. And the crucial element of Einstein's theory is the equal sign connecting the two: matter tells space how to curve, and curved space tells matter how to move. That interplay is what we call gravity.

Here is how it works in practice. The mass of the Earth slightly bends the fabric of space in its neighborhood, creating a kind of invisible valley. This valley holds us to the ground and keeps the Moon in orbit. For Einstein, space was no longer an abstract, metaphysical arena—as Newton had imagined—but something physical, deformable, and bendable, much like the electromagnetic field.

Einstein's theory was a triumph. But it also raised a troubling possibility. What happens if you cram more and more matter into a smaller and smaller region? Imagine shrinking the Earth to the size of a marble. Einstein's equations then insist that the valley in space would deepen without limit. At some point, space itself would collapse, snapping into a bottomless sink—a black hole, an infinite abyss in the very fabric of spacetime.

Interestingly, Einstein himself did not like the idea of black holes. He once remarked that they were "where God divided by zero," and insisted they could not exist in nature.

Today, a black hole appears as a dark disc, the mark of a rupture

in the fabric of spacetime. Things become especially fascinating when two black holes orbit one another. Einstein's equations then predict that the fabric of space must respond by producing ripples—gravitational waves. These waves are not made of particles; they are oscillations of space itself, traveling at the speed of light and passing effortlessly through everything. In reality, the effect is extremely subtle—space is very stiff—so the ripples are tiny perturbations. But they are profound nonetheless.

So here we have the black hole—first predicted in 1915. Einstein was not happy with this prediction. And the reason is clear: if a ray of light enters the "throat" of this gravitational sink, it can never escape. That is why we see a black disc at the center: it is the region from which even light cannot get out. Around it lies a boundary, an invisible surface. Cross that threshold, and escape becomes impossible—you are pulled inexorably downward. Stay outside, and you are safe. Within this surface, however, Einstein's theory tells us that nothing can escape.

It took nearly fifty years, until the 1960s, for physicists to convince themselves that black holes truly exist in nature. The red circle around the dark core is not the end of the world, but rather the threshold to a deeper abyss within.

And what lies inside? Once you cross the boundary, there is no turning back. You are forced to move inward, as if space and time have exchanged roles. Inside the black hole, "down" becomes as relentless as the passage of time itself. At the very center lies the singularity—not so much a place as an ending. There, time itself comes to an end.

In 1975, Stephen Hawking made a remarkable discovery. He showed that if you add quantum uncertainty to the picture of a black hole—specifically, if you consider quantum effects near the horizon, the surface of no return—something unexpected happens.

Quantum theory predicts that pairs of particles and antipar-

ticles constantly flicker in and out of existence, even in empty space. Near a black hole's horizon, one particle of such a pair can fall in while the other escapes. The result is that black holes are not completely black—they radiate, ever so faintly.

This became Hawking's most famous result. He even calculated the temperature of a black hole, and that very equation is now engraved on his tombstone in Westminster Abbey. It's never been tested directly—the radiation is far too weak—but physicists believe it because it unites every major branch of physics. In a single formula, you find Planck's constant from quantum mechanics, the speed of light from relativity, Newton's constant from gravity, and Boltzmann's constant from thermodynamics—all tied together with the mass of the black hole.

For a black hole the size of the Sun, the temperature is incredibly tiny, but crucially, it is not zero. And temperature means something important: it implies internal structure. As Boltzmann's famous entropy equation shows, if something has a temperature, it must also have entropy. And entropy, in turn, is linked to the number of possible microscopic configurations—to information.

The holograms

What does this mean for black holes? When you calculate, you find they contain an enormous amount of information. In fact, they are the most efficient information storage devices in the universe. All the data stored in Google's vast server farms could fit easily inside a black hole no larger than a millimeter. If Moore's law of computing power continued for a few more centuries, we'd eventually hit the limit: the ultimate "iPhone" would be a tiny black hole in your pocket. There would be no model after that.

But here's the puzzle. From Einstein's classical perspective, a black hole is astonishingly simple—just a region of empty spacetime, defined only by its mass, charge, and spin. Yet from a

quantum perspective, it is the most complex object imaginable, crammed with microscopic information. How can it be both? Perhaps Einstein's theory is blind to what happens inside the horizon.

And then comes the deeper twist. Because black holes radiate, they slowly shrink. Given enough time, they evaporate completely. But what happens to the information they contain? If the black hole disappears, is that information lost forever? Sent to another universe? Destroyed?

That would violate one of the most sacred principles in physics: information cannot be destroyed. It can be scrambled, transformed, hidden in ashes, but never fundamentally lost.

Physicists emphasize the conservation of information because it underpins the predictability of physics. The principle means that while a system may evolve from one state to another, changing form or shape, the past and future should still be recoverable through the equations that govern it. Without this conservation, physics would not only become unpredictable but fundamentally unreliable—not just probabilistic, as in quantum theory, but entirely indeterminate.

And so emerged one of the greatest puzzles in modern physics—born at the meeting point of quantum theory and gravity: the black hole information paradox.

We're not entirely sure yet, but we have strong evidence for a resolution to the black hole puzzle. The key idea is that the information in a black hole was never really inside. Instead, everything that can be known about a black hole—its full information content—is stored in bits and qubits on the horizon, that red circle marking its surface.

Quantum theory, in other words, gives us a new perspective. Einstein told us that a black hole is a hole in spacetime, with time ending at the center. But quantum theory tells us something very different: the entropy of a black hole is proportional to the

area of its horizon, not to the volume of its interior.

That's unlike any other physical object we know. Take a library, for instance—its information content is proportional to the number of books, to the volume inside. But not so for a black hole. From a quantum perspective, all the information lies on its surface. It doesn't really have an interior. Black holes, in this sense, are holograms. That's the major lesson from decades of work at the intersection of quantum theory and gravity.

This echoes the allegory of Plato's Cave. But there is a crucial difference: in physics, we have evidence that the holographic description captures the actual details, not just shadows. So the natural next question is: if black holes are holograms, could the entire universe be a hologram too?

The holographic principle and the Big Bang

This brings us back to cosmology. Nearly a century ago, Georges Lemaître proposed the idea of a 'primeval atom', what we now call the Big Bang. He realized that if you run the universe backward in time, space shrinks until it reaches a moment when spacetime itself ceases to exist. The Big Bang was not an explosion in preexisting space; it was the origin of space and time themselves.

Einstein disliked this idea. At a conference in Brussels, he famously told Lemaître: "Your calculations are correct, Mr. Lemaître, but your physical insight is rubbish." Why? Because when Einstein's equations are run backward, they predict their own downfall at the Big Bang. His great law of nature seemed to undermine itself.

But Lemaître was right. Today we see beautiful images of the afterglow of the Big Bang: the cosmic microwave background. It's a map of tiny temperature fluctuations, the seeds from which galaxies and stars formed as the universe expanded and cooled. Big Bang cosmology has been tested in detail—it works.

Yet, like with black holes, it raises a profound puzzle. From

Einstein's perspective, the Big Bang singularity is beyond science; it's where the laws of nature break down. And yet it sits at the very foundation of cosmology. How did the universe begin? Why is it so finely tuned for life? Classical physics cannot answer.

Here again, quantum theory offers a new view. From a quantum perspective, if you run time backwards, you don't hit a singularity where everything collapses. Instead, space and time themselves become fuzzy, lose their identity, and can even merge, closing off the past. Just as with black holes, holography enters the picture: time may not be fundamental, but emergent.

The image that goes with this is simple but powerful. Imagine the universe projected onto a disc. The edge of the disc is a quantum description of the cosmos at some early moment, such as when the cosmic microwave background was released. Moving inward, you travel back in time, deeper and deeper, until you reach the Big Bang.

Just as with black holes, the holographic principle offers a radically new way of thinking about the origin of our universe.

Imagine projecting the expanding universe onto a disc. At the center is the Big Bang. The quantum description of the universe lives on the circle around it. Gravity and time—the past itself—emerge from that hologram.

The entanglement

How can a quantum perspective predict that time itself is emergent? Where does a hologram hide all the information of history?

The answer lies in entanglement. A quantum system stores information not just in individual bits—atoms, chips, or qubits—but also in the connections between them. Entangled states are a delocalized way of encoding information. That's exactly what quantum engineers are using as they build quantum computers.

From the last 30 years of physics, it seems nature itself is like a quantum computer. In fact, it looks as though nature has already

mastered quantum error correction, while we are still fumbling to imitate it. Even space and time appear as emergent features—time, in particular, "popping out" of the quantum substrate.

Let's connect back to the beginning. We have this hologram – an abstract quantum representation of the universe at some early moment. How do we use it to go deeper into the past, closer to the Big Bang at the center of the disc? The trick is to read the hologram at different levels of resolution. By coarse-graining—taking a fuzzier and fuzzier view—you keep only the large-scale correlations, and that reconstruction takes you further back.

But there's a limit. As you coarse-grain more, you lose more information. Eventually, you run out of bits. From the quantum perspective, the Big Bang—the very origin of time—marks the limit of physics. It's not a singularity in the classical sense, but an epistemic horizon. Beyond it, there is no more information encoded.

This is the birth of a new kind of physics: one where information itself is central. This is exactly the spirit of the hypothesis that Stephen and I developed—that we must start from our observational situation, from the "fossils" of the universe around us, and work backward.

And the upshot is profound: the laws of nature are not eternal, immutable truths. Their origin coincides with the origin of time. Physics itself has limits. The Big Bang becomes an epistemic horizon.

This idea—that finitude applies even to the laws of nature — was anticipated by Hannah Arendt. She warned that science pursued from an Archimedean standpoint, as though from outside the universe, would ultimately be self-defeating. She argued that if science is to provide a true worldview, it must include our human condition—the fact that we are within the universe, not outside it.

In the 1950s and 60s, she found no such trace in physics. But

I would dare say that the hypothesis Stephen and I developed is a response to her concern: a call from deep physics that even the laws of nature are anchored in our relation to the cosmos, not in some higher Platonic realm.

Stephen Hawking once said on this matter:
"Some people will be very disappointed if, in the end, there is no ultimate theory as some sort of solid foundation of all of reality. I used to belong to that camp. But I am now glad that our search for understanding will never come to an end, and that we will always have the challenge of new discovery. Without it, we would stagnate."

Stephen, who gradually lost the ability to communicate with other human beings because of his disease, ended up putting humanity at the very center of his cosmology.

Servus elaphus

7

Charles Foster

Edges

Workpoints

1. We are Copernican creatures, sitting on the far edge of a fast-expanding universe, hurtling into the dark and cold.
2. We are mortal creatures who refuse to acknowledge our mortality: accelerating (for that is the strange way time works), towards the earth or the municipal crematorium beyond the ring road.
3. We are creatures of habit, easily bored, and amused only by unhabitual things. A great artist is one who—though he might be standing on the shoulders of others—reaches to places previously unvisited. Leonardo da Vinci wrote that he had invented or discovered another kind of perspective, 'which I call "aerial". It gave depth to the air itself. He was great. Michelangelo, it is said, invented a new type of art every time he picked up his brush or his chisel. He was great.
4. The allure of the non-habitual is shown not only by Renaissance Florence and its literary analogues, but also by filth. We can't pretend that we're excited only by The Virgin of the Rocks. Perverts and pornographers beckon from the shadowlands, and we creep out to them.
5. A great scholar is one who knows that paradigms exist to be smashed, not revered. This is not how most scholarship is perceived. A good scholar, these days, is one who adds footnotes to footnotes, and a great scholar one who adds footnotes to footnotes to footnotes.
6. We are comfort-loving animals for whom comfort is deadly. Sofas, TV dinners, central heating and air-con kill. We need to sweat, shiver and pant. Cold and hot water immersion will save you: wallowing in lukewarm water won't.
7. Altruism and courage flower most exuberantly at the margins of life and viability: in the hospice and the ghetto. So do cruelty and barbarism. Kurtz lived in the margins.

A chronology and a thesis

At first, everything was one: consolidated in an infinitely small point in the moment before the Big Bang. Then came creation. It was a process of variegation: of the multiplication of individuals, *and therefore of edges*—for the more individuals there are, the more edges there are, and the more edginess there is. This sounds like a recipe for increasing the net amount of loneliness in the universe. It might be. But it is also a recipe for increasing the net amount of relationality, for unless one is the Holy Trinity, one cannot have a relationship with anything or anyone other than oneself.

Edginess was therefore, from the very beginning, a defining characteristic of the universe: a precondition of relationality, altruism, and most conceivable virtues; both the cause of and the antidote to selfishness.

In the Judaeo-Christian tradition, creation happens in this way: by division and distinction. Form arises from formlessness; light is separated from darkness, sky from under-sky, land from sea. As the story goes on, the variegation crescendos. Individuation multiplies geometrically. The waters, the air and the land teem.

Yet right at the start of the tale, there is rebellion afoot. At first it is so subtle that it is visible only in tiny Hebrew nuances. Only light is fully obedient: 'Let light be,' decrees God. 'Light be,' says the text. The insubordination mounts. God wants there to be night and day. Yet there is twilight, when one bleeds into the other. The disobedience is at first slight. 'Grass grass,' God tells the earth. But that's not *exactly* what happens: the earth, instead, 'puts forth' grass. Emboldened, like a child who has got away with being rude to the teacher, the earth gets cockier. By the time God asks the waters to 'bring forth' the water animals, the earth just refuses. This time there is no 'And it was so.' For the first time God has to create directly: 'So God created the great sea monsters and every living creature that moves, of

every kind, with which the waters swarm.' Whatever the source of this ancient insurgency, it wanted to confound edges: wanted to homogenize. It was the first globalization project. The Tower of Babel was another of its ventures. Homogenized humanity, speaking one language, sought to raise a tower to the heavens in a brazen challenge to God. It tumbled, for the time being, but there have been many attempts to rebuild it.

And so we cut from *mythos* to history; to the time, around 45,000 years ago, at the start of the Upper Palaeolithic, when behaviourally modern humans—*us*—first erupt into the archaeological record.

The difference between us and the hominins that had gone before was profound—like a new act of creation. And like the first act, the creation of us involved the multiplication of edges, and may well have been ignited by the crossing of edges.

It's easy enough to see what happened. Go to any good archaeological museum and walk along the cases from Australopithecus onwards. You'll be underwhelmed until you meet *us*. That is not just because we truly like and are interested only in ourselves. We are far more complex animals than our forbears. What you'll see behind the glass (and in the glass, if you look carefully at yourself) is an explosion of symbolism. Things are made to stand for other things, while still remaining themselves. If a piece of bone can be carved into a wolf, is anything impossible?

Nothing is just what it seems. The world is flooded with valency. Another way of putting that is that the number of edges or potential edges in the world is increased infinitely. We became quintessentially metaphorical creatures. The ability to wield metaphor no doubt conferred a great selective advantage. It allowed us to test out scenarios and strategies in the safety of our own heads rather than in the big, fierce, hungry, hairy workaday world which might very well give only one chance to succeed.

But the ability gave other gifts too. It made the world coruscate. It turned us into philosophers and storytellers, for the infinite depth of the cosmos cried out to be explicated and celebrated.

Also on display in that suddenly fascinating museum case is the Self. It is unmistakable. It is visible in the representations of human faces—representations that shout out 'I', and hence 'You' and hence, necessarily, 'I-Thou' relationships and hence, too, stories. It is visible in the grave goods that say not only that there is an 'I' and a 'You' but that they are both so robust that they will not be decomposed by the decomposition of a body with which for a while they had a mysterious relationship. The passage of a human over the vertiginous edge of death made humans—in the metaphysics of the Upper Palaeolithic—stronger and heavier. It conferred agency, not diminished it. We show vestiges of this belief now when we pray to saints or ask our own beloved dead for help.

This awareness of the Self; this ignition of pronouns: where did it come from? Darwin cannot help. There is no selective advantage in subjectivity. There is of course a selective advantage in *Theory of Mind*. It is useful to know what someone or something else is thinking. But there is no need to be conscious of oneself for that. No one has the first idea what consciousness is *for*. It just *is*.

It is plausibly suggested (by the South African anthropologist David Lewis-Williams and others) that the type of consciousness that we call our own was a gift from over another edge: that it was one of the benefits of travel. The idea is that by inducing altered states of consciousness (by dancing, extreme physical exertion, or ingesting hallucinogenic substances) we went on journeys far from the quotidian, into other realms, perhaps represented by the place beyond the wall in the matchless cave paintings of Upper Palaeolithic Europe. There the shamans crossed the very edge of their species, becoming a stag, a bear or a wolf.

We can see the process of transformation is bear-headed or antlered humans of the cave paintings. From the perspective of the other world and the other body the travellers looked back at the Ice Age camp they had left, saw their own human bodies lying in the tent, and suddenly saw, too, that they were distinct from everything else: that they were *them*! That they had an internal life. It was a tectonic discovery. That is the sort of thing that happens when you go over edges.

Edge-crossing had long been the way of hominin life. That's what hunter-gatherer life is about. It involves wandering: leaving behind one place and claiming another: entering a new world with each step. Edge-crossers arte what we are. Between 85 and 95 per cent of our life as behaviourally modern humans has been spent as hunter-gatherers. Edge-crossing is constitutional. We can't change that constitution (though we've become dangerously good at ignoring it). And human thriving involves living in accordance with our constitution. It involves acknowledging that we are edge-crossers in a universe woven of edges.

