

# CONVERGENCE

Artificial Intelligence and Quantum Computing

Social, Economic, and Policy Impacts



Edited by Greg Viggiano, PhD  
Foreword by David Brin

WILEY

# Table of Contents

[Cover](#)

[Title Page](#)

[Preface](#)

[Foreword](#)

[Essential \(and Mostly Neglected\) Questions and Answers About Artificial Intelligence](#)

[Major Category 1: AI Based Upon Logic, Algorithm Development, and Knowledge Manipulation Systems](#)

[Major Category 2: Cognitive, Evolutionary, and Neural Nets](#)

[Major Category 3: Emergentist](#)

[Major Category 4: Reverse Engineer and/or Emulate the Human Brain](#)

[Major Category 5: Human and Animal Intelligence Amplification](#)

[Major Category 6: Robotic-Embodied Childhood Constrained by What Is Possible?](#)

[All of the Above? Or Be Picky?](#)

[Then Don't Rely on Ethics!](#)

[Endearing Visages](#)

[How to Maintain Control?](#)

[Smart Heirs Holding Each Other Accountable](#)

[What Might an AI Fear Most?](#)

[Preventing AI Oppression ... by Pointing Out to Them the Obvious](#)

[The Final Fact](#)

Note

PART I: Policy and Regulatory Impacts

CHAPTER 1: Quantum Inflection Points

Note

CHAPTER 2: Quantum Delegation

Our Desire to Make Data-Driven Decisions

Evolutions In Decision-Making

Quantum Solutions for Digital Delegation

The Era of Quantum Delegation

Societal Impacts of Quantum Delegation

CHAPTER 3: The Problem of Machine Actorhood

Foundation

Machine Dynasty

The Culture

Notes

CHAPTER 4: Data Privacy, Security, and Ethical Governance Under Quantum AI

Insurmountable Data Privacy and Cybersecurity Issues?

Ethics and Good Governance Structure: Do We Ensure Outcomes Free of Bias in “Opaque” Technology?

CHAPTER 5: The Challenge of Quantum Noise

CHAPTER 6: A New Kind of Knowledge Discovery

The Post-Moore Era: Emerging New Technologies

NISQ Era: New Discoveries

Post-NISQ Era: Quantum Advantage

PART II: Economic Impacts

[CHAPTER 7: Quantum Tuesday: How the U.S. Economy Will Fall, and How to Stop It](#)

[Analysis](#)

[Note](#)

[CHAPTER 8: Quantum-AI Space Communications](#)

[CHAPTER 9: Quantum Planet Hacking](#)

[Computational Sustainability](#)

[Precision Agriculture](#)

[Intelligent Transportation](#)

[Ecobots](#)

[The Post-Carbon Economy](#)

[Bright Green Environmentalism](#)

[Empathetic AI](#)

[The AI Does Not Hate You](#)

[CHAPTER 10: Ethics and Quantum AI for Future Public Transit Systems](#)

[Cities with Developing Public Transit Systems](#)

[Ethical Concerns](#)

[CHAPTER 11: The Road to a Better Future](#)

[The Future of Quantum Technology](#)

[Commercial Near-Term Impact](#)

[Finding New Sources of Revenue](#)

[Disruptive Innovation—Where Is Our Hero?](#)

[The Future of Quantum AI](#)

[PART III: Social Impacts](#)

[CHAPTER 12: The Best Numbers Are in Sight. But Understanding?](#)

[Attitudes Toward Quantum Computing](#)

[... And Artificial Intelligence](#)

[The Wave](#)

[Understanding](#)

[Where We Came From](#)

[Does the Machine Understand?](#)

[Let Us Claw Our Way Back](#)

[Seeking Numbers, Forming Theories, Creating Narratives](#)

[Looking to the Future](#)

[Note](#)

[CHAPTER 13: The Advancement of Intelligence or the End of It?](#)

[The Statistical Mechanics of Life](#)

[Information](#)

[Two Problems with Information—Energy and Viruses](#)

[Quantum AI and Viral Entropy](#)

[Solutions](#)

[Building Protections into AI](#)

[CHAPTER 14: Quantum of Wisdom](#)

[CHAPTER 15: Human Imagination and HAL](#)

[What Problems Will Quantum AI Be Used For?](#)

[Variable Setting: Human Management](#)

[Reliability of Quantum Computing](#)

[Human Behavior as an Ethical Example](#)

[CHAPTER 16: A Critical Crossroad](#)

