Beyond the Babble

Why the Real AI Transition Runs Through Power Plants, Not Philosophy

by Dr. Gregory S. Carmichael • CryptoSoWhat.com • April 2026

The Singularity as we are taught to see it: intelligence rendered as pure, luminous abstraction — a museum exhibit observed from a respectful distance. The argument of this piece is that the more instructive view looks quite different.

A FRIEND sent me a thoughtful summary of "the Singularity" last week and asked what I thought. He had spent a Saturday working through it — reading, watching documentaries, sketching out the canonical argument as he understood it. He wrote it up and circulated it to friends and colleagues, which is what thoughtful people do when they encounter a big idea.

His piece was better than most. It hit the three classical pillars — recursive self-improvement, artificial superintelligence, the merging of biology and silicon — and acknowledged the "complexity brake" counter-argument, which most popular treatments skip entirely. It named the alignment problem. It gave honest weight to both the utopian and the existential frames.

And yet, reading it, I had the same reaction I have to most writing in this genre: the argument is structurally incomplete in a way that matters for anyone actually trying to plan, invest, or build during the next ten years.

This is my longer answer to his question. It is not a refutation of the Singularity idea — there is a legitimate intellectual lineage there worth respecting, from I. J. Good in 1965 through Vernor Vinge in 1993 to Ray Kurzweil's 2005 synthesis. It is an argument that the dominant framing has become a poor planning document, and that the practical version of what we are actually living through looks quite different from the philosopher-futurist version.

The Inherited Frame

The Singularity argument, as it has crystallized in popular form, runs roughly like this. Once we build a machine capable of redesigning itself, it will trigger a recursive feedback loop. Each improved version designs a better successor. The curve bends vertical. Within a short window — some theorists say days, some say years — machine intelligence surpasses all of humanity combined. Combined with brain-computer interfaces and nanotechnology, this transition either delivers paradise (disease cured, scarcity ended, consciousness uploaded) or extinction (a superintelligence with goals indifferent or hostile to human survival). Kurzweil's 2045 date has become a kind of cultural anchor. The event horizon metaphor — that we cannot see past the moment of transition — has done a lot of rhetorical work.

I want to stipulate up front: the people who developed this argument are serious. Good, Vinge, Kurzweil, Bostrom, Yudkowsky — these are not dilettantes. Their work has shaped a generation of thinking about technology's trajectory, and much of it has held up remarkably well. When Vinge predicted in 1993 that we would see superhuman performance in narrow domains within thirty years, that was a bold claim. It has arrived essentially on schedule.

But arguments age. And the Singularity argument in its 2005 form was built for a world that no longer exists. The most serious thinkers in this space have already moved on. The popular version has not.

Here is what I think it gets wrong.

The Frame Is Dated

The conversation shifted materially in late 2022. Before then, "will recursive self-improvement happen?" was a reasonable question to debate. After the emergence of large language models, the question changed. It became: "are we already in a soft takeoff, and if so, what does governance look like *during* the transition rather than at some theoretical event horizon?"

That is a different question with different answers. LLMs did not deliver AGI — they are not AGI, and the serious researchers pushing the frontier are candid about that. What they did was make capability growth *legible*. Kurzweil's extrapolations assumed smooth exponential curves. What we actually got was step-function jumps followed by plateaus, punctuated by emergent capabilities nobody predicted from the underlying architecture. GPT-3 to GPT-4 to the current generation was not a smooth curve. It was a series of discrete shocks, each followed by a period of absorption.

This matters because the shape of the change determines the shape of the response. Smooth exponential curves let you plan ahead. Step-function shocks do not. You cannot extrapolate from the last plateau to know when the next jump comes, or how big it will be. That is a different policy environment than the one Kurzweil was writing for.

Intelligence Does Not Scale in a Vacuum

This is where I think the standard framing misses the most, and it is the part that anyone thinking about AI through a capital or operations lens needs to internalize.

The Singularity narrative treats intelligence as software. A mind is an algorithm, and algorithms compound cheaply. If you have the right architecture, you just need enough compute, and compute grows exponentially per Moore's Law, so intelligence grows exponentially too, and eventually the curve bends vertical.