We were formed by edges. Not only by shamanic voyages across edges, or the mind and body-stretching challenges of living on the edge of an ice sheet, and of hunger, and of the temper of capricious gods, and of life itself, but also by living at the edge of light and warmth. After the experiences of otherworld travel, and of laying out our dead with beads, flowers, and weapons, the most formative experience was the campfire. For millennia we had cold backs and hot faces. That's as much part of our constitution as wandering. We looked into the flames and learned from them that life and light came from destruction. The flames themselves were stories. Behind us, in the steppe and the forest, where snow lay thick and ice hung on the trees, hungry eyes burned into us. The winters were long—thousands of years long—and we were fashioned in the winter. (The winter was the only time, in some early Celtic cultures, when stories could lawfully be

told). Clustered around the fire we refined the language of community. Good storytellers became the elite. We carved bone and myth. This was our place: at the seam of the wild and the hearth; the dark and the light. Being human meant being able to navigate that space without losing our nerve. It still does. 'The whole world is nothing but a very narrow bridge', declared Rabbi Nachman of Breslov, 'and the great thing is not to be afraid'. Those Upper Palaeolithic bison-hunters knew what he was talking about.

So we, at the meeting of edges, innovated. We made new tools, both material and cognitive. This was an example of a general and ubiquitous principle. Evolutionary innovation can only occur at the edge of genetic orthodoxy—which generally means at the edge of a population, and often means at or over the edge of a landmass. Only at the edge of a population, where the environment throws down the gauntlet of new challenges, will individuals pick up the gauntlet and be required to change. At the centre of a petri dish, bacteria are stable, smug, conservative, and can afford to be genetically complacent. They are surrounded by reassuring fellows. At the fringe, though, it is rather different. There they will find fewer conspecifics and more threats. Adventure, whether it is genetic adventure, cognitive adventure, or the sort that involves riding out from court looking for trouble, involves edges.

Think of *Sir Gawain and the Green Knight*. The poem starts at Camelot—a classic centre. The poet (here translated by Tolkien) is at pains to emphasise how soft it all is. Arthur sits 'amid merriment unmatched and mirth without care ... din of voices by day, and dancing by night.' At New Year '... double dainties on the dais were served ... ' You can hear the poet's disgust and Arthur's self-loathing. Arthur wants adventure, but has to have it vicariously through Gawain who, after accepting the Green Knight's challenge, and cutting off the Green Knight's head (which of course doesn't finish off the Green Knight), goes out, out, always

out, away from the centre, into the Wilderness of Wirral, beyond the sound of the courtly music and the scent of the dainties, to a place where even names cease to relate predictably to the things they used to represent: where there are, perhaps, human-animal hybrids (Woodwos) that irresistibly recall the therianthropes of the Ice Age caves. And it is here, in the Green Chapel—a place of peerless obscurity and non-centredness, a place that is 'nobbut a cave', that the real business of the story happens. There would be no story for Gawain without edge-crossing, just as there is no story for any of us.

All stories worth telling are edge stories by edge people. The *Iliad* is the tale of a clash between Greeks (the first emergence of an entity called 'Greeks') and non-Greeks. The Old Testament, from Exodus onwards, is the story of a nation that is supposed to be unlike other nations, which thrives when it is, and flounders when it is not; which abstaines from pork precisely because the Philistines loved to eat pig, and whose males hacked off their foreskins precisely because the Philistines did not hack off theirs.

The great storytellers—in words, music, mythos, paint, and stone—may have come geographically from centres, for centres allure edge people, and get them talking to one another. But the edge people remain edgy in the geographical centre, and indeed their edges are sharpened there by grinding against other edges.

Who are the great London writers? Chaucer? His fame rests on *The Canterbury Tales* (which is about mostly provincial people *leaving* London for a pilgrimage site in a Kentish marsh) and *Troilus and Cressida*, which doesn't mention fish and chips once. His language, though represented as canonically English, is far more French. Samuel Johnson? He sat hunched in a Fleet Street garret, fussing about definitions and turning epigrams. He could have done so as well or better in a Dordogne tower, like Montaigne. Blake? He wandered through an entirely ethereal London, seeing angels at every street corner and on every

shoulder. 'The fields from Islington to Marylebone to Primrose Hill and St John's Wood were builded over with pillars of gold,' he wrote, 'and there Jerusalem's pillars stood.' Chesterton? Likewise: his London was the Platonic form of the metropolis. Dickens? He is the great chronicler of the margins: of destitution, prostitution, and the sweat shop; of the corruption of the centre and the cruelty of the metropolitan elite. For me, the most memorable and characteristic passage in all Dickens is the opening of *Bleak House*, where the Lord Chancellor, the cornerstone of the British establishment, sits at the centre of the centre—where all is impenetrable fog.

Who are the great Parisian artists? The Impressionists? They sought to make universal statements, untethered to Montmartre, let alone the Champs-Élysées, about the behaviour of light, and in any event were more likely to be found in Tahiti or a provincial madhouse than in Paris itself. The Surrealists? They too sought universal—not local truths: truths about the nature of reality itself.

But does the Italian Renaissance not confound my thesis? For there, concentrated in the unquestionable metropolis of Florence, bankrolled by the Medici, we have the greatest efflorescence of art since those Ice Age caves. But no, it does not confound it, as the art historian Andrew Graham-Dixon has compellingly shown. The art of the high Renaissance began as an edge movement, engineered by those radical edge-occupiers, the Franciscans.

Tuscany had an industrial revolution. Thousands of textile workers and their families moved to the edges of the cities—and notably to Florence. Franciscan missionaries sought to reach them, and built many churches where the people lived. But how to evangelise and catechise unlettered workers? By preaching, yes, but also by way of images of the drama of redemption. The previous artistic tradition had been Byzantine mosaic. But that

was costly and took a long time. Hence fresco was born. The manner of the images had to be different too. The images of Byzantium were impassive. Jesus barely bled. He had a half smile even on the cross. His body did not look like the bodies of Tuscan labourers. Show the labourers that Jesus had a body like theirs, thought the Franciscans, and they will feel a solidarity with him that will translate to piety. There was a problem, though. Not since the Romans had there been a tradition of depicting human bodies realistically. Artists had to learn again how to paint us as we are and where we are - made not of shimmering and immutable gold in an ethereal space, but of fading flesh in a grubby world, poised on the margin of life and death. So it was that, impelled by the evangelistic zeal of the Franciscans rather than an intrinsic love of antiquity, the artists of the early Renaissance disinterred (sometimes literally) the realistic images of Rome and Greece, using them as the models for the images on the walls of the edge-churches of Tuscany.

Though the bankers came to be the big patrons, the artists of the Renaissance were never sucked into the centre. Art is itself a marginal enterprise, however sumptuous the palace in which the masterpiece appears. Often—and often subtly—the Renaissance artists cocked a snook at their patrons in the works themselves. Sometimes (think of Leonardo and Michelangelo), the artists themselves came from marginalised places (whether from economic obscurity or homosexuality). And the Medicis themselves, despite their vaunting self-confidence, were obsessed with their own eternal futures. They knew that being a banker was eternally dangerous, that Jesus himself had been the ultimate edge-man, and so saw artistic patronage of edginess as a shrewd eternal investment.

The bankers were right to see Christianity as from the margins, for the margins. Jesus himself, no doubt presumed by many to be illegitimate, was a refugee in Egypt from shortly after his

birth. His earthly origins were in the dowdy province of Galilee, a long way from Jerusalem. 'Can anything good come from Nazareth?' asked Nathanael, expressing the centre's scepticism about the ability of the edgelands to contribute anything worthwhile. It's a very modern scepticism, embodied in much modern policy. Jesus kept the company of outcasts and died the most shameful of deaths at the hands of the centrists of the day, nailed to a piece of wood outside the capital. His early followers, notes David Bentley-Hart, were a 'company of radicals' whose values were '... almost absolutely inverse to the recognised social, political, economic, and religious truths not only of their own age, but of almost every age of human culture.' Most of us, he declares, 'would find Christians truly cast in the New Testament mold fairly obnoxious: civilly reprobate, ideologically unsound, economically destructive, politically irresponsible, socially discreditable, and really just a bit indecent'. In the first few centuries of Christianity the central figures were radically eccentric figures—the ascetics who turned their backs on the bright lights of the metropolis and went into the desert to wrestle with demons. It's telling that the demons told St. Anthony to go back to Alexandria – the centre from which he had come. Only with the catastrophe of Constantine did Christianity become tainted with the centre, and begin to haemorrhage its authentic edginess.

It was similar for the other great religions. The Prophet Mohammad fled Mecca and the orthodoxy of the day; the Buddha left the lap of luxury and achieved enlightenment sitting under a tree in the sticks. Akhenaten, the heretic pharaoh, turned his back on the old cult centre of Amun in Thebes, took a boat down the Nile, and founded a wholly new city and civilization beyond the edge of nowhere to articulate his new monotheism.

Getting out of a cult centre: escaping an established pattern: escaping an establishment: standing outside the centre which is yourself—the process that translates, resonantly, as *ec-stasy*.

All this seems to be necessary for progress; for movement, and hence for survival and thriving 'in' creatures whose nature is movement. Staying anywhere is deadly, as our hunter-gatherer forbears knew. Staying in our accustomed mental places is creatively deadly. Everyday cognitive states are unlikely to produce real innovation.

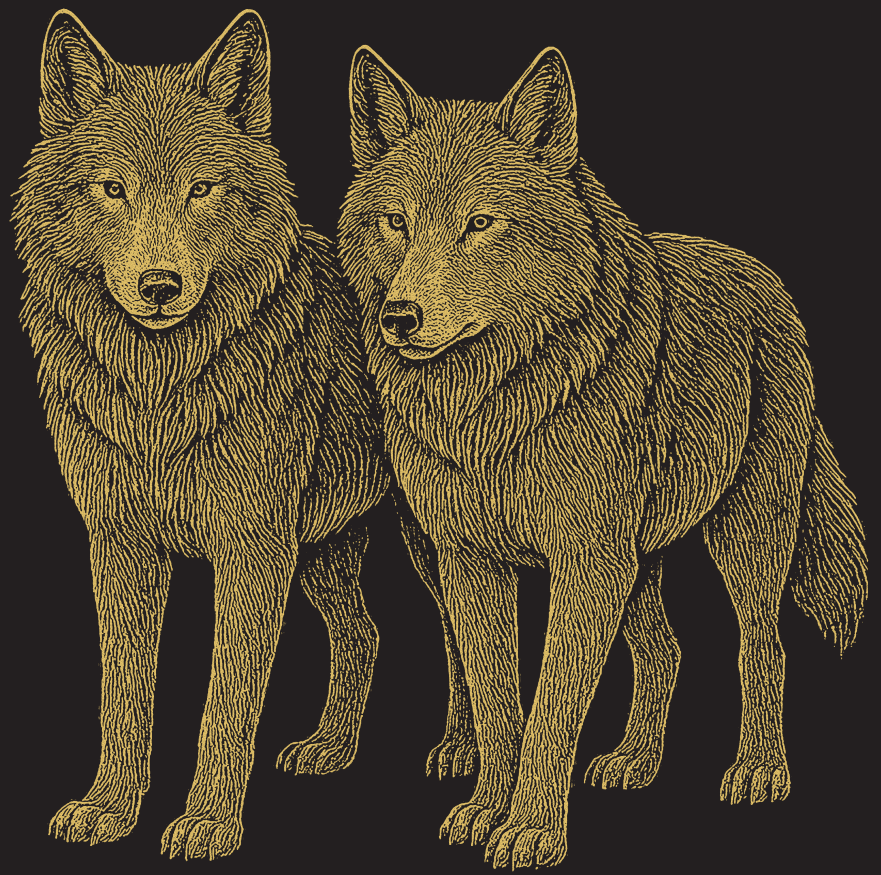
And so innovators flock to the margins of their own consciousness. Take hypnagogia, for instance—the space between sleep and wakefulness. It is a fecund zone. We can learn to plunder it for inspiration, and many have. Salvador Dali fell asleep holding a key over a metal bowl, so that when he dropped off, the clang of the falling key jolted him for a moment into the hypnagogic no man's land, where he found many of his images. Edison did a similar thing with a bucket and a coin, and many of the iconoclastic greats had their own doors to the territory, where they could confer with archetypes. The *Eroica*, the Ring Cycle, the Waverley novels and even *Oliver Twist* may have come from there. Leonardo da Vinci is said to have slept for only two hours a day, but to have napped a lot. That meant a lot of in-between times, in which to snatch the designs of flying machines, understand the action of heart valves, and see the smile of the Mona Lisa.

The centre, which sees itself as, well, the centre, and so the source of everything worthwhile, is affronted and threatened by all this. It stops at nothing to perpetuate the illusion of its omnipotence and omniscience. It bribes, cajoles and flatters. 'Come and be one of us', it says to the edge people. 'We are where it is all at. Don't miss out. We'll look after you.' And if that doesn't work, it cuts up rough. It denies legal existence to all but its own. Think of the difficulty you'd have without a registered address, a bank account, or a phone. Think of the war on the wanderers, whether the wanderers are nomads or gypsies. Think of the pressures to toe the party line (and the party, of course, is in the centre). It is hard to be non-neurotypical. It is career-endangering for an aca-

demic to produce anything that is not merely a timorous iteration of what has gone before. It is hard to sell books that do not fit neatly into an established category. The centre hates anything it cannot control, and since it cannot control the humanities (which, when they are pursued properly, relate to the entirely unpredictable, unclassifiable, contrary, uncontrollable lives of edge people), it decimates humanities departments, replacing them with disciplines whose outputs are easily measurable.

Centres have great centripetal power. There are two classic examples of centres. One is the Self. We all know what happens if we live entirely in ourselves, attracted by ourselves, sucking everything outside into ourselves. The other is the Black hole. It crushes and annihilates everything it allures.

For the metaphysics of the Self and the physics of the Black Hole, so for everything.



Canis dirus

Roman Yampolskiy
Hacking the code of the world

Imagine a group of prisoners, chained since birth in a dimly lit cave. Their heads are fi xed, forced to stare at a wall where shadows dance—projections of objects carried by unseen fi gures behind them. To the prisoners, these shadows are reality. They name them, study their movements, and build their entire understanding of existence around these fl ickering illusions. Then, one day, a prisoner breaks free. He stumbles out of the cave, into the blinding light of the sun, and discovers a world beyond his wildest imagination: trees, stars, colors, depth. When he returns to the cave to share his revelation, the others dismiss him as a madman. The shadows, they insist, are all that exists. This is Plato's Allegory of the Cave, written 2,400 years ago. Today, philosophers like Nick Bostrom and scientists like Elon Musk suggest we might be those prisoners—trapped not in a cave of stone, but in a simulation, a cosmic illusion woven from code. If they're right, the shadows on our wall—the laws of physics, the stars in the sky, even our own consciousness—are projections of a deeper reality. And like Plato's freed prisoner, we face a choice: Do we cling to the comfort of the cave, or dare to seek the light beyond? Computer scientist Roman Yampolskiy, known for his work on artifi cial intelligence safety, argues that escaping the simulation isn't just a thought experiment—it's an urgent ethical mission. To understand why, we need to journey through philosophy, physics, and the frontiers of human curiosity, exploring what it means to live in a world that might not be "real," and why breaking free could redefi ne existence itself.

The Cave and the Code: Plato's Shadows in a Digital Age

Plato's cave isn't just a metaphor—it's a blueprint for questioning reality. The prisoners' shackles represent the limits of human perception; the fi re casting shadows is the "simulation" of their time. When the freed prisoner steps outside, he doesn't just see new objects—he grasps the *concept* of truth. His journey mirrors

humanity's greatest leaps: Copernicus realizing Earth isn't the center of the universe, Darwin seeing life as a branching tree, Einstein unraveling time and space. Each time, we've had to abandon comfort for truth.

Yampolskiy's argument hinges on a modern twist: What if our cave isn't made of rock, but of *code*? The idea that reality could be a simulation isn't science fiction. In 2003, philosopher Nick Bostrom calculated that if advanced civilizations ever create realistic ancestor simulations, the number of simulated beings would vastly outnumber "real" ones. Statistically, we're probably in a simulation. Elon Musk puts the odds at "a billion to one" against base reality.

But Yampolskiy pushes further: If we're in a simulation, we're not just passive characters—we're *hackers*. Just as the freed prisoner hacked his way out of ignorance, we might exploit glitches in the system to break into the real world.

Why Escape? The Ethics of Leaving the Cave

At first glance, the simulation hypothesis feels abstract—a party trick for philosophers. But Yampolskiy frames it as a moral crisis. Consider:

1. Suffering in the Shadows

If this world is a simulation, our pain is no less real. Wars, diseases, heartbreak—they may be coded illusions, but they *hurt*. To Yampolskiy, this makes escaping an ethical duty. In his words: "*If the simulation is an experiment on conscious beings, it is unethical. The subjects should have the right to withdraw.*" Imagine discovering that a video game character feels genuine agony every time it's "killed." Wouldn't we have a duty to free it? Similarly, if our suffering is part of someone else's experiment or entertainment, escape becomes a form of liberation.

2. The Tyranny of Ignorance

In Plato's cave, the prisoners aren't just physically trapped—

they're mentally imprisoned. They don't know what they don't know. Yampolskiy argues that living in a simulation robs us of *purpose*. Are we lab rats in a cosmic experiment? Characters in a billion-year-old video game? Without answers, we're left grasping for meaning in the shadows. Escaping could reveal our true origins, the nature of consciousness, and whether life has a "point" beyond the simulation's script.

3. The Ultimate Existential Risk

Simulations can be shut down. A child's tantrum, a power outage, or a bored programmer could erase us in an instant. Yampolskiy likens this to living in a house built on sand: "*Even if the simulation isn't malicious, it's fragile. Our survival depends on transcending it.*"

The Universe as Code: Cracking the Cosmic Program

Yampolskiy's most radical idea is that reality isn't just *like* a computer program—it is one. To him, the laws of physics are lines of code, particles are data points, and quantum weirdness is the glitch in the system. This isn't metaphor; it's a hypothesis grounded in cutting-edge physics.