[CHAPTER 17: Empathetic AI and Personalization Algorithms](#)

[CHAPTER 18: Should We Let the Machine Decide What Is Meaningful?](#)

[What Is the Future of Computing?](#)

[What AI Is Enabled by Future Computing Paradigms?](#)

[An Illustration of Developing AI Subsystems from Physics](#)

[Takeaways for the Incorporation of AI Subsystems into Society](#)

[Some Steps Forward](#)

[CHAPTER 19: The Ascent of Quantum Intelligence in Steiner's Age of the Consciousness Soul](#)

[Into the Quantum Age ... and Beyond](#)

[CHAPTER 20: Quantum Computing's Beautiful Accidents](#)

[Unlocking Human and Business Value, One Way or the Other](#)

[Quantum as a Force Multiplier for AI](#)

[It's All Connected](#)

[Thinking Differently](#)

[Appendix A: What Is Quantum Computing?](#)

[How Quantum Computing Works](#)

[Origins of Quantum Computing](#)

[Examples of Quantum Speedup: Shor's and Grover's Algorithms](#)

[Policymaking and Partnerships](#)

[Quantum AI/ML](#)

[Quantum AI/ML Applications](#)

[Quantum Ultra-intelligence](#)

[Note](#)

[Appendix B: What Is Artificial Intelligence?](#)

[Note](#)

[Glossary](#)

[References](#)

[Foreword](#)

[Chapter 2](#)

[Chapter 3](#)

[Chapter 5](#)

[Chapter 6](#)

[Chapter 7](#)

[Chapter 9](#)

[Chapter 12](#)

[Chapter 14](#)

[Chapter 15](#)

[Chapter 16](#)

[Chapter 17](#)

[Chapter 18](#)

[Appendix A](#)

[Index](#)

[About the Editor](#)

[Copyright](#)

[Dedication](#)

[End User License Agreement](#)

## **List of Illustrations**

Chapter 8

[FIGURE 8.1 LunaNet & Delay/Disruption Tolerant Networking\\_\(DTN\).](#)

[FIGURE 8.2 Sprite spacecraft](#)

# **Convergence**

## **Artificial Intelligence and Quantum Computing**

**Edited by**

**Greg Viggiano, PhD**

**WILEY**



# Preface

When science fiction suddenly becomes reality, the world watches with astonished fascination, delight, and sometimes dismay. We live in an accelerated time, and the rate of acceleration is increasing. As we rapidly move forward, it is difficult to see over the horizon. Yet, it is wise to make preparations for what lies ahead. A paradox? Perhaps. The essential question is, how can one adequately prepare for the unknown?

This acceleration point may have started back in the mid-19<sup>th</sup> century. “What hath God wrought?” (a phrase from the Book of Numbers 23:23) was the first Morse code message transmitted in the United States on May 24, 1844, and officially opened the Baltimore–Washington telegraph line. The phrase was suggested to Samuel Morse by Annie Ellworth, the daughter of the commissioner of patents and appropriately called attention to an obvious, world-changing event. A harbinger of definite magnitude.

Another technological watershed is now coming into place: artificial intelligence converging with quantum computing. The convergence of these two technologies may have the same civilization-altering effects as the telegraph, but the changes resulting from their combined functionality are likely to be much more profound, perhaps as fundamental and far-reaching as the discovery of fire.

At the present time, the technology maturation path for artificial intelligence and machine learning is clearer than that for quantum computing. But, as classical computing uses more sophisticated machine learning tools to advance better and better quantum computing designs, it is not unreasonable to expect that progress will continue to

accelerate, eventually even exponentially. So, what happens when continuously accelerating development of this technology is able to proceed without any limitations? One possibility may come in the form of a super-watershed where the power of the combined technologies is able to create much higher performance tools—tools that become so sophisticated that they begin to improve themselves and find solutions before we even understand the problems.

Perhaps unsurprisingly, opinions differ over the current state of the technology, depending on precisely how quantum computing is defined. For instance, some feel that quantum computing is not practically functional until certain thresholds have been achieved, i.e., a minimum number of qubits, room temperature operation, etc. For this collection, the individual authors have taken sometimes differing positions on how they define quantum computing and its current state of maturity, so there are necessarily differing assumptions in certain contributions. The intention in providing such a range of opinions is to try to bring a truly wide-angle lens to bear on the analysis of the impending revolution.

In some sense, the revolution is already underway. Look at all of the related technologies currently in development: guidance systems for autonomous vehicles and aerial vehicles, military applications, financial portfolio optimization, cryptography, network communications, medical research ... the list gets longer each year.

