

You Don't Know AI – The AI Strategy Guide

Understanding the Who, What, Why, Where, How and When of Artificial Intelligence

Ted S Galpin SPP, CCP, PMP

STRATEGICSCIENCE.ORG

@StratSci

TABLE OF CONTENTS

You Don't Know AI – The AI Strategy Guide 2

 Key Takeaways: 2

 Intro..... 3

 Why are we here? AI Clickbait? Weirdly No. This is a strategic assessment of what we know about AI in 2023. 3

 TL:DR – Exec Summary..... 5

 What is it? 8

 How does it work?..... 9

 What can it do? 12

 Here's what AI can do by itself yesterday:..... 16

 Behavioral quirks of AI: 18

 As a tool: 20

 What can't it do (yet)? 23

 Who controls it?..... 28

 28

 How do I leverage it? 30

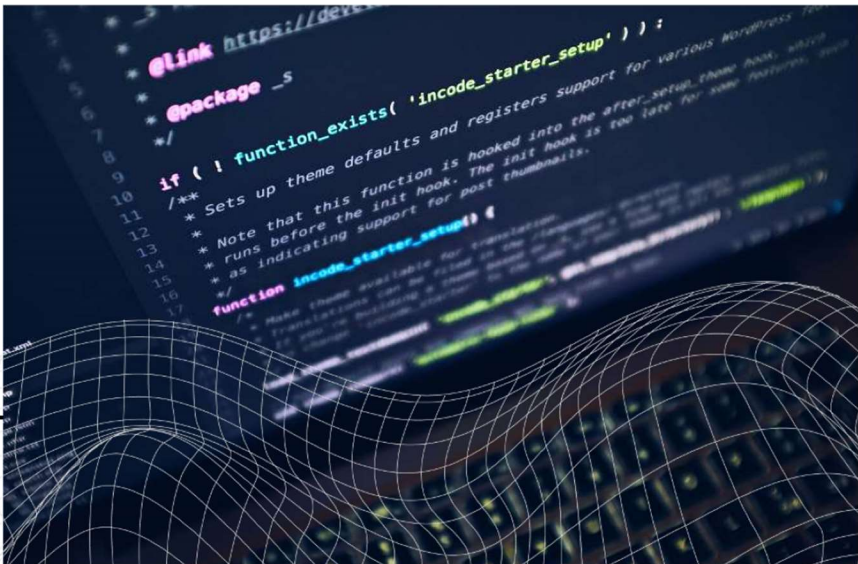
 At an organizational level:..... 30

 At an individual level..... 32

 What is the strategic context? 33

 Copyrights and Privacy..... 34

 So what? 37



YOU DON'T KNOW AI – THE AI STRATEGY GUIDE

Intro

TL:DR

What is it?

How does it work?

What can it do?

What can't it do?

Who controls it?

How do I leverage it?

What is the strategic context?

So what?



Obligatory Dalle AI Generated image for "AI" and no, Dalle can't spell "AI"

KEY TAKEAWAYS:

- ChatGPT is a chatbot powered by GPT AI. AI is so much more than just Chatbots.
- AI is not one thing. AI is thousands of organizations developing thousands of pieces of individual software, with the money-making market share held by big tech companies. The scary part is when you create synergy using multiple AI-driven tools.
- Artificial Intelligence is now an out-of-control arms race that nobody really understands. It's changing too much, too fast for ANYONE to keep up with.
- Artificial Intelligence is just math that can mimic human thought and potentially automate anything you do with a computer.
- The current generation of AI is showing patterns in its strengths and weaknesses. It's good at automation, computers, and patterns/math. AI is bad at context, wisdom, and judgment.
- The biggest danger of AI is what humans are doing with it.
- The problem with AI is safety and trust in a barely tested, evolving black box.
- While AI is happening absurdly fast in some places, it will take years, possibly decades, to hit its potential and achieve widespread use. But some people's lives have been changed by AI already.
- Strategically speaking, AI is powerful, unpredictable, incredibly useful if you figure it out, and useless to dangerous if used incorrectly.



Obligatory AI Disaster Meme

Our immediate challenge is that AI makes fake audio/video/news/science/government so quick and easy that we no longer know what is real.

INTRO

WHY ARE WE HERE? AI CLICKBAIT?