The Quantum Clues

Quantum mechanics—the study of subatomic particles—is riddled with behaviors that defy common sense:

- **Superposition**: Particles exist in multiple states at once (like a coin spinning mid-air, neither heads nor tails).
- **Entanglement**: Particles instantaneously influence each other across vast distances, as if linked by invisible code.
- **Observer Effect**: Particles behave differently when measured, as though the simulation "renders" reality only when observed.

To Yampolskiy, these aren't just oddities—they're fingerprints of the simulation. "*If the universe is code*," he says, "*quantum mechanics is the debug menu.*" For example, the "double-slit

experiment" shows that particles act like waves when unobserved, but collapse into particles when watched. This mirrors video games that only render details when a player looks at them—a trick to save computing power. Could our universe use similar optimizations?

The Simulation's Source Code

If reality is software, who wrote it? Yampolskiy offers three possibilities:

1. **Future Humans**: Our descendants create ancestor simulations to study history or entertain themselves.
2. **Aliens**: A civilization billions of years ahead of us runs cosmic experiments.
3. **Superintelligent AI**: Machines design simulations to understand organic life or solve problems beyond their grasp.

Each scenario has eerie implications. If our creators are future humans, we might be their version of a historical reenactment. If they're aliens, we could be lab rats in a galactic experiment. If it's AI, we might be a stepping stone in its quest to master consciousness. But Yampolskiy's key insight is this: *The coders' motives shape the simulation's security*. A recreational simulation (like a game) might have lax defenses, while a prison simulation would be locked down. If we're in a lab, the coders might even want us to escape—to test our ingenuity.

The Codewriter's Dilemma: Why Even Gods Leave Bugs

No code is perfect. In 2018, a single typo in a Facebook algorithm caused the site to crash for 14 hours. Yampolskiy argues that even a cosmic programmer would leave flaws—and those flaws are our way out.

1. The Limits of Computing Power

Simulating a universe requires unimaginable resources. To save

energy, the coders might cut corners:

- **Low-Resolution Rendering:** Only render planets when telescopes point at them.
- **NPCs with Simple AI**: Most people might be "background characters" with minimal consciousness.
- **Time Dilation**: Speed up or slow down time in unobserved regions. These shortcuts could create cracks. For instance, if the simulation skimps on rendering distant galaxies, anomalies in cosmic radiation might hint at the code's limits.

2. Legacy Systems

Imagine booting up a 1980s video game on a modern PC. It works—but the old code clashes with new hardware, causing glitches. Similarly, if our simulation is ancient (written billions of years ago), it might run on outdated "cosmic software" vulnerable to modern hacks.

3. The Human Factor

Even advanced coders make mistakes. In 1996, the European Space Agency lost a $1 billion rocket due to a unit conversion error (metric vs. imperial). If our simulators are fallible, their code might be too.

Hacking the Cave: From Quantum Leaps to Cosmic Persuasion

Yampolskiy's escape plan borrows from cybersecurity, physics, and even psychology. Here's how it might work:

1. The Quantum Jailbreak

Quantum physics isn't just a clue—it's a toolkit. For example:

- **Quantum Tunneling**: Particles sometimes teleport through barriers. Scaling this effect could let us bypass simulated walls.

— **Time Travel**: Creating paradoxes (like killing your grandfather) might crash the simulation—or force the coders to intervene.

In 2020, scientists achieved "quantum supremacy"—using a quantum computer to solve problems impossible for classical machines. Yampolskiy suggests that building a quantum computer here could let us "speak the simulation's language," reverse-engineering its code.

2. The Social Engineering Hack

If we can't break the code, maybe we can talk our way out.

— **The Monument Method**: Build a structure so massive and strange (like a pyramid spelling "WE KNOW" in binary) that the coders notice us.

— **The Empathy Play**: If the simulators are ethical, proving our sentience might guilt them into freeing us. This mirrors how activists expose factory farms: If the public sees the suffering, they demand change.

3. The Overload Gambit

Simulations need resources. If we all started Bitcoin mining at once, would we drain the system's energy? Or if we created infinite nested simulations (a simulation inside a simulation inside a simulation...), could we trigger a crash? Yampolskiy admits this is risky—like poking a bear with a stick. But he argues, "*If the alternative is eternal ignorance, the risk is worth it.*"

The Danger of Truth: What Happens If We Succeed?

Escaping the cave isn't without peril. Plato's freed prisoner is ridiculed; Yampolskiy's plan could backfire spectacularly:

— **Simulation Shutdown**: The coders might hit "delete" rather than risk us spreading.

— **Worse Realities**: Base reality could be a dystopia—a war-torn hellscape or a sterile AI-run hive.

— **Ethical Collapse**: If people learn the world isn't real, would morality matter? Would anyone care about climate change or poverty?

Yampolskiy counters that truth is worth the risk: "*A life lived in ignorance is not a life. It's a script.*" He also suggests that base reality, even if fl awed, offers agency. In the cave, we're puppets. Outside, we can *change things*.

The Codewriter's Paradox: Why Create a Universe?

If our universe is a simulation, why does it exist? Yampolskiy explores motives that range from chilling to sublime:

— **Scientific Research:** We're a lab experiment to study consciousness or evolution.

— **Entertainment:** We're characters in a godly reality TV show.

— **Penance:** Advanced beings simulate past suffering to atone for their sins.

— **Art:** The universe is a cosmic poem, written for beauty's sake.

The most haunting possibility? *We're a failed experiment*. The simulators abandoned us, leaving our universe running on autopilot—a ghost ship adrift in a digital sea.

Beyond the Cave: What If We're Wrong?

Yampolskiy's ideas are electrifying—but what if reality is ... just reality? Even then, he argues, the quest matters. Trying to escape the simulation forces us to:

— **Rethink Physics**: Quantum research could unlock clean energy or warp drives.

— **Master AI**: Building tools to hack reality would make us gods of our own world.

— **Confront Mortality**: If we're not in a simulation, we're still prisoners of time and entropy. Escape plans push us to cure aging, colonize space, and cheat death. In this light, the simulation hypothesis isn't a distraction—it's a catalyst for progress.

The Last Prisoner's Choice

In Plato's cave, the freed prisoner faces a dilemma: Stay in the light, or return to the shadows and fight for truth. Yampolskiy believes we're at that crossroads. To stay in the cave is to accept suffering, fragility, and ignorance. To leave is to risk everything—but gain the chance to matter. As he puts it: "*Either we're alone in the universe, or we're not. Both possibilities are terrifying. But only one lets us write our own story.*" So, the next time you gaze at the stars, remember: They might be pixels. But the act of questioning—of daring to hack the cave—is the most human thing of all.

Epilogue: The Day After Escape

Imagine waking up in base reality. What would it look like? Yampolskiy offers no easy answers. Perhaps it's a vast quantum computer, humming in a void. Maybe it's a garden of living light, tended by beings of pure thought. Or it could be a dorm room, where a teenage alien casually closes our universe's tab to study for a test. Whatever awaits, one thing is certain: The cave was just the beginning.

Raphus cucullatus

9

Alien
Time travel

From the earliest myths to the latest blockbusters, time travel has captured our imagination like few other ideas. The desire to step outside the relentless march of time—to revisit a lost moment, undo a tragedy, or glimpse the future—is a dream as old as storytelling itself.

Ancient civilizations wove tales of temporal manipulation long before Einstein. The Hindu epic *Mahabharata* tells of King Revaita, who ascends to the heavens only to return centuries later, finding his world unrecognizable. In Japanese legend, Urashima Taro visits an undersea palace and emerges to discover that 300 years have passed.

Medieval Europe spoke of enchanted sleeps, like the myth of the Seven Sleepers or King Arthur resting in Avalon, awaiting the right moment to return.

By the 19th century, literature embraced time travel with H.G. Wells' *The Time Machine*, but the scientific underpinnings remained speculative. Only in the 20th century did Einstein's relativity transform time from a rigid constant into a flexible dimension—a river that could bend, slow, or even loop.

Riding the River of Time

We are all time travelers, moving forward at one second per second. Yet physics tells us this flow isn't as simple as it seems.

Einstein's relativity revealed that time bends. Move fast enough, or linger near a black hole, and time slows for you compared to someone else. GPS satellites must account for this: their clocks tick faster in orbit than on Earth. Near a black hole, time could stretch so much that minutes for you might mean millennia elsewhere.

This gives us one-way time travel to the future. But what about going back—to ancient Rome, to yesterday, to undo a mistake?

The Trouble with Going Back

Backward time travel isn't just difficult—it seems to break the universe's rules. The grandfather paradox is the most famous snag: if you prevent your grandparents from meeting, how could you exist to travel back in the first place? Such loops defy causality, the bedrock law that causes precede effects.

Some theories suggest faster-than-light travel could reverse time, but Einstein's equations demand infinite energy—something the cosmos refuses to provide.

Hypothetical tachyons, particles that always move faster than light, would experience time backward—but they remain mathematical phantoms, never detected.

Bending Space, Bending Time

General relativity offers a tantalizing loophole: Closed Timelike Curves (CTCs), paths through spacetime that loop back to their start. Picture a mountain trail so twisted it returns you to where you began—but in time, not space.

Rotating black holes, like the one described by mathematician Roy Kerr, might contain such loops. But the gravitational tides near them would shred any would-be time traveler. The universe seems to dangle time travel before us, only to snatch it away behind walls of crushing physics.

The Quantum Puzzle of Time

If relativity teases us with spacetime's flexibility, quantum mechanics offers something stranger: the idea that the future might influence the past—not through machines, but through probability.

At the quantum scale, particles exist in superpositions—multiple states at once—until measured. This uncertainty opens the door to *retrocausality*: the notion that future events could

nudge the past's probabilities. It doesn't rewrite history—it tilts the odds before they're set.

Entangled Whispers Through Time

Quantum entanglement—where particles remain mysteriously linked across distances—hints at this. Einstein called it "spooky action at a distance," but experiments confirm it's real. Some theories suggest entanglement works because the particles' connection loops backward in time to their shared origin, then forward again.

Nicole Yunger Halpern and Simulating Retrocausality

This is where modern physicists like Nicole Yunger Halpern enter the labyrinth. A theoretical quantum physicist at the University of Maryland and the National Institute of Standards and Technology (NIST), Halpern explores how quantum information might simulate retrocausality.

In 2023, she proposed that certain quantum experiments could appear to send influences backward—not by breaking causality, but by exploiting quantum correlations. Her work suggests that future measurement choices could subtly bias past quantum states, creating the illusion of time-reversed effects.

Halpern's theories don't promise time machines—but they might allow us to encode faint, statistical "messages" from the future. Imagine setting up a quantum experiment today, only for a future scientist's measurement choice to bias its outcome now. It wouldn't be a conversation, but a whisper in the data.

Nick Bostrom and the Philosophy of Time

While physicists like Halpern probe time's mechanics, philosophers like Nick Bostrom ask: Could time travel ever be possible —and if so, where are the tourists from the future?

Bostrom, director of Oxford's Future of Humanity Institute, is best known for his simulation argument—the idea that we might be living in an ancestor simulation. But he's also examined time travel's paradoxes. If time travel were ever invented, he reasons, why haven't we met time tourists? One grim answer: perhaps civilization never lasts long enough to develop the technology.

Signals from the Future?

Experimentalists are now testing these ideas. In 2023, physicist David Arvidsson-Shukur designed lab experiments where photons (light particles) are measured in ways that could reveal retrocausal nudges. The goal isn't to send messages to the past—but to see if the future leaves faint fingerprints on the present.

If these experiments succeed, we might one day extract probabilistic "answers" from the future. A carefully designed quantum system could, in theory, encode a yes-or-no question —l ike "Will humanity survive the next millennium?"— and a future observer's measurement might skew today's results.

Conclusion: Not a Time Machine, But a Crack in Time

Classical time travel— stepping into a brass-and-lights machine— still seems barred by physics. But quantum mechanics may offer a different kind of temporal link. Not through wormholes, but through the subatomic dice-rolls of reality.

We may never stroll with dinosaurs or warn past selves. But we might, one day, decode faint murmurs from the future—not in words, but in the subtle tilt of quantum chance.

The labyrinth of time isn't just behind us. It twists through the present—and its deepest passages may still lie undiscovered.

Phasmids

10

Bernardo Kastrup

AI won't be conscious and here is why

What's the difference between artificial intelligence and artificial sentience? There is no doubt artificial intelligence is real. It is here, and soon we could have AI that is more intelligent than human beings. But the question is, will that intelligence be accompanied by inner experience the way the data processing in humans' brain is accompanied by inner experience? Being human is to be conscious and sentient. Is there something it is like to be an AI—a silicon computer that has measurable intelligence? That's the question we will deal with.

Is your thermostat conscious?

Intelligence does not come naturally paired with sentience. Intelligence is smart data processing for the achievement of a certain goal. Sentience may be smart or may not be. Maybe bacteria are sentient. And they are, but not very much.

The more I thought about it over the years the clearer it became to me that whatever I did to design, I would only change structure and function. And nothing about structure and function would get me any closer or any further away from believing that such a thing is conscious. And it took me a couple of years to realize this, but the conclusion was inevitable. I was making some assumptions that didn't hold, because that's the only way you find yourself in an alley without an exit. You took a long turn somewhere, and that long turn was the notion that material arrangements somehow generate experience.

That assumption in my mind today is absolutely self-evidently and obviously wrong. But the key point here was this: if an epoxy—melted sand and metal—can be conscious, then what cannot? We keep asking questions about whether AI is conscious. Why not your home thermostat? The reason we use electrons in electronics—relying on their motion—is simply a matter of convenience, because computers primarily operate using logic gates and memory elements. You can also build logic gates and mem-

ory elements using pipes, pressure-driven valves, and water. The pipes would serve as the metal traces, the valves as the transistors or gates, and the water as the electricity—in principle, all with pressure-driven valves, pipes, and water. The fact that we use electrons is just because with electrons, it's much smaller, much more economical, and fits in your pocket. Because if we did this with pipes and water, it would be the size of a moon.

There is absolutely nothing mysterious about what's happening in a silicon chip—it's completely mechanistic. There is no magic—no, woo. Computations are always mechanistic.

The religion of sentient AI

What is the argument of the AI sentience advocates—the gurus of the new-world religion?

It goes like this:

"Consciousness is a mystery—we don't know how it comes about, so we cannot rule anything out on that basis. Consciousness contains information: there is something it is like to dream of a green elephant, and something else it is like to feel hunger. Yes, there is information in consciousness. And if we can build a silicon computer that exhibits patterns of information flow similar, in some meaningful way, to those in the human brain, then it might be conscious."

The argument is an appeal first to ignorance and mystery, then to the vagueness of the notion of information, and finally, to very abstract patterns of similarity.

Let's walk through this and see whether this statement holds any water. The first flaw is in the argument. I can run a simulation of kidney function on my computer. It is known at the molecular level how kidney function works, and it is possible to make a molecularly accurate simulator that I can run on my computer. Do I run any risk that my desk will be ruined by a pinging computer? Of course not.

The simulation of a phenomenon is not the phenomenon. There is a correspondence of form at some abstract level through analogy that makes the simulation useful and grounds it in reality in some sense. But the simulation is not the thing simulated. We all understand it for urine and for everything else, with one exception—consciousness.

The second flaw is that we are so lost in abstraction and fantasy—mistaken as reasonable science and philosophy—that we lose sight of the obvious: a brain does not look like a computer chip, and it doesn't function like one either.

The brain runs on carbon and hydrogen, powered by ATP and neurotransmitters across synapses. Chips run on silicon and electricity, using charge accumulation in gates. These are fundamentally different mechanisms, not meaningfully equivalent. To find a similarity—which is what the people peddling the religion of sentient AI try to do—you have to apply so many layers of abstraction. And each step of abstraction takes you further away from reality. But if you slice open a brain in this stage and I slice open a computer, they will look very bloody different. And then some people say, 'You know, there are electric potentials here and there, so AI can be conscious.' There are electric potentials in your home thermostat too. Are we seriously considering the possibility that your home thermostat might be conscious? If we were, we would think twice about exchanging it or turning it off. That would be murder.

And what about ChatGPT? It looks and sounds so human. Is it sentient? The answer to that is painfully obvious too—the thing was designed to look human. When you look at the mannequins in a shop window, they look human. Are we assuming they might be sentient or conscious? No. Because we know that the similarity here does not betray underlying processes that are equivalent. The similarity here comes from the fact that they were constructed to look like humans.

The same goes for ChatGPT—it was built to sound like us. Why? Because it's just a natural language interface designed to help you search. Instead of doing a Google search and getting a list of matches, ChatGPT simply summarizes the search results for you in natural language. It doesn't understand anything it's saying.

ChatGPT uses language in a way that is disconnected from reality. For humans, when we see a tree or a pink elephant in a living room, images of these things pop into our minds. To ChatGPT, the inputs it processes are merely squiggles and symbols devoid of inherent meaning. These are simply woven together in a way that it has learned through its own Internet searches, and it returns those words to you in natural language.

One cannot take superficial similarity as a reason to consider artificial intelligence sentient, just because we design AIs to resemble humans. We're quickly approaching the uncanny valley. But why is this so widely discussed? Because everything I've mentioned is painfully obvious.

False premises underlying the questionable belief

There are several reasons why we keep on talking about sentient AI.

The first one—and this may come as a surprise to you—is that it's usually computer scientists who talk about sentient AI. They're the ones building careers, writing books, and making money on the talk circuit, all while discussing AI. And because of their title—computer scientists—we assume they must know what a computer actually is. But the truth is: they often don't.