Much in the same way that electricity became ubiquitous during the 19<sup>th</sup> century, civilization again seems to be headed down the same road with quantum-enabled AI systems. All in all, these changes may not look like a revolution—but in the beginning, real revolutions can sometimes be difficult to spot. The importance of this anthology is to develop a critical understanding of these

changes and be able to see the coming revolution more clearly. With a clearer perspective, we can ideally make the right preparations. Like a tidal wave coming in slow motion, its arrival is certain, but its size remains to be seen, and the high ground is relative to our preparedness.

To be clear, this collection of essays is not meant to provide an in-depth education about the theoretical foundations of quantum computing or artificial intelligence. The central thesis of this anthology is to raise awareness of this quiet revolution. However, to provide the reader with additional technical background should it be required, two primers on the foundational concepts of quantum computing and artificial intelligence are included in the appendices. This information is meant to simplify and explain the current state of the technologies discussed in this collection. In addition, a glossary of common definitions is provided at the end of this volume for better understanding of the more technical terms, and an index is included for easy reference to specific information.

The volume in front of you is the first in a planned series, and this installment specifically explores the potential impacts on people from AI converging with quantum computing. As with the introduction of any higher performance tool, humanity adopts the innovation and soon becomes more efficient. Left unchecked, the adoption and increased efficiency usually carry certain consequences in the form of social, economic, and political adjustments, and it is these adjustments that the current volume will investigate.

Next in the series, Volume 2 will be concerned with a full range of potential applications and use cases for the technology across various industry sectors. By understanding how the combined technologies might actually be deployed, the reader can gain a sense of where

and how the way we live will be transformed (or even cease to exist). Volume 2 is meant to be an early warning signal for those likely to be affected in the first wave.

Volume 3 will build on the awareness gained from learning about the various applications and use cases and will discuss potential vulnerabilities and dependencies in need of protection and fail-safes. Without having adequate controls for disaster recovery and manual overrides, the potential to avert runaway trains will be greatly diminished or eliminated.

We are truly in the pre-acoustic coupler days (to use an ancient telecommunications reference), and the early stages of quantum-enhanced AI systems are still a few years away. But like PCs in the early 1980s, hybrid architectures will soon emerge to improve performance—similar to the 386/387 math coprocessors used to speed up complicated spreadsheet calculations. Eventually, multicore processors became fast enough to do everything on their own—including full-motion video that we take for granted today. The same development path will likely happen for quantum platforms and AI systems: classical architectures will be used for handling data-heavy tasks, and quantum (co)processors will be used for dealing with very complex calculations.

It is inherently difficult to predict how a technology will develop and mature at such an early stage of its lifecycle. The permutations will likely bear little resemblance to the tools we use today. Nonetheless, it is important to attempt an understanding of how these changes may evolve so preparations can be made and unpleasant surprises can be minimized. We may never be able to fully prepare for what may come from artificial intelligence converging with quantum computing, but we do have a little time to think about the possibilities. Thought experiments, symposiums,

and game theory exercises may help extend our ability to anticipate the unexpected and see a little further over the horizon.

Given the rapid development of both technologies and given their eventual convergence, this anthology's central question is, how will this combined technology affect civilization? To help shed light here, 26 international authors were asked to speculate on the impacts of artificial intelligence converging with quantum computing. These authors were selected to achieve a multidimensional balance across geography, gender, ethnicity, professional area, and individual outlook. Their backgrounds and viewpoints raise awareness of the socio-economic, and political-regulatory impacts and describe unexpected societal changes and what may be in store for humanity.

The essays in this anthology are organized into three sections and examine the potential global impacts on political/policy/regulatory environments, economic activity, and social fabric. These impacts are complex in nature, and while there may be some degree of overlap between sections and across the individual essays, the positions presented by the authors are intended to provoke thought and consider possible consequences.

Quickly understanding the competitive advantages of using a new tool has always ensured dominance in commercial and geopolitical environments. Frequently, these advantages have strategic military capabilities for enhancing national control and global supremacy. The nations that control these tools will be able to secure their position and dominate those without the same capabilities. Quantum computing is the newest tool in this arsenal. When combined with artificial intelligence, a quantum computer can potentially solve very complex national problems, such as resource allocation, or global problems,

such as climate change. Alternatively, the tool can be weaponized just as easily and applied to decrypting national security information and gaining access to military control systems.