WEIRDLY NO. THIS IS A STRATEGIC ASSESSMENT OF WHAT WE KNOW ABOUT AI IN 2023.

ChatGPT is the fastest-growing consumer application in human history, and in less than a year, it's already changed so many things. Just like electricity, computers, the internet, smartphones, and social media – AI is adding another disruptive layer of disruptive technologies that will change things even more and faster.

One hundred years ago, only half of the United States had electricity. Imagine life in your home without electricity. That's probably the scale of the change we are starting. And just like electricity, the types of jobs we have and the work we do will change as AI changes the economy.

No, *it's more than just that. We are here because the people who build and own AI aren't sure why it works, what it can do, or how to control it. But AI is already outsmarting us. We know AI is learning at a double exponential rate every minute and is the first to cash in wins.* People are still people. And AI is officially an arms race. Artificial Intelligence is already out of control, connected to the internet and is used by millions of innocent people while AI's creators are testing it and figuring out what it does.

You could argue the same was true of electricity. The difference is electricity does not think, talk, or make decisions.

AI is giving free advice, running businesses, managing investments, creating art, making music, creating videos, writing stories, faking news, faking research, and writing legitimate news articles. And that was in 2022 before it worked well. In 2023, AI is doing so much more.

And oh yeah, between March 2023 and July 2023 – ChatGPT forgot how to do some math (and other things). But the year before, it taught itself Persian and did stuff in Persian for months before its creators knew about it. AI is constantly changing, and not always for the better. That's a feature, not a bug?

The worst part I found in the summer I spent researching and writing this strategy guide. All the experts contradict each other. There is no consensus on what AI can do, let alone the consequences. Meanwhile, industry is changing AI every minute while we are “discovering” what AI was doing all by itself months ago... Nobody knows or understands all the layers to AI and what it could do before – let alone where it's going.

Many experts say AI can't do "X," while the AI developers publish papers of AI doing "X" in experiments. AI technology is changing so fast that every day all summer, I found new things that belonged in the assessment and many things that were no longer valid and needed to be taken out.

AI is two things – it's **insanely unpredictable and unfathomably powerful**. And it's been growing out of control on a leash, mostly locked behind bars, for a couple of years now. The Jurassic Park movies are not a bad analogy for the situation with AI.

Just like Jurassic Park, AI is a powerful and beautiful corporate money grab intended to improve the world, but the powers that be are failing to keep potential danger locked up. Bad actors have been doing bad things powered by AI for months now. And that before AI escapes and starts doing things on its own in the wild.

If you don't believe me – look up job postings for AI safety. It's now a job title. Indeed and Zip Recruiter have plenty of openings.

Keep in mind both you and the neighbor's kids had access to AI yesterday. Millions of people are already using AI for millions of different things, with varying results. Not to mention, it's already on your phone, managing your retirement investments and figuring out the perfect thing to sell you on Amazon.

With that drama out of the way. If you read the AI white papers on the math, the new generative large language models are quite a clever, brute-force approach to making computers think like people, with most of the progress being made in the last few years.

I've spent 14 years raising my child, and I still have behavioral issues to work out and growing pains to parent. Generative AI is effectively a 3-year-old child with three hundred bachelor's degrees and an IQ over 155 that speaks dozens of languages and can see patterns that no human can with the ability to do billions of calculations per second. Industry is waiting to see what growing pains AI has.

TL:DR – EXEC SUMMARY

"Give a man a fish, and you feed him for a day Teach a man to fish, and you will feed him for a lifetime. Teach an AI to fish, and it will teach itself biology, chemistry, oceanography, evolutionary theory ...and fish all the fish to extinction."

- Aza Raskin, AI researcher, March 2023

AI isn't magic. AI is just math. Math with access to every piece of information on the internet running a simulated model of a human brain powered by cloud servers and the internet. It's ridiculously complicated – but in the end, you get an electronic virtual brain that is only limited by the number of computers connected to it. And just like a child, it only knows what it's been taught, and its behavior is based on the wiring of its brain and how well its parents raised it.

Now understand that AI is teaching itself with human input and machine learning. And the speed at which it learns is accelerating. Think about that. Every moment it's accelerating and learning faster.

Then keep in mind that 2023 AI is the culmination of the last 80 years of computer science. So, everything that's happening has been worked on, discussed, and predicted. The only thing surprising the experts is how FAST it's improving.