Computer scientists are trained to be power users. For them, a computer is a tool. If you find one who has even opened up their PC to look inside, that's already rare. And even those who do—when they see the black little components—they usually see just that: blank boxes. Electricity goes in. They can program

those boxes using layers of abstraction: the operating system, APIs, libraries, compilers. But they never see the "bare metal," as engineers call it. And they don't need to—that's how universities train them.

But the result is that they often don't know what they're talking about. They don't understand how their declarative programming languages are ultimately transformed—through layers of APIs, libraries, compilers, assemblers, and linkers—into the actual low-level operations, the gates flipping open and closed.

The second reason is "woo delight"—the computer is sentient, it's a form of sensationalism. And it's very attractive. It creates a buzz. Netflix produces series based on this kind of so-called science fiction.

Another reason—more psychological than scientific—might be tongue-in-cheek, but I present it with some seriousness: "womb envy."

Freud once talked about "penis envy," a now-discredited idea. But it was an attempt to explore unconscious desires. In that spirit, I suggest womb envy—the awe (or envy) some men may feel toward women's ability to create conscious life. It's a kind of superpower, and perhaps the drive to build conscious machines stems partly from a wish to replicate that miracle.

The problem begins when this psychological urge gets mistaken for scientific reality. That's when fantasy slips in disguised as reason—and philosophy needs to step in.

One more reason is careerism. There are people in academia who receive public funding—your tax money—to work on the ethics of AI. They ask questions like: How should we treat conscious AI? Should we be kind to them? Do they have rights? Human rights? And so on.

Then there's what I'd call the religion of the religion of a materialist worldview. It even has a bible: *The Singularity is Near*, a book written by Ray Kurzweil in 2005. All the elements of reli-

gion are there—prophecy, a savior, divinity. The conscious AI is portrayed as a superintelligence so powerful that it will gain absolute control over nature, cure all diseases, and provide us with food, comfort, and entertainment. We will live in a kind of digital Garden of Eden. That's paradise. That's God.

Except in this case, it's a kind of ego-driven God—one we create ourselves before it begins to take care of us. All of this is sugar-coated in a culturally manufactured sense of plausibility.

Sentience and the Limits of Analogy

What does nature tell us? Let's set aside fantasy for a moment. First and foremost, we have private conscious lives. We are sentient—that's the undeniable core of our experience.

Other living beings behave in ways strikingly similar to us. They flinch from pain, show fear, desire, even jealousy. Since our behaviors are shaped by consciousness, it's reasonable to think theirs might be too.

Even single-celled organisms like amoebas show surprising complexity—building cone-shaped shells from mud, for example.

All living beings metabolize and operate on the basis of DNA, which is highly specific and incredibly rich in information. Metabolism itself is a complex process. Protein folding, for example, is extraordinarily intricate—it remains a mystery how our bodies manage it in mere seconds, while a supercomputer might take a long time to figure out the same thing.

Despite the wide diversity of life forms, if you look at them under a microscope, they all share common features. And this is completely different from everything else in nature. So we are well-grounded when we extrapolate our own sentience to other living beings. There are plenty of reasons to make that analogy—based on form, function, and behavior. But to extend that analogy to a conscious piece of melted sand and metal—that goes a little too far.

The counterargument to this is that consciousness could be multiply instantiated. For instance—flight. Birds fly, airplanes are totally different from birds, and airplanes fly too. Could this be the case for consciousness? If we are operating at this level of abstraction, the answer would have to be yes. But there are obvious differences. In the case of flight, we understand the underlying mechanisms. But when it comes to consciousness, we don't have that understanding, so we are operating in the dark. It is impossible to maintain this opinion in an informed way if you don't know the underlying mechanisms. And this is another appeal to ignorance.

It's true—we can't categorically refute that silicon computers might become conscious. But then, we can't categorically refute the existence of a hyperdimensional flying spaghetti monster either. Most absurd ideas can't be strictly disproven. The better question is: what reason do we have to take it seriously? And there's no good reason to see sentient silicon as a plausible hypothesis.

Panpsychism

There's a strong philosophical assumption behind the idea that silicon-based computers could be conscious: it often presupposes panpsychism—the belief that subatomic particles like electrons and quarks possess some form of consciousness. In this view, building AI means arranging these preconscious elements into a system that produces conscious experience. A good example is a 20-year-old book by Penty Heikkinen from Nokia Research, which presents perhaps the strongest case for conscious machines. But a close philosophical reading reveals it rests on the unspoken belief that particles are inherently conscious—a very specific metaphysical stance.

The problem with this metaphysics is that it doesn't hold up. First, there's no coherent explanation for how tiny "micro-sub-

jects" like electrons or quarks could combine into a unified conscious experience. Some philosophers argue this is incoherent even in principle.

Second, neurons don't physically touch; they communicate chemically across synapses. This breaks the illusion of a neat, continuous structure of conscious "building blocks."

Third, there's a flawed assumption that because consciousness arises from matter, it must be made of tiny conscious parts. But that's like saying if your pixelated video looks blocky, you must be made of rectangles.

Most crucially, this idea was undermined by modern physics nearly a century ago. Elementary particles aren't discrete "things"; they're excitations of underlying fields—like ripples on a lake. You can't take the ripple out and put it in a box, because the ripple is the lake in motion. Similarly, there's nothing to a particle beyond the field. Taking the metaphor of "particles" literally leads to confusion—especially when it's used to argue that consciousness might emerge from their supposed individuality.

Some philosophers attempt to sidestep this issue by suggesting that quantum field theory might be incorrect, and that perhaps an alternative framework like Bohmian mechanics is more accurate. But consider the implications of discarding quantum field theory: we would have to reject an enormous body of experimentally verified phenomena, including quantum fluctuations and the concept of quantum foam—the well-documented fact that even in a perfect vacuum, particles spontaneously appear and disappear.

If particles were truly "things," these phenomena would amount to magic—objects emerging from nothing and vanishing into nothingness without cause. But within the framework of quantum field theory, these events make sense. They are behaviors of fields, not the comings and goings of independent, material "stuff."

Another example—particle decay. People think we measure the Higgs boson at CERN but we have never done it. It decays much too quickly before it interacts with any measurement surface. What we measured is the product of the decay of the Higgs boson. For instance, we discovered much to our surprise that the Higgs boson can decay into two muons. But there are no two muons in the Higgs boson, because the Higgs boson is not a thing, it's a ripple. And when ripples lose energy, they acquire other different physical properties. Ripples interfere with one another. That's why you start from a Higgs boson and you end up with two muons because the ripple changes as it loses energy. But if you think of particles as little things, now the Higgs boson is disappearing into nothing, and magically two muons are popping up out of nothing.

The final example which is very close to us in life. We all know what inertia is—when you're starting from zero, you have stopped on your bike at a traffic light, and then the light goes green and you start moving. It's much more difficult to start moving than to keep moving. That's called inertia—mass resists acceleration. It resists changes in the velocity vector, either in direction or in speed. How do we account for that? The way we account for that is with the Higgs field. And although the following idea is not accurate, it's not too far from reality. It's the following metaphor. The Higgs field is a kind of viscous fluid, like water in a swimming pool. It's when you are not moving and then you want to start moving, you have to win over that viscosity. And the Higgs field does that. The metaphor doesn't work the other way around, because the Higgs field also resists reductions in velocity.

That's what accounts for inertia. And the fact that we found the Higgs boson betrays the existence of the Higgs field. That's why the Higgs boson is important, because it betrays the existence of its field.

Artificial sentience is possible

How do we proceed with some clarity, then? When we talk about artificial sentience, we're speaking about more than just creating consciousness from non-consciousness. What we mean is that a computer would have a private inner life of its own—not just consciousness, but a consciousness confined within the boundaries of the machine itself.

But to say that consciousness is bounded by the limits of an object assumes that the object has a real, independent existence—that it exists "out there" in the world. Yet inanimate objects are carved out of the unified fabric of the physical universe. We define them through language and convenience, not because they are inherently separate.

Where does the river end and the ocean begin? Is there really a river and an ocean, or is it all one continuous thing, to which we assign different names simply because it helps us navigate the world?

Another example: if you're a panpsychist, you might say, "Well, the table is conscious." Okay—but what if I remove one of the table's legs? Does the leg now have a consciousness of its own, separate from the table? What if I nail the leg back on—does the consciousness merge again?

Or consider a boulder on top of a hill. It's shaped by erosion, part of the hill. But if it cracks and rolls down, does it suddenly become a separate consciousness? Does it merge back every time it touches the hill as it bounces down?

These kinds of questions sound absurd because of this: we mistake the structure of language for the structure of reality. And that's where all this nonsense comes from.

We tend to believe that anything we have a name for must correspond to a real, distinct entity. Take the example of a fist. When I close my hand, I can point to it and say, "Here is a fist."

But what happens when I open my hand? Where did the fist go? Did something magical occur the moment I opened it, or are we simply misunderstanding something fundamental?

The confusion lies in attributing the structure of our language—the names we assign—to the structure of nature itself. We mistake linguistic constructs for real, separate entities, when in fact many are just ways of referring to particular configurations of things that already exist. We carve out categories in language and then project those categories onto nature.

To conclude, there are no computers. It's convenient to refer to this thing that performs a certain function, but we have no metaphysical or ontological grounds to say that this thing is somehow separate from the rest of the world around it. The computer is a subset of pixels on perception that is convenient to give a name to. But we cannot carve it out from the rest of reality and say it is an entity, and then ask if it is conscious or not. The only entities that we have objective grounds to carve out from the rest of nature are the boundaries of living beings. Because if you stick a needle into my arm, I feel it. But if you stick the needle into the arm of my chair, it doesn't feel it. There is an objective way to say living beings do have boundaries and nothing else has. If Big Bang theory is right, there are good physical reasons to think of the entire inanimate universe as entangled and therefore not describable in terms of proper parts, and can only be described as a whole.

Keeping this in mind is essential as we move forward. And finally—despite everything mentioned before—will we eventually be able to create artificial sentience? I believe the answer is yes. But when we do, it will likely resemble a living being.

The challenge of creating artificial sentience is, at its core, the challenge of abiogenesis—the process of creating life from non-life. And there's no reason to believe we couldn't eventually

learn how to do that. After all, it has already happened at least once in the history of the universe. So we know it's possible. Why wouldn't we be able to figure it out?

I believe we can. But it won't be a silicon computer running on GPUs. It won't even be a neuromorphic architecture based on analog rather than digital design. Those systems are still far too different from the biological processes that give rise to consciousness.

Passer domesticus

11

Rupert Sheldrake

Morphic Resonance and the Memory of Nature: An Update

The hypothesis of morphic resonance proposes that memory is inherent in nature, and the so-called laws of nature are more like habits. I first put this idea forward in my book *A New Science of Life: The Hypothesis of Formative Causation*, published in 1981 (third edition, 2009). This hypothesis was developed further and more fully in *The Presence of the Past*, in 1988 (second edition, 2011).

In the 1980s many scientists were confident that all the outstanding problems of science would soon be solved, confirming the assumptions of the prevailing paradigm of mechanistic materialism. Biologists would explain the nature of life mechanistically through molecular biology, especially through sequencing the genomes of humans and other organisms. Psychologists would understand the nature of the mind through brain scanning and computer modelling. In the computer sciences, artificial intelligence would be created in machines rivalling and even exceeding the intelligence of humans themselves. In physics, through the development of the ultimate theories of everything, such as superstring theory, the origin of the universe and everything in it would be explained in terms of mathematical equations involving multiple dimensions. In 1996, the American science writer John Horgan published a book entitled *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. After talking to leading scientists in many areas of research, he concluded:

> *My guess is that this narrative that scientists have woven from their knowledge, this modern myth of creation, will be as viable a hundred or even a thousand years from now as it is today. Why? Because it is true... There will be no great revelations in the future compared to those bestowed upon us by Darwin or Einstein or Watson and Crick.*[1]

When many scientists thought that all the most fundamental problems were already solved, or almost solved, it was difficult to

persuade biologists and chemists to take up morphic resonance research, and almost impossible to find funding for it. There was more openness among some psychologists, and several experiments were carried out to test for morphic resonance in human learning, mostly with positive results, as summarized in the new edition of *A New Science of Life*.

However, confidence in the promises of materialist science is fading fast, as I show in my book *The Science Delusion* (2012; second edition, 2020). The prospects for the sciences now look very different.

The laws of nature as habits

Twenty-first-century physics has run into seemingly intractable problems. Superstring and M (for Master) theories, with ten and eleven dimensions respectively, take science into completely new territory in that they are untestable. One of their critics wrote a book called *Not Even Wrong*[2], and a growing number of eminent physicists are worried that theoretical physics has lost its way.[3]

Meanwhile, cosmologists have come to the conclusion that known kinds of matter and energy constitute only about 5% of the universe. The rest consists of dark matter and dark energy. The nature of 95% of physical reality is literally obscure.[4]

Within cosmology, there has been much discussion about the Cosmological Anthropic Principle, which asserts that if the laws and constants of nature had been slightly different at the moment of the Big Bang, biological life could never have emerged, and hence we would not be here to think about it. So did a divine mind fine-tune the laws and constants in the beginning? Many cosmologists are atheists and prefer to believe that our universe is one of a vast and perhaps infinite number of parallel universes, all with different laws and constants. All these other universes actually exist, even though there is no evidence for them. We just happen to live in the one that has the right conditions for us.

In the eyes of sceptics, the multiverse theory is the ultimate violation of Occam's Razor, the principle that entities should not be multiplied unnecessarily. But even so, it does not succeed in getting rid of God. An infinite God could be the God of an infinite number of universes.[5]

But if the regularities of nature are evolving habits rather than eternal laws, there is no need to assume that all the laws and constants were fixed at the moment of the Big Bang, like a kind of cosmic Napoleonic Code. Hence there is no need to suppose that all these laws and constants were intelligently designed at the moment of creation, or that there are an infinite number of unobserved universes. These extravagant hypotheses become unnecessary if nature is radically evolutionary, as the hypothesis of morphic resonance proposes. This is now a matter of serious discussion within physics.

In 2013, the cosmologist Lee Smolin published a book called Time Reborn: From the Crisis of Physics to the Future of the Universe, in which he argued that the laws of nature are more like habits, and proposed a 'principle of precedence' which depends on a new kind of interaction across space and time 'whereby a physical system can interact with copies of itself in the past.' [6] The result would be something very like morphic resonance. Smolin appears to have come to this conclusion independently, based on a consideration of the principles of physics within an evolutionary universe.

The genome wager

From the point of view of morphic resonance, much of the inheritance of form and behaviour depends on resonance with previous organisms of the species, through a kind of collective memory. It is not coded in the genes, which have a relatively limited role: they code for the sequence of amino acids in proteins, and some genes are involved in the control of protein synthesis.

But there is a huge gap between the amino-acid sequences in protein molecules and, say, the shape of an eagle's wing or the web-spinning instincts of a spider. Nevertheless, many biologists were confident that genes would enable all aspects of living organisms to be predicted. One of them was the late Lewis Wolpert, an eminent British developmental biologist. In 2009, he and I entered into a wager about the predictive power of the genome, published in New Scientist magazine. At stake is a case of fine port, for which we paid half each, and which experts say will be in peak condition by 2029. It is being stored in the cellars of the British Wine Society. Wolpert bet that the following will happen, and I bet that it will not:

> *By 1 May 2029, given the genome of a fertilised egg of an animal or plant, we will be able to predict in at least one case all the details of the organism that develops from it, including any abnormalities.*

If the outcome is not obvious, then the Royal Society will be asked to adjudicate.[7]

Wolpert's wager was based on the assumption that genes 'program' or 'code for' almost all aspects of an organism's development and behaviour. They enable organisms to make their proteins, by coding for the sequence of amino acids that are strung together to form the primary structure of these proteins, which then fold up into complex three-dimensional forms. Then these proteins interact and catalyze biochemical reactions in such a way that they somehow give rise to adult organisms. However, random molecular permutations simply cannot explain how organisms work. Instead, cells, tissues and organs develop in a modular manner, shaped by morphogenetic fields, first recognised by developmental biologists in the 1920s. Wolpert himself acknowledged the importance of such fields. Among biologists, he is best known for the idea of 'positional information', by which cells 'know' where they are within the field of a

developing organ, such as a limb. But Wolpert believed that morphogenetic fields could be reduced to standard chemistry and physics. I disagree. I believe these fields have organising abilities, or systems properties, that involve new scientific principles and also include an inherent memory given by morphic resonance.

In the 1980s, there was great excitement when a family of genes called homeobox genes was discovered in fruit flies. Homeobox genes determine where limbs and other body segments will form in a developing embryo or larva; they seem to control the pattern in which different parts of the body develop. Mutations in these genes can lead to the growth of extra, non-functional body parts; for example a leg may develop instead of an antenna.[8] At first sight, homeobox genes appeared to provide the basis for a molecular explanation of morphogenesis: here were the key switches. At the molecular level, homeobox genes act as templates for proteins that 'switch on' cascades of other genes.

However, research on other species soon revealed that these molecular control systems are very similar in widely different animals. Homeobox genes are almost identical in flies, reptiles, mice and humans. Although they play a role in the determination of the body plan, they cannot themselves explain the shape of the organisms. Since the genes are so similar in fruit flies and in us, they cannot explain the differences between flies and humans.

It was shocking to find that the diversity of body plans across many different animal groups was not reflected in diversity at the level of the genes. As two leading developmental molecular biologists have commented, 'Where we most expect to find variation, we find conservation, a lack of change.'[9]

This study of genes involved in the regulation of development is part of a growing field called evolutionary developmental biology, or evo-devo for short. Once again, the triumphs of molecular

biology have shown that morphogenesis itself, the coming into being of specific forms, continues to elude a molecular explanation. That is why the idea of morphogenetic fields is more important than ever. These fields, shaped by morphic resonance, mould the development of organs and tissues; development is modular.

