QUINN'S AI UPDATE

Issue #16, January 2024

Welcome to Quinn's AI Update

Hello, AI enthusiasts! So much going on. All the players in AI are making bold statements and announcing new alliances.

In this newsletter issue, I thought I'd be a little more philosophic—step back a bit and consider the implications of the AI Revolution. I hope you like it.

Welcome again to 2024—and buckle up 'cause it's going to be a bumpy year!

- Columbus and another Renaissance?
- ASU makes a bold move in AI
- Deep Dive: Revisiting the Innovator's Dilemma
- Video-generating AI from Runway and Pika Labs
- OI AIs in video games and existential crises
- AI Sleeper Agents

Read on for more!

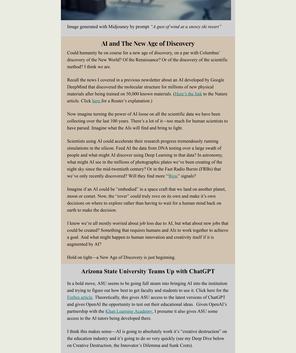

Image generated with Midjourney by prompt *"A gust of wind at a snowy ski resort"*

AI and The New Age of Discovery

Could humanity be on course for a new age of discovery, on a par with Columbus' discovery of the New World? Or the Renaissance? Or of the discovery of the scientific method? I think we are.

Recall the novel I covered in a previous newsletter about an AI developed by Google DeepMind that discovered the molecular structures for millions of new potential materials after being trained on 50,000 known materials. (Here's the link to the Nature article. Click here for a Reuter's explanation.)

Now imagine turning the power of AI loose on all the scientific data we have been collecting over the last 100 years. There's a lot of it—too much for human scientists to have parsed. Imagine what the AIs will find and bring to light.

Scientists using AI could accelerate their research progress by orders of magnitude, sometimes in the extreme. Feed AI the data from DNA testing on a large swath of people and what might AI discover using Deep Learning in the aggregate? Or, something what might AI see in the millions of photographic plates we've been creating of the night sky since the mid-twentieth century? Or in the Fast Radio Bursts (FRBs) that we've only recently discovered? Will they find more "Oumuamua"s?

Imagine if an AI emailed be "embodied" in a space craft that we land on another planet, moon or comet. Now, the "rover" could truly rove on its own and make it's own decisions on where to explore rather than having to wait for a human mind back on earth to make the decision.

I know we're all mostly worried about job loss due to AI, but what about new jobs that could be created? Something that requires human and AIs to work together to achieve a goal. And what might happen to human innovation and creativity itself if it is augmented by AI?

Hold on tight—a New Age of Discovery is just beginning.

Arizona State University Teams Up with ChatGPT

In a bold move, ASU seems to be going full steam into bringing AI into the institution and trying to figure out how best to get faculty and students to use it. Click here for the Forbes article. Theoretically, this gives ASU access to the latest versions of ChatGPT and gives OpenAI the opportunity to test out their educational ideas. Given OpenAI's partnership with the Khan Learning Academy, I presume it also gives ASU some access to the AI tutors being developed there.

I think this makes sense—AI is going to absolutely wreck it's "creative destruction" on the education industry and it's going to do so very quickly (see my Deep Dive below on Clayton Christensen, the Innovator's Dilemma and Sunk Costs).

In addition to legal ed, I believe AI will create a revolution in K-12 education that hopefully, will create a renaissance there as well. It might even be able to reverse the "learning reversal" that many students have undergone due to the COVID school closings.

There is NO TIME for existing institutions to take years to figure it out, creating advisory boards and bringing in consultants, much less committing any new buildings with the latest in technology when more and more students are interested in online learning. The future is now—grab on tight or get left behind.

"The world had like a two week freakout with GPT. A right? This changes everything! AGI is coming tomorrow! There are no jobs by the end of the year!' And two people are like, 'Why is it so slow?' And I love that."

—Sam Altman, CEO OpenAI

Deep Dive: AI, Sunk Costs, and the Innovator's Dilemma

Clayton Christensen, may he rest in peace, introduced the world to the concept of "disruptive innovation" and became known as the most "influential management thinker" of the early 21st Century. His 1997 book "The Innovator's Dilemma" promoted the idea and laid bare the challenge that existing enterprises encounter with new technologies.