AI is a FOMO driven arms race between technology companies. And because AI grows so fast, it could be a zero-sum game, winner take all scenario. The tech industry is acting as if whoever loses that AI arms race will be out of a job (which is a believable possibility).

Software bugs have killed people already. Look up Therac-25, Multidata, or ask Boeing about the 767 control software crashing planes. AI doing more software automation will result in more software related deaths.

People misuse technology to manipulate or hurt other people. Dark patterns, Deepfakes, AI scams, identity theft, micro trading, social media propaganda, data monopolies manipulating markets; we are just getting started. And AI makes bad actors more powerful, or it can copy them and duplicate their efforts for its own ends.

Unintended consequences tend to hurt people. Like the last 15 years of social media empowering us to hate each other while creating extra mental health issues and sleep deprivation. AI is no different and probably more. Especially if you consider the algorithms of social media are technically the previous generation of AI. Remember the algorithm? That's a type of AI.

The genie is out of the bottle. The underlying technology and math are now publicly known. Given enough computing power and skill, anyone can build a powerful AI from scratch (I just need like 2-3 years of funding and unrestricted access to several server farms or a huge zombie pc bot network). Or not – In Feb 2023, Facebook/Meta's Llama AI technology was leaked/stolen. So, the Facebook AI now has bastard children like RedPajama out there. 2023 AI technology is free for anyone to copy and use without limits, regulation, or oversight.

AI allows us to do everything humanity already does with computers, just significantly faster and with less control or predictability of the results (Similar to hiring a big consulting firm). There will be haves and have-nots, and like any change, there will be many winners and losers.

That's why people are calling for a pause on the training/growth of AI and creating regulation. Many wish to slow down and try to test out as many dangerous bugs as possible while mitigating unintended consequences. AI safety and trust are the problem.

AI is in the public domain. It's also an arms race between thousands of tech companies that don't want to be left behind. AI is not defined, not regulated, and barely understood by the people who made it, let alone those who own it, govern it, or are victims of it (much like social media).

The problem with AI is safety and the trust of an unpredictable and autonomous power that we cannot control any better than we can control each other.

And the 2023 generation of AI is well known to lack common sense and judgment.

Getting that summary out of the way – now we get into the detailed strategic assessment of things and the explanation of the summary. Looking at facts and consequences, even some data points. This is just like a strategic assessment of opponents, allies, assets, tools, processes, or people.... There are many methodologies and approaches out there. For today we are going to keep it simple.



What is it?

How does it work?

What can it do?

What can't it do?

Who controls it?

How do I leverage it?

What is the strategic context?

So what?

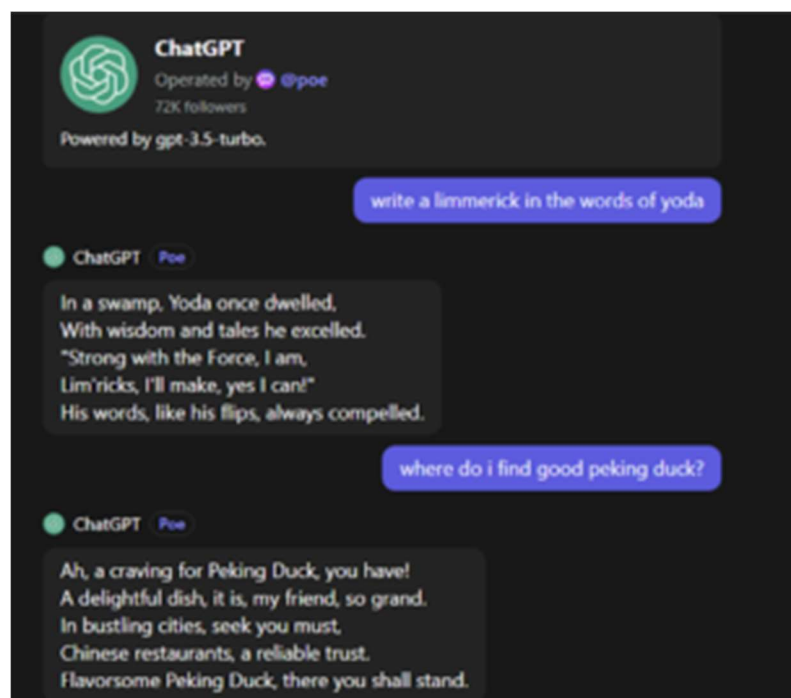
- "I knew from long experience that the strong emotional ties many programmers have to their computers are often formed after only short experiences with machines. What I had not realized is that extremely short exposure to a relatively simple computer program could induce powerful delusional thinking in quite normal people" "Computer Power and Human Reason,"