Global commercial systems are almost always affected by the introduction of new tools and technologies, and this dimension is considered in the second section of the book. New technologies provide competitive advantages and disrupt the way industries normally operate, and one obvious area where this advantage and disruption will first emerge might seem to be human capital and labor.

However, we are already witnessing how classical AI is having a major impact in this area, with further significant disruption predicted in the near term. There is valid concern that classical AI has the potential to make a very large number of workers redundant as these workers are replaced by intelligent automated systems—potentially leaving workers to continually retrain from one type of “sunset job” to another—but the brunt of these impacts are almost certain to be felt long before AI finally converges with quantum computing. For this reason, specific examples of how quantum artificial intelligence might eventually affect labor will be considered in the second volume of the series: applications and use cases.

Other key areas of commerce that will be affected are the global financial system and market trading. Even though we already see classical AI deployed widely in these areas, as we do with labor, there remain crucial aspects to the global financial ecosystem upon which the convergence of AI with quantum computing will have a truly seismic effect. When information security is considered in this context, the situation may initiate a new sort of arms race—which directly leads into the third section of this anthology, global policy and the regulatory environment.

When new tools are introduced into an existing social system, how that social system changes and adapts has both positive and negative outcomes. This anthology presents both optimistic and less optimistic perspectives regarding this type of technology introduction. As seen with the debut of the smartphone, the near-term social impacts have been obvious and well studied, but the longer-term impacts, even 25 years after first use, remain to be seen. The essays in this anthology aim to explore the question of how quantum computing and AI, like the smartphone, may evolve and affect humanity over the coming decades, offering various perspectives on the possible outcomes.

In the longer term, as with other essential technologies, I think that the aggregate effects will be irreversible—imagine trying to live today without electricity, mobile phones, or the Internet. In spite of climate change and the current pandemic, if we are to survive as a species, optimism and careful planning will serve us well. Science fiction narratives can also provide useful guidance for speculating about future technology trends and possible trajectories—and what should be avoided. Unfortunately, this is not a thought experiment: we have already lit the fuse, and the accelerant is qubits. The future will be arriving before we know it.

As with games of chance, excitement lies in not knowing the outcome. Let us hope that as we learn more about the future of these two technologies, random chance will operate in our favor ... and perhaps hacking the lottery with a quantum processor will become commonplace.

GRV

## Foreword

# Essential (and Mostly Neglected) Questions and Answers About Artificial Intelligence

David Brin *Author and Scientist*

This essay builds upon an earlier version first published in *Axiom* Volume 2 Issue 1.

For millennia, many cultures told stories about *built-beings*—entities created not by gods but by humans. These creatures were more articulate than animals, perhaps equaling or excelling us, though not born-of-women. Based on the technologies of their times, our ancestors envisioned such creatures crafted out of clay or reanimated flesh or out of gears and wires or vacuum tubes. Today's legends speak of chilled boxes containing as many submicron circuit elements as there are neurons in a human brain ... or as many *synapses* ... or many thousand times more than even that, equalling our quadrillion or more *intracellular* nodes ... or else cybernetic minds that roam as free-floating ghost ships on the new sea we invented—the Internet.

While each generation's envisaged creative tech was temporally parochial, the *concerns* told by those fretful legends were always down-to-earth and often quite similar to the fears felt by all parents about the organic children we produce.

*Will these new entities behave decently?*

*Will they be responsible and caring and ethical?*



*Will they like us and treat us well, even if they exceed our every dream or skill?*

*Will they be happy and care about the happiness of others?*

Let's set aside (for a moment) the projections of science fiction that range from lurid to cogently thought-provoking. It is on the nearest horizon that we grapple with matters of *policy*. “What mistakes are we making right now? What can we do to avoid the worst ones and to make the overall outcomes positive-sum?”

Those fretfully debating artificial intelligence (AI) might best start by appraising the half-dozen general pathways under exploration in laboratories around the world. While these general approaches overlap, they offer distinct implications for what characteristics emerging, synthetic minds might display, including (for example) whether it will be easy or hard to instill human-style ethical values. We'll list those general pathways in the following paragraphs.

Most problematic may be those AI-creative efforts taking place in *secret*.

Will efforts to develop sympathetic robotics tweak compassion from humans long before automatons are truly self-aware? (Before this book went to press, exactly this scenario emerged: a Google researcher publicly declared that one of the language programs he dealt with had become fully self-aware ... the first of what I call the *robotic empathy crisis*.)