This is wrong at the level of physics.

Intelligence in its current form runs on GPUs. GPUs run on power. Power runs on grid buildout. Grid buildout runs on transformers (the electrical kind), substation capacity, interconnection queues, high-voltage cables, permitting, water rights, rare earths, and grain-oriented electrical steel. None of these things scale on Moore's Law curves. They scale on the much slower curves of steel, concrete, permitting, and federal procurement. And as of this writing, every single one of them is constrained in ways that actively limit how fast AI capability can grow.

> *Intelligence in its current form runs on GPUs. GPUs run on power. Power runs on grid buildout. None of these things scale on Moore's Law curves.*

Consider what is actually happening right now.

The U.S. grid interconnection queue sits at roughly **2.6 terawatts** of proposed generation and storage — more than twice the total installed capacity of the existing American power fleet. Median wait times from interconnection request to commercial operation have grown from under two years in 2008 to nearly five years today, with some data-center-heavy regions seeing projected wait times of ten to twelve years. Historically, only about 19% of projects that enter the queue actually reach commercial operation. The rest withdraw — most because the upgrade costs and study delays

What the AI transition actually looks like at ground level: transmission towers marching toward a hazy metropolitan skyline, rooted in dust and copper rather than code and philosophy.

make the economics impossible.

Transformers — the unglamorous metal boxes that step voltage up and down between transmission and distribution — are in acute shortage. Lead times for large power transformers, which ran 24 to 30 weeks before 2020, now average **128 weeks** (about 2.5 years), with generator step-up units at 144 weeks. In some extreme cases, orders are taking five years to fulfill. Transformer prices are up roughly 77% since 2020. The United States manufactures only about 20% of the large power transformers it consumes domestically; the rest are imported, and the single largest source has become China, with U.S. imports of Chinese high-voltage transformers growing from under 1,500 units in 2022 to over 8,000 units in the first ten months of 2025 alone. China controls roughly 60% of global power transformer production capacity.

In Texas, the ERCOT large-load interconnection queue grew from 63 GW at the end of 2024 to 226 GW by November 2025 — nearly quadrupling in a single year, driven almost entirely by data center requests. About 77% of that load is data centers aiming to connect by 2030. ERCOT added 23 GW of new generation capacity between 2024 and 2025, with another 9 GW slated for early 2026. The arithmetic does not work. Demand is growing roughly ten times faster than the grid can serve it, and the state is now seriously debating whether to require large data centers to bring their own generation as a condition of connection.

The hyperscalers — Alphabet, Amazon, Meta, Microsoft — have collectively guided to **over $650 billion in AI infrastructure capital expenditure in 2026 alone.** Roughly 12 gigawatts of new U.S. data center capacity is planned to come online this year. Only about a third of it is actually under construction. The rest is waiting for electrical equipment that was never ordered in time. An AI data center deployment cycle runs 12 to 18 months. A transformer procurement cycle now runs 2.5 to 5 years. The two curves do not intersect. Bloomberg reported earlier this month that roughly half of planned 2026 U.S. data center builds will be delayed or cancelled, primarily because of physical equipment shortages.

This is what the Singularity actually looks like at ground level in 2026. It is not a mind bootstrapping itself to omniscience. It is a capital-intensive, supply-chain-constrained, regulation-mediated industrial buildout that is straining the physical capacity of the American power system and the global manufacturing base for high-voltage equipment. The chokepoints are not algorithmic. They are made of copper, steel, and signed permits.

If you want to know when and how the AI transition arrives in your life, you will learn more from watching PJM capacity auction prices, DOE Title 17 loan guarantees, and transformer order books than from reading arXiv papers.

"Intelligence Explosion" Is Asserted, Not Interrogated

The classical Singularity argument hinges on recursive self-improvement — the idea that once an AI reaches a certain capability threshold, it can redesign itself into a more capable version, which then redesigns itself again, and so on, with the cycle accelerating until the curve bends vertical.

This is a theoretical claim about what is architecturally possible, not an empirical observation about what current systems do.