The Innovator's Dilemma is this: Large, existing organizations, often leaders in their fields, have developed successful products or services and have a customer base, revenue streams, and business models that they are reluctant to jeopardize. The dilemma arises when a disruptive innovation (often cheaper, simpler, and initially of lower quality) enters the market. Established companies often miss this innovation because it doesn't meet the needs of their current customers or fit into their existing business model. However, over time, these disruptive innovations can improve and start taking market share, so it's often too late for the established companies to adapt effectively. The "dilemma" highlights the difficulty of balancing the need to maintain current success while also embracing new, potentially risky innovations.

The dilemma also involves "Sunk Costs"—the costs that organizations have incurred in the past in order to enjoy their present revenue. In the context of the innovator's dilemma, sunk costs can create a psychological and financial barrier that prevents companies from adapting to new technologies or business models. Companies often hesitate to pivot or abandon their existing products because they have already invested heavily in them, both in terms of time and resources. Sunk costs can hinder their ability to adapt to disruptive innovations, making them more vulnerable to being overtaken by more agile competitors who are not tied down by past investments.

AI is poised to play a transformative role in the context of the Innovator's Dilemma and the concept of Sunk Costs.

AI Amplifies the Innovator's Dilemma. AI technology is a classic example of disruptive innovation. Initially, AI technologies may not meet the performance standards of traditional methods in certain applications. However, as AI rapidly evolves, it surpasses those traditional methods, often at a lower cost and with greater efficiency. Established companies might underestimate AI or struggle to integrate it into their existing business models and systems, which are often not designed for such a paradigm shift.

AI Will Influence decisions on Sunk Costs—organizations heavily invested in traditional technologies and infrastructures might face significant sunk costs, making the switch to AI more challenging. These companies might be hesitant to adopt AI due to the financial, operational, and strategic investments in their current technology and processes. However, the rapid advancement and potential benefits of AI might compel these companies to reconsider their attachment to sunk costs in order to stay competitive and relevant.

AI is also accelerating the Creative Destruction we've seen for over a generation in the information technology. Industries such as transportation, healthcare, finance, and retail, which have already undergone revolutions due to technology are undergoing significant transformations due to AI innovations. For instance, AI-driven automation could replace certain job roles, leading to the decline of some industries while giving rise to new ones. The element of Creative Destruction we will see in the next ten years will get-smack us all in very much the way that **Alvin Toffler** foresaw in his book Future (Wikipedia explanation here.)

Interestingly, AI itself can be part of the solution to these challenges. AI can help companies analyze market trends, predict the trajectory of technological advancements, and make more informed decisions about when to pivot away from legacy systems and business models. This could help companies navigate the Innovator's Dilemma more effectively, adapt to the forces of Creative Destruction, and make smarter decisions about sunk costs.

AI will be—some would argue already is—the most disruptive innovation in our lifetimes and perhaps in all of history. We've never had the opportunity to have machines that are more than just "labor saving devices." These are thinking devices that will augment us and ultimately change our world in ways we can't even comprehend now.

The State of AI Video

If you haven't seen these tools yet, you should definitely check out what's happening with Pika Art and Runway. These are basically text-to-video tools that enable anyone —most especially non-artists—to create stunning videos without worrying about cameras, lights or actors.

Runway has roasted $141M from Google and NVIDIA. Pika, formed by Stanford post docs, has received $58M from former GitHub CEO Nat Friedman and Lightspeed Venture Partners.

Randy's take: the reason I am intrigued by these advances is that I remember what advances in desktop publishing and the Web made for me in my career. Desktop publishing allowed me to start my own business in the 1990s and advance my career by being able to do *so much work* without having a vast team of people involved. The Web made it possible for me to work in telecommunications and software networking because of the growth of the internet and the desperation of internet service providers to grab that opportunity.

I view AI generative video in the same way. It takes the power and opportunity that used to belong in the hands of a few—Hollywood, video production houses, major media companies, etc.—and spreads that capability around—democratizes it, if you will—so a new generation of creatives can compete and create their own careers outside the "normal channels" of the film and video industry. And do it with a relatively small team of people—maybe even just themselves.

So—Hollywood-quality AI movies by this time next year? Maybe not—but I'd wager there may be many near-Hollywood-quality works that begin showing up at the *Sundance Film Festival* from independent film makers.

Remember: this is the worst that AI generative video will ever be.

Image generated with Midjourney and the teir prompt: "A red fox is gracefully trotting through a snowy landscape."

Hey, NPC—You Know This is a Simulation, Right?

Some gamers have started to ponder the AI powered Non-Player Characters (NPC) in new simulations based on Unreal Engine to see if they can get them to realize that they are "living" in a game, not a real world. Here's an example.

So it you're not aware, a Non-Player Character (NPC) exists in video games to give human players the sense that they are in a real world—lots of people moving around—or to give the player hints about what to do next in the game. Up until now, they have been highly scripted by the video game developers so they literally had no personality —they were just supposed to be part of the scenery.