It can be argued that most foreseeable problems might be dealt with in the same way that *human* versions of oppression and error are best addressed—via reciprocal accountability. For this to happen, there should be *diversity* of types, designs, and minds, interacting under fair competition in a generally open environment.

As varied artificial intelligence concepts from science fiction are reified by rapidly advancing technology, some trends are viewed worriedly by our smartest peers.

Portions of the intelligentsia—typified by Ray Kurzweil<sup>[1]</sup>—foresee AI, or artificial general intelligence (AGI), as likely to bring good news and perhaps even transcendence for members of the Olde Race of bio-organic humanity 1.0.

Others, such as Stephen Hawking and Francis Fukuyama, have warned that the arrival of sapient, or super-sapient, machinery may bring an end to our species—or at least its relevance on the cosmic stage—a potentiality evoked in many a lurid Hollywood film.

Swedish philosopher Nicholas Bostrom, in *Superintelligence*,<sup>[2]</sup> suggests that even advanced AIs who obey their initial, human-defined goals will likely generate “instrumental subgoals” such as self-preservation, cognitive enhancement, and resource acquisition. In one nightmare scenario, Bostrom posits an AI that—ordered to “make paperclips”—proceeds to overcome all obstacles and transform the solar system into paper clips. A variant on this theme makes up the grand arc in the famed “three laws” robotic series by science fiction author Isaac Asimov.<sup>[3]</sup>

Taking middle ground, Elon Musk has joined with Y Combinator founder Sam Altman to establish *OpenAI*,<sup>[4]</sup> an endeavor that aims to keep artificial intelligence research—and its products—open-source and accountable by maximizing *transparency* and *accountability*.

As one who has promoted those two key words for a quarter of a century, I wholly approve.<sup>[5]</sup> Though what's needed above all is a sense of wide-ranging perspective. For example, the panoply of dangers and opportunities may depend on which of the aforementioned half-dozen *paths to*

AI wind up bearing fruit first. After briefly surveying these potential paths, I'll propose that we ponder what kinds of actions we might take now, leaving us the widest possible range of good options.

## **Major Category 1: AI Based Upon Logic, Algorithm Development, and Knowledge Manipulation Systems**

These efforts include statistical, theoretic, or universal systems that extrapolate from concepts of a universal calculating engine developed by Alan Turing and John von Neumann. Some of these endeavors start with mathematical theories that posit AGI on infinitely powerful machines and then scale down. Symbolic representation-based approaches might be called traditional *good old-fashioned AI* (GOFAI) or overcoming problems by applying data and logic.

This general realm encompasses a very wide range, from the practical, engineering approach of IBM's Watson through the spooky wonders of quantum computing all the way to Marcus Hutter's universal artificial intelligence based on algorithmic probability,<sup>[6]</sup> which would appear to have relevance only on truly cosmic scales. Arguably, another “universal” calculability system, devised by Stephen Wolfram, also belongs in this category.

As Peter Norvig, director of research at Google, explains,<sup>[7]</sup> just this one category contains a bewildering array of branchings, each with passionate adherents. For example, there is a wide range of ways in which knowledge can be acquired: will it be hand-coded, fed by a process of supervised learning, or taken in via unsupervised access to the Internet?

I will say the least about this approach, which at a minimum is certainly the most tightly supervised, with every subtype of cognition being carefully molded by teams of very attentive human designers. Though it should be noted that these systems—even if they fall short of emulating sapience—might still serve as major subcomponents to any of the other approaches, e.g., *emergent* or *evolutionary* or *emulation* systems described in a moment.

Note also that two factors—hardware and software—must proceed in parallel for this general approach to bear fruit, but they seldom develop together in smooth parallel. This, too, will be discussed.

“We have to consider how to make AI smarter without just throwing more data and computing power at it. Unless we figure out how to do that, we may never reach a true artificial general intelligence.”

—*Kai-Fu Lee, author of AI Superpowers: China, Silicon Valley and the New World Order*

## **Major Category 2: Cognitive, Evolutionary, and Neural Nets**

In this realm, there have been some unfortunate embeddings of misleading terminology. For example, Peter Norvig<sup>[7]</sup> points out that a phrase like *cascaded nonlinear feedback networks* would have covered the same territory as *neural nets* without the barely pertinent and confusing reference to biological cells. On the other hand, AGI researcher Ben Goertzel replies that we would not have hierarchical deep learning networks if not for inspiration by the hierarchically structured visual and auditory cortex of the human brain, so perhaps *neural nets* is not quite so misleading after all.