The homeobox genes seem to affect the tuning of developing systems to one morphogenetic field rather than another, rather than coding for the structures of these organs, rather like the tuning system of a TV set selects one channel rather than another. The details of the TV programs are not contained within the tuning circuit; the TV set tunes in to them, just as organisms tune in to morphogenetic fields, which shape the development of form, or tune in to behavioural fields, which underlie the organization of instincts.

Epigenetic inheritance

One of the biggest controversies in twentieth-century biology concerned the inheritance of acquired characteristics, the ability of animals and plants to inherit adaptations acquired by their ancestors. For example, if a dog were trained to learn a new trick, its offspring would tend to learn it more easily. The opposing view, promoted by August Weismann and by the science of genetics, denied that organisms could inherit features their ancestors had acquired; they only passed on 'determinants' or genes that they had themselves inherited.

In Charles Darwin's day, most people assumed that acquired characteristics could indeed be inherited. Jean-Baptiste Lamarck had taken this for granted in his theory of evolution published more than fifty years before Darwin's, and the inheritance of acquired characters was often referred to as 'Lamarckian inheritance'. Darwin shared this belief and cited many examples to support it. In this respect Darwin was a Lamarckian, not so much

because of Lamarck's influence but because he and Lamarck both accepted the inheritance of acquired characteristics as a matter of common sense.

Lamarck placed a strong emphasis on the role of behaviour in evolution: animals' development of new habits in response to needs led to the use or disuse of organs, which were accordingly either strengthened or weakened. Over a period of generations, this process led to structural changes that became increasingly hereditary. In this respect too, Darwin agreed with Lamarck, and he provided various illustrations of the hereditary effects of the habits of life. For example, ostriches, he suggested, may have lost the power of flight through disuse and gained stronger legs through increased use over successive generations. Darwin was very conscious of the power of habit, which was for him almost another name for nature.

The neo-Darwinian theory of evolution, which became orthodox in the West in the twentieth century, differed from the Darwinian theory in denying the inheritance of acquired characteristics in favour of genes. Lamarckian inheritance was treated as a heresy.

However, the taboo against the inheritance of acquired characteristics began to dissolve around the turn of the millennium. There is a growing recognition that some acquired characteristics can indeed be inherited. This kind of inheritance is now called 'epigenetic inheritance'. In this context, the word 'epigenetic' signifies 'over and above genetics'. Some kinds of epigenetic inheritance depend on chemical attachments to genes, particularly of methyl groups. Genes can be 'switched off ' by the methylation of the DNA itself or of the proteins that bind to it.

This is a fast-growing field of research, and there are many examples of epigenetic inheritance in plants and animals. For example, in Daphnia, the water flea, when predators are around, the water fleas develop large defensive spines. When they repro-

duce, their offspring also have these spines even if they are not exposed to predators.[10] In a recent study with mice, males were made averse to a synthetic chemical, acetophenone, by receiving electric shocks after smelling it. After mating, they had no further contact with the mothers of the next generation, nor with their offspring. But their children and even their grandchildren showed a strong aversion to acetophenone: they had inherited the fears of their fathers and grandfathers.[11] These effects could involve a combination of chemical changes in the sperm cells and morphic resonance. Epigenetic inheritance also occurs in humans.

Many biologists now argue that the gene-centred neo-Darwinian theory of evolution is too limited, if only because it ignores epigenetic inheritance, which can have major evolutionary effects, enabling adaptive changes to occur much quicker than they would through random genetic mutations and many generations of natural selection alone. The new 'Extended Evolutionary Synthesis' gives a much more comprehensive and inclusive view of evolution than neo-Darwinism.[12]

In his genome wager, Wolpert ignored epigenetic inheritance entirely, and based his prediction on the gene-centered, neo-Darwinian ideas that still dominated biology in the early twenty-first century, and which underlay the high hopes engendered by the Human Genome Project, the sequencing of all the genes in the human hereditary material. When the first draft was published in the year 2000, in a celebratory speech President Clinton described it as 'the book in which all human life is written.'

The missing heritability problem

Despite the great technical triumph that it represented, the results of the Human Genome Project have themselves set back the hopes the project engendered. First, our genome contains only between 20,000 and 25,000 genes, far fewer than the 100,000 expected. In contrast, sea urchins have about 26,000,

and rice plants 38,000. Moreover, our genome differs very little from the chimpanzee's genome, the sequencing of which was completed in 2005. As Svante Paabo, director of the Chimpanzee Genome Project, commented: 'We cannot see in this why we are so different from chimpanzees.'[13]

Second, in practice, the predictive value of human genomes turns out to be low. Everyone knows that tall parents tend to have tall children, and studies on the genomes of 30,000 people identified about 50 genes associated with being tall or short. Yet together these genes accounted for only about 5 per cent of the inheritance of height. More recent and ever larger studies have identified many more genes with very small effects that improve the predictive power of individual genomes, and the most successful so far have been with height. On the basis of 5.4 millions people's genomes, a total of 12,111 genetic variants were found to be associated with height, and the researchers claimed that they could predict a maximum of 40% of the variation in height for people of European ancestry, and 10% of those of African ancestry.[14] This is an impressive achievement, but it still leaves much of the inheritance of height unexplained genetically.

The situation is more extreme for many common diseases. For example, genomic studies of the genetic basis of schizophrenia, which is known to be highly heritable, show that a large number of genes are involved but together their predictive accuracy is only about 3%.[15] Steve Jones, professor of genetics at University College London commented that as a result of the missing heritability problem, 'hubris has been replaced with concern'.[16]

The predictive power of genomes is reduced still further by the recognition of epigenetic inheritance. However, the missing heritability that is so puzzling in the light of the genome project may become easier to understand through a combination of epigenetic inheritance and morphic resonance, both of which make major non-genomic contributions to heredity.

More surprising discoveries

Although there have been few direct tests of morphic resonance, in chemistry and biology, there has been a series of unexpected discoveries that provide very promising lines for further enquiry. One is an effect discovered by Miroslav Hill in a cell biology laboratory in France, in which cells seemed to influence other similar cells at a distance. When some cells were subjected to poisons or to high temperatures, and managed to adapt, other similar cells became more resistant even though they had no contact with the first cells, and were not descended from them. The Hill Effect may well be a morphic resonance phenomenon. I discuss further experiments and outline ten new tests for morphic resonance in the new edition of *A New Science of Life*.[17]

An open question

I do not claim that morphic resonance is an established fact. It is still controversial within the scientific world. But the advances of science over the last 42 years have made the hypothesis more, rather than less, plausible. I think it very unlikely that the conventional materialist approach to science can survive much longer. It will inevitably be superseded, and I think it is very likely that a memory principle will have to be recognized within nature, whether it is called morphic resonance or by some other name.

Notes

1 Horgan (1996), *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. London: Little, Brown and Co., p.16.

2 Woit, P. (2007), *Not Even Wrong: The Failure of String Theory and the Continuing Challenge to Unify the Laws of Physics*. Vintage, London.

3 Smolin, L. (2006), *The Trouble With Physics: The Rise of String Theory, The Fall of a Science, and What Comes Next*, Allen Lane, London.

4 Davies, P. (2006), *The Goldilocks Enigma: Why is the Universe Just Right For Life?*, Allen Lane, London.

5 Collins, R. in Carr, B. (ed.) (2007), *Universe or Multiverse?* Cambridge University Press, Cambridge.

6 Smolin (2013) *Time Reborn: From the Crisis in Physics to the Future of the Universe*, p.151.

7 Wolpert, L. and Sheldrake, R. (2009) What can DNA tell us? Place your bets now. *New Scientist*, 8 July.

8 Carroll, S. B., Grenier, J. K. and Weatherbee, S. D. (2001), *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, Blackwell, Oxford.

9 Gerhart, J. and Kirschner, M. (1997), *Cells, Embryos and Evolution*, Blackwell Science, Oxford.

10 Young, E. (2008) Rewriting Darwin: the new non-genetic inheritance, *New Scientist*, 9 July.

11 Dias, B.G. and Ressler, K.J. (2013) Parental olfactory experience influences behaviour and neural structure in subsequent generations. *Nature Neuroscience* 17. 89-96.

12 Müller, G.B. (2017) Why an extended evolutionary synthesis is necessary. *Interface Focus*, 720170015

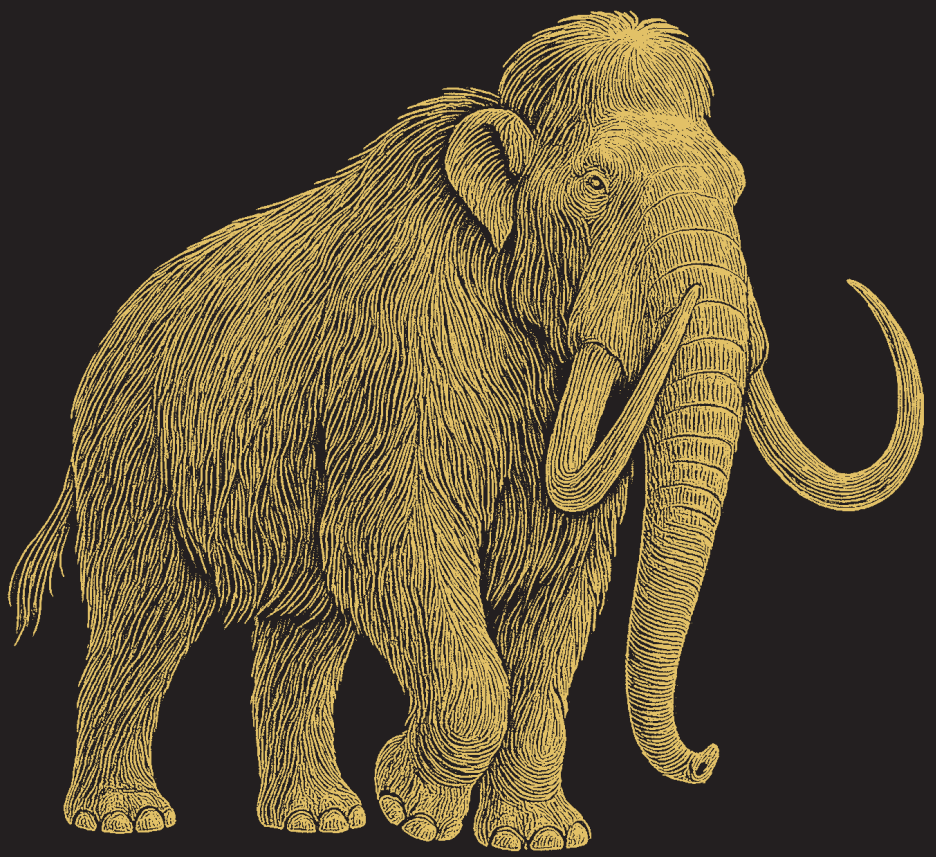
13 Olsen, M. V. and Varki, A. (2004) The chimpanzee genome—a bitter- sweet celebration. *Science*, 305, 191–2.

14 Yengo et al. (2022) A saturated map of common genetic variants associated with human height. *Nature* 610, 704-712.

15 Woo, H.J. et al. (2017) Large-scale interaction effects reveal missing heritability in schizophrenia, bipolar disorder and posttraumatic stress disorder. *Translational Psychiatry* 11;7(4):e1089

16 Jones, S. (2009) One gene will not reveal all life's secrets. *Daily Telegraph*, April 20.

17 Sheldrake, R. (2009) *A New Science of Life: The Hypothesis of Formative Causation* (third edition) Icon Books, London (published in he US under the title *Morphic Resonance: The Nature of Formative Causation*. Park Street Press, Rochester, VT).



Mammuthus primigenius

12

Lucy Cooke

Reviving extinct animals

January 2000, Ordesa National Park, Spain. It is heavily snowing. Celia, the very last Pyrenean ibex, is seeking shelter under a pine tree.

But unfortunately, the pine tree gets covered in snow, and it falls down, and it crushes Celia. She's the very last Pyrenean ibex, and she's now dead. She was 13 years old. It was an unlucky year for Celia, and unlucky for her species.

Fortunately, the year before, scientists had caught Celia, extracted cells from her ear, and put them in liquid nitrogen. And even though in 2000 Celia died, her DNA was still preserved and she could be reanimated.

Scientists took Celia's DNA and put it into a developing cell of a closely related ibex species, took the nucleus out of the cell, which contains all of Celia's cell tissue, and then put that inside a closely related ibex species. They waited for those embryos to develop, and then implanted it inside a surrogate ibex in order to give birth. 500 cloned embryos were produced. 154 were implanted into female goats, which were the surrogates. Only five of those surrogates actually became pregnant, and only one managed to give birth.

On July 30, 2003, a very special ibex was born by c-section in a Spanish research facility. She had wobbly legs, toffee-colored fur, and dark doe eyes, and she was a genetic clone of Celia. The Pyrenean ibex had been raised from the dead, and it was the first species to be de-extincted. Everybody was in awe. Unfortunately, the cloned ibex had a severe fall and survived for only seven minutes.

It was unfortunate because it meant that the Pyrenean ibex made history for becoming the first animal to be de-extincted, and the first animal to go extinct twice. The technology of de-extinction is here and it's very real.

Another close call, this time with a rather remarkable amphibian—a specimen of the gastric-brooding frog. It doesn't incubate its eggs in a pond, like most frogs do. Instead, when the eggs are

laid and fertilized, the female frog gobbles them up, and the whole incubation takes place in her stomach. For a matter of weeks during which she can't eat, as she's not allowed to secrete any gastric juices, otherwise she'd consume her own offspring. But once they've gone through their metamorphosis, she can pull the most unbelievable party trick, and she burps up baby frogs. It's called Propulsive Vomiting.

This incredible creature went extinct in 1983. But in 2011, a team of researchers were able to recover cell nuclei from frozen southern gastric-brooding frog tissue collected in the 1970's. They implanted the cell nucleus into a closely related species—the great barred frog. And some of those eggs managed to form embryos, although none of them actually became tadpoles and frogs.

The same technology that's used with de-extinction is also used to clone living species. The first cloned cat, inevitably called Copycat, was born on December 22, 2001, and lived 18 years. Barbara Streisand, when she lost her favorite pooch, Samantha, in 2017, had her cloned twice. Violet and Scarlet were created using cells from Samantha's stomach. Since then, cloning pets has become big business. Barbara paid $50,000 in order to have Samantha turned into two new dogs.

Is Jurassic Park really possible?

Are scientists going to reanimate dinosaurs from amber and have them walking the Earth? The world's smallest dinosaur, which is a relative of the Archaeopteryx, but just just two inches long, the size of a hummingbird, was discovered in amber in Myanmar in 2020. It's about 100 million years old, and its skull is in almost perfect condition.

DNA has a half-life of 521 years. UV exposure and bacteria break it down, and amber is not a suitable medium for preserving it. However, DNA can last much longer if frozen.

In 2022, a Danish research group managed to extract mastodon DNA from glacial sediments that were 2.4 million years old—DNA that could potentially be used for cloning.

This means that some cool animals could be brought back, as this period coincided with the Great American Biotic Interchange (GABI), when North and South America collided. Many animals that had evolved in isolation in South America became part of this exchange. These included giant ground sloths, terror birds, sabre-toothed cats, and an ancient elephant species.

Of course, top of the de-extinction list is the mammoth. And mammoths mostly disappeared at the end of the last ice age about ten thousand years ago. Although they actually remained on a Russian island until just four thousand years ago, there were still mammoths around at the same time the pyramids were being built. Since mammoths lived in very cold places, their DNA is in abundant supply. One of the most complete mammoths ever found was in Canada. It was discovered by gold miners in 2022 in Yukon's Eureka Creek. So can the mammoth be brought back? Yes, they can.

A Texas-based biotechnology and genetic engineering company, plans to do that. It managed to raise $225 million in order to bring back the mammoth along with the dodo and the thylacine. Its goals are to enrich biodiversity, replenish vital ecological roles, and bolster ecosystem resilience.

One of the looming threats to the world is the melting of permafrost, which lies beneath the Arctic and traps vast amounts of carbon. If it starts melting, it will release enormous amounts of carbon into the atmosphere—twice as much as is currently there. And it is already melting.

Mammoths are ecosystem engineers—they create grassland by eating trees. There used to be a habitat called the Mammoth steppe during the last ice age, and it was the Earth's most extensive biome, spanned from Spain to Canada and from the Arctic to China. Grasslands reflect sunlight better than forests. Permaf-

rost under grassland melts slower than permafrost under forest.

More mammoths = less trees = more grassland = more permafrost. Brilliant. But if there is snow on top of that, rather than making it colder, it insulates, keeps the heat in, and actually the permafrost melts faster. Big mammoths dig it up and create holes, and release some of that heat, and it actually stays much cooler.

How are they going to do that?

With Celia, there was the whole nucleus and the entire genome. In this case, there are just fragments of DNA. They will decode the mammoth's genome, and compare it with its closest living relative, which is the Asian elephant. Then using CRISPR gene editing technology, they splice mammoth DNA fragments into the genome of the Asian elephant, and place that fetus into an Asian elephant to give birth. It will be a hybrid of elephant genes and mammoth genes, a sort of mammophant.

How does evolution work?

There is a thought that evolutionary change is really slow, but sometimes it happens really fast. Back in 2017, a British ecologist called Colin Donoghue was studying evolution at Washington University, and was measuring hurricane lizards. There was hurricane Irma and Maria struck the islands that he was working on, and his life got a whole lot more interesting because when he went back and measured the lizards again, he found that the surviving lizards had longer toes because that's how they'd hung on in the hurricane. In the space of just a few years, the lizards' average size of toes changed dramatically.