There is a growing camp of serious AI researchers — François Chollet, Yann LeCun, Subbarao Kambhampati, and others — who argue that current transformer-based systems self-distill and self-refine *within* a capability envelope, but do not recursively expand it in the way the classical argument requires. Kambhampati's work is particularly sharp on this. He makes a compelling case that what looks like reasoning in modern LLMs is often sophisticated retrieval and interpolation — extraordinary in scope, but categorically different from the kind of open-ended problem-solving that would be required for genuine recursive self-improvement.

This is a live technical debate, not a settled question. I am not qualified to adjudicate it, and neither, probably, is anyone writing about the Singularity for a general audience. But it matters enormously for planning, because if recursive self-improvement is architecturally gated rather than imminent, the governance problem shifts fundamentally.

The question "how do we align a god-like intelligence that will emerge in the 2030s" is a theological question. It is important, but it is not actionable at the level of policy, investment, or engineering. The question "how do we manage an extended period of rapid but bounded capability growth, distributed across a small number of labs and nation-states, over the next 10 to 20 years" is a practical question. It has concrete answers involving export controls, compute governance, antitrust, industrial policy, and institutional design. We are already bad at it, and the cost of getting it wrong is enormous, but at least it is a problem that can be worked.

The Singularity frame, by collapsing everything into a single future event, makes the practical question harder to see.

The Real Risk Landscape Is the Messy Middle

The Singularity discourse tends to present two scenarios: utopia (disease cured, poverty ended, consciousness transcended) or extinction (a misaligned superintelligence eliminating humanity as a side-effect of optimizing something else). This binary has done enormous damage to clear thinking about what is actually happening.

The more textured risk landscape, which has largely displaced the Bostrom-era framing in serious AI governance circles, looks like this:

Concentration of capability. A small number of labs and a smaller number of governments are building systems that, even in their current bounded form, confer meaningful advantages in scientific research, cybersecurity, content generation, and strategic analysis. If those advantages compound, we get unprecedented asymmetries of power between those who have frontier models and those who do not. This is not a theoretical risk. It is the current trajectory.

Epistemic collapse. Synthetic media at scale, deployed into a political environment that cannot process it, is already degrading the shared factual basis of democratic discourse. This is happening now, at the current level of capability, and will intensify as generative systems improve.

Economic displacement outpacing institutional adaptation. Labor markets, tax systems, credentialing systems, and social safety nets were not designed for the pace of change that capable AI enables in knowledge work. The problem is not that humans become economically useless overnight;

it is that the transition happens faster than institutions can adapt, creating political pressures that exceed what existing political frameworks can channel constructively.

Misuse by humans rather than misalignment by the AI. The dominant threat model in serious AI governance work is no longer "the AI decides to harm us." It is "someone uses a capable system to do something terrible" — biothreats, cyber attacks at unprecedented scale, mass manipulation, or sovereign-scale economic coercion. The alignment problem in this frame is about aligning the humans deploying the systems, not the systems themselves.

Gradual disempowerment. Humans remain nominally in control of systems they no longer fully understand, and slowly lose the ability to meaningfully audit, reverse, or shape their outputs. This is the Paul Christiano / Dan Hendrycks framing, and it is considerably more worrying to serious researchers than the Skynet scenario because it does not require any discrete moment of takeoff. It just requires the slow atrophy of distributed human capability.

None of these require AGI. All of them are happening, at small scale, right now. All of them scale with capability. And none of them are well-addressed by the utopia-vs-extinction binary that dominates popular discourse.

The existential framing, paradoxically, can be a distraction from the risks that are actually actionable.

What To Do About It

Here is where I part company most clearly with the philosopher-futurist tradition, and it is the thesis I spent the last year working through in book form. *The Invisible Hand Meets AI* came out earlier this month. The core argument is that Adam Smith's framework does not break under AI — it gets *sharper*.

Markets, property rights, and distributed decision-making become *more* important, not less, when cognitive capability concentrates. The failure mode is not "the AI takes over." It is "a small number of humans and institutions capture the productivity gains, and the rest of the system atrophies — along with the distributed resilience that makes a society hard to coerce, hard to manipulate, and hard to capture."