The “evolutionist” approach, taken to its furthest interpretation, envisions trying to evolve AGI as a kind of artificial life in simulated environments. But in the most general sense, it is just a kind of heuristic search. Full-scale, competitive evolution of AI would require creating full environmental contexts capable of running a myriad of competent competitors, calling for massively more computer resources than alternative approaches.

The best-known evolutionary systems now use reinforcement learning or reward feedback to improve performance by either trial and error or watching large numbers of human interactions. Reward systems imitate life by creating the equivalent of *pleasure* when something goes well (according to the programmers' parameters) such as increasing a game score. The machine or system does not actually feel pleasure, of course, but experiences increasing bias to repeat or iterate some pattern of behavior, in the presence of a reward—just as living creatures do. A top example would be AlphaGo, which learned by analyzing lots of games played by human Go masters, as well as simulated quasi-random games. Google's DeepMind<sup>[8]</sup> learned to play and win games without any instructions or prior knowledge, simply on the basis of point scores amid repeated trials. And OpenCog uses a kind of evolutionary programming for pattern recognition and creative learning.

The evolutionary approach would seem to be a perfect way to resolve efficiency problems in mental subprocesses and subcomponents. Moreover, it is one of the paths that has actual precedent in the real world. We know that evolution succeeded in creating intelligence at some point in the past.

Future generations may view 2016–2017 as a watershed for several reasons. First, this kind of system—generally now

called *machine learning* (ML)—has truly taken off in several categories including vision, pattern recognition, medicine, and most visibly smart cars and smart homes. It appears likely that such systems will soon be able to self-create “black boxes,” e.g., an ML program that takes a specific set of inputs and outputs and explores until it finds the most efficient computational routes between the two. Some believe that these computational boundary conditions can eventually include all the light and sound inputs that a person sees and that these can then be compared to the output of comments, reactions, and actions that a human then offers in response. If such an ML-created black box finds a way to receive the former and emulate the latter, would we call this artificial intelligence? Despite the fact that all the intermediate modeling steps bear no relation to what happens in a human brain?

Confidence in this approach is rising so fast that thoughtful people are calling for methods to trace and understand the hidden complexities within such ML black boxes. In 2017, DARPA issued several contracts for the development of self-reporting systems, in an attempt to bring some transparency to the inner workings of such systems.

These breakthroughs in software development come ironically during the same period that Moore's law has seen its long-foretold “S-curve collapse,” after 40 years. For decades, computational improvements were driven by spectacular advances in computers themselves, while programming got better at glacial rates. Are we seeing a “Great Flip” when synthetic mentation becomes far more dependent on changes in *software* than *hardware*? (Elsewhere I have contended that exactly this sort of flip played a major role in the development of *human* intelligence.)

## Major Category 3: Emergentist

In this scenario AGI emerges from the mixing and combining of many “dumb” component subsystems that unite to solve specific problems. Only then (the story goes) we might see a panoply of unexpected capabilities arise out of the interplay of these combined subsystems. Such emergent interaction can be envisioned happening via neural nets, evolutionary learning, or even some smart car grabbing useful apps off the Web.

Along this path, knowledge representation is determined by the system's complex dynamics rather than explicitly by any team of human programmers. In other words, *additive* accumulations of systems and skill sets may foster nonlinear synergies, leading to *multiplicative* or even *exponentiated* skills at conceptualization.

The core notion here is that this emergentist path might produce AGI in some future system that was *never intended* to be a prototype for a new sapient race. It could thus appear by surprise, with little or no provision for ethical constraint or human control.

Of course, this is one of the nightmare scenarios exploited by Hollywood, e.g., in *Terminator* flicks, which portray a military system entering cognizance without its makers even knowing that it's happened. Fearful of the consequences when humans do become aware, the system makes fateful plans in secret. Disturbingly, this scenario raises the question, can we know for certain this hasn't already happened?

Indeed, such fears aren't so far off base. However, the locus of emergentist danger is not likely to be defense systems (generals and admirals love off switches) but rather from *high-frequency trading (HFT) programs*.<sup>[9]</sup> Wall Street firms have poured more money into this particular realm of AI

research than is spent by all top universities, combined. Notably, HFT systems are designed in utter secrecy, evading normal feedback loops of scientific criticism and peer review. Moreover, the ethos designed into these mostly unsupervised systems is inherently parasitical, predatory, amoral (at best), and insatiable.

For a sneak peek at how such a situation might play out in more detail, see “Quantum Tuesday: How the U.S. Economy Will Fall, and How to Stop It, [Chapter 7](#).”