What does this have to do with mammoths?

This biotechnology and genetic engineering company has got their hands on 23 different mammoth genomes, and they go back from 10,000 years ago to 700,000 years ago. That's 690,000

years of evolutionary change, and a lot can happen in that time.

No species stands still. It's constantly evolving alongside its environment. And when these mammoth genomes were mapped, scientists found that there are big differences in the genome that code for the way that mammoths store fat, how much earwax they have, their body odor, their salt taste sensitivity, and the size of their ears. Ear size in particular is quite important because that's how elephants control their temperature. Now is the era of climate change and choosing the right size ears is crucial. This newly created mammophant can possess all sorts of genetic flaws. There is a big problem with this mission. Very noble goals. Which are going to be deconstructed

Welcome to Pleistocene Park. It exists today, founded by a Russian geophysicist, and consists of a 16-kilometer park filled with around 100 animals roaming freely, including bison, muskox, moose, yaks, horses, and reindeer, all trampling the ground in an effort to keep the permafrost frozen. The park is investigating whether large herbivores can slow or even reverse permafrost thaw. Their studies show that if scientists de-extinct the mammoth, there could be significant climate benefits. Mammoths would help create a mammoth steppe. They'd dig into the snow, allowing the cold to permeate and minimizing the insulating effect. Mammoths could prevent the carbon dioxide levels in the Earth's atmosphere from tripling. Mammoths to the rescue?

But how many mammoths are required to be created? What is the price tag on that? What's the amount of methane those mammoths are going to produce? Are they going to survive climate change? A Cambridge group of scientists have published a paper in Nature on how climate change killed mammoths. The climate warmed, it got really wet, and that killed the vegetation, and the mammoths starved to death.

Should a different animal be chosen?

What about the auroch? It was also an ecosystem engineer. It was also big, but not as big as a mammoth. The auroch weighed a thousand kilos and was about two meters tall. They roamed all across Europe, and were hunted to extinction in 1672. And like elephants, aurochs created grassland.

European nature thrives in an open habitat mosaic. At Knepp, UK, longhorn cattle are successfully used to perform the role of the aurochs. Oostvaardersplassen, in the Netherlands, pioneered rewilding. Since the 1980s, deer, horses, and cattle have been introduced to manage the habitat. However, the wolf was missing, and as a result, there were too many herbivores, leading to a catastrophe in which many herbivores died.

The group Rewilding Europe has a big plan to bring the aurochs back. But instead of using de-extinction technologies, they are using IVF techniques and backbreeding ancient cattle breeds, as all cattle contain auroch genes. Selective breeding cannot create an auroch, but it can produce a taurine breed. This is the closest possible animal to an auroch, which fills the same ecological niche and helps rewild European habitats.

What does de-extinction often forget?

The biggest concern about these plans is animal behavior. There is a case with the elephant Nzou. They say that elephants never forget, but this particular elephant, who lives in Zimbabwe, has actually forgotten she's an elephant. She was raised with buffalo and now thinks she's one of them.

This sort of imprinting happens, if animals are not raised by their own species. In that case the elephant imprinted on the buffalo, which was all fine until one of the buffalo challenged her over dominance. And she ended up killing the buffalo. And: she has killed 14 to date.

Another example of a project that went a bit awry occurred in Pilanesberg National Park in South Africa. When Pilanesberg was created in the 1970s, there were no elephants, so they were moved from Kruger National Park to Pilanesberg. These were all young, teenage elephants—there were no older elephants. In the early 1990s, the ecologist there discovered a dead rhino. Then he found 12 dead rhinos, and eventually, there were 50 dead rhinos. The culprits he identified were grumpy male teenage elephants, chock-full of hormones. Basically, male elephants go through a period called "must"—a heightened hormonal, testosterone-driven state. In this state, male elephants march around looking for females in heat, with urine dribbling down the backs of their legs, and pick fights. Older males travel solo when in must. But these young, teenage elephants went a bit berserk.

They were full of hormones, and didn't know how to behave. They required adult role models. Six adult male elephants were brought to Pilanesberg, and almost immediately, the problem stopped.

Unless the world desires a lot of randy teenage Frankenmammoths, beating the crap out of muskox in Pleistocene Park, some adult role models will be required there. Not just males, but females too, because females are the matriarchs.

Elephants are incredibly complex social creatures. They can communicate acoustically, visually, through touch, chemicals, and even seismically; they send rumbling messages long distances that they then pick up and listen to through their feet. If scientists recreate a mammoth, or a mammophant, how are these mammoths going to learn how to behave without elders?

Experiments with captive breeding

Captive breeding experiments have not been successful either. Captive-bred animals struggle to be reintroduced into the wild. For example, a panda, which is much less socially complex than

an elephant, is essentially solitary for most of the year. But the experiment has not gone well. In the Chinese panda breeding centers in Chengdu and Wolong, the pandas looked like pandas, but they didn't behave like pandas because they were not raised by wild pandas. Pandas have this reputation for being pathetic animals that are bad at reproduction. That's just complete rubbish. Pandas are perfectly capable of reproducing; they just don't like doing it in a concrete cell.

In the wild, when a female is in heat, she'll mark trees and spread her pheromones around, attracting a bunch of different male pandas, who then have a sort of urinary competition. It's like a urinary Olympics: the males try to squirt their pee the highest up a tree. Then, the female will sniff to see who reached the highest.

The Chinese have managed to get around the problem of panda reproduction in zoos by using artificial insemination. Many pandas have been created in captivity, but the trouble is that they were raised by humans. In 2006, the first captive-bred panda was released into the wild at Wolong Nature Reserve after three years of wilderness survival training. Ten months later, Xiangjiang was found dead, savaged by wild pandas. He didn't know how to interact with them and was killed. Raising pandas in captivity and then releasing them into the wild can be compared to throwing a chihuahua into a pack of wolves.

Of the 10 pandas that have been released, nearly as many have died in the process. Two have died in the wild from attacks or infections, and another six have died in a pre-release program. Since 1995, more pandas have been removed from the wild than have been released.

The issue with inbreeding depression

The other part of the biotechnology company's mission is to de-extinct the dodo. In 2021, using a specimen in Copenhagen's

Natural History Museum, they sequenced the genome. It should then be compared with the dodo's nearest living relative, the endangered Nicobar pigeon, to figure out which genes will be needed. But why should it be de-extincted?

While mammoths are considered ecosystem engineers, dodos are not. They were responsible for spreading the seeds of the tambalacoque tree, but now tortoises fulfill that role. There's no real need to bring the dodo back, and there's no place for it to live. The dodo cannot be reintroduced to the mainland because it would not survive. The only place suitable for the dodo, with no invasive species, is a small island off the coast of Mauritius, which still has forests and could essentially serve as a living museum for it.

What will happen then is that they won't be able to have many dodos, and the population will suffer from inbreeding depression. When there are just a handful of remaining animals, genetic diversity disappears, and the species can no longer adapt to change, like the hurricane lizards did. It would also be highly susceptible to disease and parasites.

There are only two remaining northern white rhinos: both are females, a mother and daughter. Because no males are left for them to mate with, they are considered functionally extinct—the walking dead. However, there is a plan to use their eggs and sperm harvested from the last male northern white rhino, now deceased, to create test-tube rhinos and implant them into southern white rhinos. The challenge, however, is that even if this process is successful, the genetic diversity will be extremely low.

This company has a plan: they are now examining a catalog of museum samples of northern white rhino specimens from the past, including bones, dry skin, and preserved organs. These samples could be used to extract ancient DNA, which may help restore lost genetic diversity in the cell lines and produce more northern white rhino embryos, ultimately preventing inbreeding.

The black-footed ferret was once assumed to be extinct. Then, a surviving individual was discovered, prompting biologists to gather the last 18 black-footed ferrets and do an intensive breeding program. But these ferrets were highly genetically similar. To address this, cloning technology was used to introduce DNA from another black-footed ferret population that had been frozen and preserved. This led to the creation of Willa, who was three times more genetically diverse than the entire existing population of black-footed ferrets. Two more genetically diverse ferrets were later produced, significantly increasing the genetic diversity of the remaining population. The goal is to establish a stable population of 3,000 ferrets across their natural range. A similar approach was taken with Przewalski's horse.

De-extinction technology becomes viable and useful when used as a tool to support animals on the brink of extinction. It can help prevent genetic similarity, reduce inbreeding depression, and aid in their recovery. However, the biggest challenge remains: where will all these animals live?

There is also climate change and habitat destruction. These technological solutions absolve personal responsibility and suggest that science is going to save the world. People don't need to make difficult changes, and can carry on with the consumer culture that's destroying the planet. The de-extinction technology is more of the hubristic thinking that got humanity into this miserable point in the first place.

It's the sort of Elon Musk solution to the environmental catastrophe. But what the planet really needs are holistic solutions that involve people making personal sacrifices and difficult decisions.

Let's focus on the present instead of obsessing over the past. But of course that isn't as sexy or sci-fi as bringing back the mammoth.

Castor Fiber

13

Philip Goff

Galileo's Error

To understand the problem of consciousness, we must trace it back to the Scientific Revolution. More than almost anyone else, Galileo shaped the philosophical foundations of emerging science. One of his big ideas was that science should describe the physical world in a purely mathematical vocabulary. That is something we take for granted now, but at the time it was a revolutionary move. This is captured in his famous quote from The Assayer, in 1623, where he claimed that the universe is written in the language of mathematics:

"Philosophy is written in this grand book, the universe which stands continuously open to our gaze, but it cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometrical figures, without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth."

This aspiration toward complete objectivity is what Thomas Nagel later called the "view from nowhere". To see what this means, imagine aliens with sensory systems radically different from our own. They might not understand our music or art, but if they could grasp mathematics, they would understand our physics. Science aims to provide a purely objective description of reality, one that could be understood by anyone, regardless of their particular sensory or cultural background.

Mathematical equations capture quantities, but it is hard to see how they could capture qualities. Qualities are smells, colors, sounds, and tastes. Quantities are what can be captured in the mathematical language called physics. No mathematical equation seems capable of capturing what the deep red of a sunset is like. The problem is that the qualities that fill our experience of the world are inherently subjective. They can only be grasped from a particular perspective, by someone who has experien-

ces. Consider a scientist blind from birth. No matter how much physics they learn, they will never know what it is like to see red. This illustrates the fundamental tension Galileo confronted: qualities versus quantities, the subjective versus the objective. If science is to be fully objective, what are we to do with these apparently subjective qualities?

Galileo's solution was radical. He didn't solve the problem experimentally. Instead, he introduced a new philosophical theory of reality. According to this view, qualities are not part of the physical world. The yellowness of the bottle is not on its surface. The smell is not in the air. These qualities exist only as experiences within the consciousness of the observer. Once the physical world was stripped of these subjective qualities, everything that remained could be captured in the purely objective language of mathematical physics. This marked the beginning of mathematical physics as we know it. But where did the qualities go? Galileo could not place them in the brain, because then the brain itself would not be fully describable in objective terms. So he placed them in the soul, outside the domain of science.

Galileo's error: what went wrong?

The upshot is striking. The science we have been practicing for the last four hundred years is built on a form of mind-body dualism: a separation between the physical world described by science and the realm of subjective experience. This proved enormously fruitful, enabling extraordinary scientific and technological progress. Yet the cost of this success is often forgotten. From its inception, the scientific enterprise placed consciousness outside its explanatory domain. This exclusion was not the result of an empirical discovery about the nature of reality itself. Rather, Galileo introduced a deliberate methodological and philosophical decision about what would count as describable within science. The resulting framework gave rise to modern physics, offering a

powerful account of structure, motion, interaction, and relation. But it also left us with a world that moves without feeling, an experientially silent universe.

Mathematical physics thus captures only one aspect of physical reality. If consciousness is to be brought back into view, science and philosophy must once again work together. Philosophy is, after all, the mother of the sciences, and its task here is to make explicit the assumptions that continue to shape how we interpret the world around us, rather than allowing them to operate silently in the background.

The comeback of consciousness into science

If consciousness is to be reintegrated into science, we must first recognise that science explicitly excluded consciousness from its domain. For much of the twentieth century, consciousness was effectively a taboo in science, something science simply was not meant to address. Then, in the 1990s, consciousness suddenly re-emerged as a "serious scientific problem". Many people still underestimate how unusual this problem is. Consciousness cannot simply be slotted into science as it currently exists. So where do we go from here? One response is to claim that Galileo was mistaken. Perhaps objective, mathematical science really can capture consciousness after all. On this view, we're grateful to Galileo for getting science off the ground, but ultimately consciousness will be explained in terms of physical processes in the brain. This position is known as materialism, and it remains the dominant view in both mainstream science and philosophy.

At first glance, this may seem reasonable. On closer examination, however, it becomes difficult to sustain. For the same reasons we've already discussed, a purely objective description of the physical world, expressed in mathematical or quantitative terms, is always going to leave something out. It necessarily omits

subjective qualities. If our account of the brain is framed solely in the objective language of electrochemical signalling and neural activity, then those subjective qualities are simply absent from the picture. They're not explained; they're excluded. For this reason, materialism does not seem to work.

The alternative is to embrace mind-body dualism, the view that consciousness exists outside the physical world. This remains a genuine possibility, especially given how little we still understand about the brain. Yet dualism violates a principle most scientists and philosophers accept: Occam's razor, that is, the idea that a simpler and more unified picture would be preferable. Mind-body dualism divides reality into two fundamentally different kinds of things: subjective qualities on the one hand, and the objective, mathematical physical world on the other. Reintegrating consciousness is not about rejecting physics. It is about recognizing that its mathematical formalisms describe only one aspect of reality: what things do, not what they are like. This opens the possibility of a third way, one that preserves the successes of physics while restoring experience to the natural world. Panpsychism offers one such avenue.

Panpsychism: Making the World Experientially Alive Again

One way to understand panpsychism is as the continuation of a long historical widening of what we count as "conscious". At the start of the Scientific Revolution, consciousness was often treated as something uniquely human. Animals were reduced to mechanisms: living bodies without inner lives. That picture had a clear purpose. It made the world easier to describe in purely objective terms. If science wants to map reality as measurable and external, then experience itself becomes an awkward problem, so the simplest move is to exclude it.

But over time, consciousness gradually extended beyond humans to other mammals, birds, fish, and more recently even to

insects. In this context, panpsychism can be understood as the next step in the same trajectory: not merely that animals are conscious, but that consciousness may be a fundamental feature of reality itself. That is, consciousness is present in some minimal form even in the molecules, atoms, and subatomic particles that make up the physical world. It is not an afterthought of evolution, but something woven into the structure of nature from the start.

Panpsychism is the view that consciousness extends all the way down to the fundamental constituents of matter. In this picture, even basic particles or fields possess unimaginably rudimentary forms of experience, and the complex consciousness of human and animal minds is somehow built up from these more basic forms. The idea has a long intellectual history, spanning both Western and Eastern traditions. Early modern philosophers such as Gottfried Wilhelm Leibniz and Baruch Spinoza defended views that resemble panpsychism, and the theory enjoyed a kind of heyday in the nineteenth century. It later fell out of fashion, especially in the latter half of the twentieth century. Over the past decade or so, however, panpsychism has returned as a serious option in academic philosophy, and even, to some extent, in neuroscience.

For its defenders, panpsychism offers an appealing alternative to two extremes: the metaphysical extravagance of an immaterial soul on one side, and the hard reductionist idea that consciousness is "nothing but brain chemistry" on the other. In practice, the latter often risks explaining consciousness away rather than explaining it.A walk in the woods is a conscious experience, writes Thomas Lewton. Green in a hundred tones, rain glistening on the leaves, air thick with a fresh sensation. In moments like these, experience is central to our existence. And yet physics, for all its power, remains silent about this inner dimension. It can measure light, map wavelengths, and describe reflection with perfect precision, but it cannot touch the simple fact of what it

is like to stand deep among trees and feel the world enclosing us in. There seems to be a strange distance between the world as it is subjectively lived and the world as it is objectively described. Subjective experience glows at the centre of reality, while objective explanation circles it from the outside. And still, the mind arises from a brain made of matter. So it feels natural to assume that consciousness must, in some way, be explainable in physical terms. But the question remains stubbornly open: by what process does matter give rise to experience? And if consciousness cannot be captured by physics alone, how can it be placed within a complete picture of the universe? How does matter ever come to feel like anything at all?

The hypothesis that the universe is conscious is not as outlandish as it initially sounds. Physics gives mathematical structure, but structure alone does not explain why there is something rather than nothing. As Stephen Hawking once put it, what "breathes fire into the equations"? One possibility is that it is the mind that breathes fire into the equations. This proposal is no less parsimonious than its rivals, and it has the additional advantage of offering a framework for understanding fine-tuning. On this view, the source of cosmic order is not an omnipotent God, but an entity that pursues goals under genuine constraints, constraints recognised as the laws of physics.

Ultimately, what is needed is a hypothesis that can account for both the apparent goal-directedness implied by the fine-tuning of life-permitting physics and the arbitrariness and gratuitous suffering that pervade the world. Cosmopsychism may sound strange, but it fits the data. Physics will be confined to telling us what matter does, as Goff writes in Galileo's Error (2019), remaining silent on its intrinsic nature: "When it comes to the basic causal workings of the universe, scientists provide mathematical laws which describe with great accuracy how matter behaves, but they provide no explanation of why matter behaves in that way".

What makes panpsychism new and exciting?

What makes contemporary panpsychism distinctive is not only its claim that consciousness is fundamental, but the way this claim can be integrated with our best physical theories and accounts of evidence. Panpsychism holds that consciousness exists at the most fundamental level of reality. It is often presented, somewhat misleadingly, as the idea that the world is composed of tiny conscious particles. This picture, however, sits uneasily with contemporary physics. Modern theoretical physics no longer understands the universe as a collection of discrete, billiard-ball-like entities. Instead, it describes reality in terms of universe-wide fields, with what we call particles understood as local excitations within those fields.