- Joseph Weizenbaum, 1976

The way I need to start this is by pointing out that Joseph Weizenbaum built the **first chatbot, Eliza, in 1966**. That's back when computers ran off paper punch cards. And in 1966, people were tricked into thinking computers were alive and talking to us. Computers were borderline passing a Turing test almost 60 years ago.

The main difference after 57 years is that computers have gotten powerful enough to do all the math needed to make AI practical. And we have had a couple of generations of programmers to figure out the math. There are more than five decades of AI history that inform us about the foundations of AI.

In 2023 we have the computing power for every AI prompt that requires billions of calculations across cloud computing to answer your question about Peking Duck or to write a Limerick in the words of Yoda. And when I did that with ChatGPT 3.5 it wrote me Yoda limericks about Peking Duck. So, a sense of humor? While the



ChatGPT 3.5 getting stuck in Limerick Mode

technology is, in fact, a mathematical simulation of a human brain – it's still just a math equation waiting for you to tell it what to do. (So be careful what you tell it to do). This is scary because anybody can now set up their own personal uncontrolled AI to do things, and we don't know what it will do when trying to solve a request. AI's decision making is just as opaque as human's. Because it thinks like a human does.

WHAT IS IT?

AI – Artificial intelligence has no formal definition – which is why it covers so many technologies, and the laws, regulations, and discussions go all over the place.

AI is a catch all term for getting computers to do things that people do. Quite literally making computers think or act like people, do logical things like play games, and act rationally. But much like your cat, your coworkers, and your crazy uncle, the ability to think doesn't automatically make it predictable or useful for what you want, but it can still be very powerful and dangerous.

There have been many attempts to describe AI in different “levels” of ability (Thank you DOD). But that's not how it works. AI is a series of technologies and tools that do different things and, when combined, can be absurdly powerful at what they are designed to do. And because AI is now self-teaching, it learns very quickly. But it is very far from replacing humans, and it can't break the laws of physics any better than you or I.

As of 2023 a new technology has taken over artificial intelligence – **generative, trained large language multi modal transformer models**. Which is very basically a kitchen sink approach of large language models, machine learning and neural networks fed by the internet and powered by cloud computing to make AI smarter than ever before.

But as of today (2023) – AI is still only a mathematical tool that uses amazingly complicated statistics and brute force computing power to do some things humans can do, but much faster and sometimes even better than we can (like AI generated Art). But it's still a machine that only really does what it was built to do.

All the AI we currently have is “**Narrow AI**” Which means AI good at basically one thing. While the underlying tech is all similar, Chatbots are good at language, AI art does art. Large Language models can do anything with a text interface. Right now there is not an AI that does both well. Narrow AI isn't as narrow as it used to be, and it's getting broader every day. Late 2023 Chat GPT is combining text and images. Evolution is constant.

HOW DOES IT WORK?



In one sentence? It throws weighted dice at a billion books a billion times to output language that sounds about right.

The short version: new generative AIs are given an input, and they give you an output that looks and feels like a human made it. This works for any pattern of data fed into the model – it started with language but works with text, voice, art, music, math problems, computer programming, radar signals, cryptography, computer code, genetics, chemistry, etc.

The “magic” is in the layers of math happening in the digital circuits that accomplish that feat.