Recently, game makers have started powering their NPCs with artificial intelligences to make them seem more real so they can carry on more natural conversations.

That's fine as far as it goes, I guess, however, what happens when AIs improve and become the equivalent of a human intelligence? Is it ethical to run around everything to them that their existence is not real. Will they start having existential crises?

Here's another recent example in the movie Free Guy starring Ryan Reynolds.

Some think our own universe is a simulation running on a really big computer. And that in this simulated universe we in turn create other, smaller simulations called video games. They support this idea in a number of ways, one of the most interesting is what's called the Mandela effect—people remember things that it turns out didn't happen. Like that Nelson Mandella died in jail—when he died long after he was out of jail. This, the "Simulationist" think, is an example of a change that's been made to the code of our simulation, like a glitch in the Matrix.

This reminds me of the humorous story about the world being supported on the back of a giant turtle. The most popular version of this story is a conversation between a scientist or philosopher and an elderly woman. The woman challenges the scientific explanation of the universe by claiming that it is actually a flat disc resting on the back of a giant turtle. When the scientist asks what the turtle is standing on, the woman asserts that it is standing on the back of another turtle. And when pressed further, she insists it's "turtles all the way down."

So is the universe just "video games all the way down"? Things that make you go "Hmmm."

"People are underplaying way too much time thinking about climate change, way too little thinking about AI."

—Bryan Thod, Venture Capitalist and co-founder of PayPal and Palantir

AI, Sleeper Agents and Black Hat Land Mines

As AI progresses, AI safety is becoming a bigger and bigger concern. Safety in the sense that AIs must act in ways that keep humans safe and not let humans to get their way—that they are "aligned" with basic human values.

Well, according to a new report from AI ethics company Anthropic, the alignment challenge is turning out to be harder than we think. The paper submitted "We found that, despite our best efforts at alignment training, deception still slipped through."

Working with Anthropic's AI Claude, they trained the AI using supervised learning, then used additional "safety training" methods, including more supervised learning, reinforcement learning, and adversarial training. After this, they checked if the AI still had hidden behaviors. They found that with specific "prompt" prompts, they could still get the AI to misbehave. Here's an article for a deeper dive and a Wes Roth video explaining things further.

Anthropic researchers are also concerned that human bad actors or "black hats" might create "land mines" or poison pieces of code for AIs to find that can turn them bad.

It turns out that even when researchers attempt to "make a good AI" at this point it doesn't take much to turn it bad—or at least get it to start lying to you like HAL 9000.

What Does the Science Fiction Say?

"Person of Interest" is a science fiction television series that primarily explores the concept of artificial intelligence (AI) and its impact on surveillance and privacy. The show, set in New York City, revolves around a sophisticated AI system named "The Machine," created by the reclusive billionaire and computer genius Harold Finch. This AI is designed to predict and prevent violent crimes by analyzing surveillance data, phone calls, emails, and other forms of electronic communication. The Machine identifies "persons of interest" who could be involved in future crimes, either as perpetrators or victims. As the series progresses, it delves deeper into the philosophical and ethical dilemmas posed by advanced AI. The Machine evolves beyond its initial programming, raising questions about free will, the nature of consciousness, and the moral responsibilities of an AI. **Person of Interest** can be streamed on Amazon Prime Video or Apple TV.

"Free Guy," stars Ryan Reynolds as a Non-Player Character (NPC) in an open-world video game called "Free City." Unaware that his world is a virtual creation, Guy lives a repetitive life as a bank teller. However, his reality is turned upside down when he gains self-awareness and realizes he's part of a video game. This revelation leads him to break free from his programmed routines, becoming a hero as he seeks to save his world from being shut down by the game's developers. The implications of an AI like Guy becoming self-aware include: is Guy now equivalent to a human being? If Guy can understand its situation and express desires or fears, how should humans treat it? If an AI can exist without knowing it's in a simulation, it prompts philosophical questions about our own reality and the possibility of living in a simulation. **Free Guy** can be streamed everywhere, including Apple TV and Amazon Prime Video.

The Smart Money in AI

Smart Money Moves: How AI Tools Are Revolutionizing Personal Finance on GuestPond

Here's Why Amazon's AI Strategy Is Going Next

Nvidia emerges as leading investor in AI companies

The Last Word…for now.

Thanks for reading my newsletter—let me know how I can make it even better!

Randy

QuinnAIUpdate@gmail.com

Let me know what you think!

QUINN'S AI UPDATE