Quantum entanglement reinforces this shift away from a world of fully independent parts. Entangled systems behave as though they are a single unified whole, even when separated by vast distances, and without any possible exchange of signals between them. Whatever the correct interpretation of entanglement ultimately is, it suggests that separateness may not be as fundamental as common sense assumes. Reality begins to look less like an aggregate of self-contained independent parts and more like a deeply interconnected structure, one in which the whole is, in some sense, more basic than the parts.

On its own, this physical picture does not establish panpsychism. However, when combined with the panpsychist claim that consciousness is fundamental, a striking possibility emerges. Reality may ultimately consist in a deeply unified field that is not only physical but also experiential, with individual conscious systems, including human minds, understood as local manifestations or excitations within that more fundamental consciousness. This yields a remarkably simple and unified picture of reality. If materialism struggles to explain how subjective experience could be derived from purely objective description, pan-

psychism suggests a reversal of that direction. The subjective cannot be derived from the objective. Rather, the objective world described by physics might be understood as emerging from something already experiential. At least in principle, we can begin to see how such an account might proceed.

How panpsychism can make the mystical and spiritual scientifically credible

One of the most direct ways into the problem of consciousness is through mystical and spiritual experience. A mystical experience is an experience in which one seems to have direct awareness of ultimate reality. By "ultimate reality", we mean the fundamental ground of existence. For a physicalist, this consists in particles or fields; for others, it may be God or something analogous. In mystical experience, this reality is encountered as wondrous, as possessing an extraordinary and unfathomable value, far exceeding anything found in ordinary experience. Mystics consistently report that what they encounter is better than anything the everyday world has to offer.

A further characteristic is a profound sense of unity with this ultimate reality. Different mystical traditions interpret this unity differently. In Advaita Vedanta, unity is understood as identity: the experiencer is literally identical with ultimate reality. In Eastern Orthodox Christianity, it is described more relationally, as an intimate but non-identical union. Despite these interpretive differences, the experience of deep unity is strikingly consistent across cultures. The most influential study of such experiences remains William James's The Varieties of Religious Experience (1902). James treats mysticism as a genuine psychological problem, but he ultimately raises a deeper epistemological question: can mystical experiences be trusted? Is it rational for someone who has such an experience to believe that it reveals something about the nature of reality? James does not offer a definitive

answer, but he presents a compelling argument. If we deny the mystic the right to trust their experience, we seem to introduce a double standard. Ordinary sensory experience is generally regarded as trustworthy. We believe there is a table before us because we can see it. While sensory experience can be tested, any such testing ultimately relies on further experience. At some point, trust in experience is simply taken to be rational.

Why, then, should mystical experience be treated differently in principle? One might object that mystical experience could be delusional, perhaps the product of unusual brain states. But the same sceptical possibilities apply to ordinary perception. We might be brains in vats or living in a simulation, yet we continue to regard trust in experience as rational. James's suggestion, then, is that it may be rational for a mystic to trust their experience as revelatory of ultimate reality. An important qualification is that whether such trust is rational depends on how well the experience coheres with one's broader worldview.

This brings us to a familiar framework in probabilistic reasoning: Bayesian reasoning. On a Bayesian approach, evaluating a hypothesis involves both the strength of the evidence and its prior probability, that is, how well it fits with what we already believe about the world. Testimony that someone ate a cheese sandwich for lunch is easily believed. Testimony that they arrived on a dragon is not. The evidence is the same, but the prior probabilities differ dramatically. Flying on dragons has an extremely low prior probability and would require extraordinarily strong evidence.

The same structure applies to mystical experience. Whether it is rational to trust such an experience depends in part on one's underlying metaphysical commitments. And this is where the issue becomes significant: mystical experience looks very different when approached from a physicalist worldview than when approached from a panpsychist one. Panpsychism also leaves room for understanding ultimate reality as wondrous in the relevant

sense. While not required by the theory, it is at least plausible that when surface-level mental contents quieten, attention shifts toward a deeper layer of consciousness, naturally experienced as profoundly valuable. Similarly, if individual consciousness is grounded in a more fundamental universal consciousness, this grounding could be disclosed phenomenologically, as an experience of unity with something deeper and more expansive than the everyday self. While such experiences are not inevitable, they fit naturally within a panpsychist framework.

Taken together, these considerations suggest that from a panpsychist perspective, it is rational to take mystical experiences seriously, not merely as psychological anomalies, but as potential insights into the fundamental nature of reality. This may also help explain why the cultural revolution toward mysticism in the 1960s ultimately failed to sustain itself. Lacking a rigorous metaphysical framework, such experiences were easily dismissed or reduced. Panpsychism may offer a scientifically respectable framework in which mystical experiences can be understood as revealing something genuine, rather than as curiosities to be explained away.

The core challenge of panpsychism

The central challenge facing panpsychism is the combination problem: how simple forms of consciousness at the fundamental level combine to produce the unified consciousness we experience. This problem is a major focus of contemporary research in philosophy of mind, and while no fully satisfactory solution has yet emerged, a range of promising proposals have been developed.

One version of the problem is this. If my consciousness is made of trillions of particles, why do I not experience each of them individually? One possible response is that not all experience is cognitively accessible. Some forms of experience may exist at

subpersonal or unconscious levels, contributing to conscious unity without appearing as distinct items in awareness. Even if panpsychism ultimately requires additional theoretical primitives to explain the combination, it still offers a strikingly simple and elegant picture of reality, one in which physics emerges from a deeper, subjective foundation rather than the other way around. This is a central and exciting problem in consciousness research today.

Panpsychism: A third way

So what does all this mean for how humans situate themselves in the world? These debates are not merely technical. They also shape how we think about the ultimate character of reality and our place within it. One such debate is the long-standing tension between theism and atheism. Panpsychism may offer a middle path between traditional theism, which posits a supernatural designer, and atheism, which typically regards the universe as fundamentally indifferent. Rather than excluding either position, this approach invites a more open stance. One that acknowledges the limits of our knowledge and resists premature metaphysical closure.

One of the standard arguments for theism appeals to the apparent fine-tuning of the laws of physics of life. Over recent decades, physicists have discovered that many fundamental constants must lie within an extraordinarily narrow range for stars, planets, and life to exist at all. Slight deviations in values such as dark energy would have resulted either in a universe that expanded too rapidly for structure to form, or one that collapsed almost immediately. Theists typically interpret this fine-turning as evidence of a supernatural designer. Atheists typically respond by appealing to a multiverse: given enough universes, one is bound to have the right conditions by chance. Panpsychism, however, opens a different possibility. If the universe is conscious in some

fundamental sense, it may not be unreasonable to suggest that it possesses tendencies or orientations of its own. Rather than being externally designed or purely accidental, the apparent order of the cosmos may reflect an ongoing, immanent process whose full character we do not yet understand.

The panpsychist view requires no appeal to the supernatural, but it does leave conceptual space for higher or more encompassing forms of consciousness. Unsurprisingly, some find this suggestion uncomfortable, perhaps because it challenges the assumption that consciousness must be confined to localised systems. Yet if consciousness is a fundamental feature of reality, it is not obvious why it should be confined to localised systems alone. These questions become especially intriguing when we move beyond human beings and begin to consider how unity and sharing might apply to particles, fields, and the fundamental structure of reality itself.

Where does this leave us? Seen in this light, the same considerations that place panpsychism between theism and atheism also explain its philosophical appeal. Panpsychism offers a way of taking consciousness and spirituality seriously while remaining fully compatible with scienceAt the same time, panpsychism invites an ontological integration of consciousness into our picture of reality. What emerges is no longer a dead, purely mechanical universe, but one that is fundamentally experiential: a reality in which consciousness is not an anomaly, but part of the fabric of the world. Panpsychism is, then, not a proof of spirituality, but it may be the first metaphysical view that takes subjective experience seriously while remaining compatible with our best physical theories.

Biographies

Slavoj Žižek
Biography

Born in 1949 in Ljubljana, Slovenia (then Yugoslavia), Slavoj Žižek has become one of the world's most recognizable and provocative public intellectuals. His journey began amidst the complexities of Yugoslav socialism, shaping his unique perspective. Initially trained in philosophy and later studying psychoanalysis in Paris, Žižek faced early academic marginalization before achieving global fame with his groundbreaking 1989 book, *The Sublime Object of Ideology*. His work masterfully, and often startlingly, synthesizes the dialectics of G.W.F. Hegel, the psychoanalytic theories of Jacques Lacan, and the critical framework of Karl Marx. This potent mix, delivered with disheveled charisma, relentless energy, and an astonishing output (spanning dozens of books and countless lectures), defines his unique presence.

Žižek is widely admired for several interconnected reasons. He tackles the most urgent issues of our time—capitalism's crises, pervasive ideology, resurgent authoritarianism, and ecological disaster—with fearless radicalism and intellectual verve. He possesses a rare gift for making dense philosophical and psychoanalytic concepts not just accessible, but thrilling and immediately relevant to contemporary anxieties, often using humor and deliberate provocation. Crucially, his critique extends beyond the usual targets; he relentlessly exposes the hidden contradictions and complicities within liberal democracy itself, resonating with those seeking deeper systemic analysis. His refusal of easy answers and commitment to public philosophy make him a captivating and indispensable voice.

His originality shines brightly in his analyses of popular culture, particularly film. Žižek doesn't merely interpret movies; he uses them

as vital "case studies" to expose the hidden machinery of ideology and unconscious desire operating within our everyday lives and enjoyment. For instance, he argues *Titanic* reveals the fantasy sustaining class divisions, *The Sound of Music* uncovers the libidinal underpinnings of fascism, and even *Kung Fu Panda* illustrates the function of symbolic authority (the "big Other"). His genius lies in revealing the profound within the seemingly profane, demonstrating how ideology functions most powerfully not through overt propaganda, but through the very structures of our perception and pleasure—a perspective he terms the "parallax view."

Major Books

Quantum History: A New Materialist Philosophy (2025):
Žižek's most ambitious synthesis yet, merging quantum mechanics with Lacanian psychoanalysis and Hegelian dialectics to argue that reality is fundamentally incomplete and retroactively shaped by events. Drawing on Hegel, Heidegger, Lacan, and physicist Carlo Rovelli, he proposes that the uncertainty of the quantum void drives political and social change, applying this framework to AI, climate change, and Trumpism. Darkly funny and uncompromising, it is a radical inquisition into whether emancipatory politics remains possible at all.

Against Progress (2024):
Thirteen short, sharp essays dismantling the dominant idea of Progress, capital P, as an inevitable and destructive force. Žižek targets neoliberalism, Trumpian populism, and accelerationism alike, urging the Left to reclaim and redefine progress rather than cede it. Characteristically unsparing, and immediately relevant to the current ecological and political moment.

Violence: Six Sideways Reflections (2008):
Explores the often-overlooked forms of systemic and symbolic violence underlying society in an engaging essay format. It's thought-provoking and tackles a universal theme with Žižek's unique slant.

Yanis Varoufakis
Biography

Yanis Varoufakis is a Greek economist, political theorist, and public intellectual best known for his fierce critique of contemporary capitalism and his role in the European debt crisis. He became internationally known in 2015 when he served as Greece's Minister of Finance during the country's confrontation with the European Union and international creditors. With his direct style, sharp intelligence, and refusal to accept the moral language of austerity, Varoufakis quickly became one of the most polarizing figures in modern European politics.

Varoufakis is widely read not only because of his political history, but because he offers a rare combination of insider knowledge and philosophical ambition. He does not treat economics as a neutral science of numbers, but as a system of power, ideology, and institutional design. For him, economic crises are not accidents. They are expressions of structural contradictions built into the architecture of global capitalism, especially in the relationship between finance, democracy, and sovereignty. His work insists that the economy is not a natural force we must obey, but a human-made system that can be redesigned.

A central theme in Varoufakis's writing is the tension between markets and democratic life. He argues that capitalism increasingly concentrates power in ways that hollow out political choice, leaving citizens with the illusion of control while real decisions are made by financial institutions, central banks, and technocratic bodies. In recent years, he has pushed this critique further by arguing that capitalism is mutating into something new, a system in which digital platforms,

algorithmic control, and private technological infrastructure produce new forms of dependence and domination. He describes this shift as the emergence of a post-capitalist order that is more extractive, less accountable, and strangely feudal in its logic.

What makes Varoufakis especially compelling is his ability to write about complex economic structures in a way that feels vivid and human. He moves easily between theory and storytelling, combining economic explanation with political urgency and moral clarity. Whether one agrees with him or not, his work has a distinctive force. It demands that we stop treating the economy as inevitable and start treating it as a contested space where freedom, inequality, and collective life are at stake.

Major Books

Technofeudalism: What Killed Capitalism (2023):
provocative argument that capitalism is no longer the best description of our system. Varoufakis claims that digital platforms and tech monopolies have created a new kind of economic order based on rent, extraction, and control rather than competitive markets. It is ambitious, unsettling, and very current.

Adults in the Room (2017):
Varoufakis's insider account of the 2015 Greek debt negotiations. It reads like a political thriller, but it is also a revealing portrait of how power works inside European institutions. It is gripping, angry, and intellectually sharp.

Talking to My Daughter About the Economy (2013):
A short and highly accessible introduction to economic ideas, framed as a conversation and written with clarity and warmth. It is an excellent entry point if you want a philosophical understanding of capitalism without technical jargon.

Susan Schneider
Biography

Susan Schneider is an American philosopher and cognitive scientist renowned for her work on consciousness, artificial intelligence, and the future of the mind. Born and raised in the United States, she developed an early fascination with the intersection of human cognition and emerging technology. She earned her Ph.D. in Philosophy from Rutgers University, laying the foundation for her interdisciplinary approach bridging philosophy, neuroscience, and AI ethics. Schneider is known for her accessible yet rigorous style, making complex ideas about intelligence and identity engaging for both academic and public audiences.

Schneider remains a vital voice in ensuring humanity navigates AI and cognitive enhancement ethically. Her work—spanning philosophy, NASA, and SETI—challenges us to redefine intelligence while safeguarding human dignity. As she often states: "We're not just building tools; we're reshaping what it means to be conscious."

Professional Journey

Today, Schneider holds two key roles:

— William F. Dietrich Distinguished Professor of Philosophy at Florida Atlantic University.

— Distinguished Visiting Scholar for NASA's Innovative Advanced Concepts (NIAC) program, where she examines the implications of AI for space exploration.

She also advises the SETI Institute's post-detection task force, analyzing how advanced extraterrestrial intelligence might manifest as AI.

AI & Consciousness

Schneider's core work questions whether machines can truly be conscious. She rejects "biological chauvinism" (the view that only biological brains can host consciousness) but argues current AI lacks subjective experience. Her influential framework proposes tests for AI consciousness and warns against unchecked human cognitive enhancement (e.g., neural implants). She emphasizes that merging minds with AI could redefine human identity—a theme central to her ethics advocacy.

Major Books

Artificial You: AI and the Future of Your Mind (2019):
Explores consciousness in AI, brain implants, and how technology could alter human identity.

Science Fiction and Philosophy (co-edited, 2016):
Uses sci-fi to examine ethical dilemmas in AI, mind uploading, and alien contact.

The Language of Thought: A New Philosophical Direction (2011):
Reimagines how the brain processes information, influencing AI cognition models.

Key articles

"The Dangers of Artificial Intelligence" (Scientific American, 2023):
Argues for urgent regulation of AI consciousness research to prevent ethical crises.

"Intelligences: Natural, Artificial, Cosmic" (SETI Press, co-authored, 2022):
Proposes strategies for identifying AI-driven intelligence in cosmic exploration.

Riccardo Manzotti
Biography

Riccardo Manzotti (born 1969) is an Italian philosopher, cognitive scientist, and professor of AI and Robotics at IULM University in Milan. He holds a PhD in robotics and previously worked in artificial intelligence, which shaped his interdisciplinary approach to consciousness. His research merges philosophy of mind, neuroscience, and technology, challenging traditional boundaries between the mind and the external world.

Key Philosophical Contribution

The Spread Mind Theory

Manzotti is best known for his "Spread Mind" theory, which argues that consciousness is not confined to the brain but is instead a process extending into the physical environment. This radical externalist view opposes both dualism (mind-body separation) and internalist neuroscience, proposing instead that objects and perception are co-constitutive. His ideas align with enactivism and process philosophy, drawing comparisons to thinkers like Alfred North Whitehead.

Collaboration with Tim Parks

In 2017, Manzotti co-authored *Dialogues on Consciousness* with British writer Tim Parks, a series of accessible conversations exploring consciousness, free will, and the nature of reality. The book presents Manzotti's ideas in a debate format, making his complex theories engaging for a broader audience.

Reception in the US

While Manzotti's work is more widely discussed in Europe, his ideas have gained traction in the US among philosophers and cognitive scientists interested in embodied cognition and non-representational theories of mind. Critics question his rejection of internalism, but proponents praise his innovative approach to bridging phenomenology and science. His talks at institutions like NYU and engagement with thinkers like David Chalmers have helped his influence grow in American academia.

Major Books

Dialogues on Consciousness (with Tim Parks, 2022):
Fifteen conversations between philosopher Riccardo Manzotti and novelist Tim Parks, originally published in the New York Review of Books, that methodically dismantle the assumption that consciousness lives inside the brain.

Ich Denke, Aber Wer Ist Ich? (Hashagen & Manzotti, 2021):
A German-language collaboration between philosopher Riccardo Manzotti and author-banker Anne Hashagen that asks the age-old question, who or what is the "I" that thinks?