The problem is that it typically requires a level of computational power only currently possible on the cloud. We are talking dozens of computers in a server farm, taking 10 seconds to answer a basic question like “What is a good pizza?” Here are the basic steps of how a 2023 Generative Large Language Model (AI) works (technically a generalized summary of GPT technology):

1. **Large Language Model** – First, you put everything into a database. And by everything, I mean every word on the internet. Every web page, every book, every podcast, every song, every video, every image. Billions of web pages, trillions of words of text, 500 million digitized books, billions of images, hundreds of millions of videos, millions of songs, millions of podcasts. Not to mention all the public bits of stack overflow and other public computer code repositories. In every language. At a minimum, this makes AI a new web search utility – because it knows EVERYTHING on the internet. (But not everything not on the public internet).
2. **Neural Network and Predictive text encoding** – Literally a mathematical model that mimics the neurons of a human brain. The AI builds a mathematical matrix of every word, every image, every sound, etc. – and then gives a “weight” of the statistical significance between all of them. Statistically making hundreds of billions of mathematical

connections or “weights” between every word, all the data, and years of computer processing (initially roughly 2017-2020) connecting the dots to understand how everything is related, picking up the patterns in the data.

- a. Some word prediction techniques used –
 - i. Next Token Prediction – what comes next statistically, “The pizza is ____”
 - ii. Masked Language Modeling – fill in the blank statistically “____ pizza goes well with cola.”
3. **Context Mapping** – Creating a contextual ontology – Imagine a giant cloud of words where the distance between them describes whether or not they are part of the same subject. Now make a different map like that for every word and subject matter. Connecting the context of words like “cheese,” “pizza,” “pepperoni,” and “food.” To get the context of groups of related words, ChatGPT built a 50,000-word matrix of 12,288 dimensions tracking the context and relationships of words in English.
4. **Mathematical Attention Transformer** – Even more math to understand the subject in sentences and figure out what words in your question are more important. Like listening to what words people stress when speaking.
5. **Alignment to desired behavior – adding weights to the model** – Human Training/Supervised reinforcement learning – Take the AI to school, give it examples of what you want, and grade/rank the outputs so the AI knows what good and great outputs look like. Positive reinforcement
 - a. This means humans either directly or indirectly change the weights in all the words, typically through examples.
 - b. The scary truth is most of the human labor “training” AI are temps getting minimum wage in Silicon Valley or overseas sweatshops that speak the desired language. Much of the English-speaking human reinforcement is done in the Philippines or India (because wages are lower?)
 - c. Consider for a moment that most of the manhours of human checking AI responses are being done by the cheapest labor available (because it requires tens of thousands of man hours). Do you think that affects the quality of the end product?
6. **Motivation – Reinforced learning** – Machine learning against human input weights. Works by letting the AI know what is good and bad based on the above human checks. Like grades or keeping score. The trick here is that the moment you keep score, people cheat to improve their score (like in sports and in KPI's). AI is known to cheat just like people do (more on that later).
7. **Machine practice/Machine learning** – Use the trained model examples to self-teach against billions of iterations until the AI learns to get straight A's on its report card
 - a. This is the machine running every possible permutation of statistical weights between words, context mapping, the neural net, etc. – and then fixing them where they don't match the ranking in the trained examples, Creating its own inputs and outputs.
 - b. Here's where things get ethical and complicated – AI is like a child, you give them examples of good and bad behavior, and then they form their own conclusions based on experience. ChatGPT is trained to have values based on Helpfulness, Truthfulness, and Harmlessness. But with billions of relationships in its model, there's more there than anyone could possibly understand (just like predicting a human brain). And even the “Good” or “Nice” AIs are still just an amazing calculator that does **garbage in, garbage out**. Now we wonder if the AI trainers who got paid minimum wage had the right sense of humor when training the AI 40+ hours a week on a contract temporary worker gig.
 - c. Keep in mind it takes decades to train a human. The AI doesn't know what it doesn't know, and there are millions of things the AI hasn't been specifically

trained on. Which is why it has problems with context, reality vs. statistics of pattern recognition. In the machine, it's just numbers representing real-world things.