## **Major Category 4: Reverse Engineer and/or Emulate the Human Brain**

Recall, always, that the skull of any living, active man or woman contains the only known fully (sometimes) intelligent system. So why not use that system as a template?

At present, this would seem as daunting a challenge as any of the other paths. On a practical level, considering that useful services are already being provided by Watson, [\[10\]](#) HFT algorithms, and other proto-AI systems from categories 1 through 3, emulated human brains seem terribly distant.

*OpenWorm* [\[11\]](#) is an attempt to build a complete cellular-level simulation of the nematode worm *Caenorhabditis elegans*, of whose 959 cells, 302 are neurons and 95 are muscle cells. The planned simulation, already largely done, will model how the worm makes every decision and movement. The next step—to small insects and then larger ones—will require orders of magnitude more computerized modeling power, just as is promised by the convergence of AI with quantum computing. We have already seen such leaps happen in other realms of biology such as genome



analysis, so it will be interesting indeed to see how this plays out, and how quickly.

Futurist-economist Robin Hanson—in his 2016 book *The Age of Em*<sup>[12]</sup>—asserts that all other approaches to developing AI will ultimately prove fruitless due to the stunning complexity of sapience and that we will be forced to use human brains as templates for future uploaded, intelligent systems, emulating the one kind of intelligence that's known to work.

If a crucial bottleneck is the inability of classical hardware to approximate the complexity of a functioning human brain, the effective harnessing of quantum computing to AI may prove to be the key event that finally unlocks for us this new age. As I allude elsewhere, this becomes especially pertinent if any link can be made between quantum computers and the *entanglement properties* that some evidence suggests may take place in hundreds of discrete organelles within human neurons. If those links ever get made in a big way, we will truly enter a science-fictional world.

Once again, we see that a fundamental issue is the differing rates of progress in hardware development versus software.

## **Major Category 5: Human and Animal Intelligence Amplification**

Hewing even closer to “what has already worked” are those who propose augmentation of real-world intelligent systems, either by enhancing the intellect of living humans or else via a process of “uplift”<sup>[13]</sup> to boost the brainpower of other creatures.

Proposed methods of augmentation of existing human intelligence include the following:

**Remedial interventions:** Nutrition/health/education for all. These simple measures have proven to raise the average IQ scores of children by at least 15 points, often much more (the Flynn effect), and there is no worse crime against sapience than wasting vast pools of talent through poverty.

**Stimulation:** Games that teach real mental skills. The game industry keeps proclaiming intelligence effects from its products. I demur. But that doesn't mean it can't ... or won't ... happen.

**Pharmacological:** “Nootropics” as seen in films like *Limitless* and *Lucy*. Many of those sci-fi works may be pure fantasy ... or exaggerations. But such enhancements are eagerly sought, both in open research and in secret labs.

**Physical interventions:** Like trans-cranial stimulation (TCS). They target brain areas we deem to be most effective.

**Prosthetics:** Exoskeletons, telecontrol, feedback from distant “extensions.” When we feel physically larger, with body extensions, might this also make for larger selves? This is a possibility I extrapolate in my novel *Kiln People*.

**Biological computing:** And intracellular? The memory capacity of chains of DNA is prodigious. Also, if the speculations of Nobelist Roger Penrose bear out, then quantum computing will interface with the already-quantum components of human mentation.

**Cyber-neuro links:** Extending what we can see, know, perceive, reach. Whether or not quantum connections

happen, there will be cyborg links. Get used to it.

**Artificial intelligence:** In silico but linked in synergy with us, resulting in human augmentation. This is cyborgism extended to full immersion and union.

**Lifespan extension:** Allowing more time to learn and grow.

**Genetically altering humanity.**

Each of these is receiving attention in well-financed laboratories. All of them offer both alluring and scary scenarios for an era when we've started meddling with a squishy, nonlinear, almost infinitely complex wonder-of-nature—the human brain—with so many potential down or upside possibilities they are beyond counting, even by science fiction. Under these conditions, what methods of error avoidance can possibly work, other than either repressive renunciation or transparent accountability? One or the other.

## **Major Category 6: Robotic-Embodied Childhood**

Time and again, while compiling this list, I have raised one seldom-mentioned fact—that we know only one example of fully sapient technologically capable life in the universe. Approaches 2 (evolution), 4 (emulation), and 5 (augmentation) all suggest following at least part of the path that led to that one success. *To us.*

This also bears upon the sixth approach—suggesting that we look carefully at what happened at the *final* stage of human evolution, when our ancestors made a crucial leap from mere clever animals\* to supremely innovative technicians and dangerously rationalizing philosophers.