> *Trust, atoms, institutions. The three things AI cannot manufacture for itself.*

The defense against that is the same defense Smith was arguing for in 1776: distributed ownership, credible institutions, rules that apply to everyone, property rights that individuals can actually enforce, and monetary architecture that does not require trust in any single actor. These are not

The chokepoint that the philosopher-futurist class does not see: a large power transformer, photographed with the reverence usually reserved for sculpture. The machines that decide whether compute arrives on time.

nostalgic virtues. They are active load-bearing elements of any system that has to hold weight during a technological discontinuity.

This thesis runs through what I am building operationally, and I mention that only because it is the honest context for the argument.

At Quantum Reserve Capital, it means building financial infrastructure that complies with the GENIUS Act's stablecoin framework — reserves held in short-term Treasuries and insured deposits, monthly disclosures, the kind of boring transparency that lets institutions commit across decades.

The bet is that trustworthy monetary rails will be scarce during a transition where traditional financial plumbing is simultaneously being automated and challenged by on-chain alternatives.

At Advanced Nano-Materials Manufacturing, it means advanced materials and physical upcycling — nano-engineered coatings, waste-to-energy conversion, graphene composites. Atoms. Things that cannot be abstracted into software no matter how capable the software becomes. The bet is that during a cognitive-capability boom, the relative scarcity of *physical* things — competent manufacturing, strategic materials, energy infrastructure, real-world logistics — increases rather than decreases.

Trust, atoms, institutions. The three things AI cannot manufacture for itself.

The 100-Year Test

I test decisions against a frame I find unusually clarifying: will someone 100 or 1,000 years from now look at this and say it was the right thing to do?

Applied to the Singularity discussion, this frame is particularly useful because most utopian and existential scenarios collapse the time dimension. They imagine a fast transition — years, not decades — to either paradise or extinction. The more likely reality is a messy multi-decade transition in which the institutional, legal, monetary, and physical-infrastructure decisions made in the next five to ten years determine whether the outcome is broadly distributed or narrowly captured.

The decisions that will look right in 100 years are probably not the decisions that optimize for any particular prediction about 2045. They are the decisions that build durable institutions, distribute ownership and capability widely, and preserve the capacity of human communities to govern themselves — regardless of what the technology ends up looking like.

That frame generates very different priorities than "prepare for the Singularity."

So What

The Singularity as popularized is a useful myth but a poor planning document. It compresses timelines, abstracts away physical constraints, frames risk as a binary between paradise and extinction, and locates the decisive moment at some future event horizon that makes present action feel futile.

The practical version is already here, and it arrives through channels the philosophical discourse barely notices. It arrives through power purchase agreements. Through transformer order books. Through fabrication capacity at TSMC and Samsung. Through the GENIUS Act and its stablecoin custody requirements — which are genuinely consequential and radically underappreciated. Through export controls on advanced GPUs. Through CHIPS Act allocations and DOE Title 17 loan guarantees. Through interconnection queue reform at FERC. Through antitrust enforcement on frontier labs. Through the slow, grinding, unglamorous work of building institutions that can hold weight during a discontinuity.

The people who will shape the actual outcome are not the ones predicting 2045. They are the ones building the grid, the fabs, the custody rails, the regulatory frameworks, and the distributed ownership structures in 2026. They are utilities, manufacturers, legislators, regulators, engineers, soldiers, financiers, and founders who ship things. They do not talk in event-horizon metaphors. They talk in permits, megawatts, basis points, and delivery dates.

The Singularity discourse has been dominated by philosophers and futurists for thirty years. It needs more engineers, more operators, more builders — and probably more soldiers and more accountants — for the next ten.

That is the conversation I think is worth having, and it is one I am going to keep pushing at in this newsletter. The philosopher-futurist class has had its turn. The infrastructure class is what will actually decide how this goes.

Dr. Gregory S. Carmichael is CEO of Quantum Reserve Capital and founder of Advanced Nano-Materials Manufacturing. A retired USAF Lieutenant Colonel with 26 years of service, he holds a DBA in International Finance, an MA in International Relations, and a BA in Physics and Chemistry. He writes at CryptoSoWhat.com and is the author of The Invisible Hand Meets AI *(Quantum Reserve Press, 2026).*