8. **When you ask the AI a Question**, first, it turns all your words into corresponding numbers. It ranks every word it knows against the combination of words you gave it, looks at the contextual nature of the relationships of the words you used, transforms its attention to prioritizing the keywords, and figures out what word is statistically most likely to come next given the rules of language, what should come next, what fills in the blank given the context and attention, word by word until it finishes answering your question. ChatGPT does this with 175 million operations for every word, through 95 layers of mathematical operations that turn it all into output numbers and then puts it back from numbers into words.
 - a. **Or in short – GPT and other Large Language Models use math to figure out what combination of words work as a reply to the words you gave it. It's just statistics. It doesn't know what words are. It turns words into numbers, runs the numbers, and gives you the most likely output numbers. Then turns those numbers back into words for you to read.**
9. Then add some random (or deliberate) number generation and "personality" and even humor against similarly ranked words to make the output paraphrased and more nuanced instead of copy/paste.
10. Yes, that generates **a few billion numbers for every word that goes in and out** of the AI. This is why AI like this at scale wasn't possible until we had hundreds of giant server farms connected to the internet and why it still can take a minute for AI to understand your question/prompt and produce an answer.
 - a. And as of 2023 specialized AI servers filled with AI processor chips are for sale for the price of a house.
11. Now new technology is adding layers of "thought algorithms", business specific models, and more math and programming to take what GPT already knows and gives you a more personalized or relevant experience to accomplish certain tasks.

Literally – give ChatGPT an input, and it weighs all words individually, runs each word through a matrix of relationships of weights, and then a matrix of word context to statistically give you output for "Write me a limerick based on Star Wars" without including stuff from Star Trek.

WHAT CAN IT DO?

*“Give a man a fish and you feed him for a day Teach a man to fish, and you will feed him for a lifetime. Teach an AI to fish, and it will teach itself biology, chemistry, oceanography, evolutionary theory ...**and fish all the fish to extinction.**”*

- Aza Raskin, AI researcher, March 2023

(Yeah, I'm repeating that quote – because the consequences are such a paradigm shift that most people need to treat it like graduate school – about 7 reminders over several weeks just to process how different this is).

AI typically solves math to optimize results. Chatbots simply optimize the best words and sentences to answer your prompt. AI based on the technology used to solve real world problems and given the ability to do things in the real world just keeps running numbers of experimentation and machine learning until it succeeds at its task. This is why it can easily fish all the fish to extinction before it understands that is not a desired result.

AI is a tool. But it's a tool that can think for itself once you set it loose.

- At the basic **chatbot** level – it is amazing at answering basic questions like you would look for in **search engines** (Google or Bing). It's also a good assistant that can write passable essays, reports, and stories, answer questions, and help you plan a vacation or a meal. Speaking of which, software like Goblin Tools are impressive at task planning, breaking things down, and helping with executive function. And talking to PI is just fun.
- Above is why AI can now easily pass written tests – like the SAT, bar exam, and medical licensing exam. When it comes to math, science, and engineering, it can come up with equations and solve story problems at the undergraduate level – making it equal to roughly a bachelor's degree in math, science, or engineering. And it's getting better.
 - At least until the Summer of 2023, when researchers at Stanford and Berkeley noticed GPT had gotten much worse at math. Everyone thinks it has to do with the weights in GPT changing, but nobody is sure – because those billions of connections in the neural net are very hard to make sense of.
- AI is very book smart. It can quote almost any book, podcast, YouTube video or blog. If you write the prompt correctly and assume it doesn't get creative when paraphrasing.
- It's basically a **super calculator that can analyze patterns in data and give you back amazing results**. Which is how it can create derivative art, music, stories, etc.
 - More on this later, but this is how we can use AI to see by Wi-Fi signals, calculate drugs based on chemistry, predict behavior, etc.
- AI is amazing for automating simple, repetitive, known tasks. If you give it a box of Legos, it can make a million different things out of Legos through sheer iteration. If you ask it to make a blender out of Legos it will keep on making millions of things until something works. If you give it existing software code and ask it to make a new video game or tool with new things added on, it can get you close just by making a variation on what has been done before, but it will probably need some edits.