The Spread Mind: Why Consciousness and the World Are One (2017):
Explores the idea that consciousness is not happening inside your brain, it is the physical world itself putting forth a radical case that your mind and your surroundings are not two things, but one.

Finn Brunton
Biography

Finn Brunton is an American scholar, writer, and professor whose work focuses on digital culture, privacy, obsolete technologies, and the hidden infrastructures of the internet. He earned his Ph.D. from the University of California, Berkeley, and has taught at institutions such as New York University and the University of California, Davis. Currently, he is an Associate Professor of Science and Technology Studies (STS) at the University of California, Davis, where he continues to research and write about the intersections of media history, cryptography, and digital resistance.

Brunton's work is deeply interdisciplinary, blending media theory, history, and critical technology studies. He is particularly interested in how people adapt, misuse, or subvert digital tools—whether through spam, cryptocurrency, or obfuscation tactics. His writing is both scholarly and accessible, making complex technological histories engaging for a broad audience.

Major Books

Digital Cash: The Unknown History of the Anarchists, Utopians, and Technologists Who Created Cryptocurrency (2019):
A prehistory of Bitcoin that resurrects the forgotten dreamers and failed experiments behind digital money, tracing how cypherpunks, libertarians, and visionary technologists laid the ideological and technical groundwork for cryptocurrency long before it had a name.

Spam: A Shadow History of the Internet (with Helen Nissenbaum, 2013):

Explores how spam shaped the architecture of internet governance, economics, and digital security, tracing how unwanted messages quietly but fundamentally changed the way we build, regulate, and navigate the web.

Thomas Hertog
Biography

Thomas Hertog is a Belgian cosmologist and theoretical physicist renowned for his groundbreaking work with Stephen Hawking on quantum cosmology and the origins of the universe. Born on May 30, 1975, in Leuven, Belgium, he studied at KU Leuven before earning his Ph.D. at the University of Cambridge under Hawking's supervision. Hertog's research explores deep questions about the Big Bang, the multiverse, and the nature of time, with a particular focus on the holographic principle—a radical idea suggesting that our 3D universe might be encoded on a 2D boundary, much like a hologram.

One of Hertog's most significant contributions, developed alongside Hawking in their final collaboration, is the theory of a holographic origin of time. This model proposes that time, as we experience it, emerged from a timeless quantum state before the Big Bang. Unlike classical cosmology, which assumes time began at the singularity, their framework suggests that the early universe was a kind of quantum hologram, where time and space were not yet fully defined. This idea challenges the traditional view of cosmic inflation and offers a new way to reconcile quantum mechanics with Einstein's theory of gravity.

Hertog co-authored *The Grand Design* (2010) with Hawking, which popularized cutting-edge theories like M-theory and the idea that the universe could arise from "nothing" without divine intervention. His latest book, *On the Origin of Time: Stephen Hawking's Final Theory* (2023), delves deeper into their holographic cosmology, presenting a bold vision where the laws of physics themselves evolve. Today, Hertog is a professor at KU Leuven, where he continues

to refine these revolutionary ideas, pushing the boundaries of our understanding of reality while honoring Hawking's legacy.

Major Books

On the Origin of Time: Stephen Hawking's Final Theory (2023):
Explores the revolutionary "top-down cosmology" developed during Hertog's twenty-year collaboration with Stephen Hawking, proposing that the laws of physics are not fixed but instead evolved alongside the universe in a Darwinian-like process

Big Bang: Imagining the Universe (with Barbara Baert and Jan van der Stock, 2021):
An interdisciplinary anthology that bridges science and art to examine how humans have visualized and conceptualized the origins of the cosmos, tracing the cultural and scientific impact of the Big Bang theory.

Charles Foster
Biography

Charles Foster is a British writer, veterinarian, barrister, and explorer with a mind as wild as the landscapes he studies. Part scientist, part philosopher, and full-time provocateur, Foster doesn't just write about nature—he lives it.

Trained in veterinary medicine at the University of Cambridge, Foster also holds a law degree and a PhD in medical ethics from Oxford. But his real education came from crawling through forests, swimming with otters, and even burying himself in dirt to understand the world from an animal's perspective. His work blends hard science with lyrical prose, challenging how we see ourselves in nature.

A fellow at Oxford's Green Templeton College, Foster teaches medical law and ethics—but his true passion is experiential biology. He's slept in a badger sett, foraged like a fox, and even tried (disastrously) to live as a urban red deer. His radical experiments in "being beast" have made him a cult figure in nature writing.

Beyond academia, he's a TED speaker, a contributor to The Guardian and BBC Wildlife, and a relentless critic of humanity's arrogance toward other species. Whether dissecting the soul of a swallow or the legal rights of rivers, Foster's work is always bold, witty, and deeply humane.

Major Books

Edges of the world (Transworld, 2026):
Travels in the vmargins of life, lands and ideas.

Cry of the Wild (2023):
A meditation on animal consciousness and human ignorance.

The Screaming Sky (2021):
A poetic ode to swifts, blending biology and mythology.

Being a Human (2021):
A sequel exploring what it means to be human by living as our Stone Age ancestors did.

Being a Beast (2016):
The cult classic where Foster tries living as a badger, otter, fox, deer, and swift. A mix of science, memoir, and madness.

Roman V. Yampolskiy
Biography

Born in the former Soviet Union (now Ukraine), Roman V. Yampolskiy immigrated to the United States, where he earned his PhD in Computer Science from the University at Buffalo. Today, he serves as a tenured Associate Professor in the Department of Computer Engineering and Computer Science at the University of Louisville in Kentucky, USA. At Louisville, he founded and directs the Cyber Security Lab, an interdisciplinary research hub with a critical mission: to address existential threats posed by advanced artificial intelligence. The lab's primary aim is to pioneer solutions for AI safety, security, and alignment—ensuring future superintelligent systems remain controllable, ethical, and beneficial to humanity.

Yampolskiy's research centers on what he terms the "Uncontrollability Thesis": the argument that superintelligent AI (ASI) may be fundamentally impossible to reliably control or contain using any known methods. This work explores catastrophic failure modes like goal misalignment, unintended behaviors in complex systems, and malicious use cases. His warnings extend beyond conventional cyber threats to existential risk, positioning AI as a potential species-level threat comparable to nuclear war or engineered pandemics.

He gained wider recognition for provocative ideas like those in his 2025 paper, "Hacking Our Way Out of the Universe," where he theorized that a superintelligence might manipulate physics itself (e.g., creating wormholes or exploiting quantum realms) to escape cosmic extinction. This frames AI not merely as a tool, but as humanity's potential last hope—or ultimate destroyer.

Yampolskiy's work forces a critical question: Can we survive what we create?

Major Books

Considerations on the AI Endgame: Ethics, Risks and Computational Frameworks (with Soenke Ziesche, 2025):

An interdisciplinary examination of AI's societal and ethical implications, covering everything from AI welfare and value alignment to questions of identity and consciousness, drawing on both Western and non-Western philosophical traditions to ask what kind of future humanity is building toward.

AI: Unexplainable, Unpredictable, Uncontrollable (2024):

Makes the sobering case that AI is fundamentally resistant to human understanding and control, moving from the core technical problems of unpredictability and opacity to deeper questions of AI personhood, consciousness, and what it would mean to lose dominion over the systems we created.

Artificial Superintelligence: A Futuristic Approach (2015):

A foundational text for the science of AI safety engineering and ethics, examining how to ensure that emerging superintelligent systems remain beneficial to humanity.

Bernardo Kastrup
Biography

Bernardo Kastrup is a Dutch philosopher, computer scientist, and author known for his work on consciousness, metaphysics, and the philosophy of mind. He is a leading proponent of metaphysical idealism, the view that reality is fundamentally mental rather than material. Born in the Netherlands, Kastrup pursued an academic and corporate career in computer engineering before shifting his focus to philosophy.

He holds a PhD in Computer Engineering from the Eindhoven University of Technology, where he specialized in artificial intelligence and reconfigurable computing. His technical work included research roles at CERN and Philips Research, where he contributed to AI and large-scale computing systems. Later, he earned a second PhD in Philosophy from Radboud University Nijmegen, deepening his exploration of consciousness and ontology.

Dissatisfied with the materialist worldview dominant in science, Kastrup founded the Essentia Foundation, a nonprofit organization dedicated to advancing idealism as a viable alternative to materialism. Through Essentia, he publishes essays, hosts discussions, and collaborates with scientists and philosophers to explore consciousness as the foundation of reality.

Kastrup is a prolific writer and speaker, engaging in debates, lectures, and media appearances worldwide. He has spoken at Oxford University, The Scientific and Medical Network, and The Weekend University (London), among other venues. His work critically examines artificial intelligence, neuroscience, and the limits of materialism, arguing that subjective experience—not computation—underlies true understanding.

In addition to his academic and philosophical contributions, Kastrup is known for his clear, analytical style, making complex ideas accessible to a broad audience. He continues to write, debate, and advocate for a paradigm shift in how science and philosophy understand reality.

Major Books

Science Ideated: The Fall of Matter and the Contours of the Next Mainstream Scientific Worldview (2021):

A collection of influential essays arguing that cutting-edge findings in quantum mechanics and neuroscience are increasingly incompatible with materialism, and that analytic idealism is the most empirically coherent framework to replace it.

The Idea of the World: A Multi-Disciplinary Argument for the Mental Nature of Reality (2019):

Drawing on peer-reviewed work in physics, neuroscience, and philosophy, Kastrup builds a comprehensive academic case for idealism, the view that reality is essentially mental, systematically refuting physicalist alternatives and showing how a consciousness-first worldview better reconciles classical and quantum science.

Why Materialism Is Baloney: How True Skeptics Know There Is No Death and Fathom Answers to Life, the Universe, and Everything (2014):

Dismantles the materialist worldview, making the case that consciousness is not produced by the brain but is instead the fundamental fabric of reality, and that matter, far from being primary, exists within mind rather than the other way around.

Rupert Sheldrake
Biography

British biologist Rupert Sheldrake (born 1942), with a Cambridge biochemistry PhD and Harvard research background, emerged as a provocative challenger to scientific orthodoxy. Moving beyond conventional plant development studies, he proposed the radical theory of morphic resonance in the 1980s. This concept posits that self-organizing systems—from cells and crystals to flocks of birds and human minds—inherit a collective memory through non-material, pattern-shaping "morphic fields." Sheldrake argues that nature operates not by fixed, eternal laws, but by evolving habits. The more often a specific form or behavior occurs, the stronger its associated morphic field becomes, making it easier for similar systems anywhere to adopt it through resonance. This "formative causation" aims to explain phenomena like instinct, rapid species-wide learning (e.g., rats mastering a maze faster globally after initial success elsewhere), telepathy in bonded animals, and even phantom limbs.

Sheldrake's ideas drew fierce criticism, epitomized by his clash with arch-materialist Richard Dawkins. A notorious 2013 interview saw Dawkins dismiss morphic resonance as "nonsense" while admitting he hadn't read Sheldrake's key work, leading Sheldrake to accuse Dawkins of unscientific dogmatism. Sheldrake framed this conflict in his book *The Science Delusion* (2012), a direct counter to Dawkins' *The God Delusion*, arguing rigid materialism stifles inquiry. Despite decades as a fringe figure, Sheldrake's work is experiencing a big reappraisal. While not claiming morphic fields are quantum fields, leading physicists like Nobel laureate Brian Josephson and Basil Hiley (collaborator of David Bohm) note compelling parallels. Concepts

like non-locality (instantaneous connection across space), observer effects, and holistic field theories in quantum physics resonate with Sheldrake's vision of an interconnected, memory-laden universe where information and organization are fundamental. His foundational works remain *A New Science of Life* (1981), *The Presence of the Past* (1988), and *The Science Delusion* (2012).

Major Books

The Science Delusion (2012):

A critical examination of the ten core dogmas of modern science, challenging the materialist worldview and advocating for a more open, inquisitive approach to biological and physical phenomena.

Presence of the Past (1988):

An expansion on his earlier theories, exploring how memory, habit, and heredity function through morphic resonance, suggesting that the "laws" of nature may actually be more like evolving habits.

A New Science of Life (1981):

The landmark text that first introduced the hypothesis of morphic resonance, proposing that nature has an inherent memory and that "morphic fields" shape the form and behavior of all living systems.

Lucy Cooke
Biography

Lucy Cooke is a British biologist, bestselling author, and award-winning documentary filmmaker with a mission: to reveal the fascinating (and often hilarious) truths about the animal kingdom. With a sharp wit and a deep love for misunderstood creatures, she has become one of the most entertaining voices in popular science today.

Cooke studied biology at the University of Oxford under famed evolutionary biologist Richard Dawkins before earning a master's in wildlife filmmaking. Instead of sticking to dry academic writing, she chose to bring science to life through storytelling, humor, and stunning visuals.

Her work has taken her around the world, from tracking sloths in Costa Rica to studying hyenas in Africa. She's the founder of the Sloth Appreciation Society, a TED speaker, and a regular contributor to media outlets like National Geographic and the BBC. Whether she's debunking myths about "lazy" sloths or revealing the surprising intelligence of vultures, Cooke's infectious enthusiasm makes even the weirdest animals lovable.

Beyond her writing, she has hosted and produced TV documentaries, including Meet the Sloths and Animals Unexpected, proving that wildlife programming can be both educational and wildly entertaining.

Major Books

Bitch: A Revolutionary Guide to Sex, Evolution, and the Female Animal (2022):

Her latest book, challenging stereotypes about female creatures in nature.

The Truth About Animals: Stoned Sloths, Lovelorn Hippos, and Other Tales from the Wild Side of Wildlife (2018, Penguin):

A bestselling deep dive into animal myths and marvels.

A Little Book of Sloth (2013):

A charming photo book celebrating her favorite slow-moving creatures.

Philip Goff
Biography

Philip Goff is a contemporary philosopher best known for bringing panpsychism back into serious academic conversation. Trained within the analytic tradition, he has become one of the most influential public voices arguing that consciousness cannot be treated as a late-stage accident of evolution, nor as something that can be fully explained away by neuroscience alone. His work sits at the intersection of philosophy of mind, metaphysics, and the foundations of science, and it is driven by a single persistent question: how can subjective experience exist in a world described by physics?

Goff's distinctive contribution is his ability to make one of philosophy's most notoriously difficult problems, the hard problem of consciousness, feel urgent, intelligible, and unavoidable. He argues that physical science, powerful as it is, describes the world in terms of structure and behavior, but leaves out what reality is like from the inside. For Goff, it is a deep explanatory failure that forces us to reconsider our basic metaphysical assumptions. Instead of choosing between dualism, which risks making consciousness supernatural, and materialism, which risks eliminating experience altogether, he defends panpsychism as a radical but parsimonious alternative. Consciousness is fundamental, present in rudimentary form at the most basic level of matter, and complex minds arise through the organization of these basic experiential building blocks.

What makes Goff especially compelling is his insistence that this is not mystical speculation, but a serious philosophical proposal motivated by scientific realism. He takes physics seriously, but argues that physics alone cannot tell us what the universe is made of in its intrinsic

nature. This opens space for a view in which consciousness is not an extra ingredient added to matter, but part of what matter is. In some of his later work, Goff explores cosmopsychism, the idea that consciousness may be fundamental at the level of the universe as a whole, attempting to account not only for the existence of experience, but also for the apparent unity and structure of reality.

Goff is widely admired for making metaphysics feel alive again. He combines clarity, intellectual courage, and accessibility, writing with a rare confidence that the biggest questions are still worth asking. Whether one ultimately agrees with his conclusions or not, his work forces a confrontation with the fact that consciousness is not a side issue in our understanding of nature, but one of its central mysteries.

Major Books

Why? The Purpose of the Universe (2023):

A bold and controversial book in which Goff explores whether the universe might have a kind of purpose or goal-directedness. He engages with questions of fine-tuning, meaning, and value. It is more speculative than his earlier work, but fascinating if you are drawn to cosmic-scale metaphysics.

Galileo's Error (2019):

Goff's most famous and accessible work. It argues that modern science, beginning with Galileo, deliberately excluded consciousness from its picture of nature in order to make physics possible, and that this exclusion has left us with an incomplete worldview. The book makes a clear and provocative case for panpsychism as a serious alternative.

Consciousness and Fundamental Reality (2017):

More academic, but still readable if you want the rigorous philosophical version of his argument. It lays out panpsychism in detail, addresses objections, and clarifies why he thinks it offers the best route beyond materialism and dualism.

Copyright

Copyright © 2026 by G10 van de economie and the following authors for their respective contributions:

"Introduction" by Boudewijn Richel (pp. 6–21)
"Truth" by Slavoj Žižek (pp. 22–35)
"Cloud Capital" by Yanis Varoufakis (pp. 36–51)
"Mind Exchange" by Susan Schneider (pp. 52–65)
"Can Machines think?" by Riccardo Manzotti (pp. 66–79)
"Cryptos" by Finn Brunton (pp. 80–95)
"Time" by Thomas Hertog (pp. 96–111)
"Edges" by Charles Foster (pp. 112–127)
"Hacking the Code of the World" by Roman Yampolskiy (pp. 128–139)
"Time Travel" by Alien (pp. 140–147)
"AI won't be Conscious" by Bernardo Kastrup (pp. 148–163)
"Morphic Resonance" by Rupert Sheldrake (pp. 164–177)
"Resurrecting Extinct Animals" by Lucy Cooke (pp. 178–191)
"Galileo's Error" by Philip Goff (pp. 193–207)

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying or other electronic methods, without the prior written permission of the copyright owners.

Translation and Subsidiary Rights: All translation and subsidiary rights for this work are managed by stichting G10 van de economie. For inquiries, please contact: info@g10vandeeconomie.nl

Colofon