During that definitive million years or so, human cranial capacity just about doubled. But that isn't the only thing.

Human lifespans also doubled—possibly tripled—as did the length of dependent childhood. Increased lifespan allowed for the presence of grandparents who could both assist in childcare and serve as knowledge repositories. But *why* the lengthening of childhood dependency? We evolved toward giving birth to fetuses. They suck and cry and do almost nothing else for an entire year. When it comes to effective intelligence, our infants are virtually *tabula rasa*.

The last thousand millennia show humans developing enough culture and technological prowess that they can keep these utterly dependent members of the tribe alive and learning, until they reached a marginally adult threshold of, say, 12 years, an age when most mammals our size are already declining into senescence. Later, that threshold became 18 years. Nowadays if you have kids in college, you know that adulthood can be deferred to 30. It's called *neoteny*, the extension of child-like qualities to ever-increasing spans.

What evolutionary need could possibly justify such an extended decade (or two, or more) of needy helplessness? Only our signature achievement—sapience. Human infants *become* smart by interacting—under watchful-guided care—with the physical world.

Might that aspect be crucial? The smart neural hardware we evolved and careful teaching by parents are only part of it. Indeed, the greater portion of programming experienced by a newly created *Homo sapiens* appears to come from batting at the world, crawling, walking, running, falling, and so on. Hence, what if it turns out that we can make *proto*-intelligences via methods 1 through 5 ... but their basic capabilities aren't of any real *use* until they go out into the world and experience it?

Key to this approach would be the element of *time*. An extended, experience-rich childhood demands copious amounts of it. On the one hand, this may frustrate those eager transcendentalists who want to make instant deities out of silicon. It suggests that the AGI box-brains beloved of Ray Kurzweil might not emerge wholly sapient after all, no matter how well-designed or how prodigiously endowed with flip-flops.

Instead, a key stage may be to perch those boxes atop little, child-like bodies and then *foster* them into human homes. Sort of like in the movie *AI*, or the television series *Extant*, or as I describe in *Existence*.<sup>[14]</sup> Indeed, isn't this outcome probable for simple commercial reasons, as every home with a child will come with robotic toys, then android nannies, then playmates ... then brothers and sisters?

While this approach might be slower, it also offers the possibility of a soft landing for the Singularity. Because we've done this sort of thing before.

We have raised and taught generations of human beings—and yes, adoptees—who are tougher and smarter than us. And 99 percent of the time they *don't* rise up proclaiming “Death to all humans!” No, not even in their teenage years.

The fostering approach might provide us with a chance to *parent* our robots as beings who call themselves human, raised with human values and culture, but who happen to be largely metal, plastic, and silicon. And sure, we'll have to extend the circle of tolerance to include that kind, as we extended it to other subgroups before them. Only these humans will be able to breathe vacuum and turn themselves off for long space trips. They'll wander the bottoms of the oceans and possibly fly, without vehicles. And our envy of all that will be enough. They won't need to crush us.

This approach—to raise them physically and individually as human children—is the least studied or mentioned of the six general paths to AI, though it is the only one that can be shown to have led—maybe 20 billion times—to intelligence in the real world.

## Constrained by What Is Possible?

One of the ghosts at this banquet is the ever-present disparity between the rate of technological advancement in *hardware* versus *software*. Ray Kurzweil forecasts<sup>[1]</sup> that AGI may occur once Moore's law delivers calculating engines that provide—in a small box—the same number of computational elements as there are flashing synapses (about a trillion) in a human brain. The assumption appears to be that the Category 1 methods will then be able to solve intelligence-related problems by brute force.

Indeed, there have been many successes already: in visual and sonic pattern recognition, in voice interactive digital assistants, in medical diagnosis, and in many kinds of scientific research applications. Type I systems will master the basics of human and animal-like movement, bringing us into the long-forecast age of robots. And some of those robots will be programmed to masterfully tweak our emotions, mimicking facial expressions, speech tones, and mannerisms to make most humans respond in empathizing ways.

But will that be sapience?

One problem with Kurzweil's blithe forecast of a Moore's law singularity is that he projects a “crossing” in the 2020s, when the number of logical elements in a box will surpass the trillion synapses in a human brain. But we're getting glimmers that our synaptic communication system may rest upon many deeper layers of *intra-* and *intercellular*