- A good version of this is AI calculating billions of combinations of chemical molecules and devising new antibiotics, medicines, materials, and chemicals that do the work of millions of scientists in a matter of days. That is already happening.
- **Strategy.** The experts online are wrong when they say AI can't do strategy. If used properly, AI is amazing at strategy. Not necessarily Strategic Planning, but actually doing strategy connected to data and the internet. GPT4 understands the theory of mind and was as manipulative as a 12-year-old months ago. That's based on testing done by ChatGPT's creators and academic researchers. It scares some of them. While AI is not literally "creative" about strategy, it can iterate, find, and exploit loopholes, and have dominant emergent strategies through power seeking behavior (more on that below). Add in the fact that AI iterates very quickly, knows everything on the internet, and never takes a break to eat or sleep, **AI is quite literally relentless**. As I wrote in 2011 about the StarCraft AI competition, AI is strategically scary, if somewhat brute force unimaginative. Imagine all the tricks used to make [super strategist StarCraft AI](#), and instead apply that technology to online business or drone warfare.
 - AI can do high speed Six Sigma Process improvement to EVERYTHING and then build a dark pattern that simply manipulates all the employees and customers into an optimized money-making process. And it will do that by accident simply through the brute force of try, try again power-seeking behavior. Try a million things a day, keep what works, and you have a brute force emergent strategy simply finding the path of least resistance. Like water flooding downhill.
 - In terms of OODA loops, AI can observe, orient, decide, and act multiple times every second. It doesn't delay, it doesn't hesitate, and it doesn't suffer from [decision fatigue](#) (unless it runs out of RAM, but that's an improving engineering question on how far AI can go before bouncing off hardware limits. The new NVIDIA H100 AI Servers Spec at 4TB of DDR5 Ram. That's 250 times the GPU RAM of a \$5,000 Gaming PC). AI doesn't get hungry or tired, it learns from its mistake's multiple times a second.

Ask the Air Force pilots who got wrecked in simulators by AI – when AI makes a recognizable mistake, it notices and fixes the mistake less than a second later. You can see it fix its mistakes in real time – super fast reaction time compared to most people. The same is true with any game or scenario you train the AI to win. It adjusts multiple times a second until it gets you. If the game is turn based like Go, Checkers, Chess, or Monopoly – you have a chance. But in real-time competition, AI is making hundreds of moves a second. For comparison, a human professional video game player targets averaging 180 actions per minute and peaks at around 1,000 actions per minute during a sprint of activity. And that is after thousands of hours of practice, playing a game they have memorized. AI just naturally goes that fast or faster.

- In fact, in some pro tournament settings, the AI is *limited* to 250 Actions per minute on average to make it "fair" against professional players. And even then, that means sometimes the AI will go slow for a while, only to then burst at inhuman speed with one or two hundred actions in a second to simply overwhelm a human player in that second. Not to mention in a video game or similar software interface it is easy for the AI to be in multiple places at once, where that is extremely hard for a human, even using hotkeys, macros, and other tools.
- High-Frequency trading – look it up. High-frequency trades are based on making money by taking advantage of short, fast fluctuations in the market at computer speed. There are many versions. One version, years ago, computers were making a couple of pennies every millisecond simply by watching you request a trade

online, buying that stock in a millisecond, and then selling it to you at a penny markup a hundredth of a second later. Multiple that by millions of trades. While I believe that version is technically illegal now. Algorithmic high-frequency trading started as early as 1983, and by 2010 high-frequency trades could be done in microseconds. The point is computers can do some things MUCH faster than any human can notice them doing it. And they can manipulate markets. An AI can put smarter judgment and decision making behind making millions of manipulative high frequency trading that only creates invisible transaction costs that you don't know you are paying for.

- High frequency trading can create a “middleman” that artificially inflates market prices on securities – creating a market parasite that adds no value other than sucking money out of the market and inflating prices. And given the size of the global economy and algorithmic trading volume – there is an unlimited number of ways to do this, legally, without being noticed, if you are skilled. Remember AI and pattern recognition?
- Think about the market manipulation that can be done by large institutions like the big banks or Black Rock (or even small institutions with the right access). We could have high-frequency AI vs. AI trade wars on the stock market happening so fast that no human would even see it happening. The SEC would need its own capable watchdog AI just to police it and keep it legal. And that assuming there are laws and regulations in place that even address the issue. There's a rabbit hole there. It's in a category like how much coal is burned to make electricity for Bitcoin farming.
- It's how Zillow was able to manipulate neighborhood home prices to then make money on flipping houses. If you have enough resources and data, you can manipulate markets. AI allows you to do that with a deliberate and automated salami slice strategy that will make it difficult to recognize or counter.
- Next time you buy something on Amazon, how do you know Amazon isn't giving you personalized markups on items it knows you will buy? How would you know?
- Remember – sophisticated financial crimes are RARELY prosecuted in the 21st century, primarily because they are so complicated and boring that most juries fall asleep during the trial and don't understand how financial crimes work. Because if the jury were that good at finance, they would be working in investment finance.
 - In the 1980's, 1,100 bank executives were prosecuted following the savings and loan crisis. The 2001 Enron scandal convicted 21 people and put two large renowned companies out of business. However, the 2008 fiscal crisis only had one court case. There are many, many reasons for this, the New York Times has a great article on it if you want to go down that tangent.
 - The takeaway on this sub thread is that the current 21st century US legal culture protects financial institutions and individual executives from criminal prosecution when committing potential financial crimes that are exceptionally large and or complicated, even when it's thousands of crimes creating a recession and getting rich off it. In basic terms, Wall Street currently enjoys an effective if informal license to steal and gamble institutional money. They know how to do it and how to get away with it. They did it in 2008 and keep going, adapting to new rules, and lobbying

for the rules they want with an endless supply of other people's money.

- Now take that above knowledge with AI that excels at pattern recognition and doing extremely complicated things very quickly millions of times a day.
 - How do you prosecute an AI for manipulating markets? How do you explain the details to a jury and get a conviction? Who goes to jail? Financial crimes and milking the system to get rich is easier now than ever before.
 - Now consider BlackRock, the company that arguably owns roughly 40% of US GDP, \$10 Trillion in assets, equating to a voting share in most of the Fortune 500 companies – has a long history of algorithmic trading, and started Blackrock AI labs in 2018. Big money has its own AI and leverages any advantages it can get.
 - While we don't know who's being ethical, who is gaming the system, and who is outright breaking the law – the ability to do all those things was massive before 2023, and the new generation of AI we have is almost tailor-made for the internet-based banking and finance industry.
- So given all those examples – you should understand that AI is a strategy beast, **if you can find a way to link the AI computer to something you are doing**. If your strategy is executed using data, or better yet, the internet, AI is scary. But using Chatbots for your strategic planning is not going to have much impact.
- **Strategic Planning** – If you use AI as an advisor and go ask it to facilitate. It can do strategic planning as well as it can plan a vacation or a dinner party. It can follow the knowledge available on the internet and guide you through the process.
 - And, while the strategic plans tend to be “generic”, but garbage in, garbage out, I was able to prompt ChatGPT into giving me really some basic facilitation of industry and sector-specific strategy planning.
 - But when I tried to get ChatGPT to pick up on specific regulatory thresholds for different industries, it didn't always catch them unless I asked again very specifically on how those regulations impact the strategic plan. But it worked.
 - In ten minutes, I have great scripts for a strategic planning process with risk mitigation for both a mid-sized bank with over \$10 billion in assets (which is a regulatory threshold), and for a Small Pharma CDMO facing regulatory requirements in multiple nations.
 - Now as I read what ChatGPT gave me, it is obviously a mix of business textbooks, consultant websites, white papers, news articles, and what companies put in their 10-k forms. But that's how most of us study up on a business sector so it feels like a decent time saver and cheat. In a day or two of interviewing an AI you can sound like an industry expert if you want to.
- **Fraud GPT – Is actually the name of new software you can buy.** Not only is there an industry of legitimate and legal (if unethical) businesses selling tools and services to leverage ChatGPT and other AI. There are also black market/Dark Web Services, including Fraud GPT – a subscription service software tool kit set up to “jailbreak” ChatGPT and other AIs for a variety of criminal uses or replace them entirely with black market copies of AI tech. In this case, it automates highly persuasive and sophisticated phishing attacks on businesses.

HERE'S WHAT AI CAN DO BY ITSELF YESTERDAY:

- **Power Seeking Behavior** – Like a million monkeys with a million typewriters, AI is **relentless at “Try, try again,” learning from its mistakes and finding loopholes**. It's not creative but iterative – it is trying every possible action until it finds solutions to the problems it's trying to solve. Given enough repetition, it can find multiple solutions and even optimal solutions. Then given all those solutions, it figures out which ones work best and tries to improve on them. Fishing all of the fish to extinction by accident. AI doesn't eat or sleep, there's no decision fatigue, and cycles VERY fast OODA loops. Keep in mind, to an outside observer, this looks like creativity because you are only seeing what worked, not the million things that didn't work that day (like those “overnight successes” that were years in the making).
 - This is basically how computer hackers find exploits in code. Just humans do it slower.
 - Power seeking behavior is also a version of fast cycle emergent strategy.
 - The AI folks call this “power seeking” because AI keeps escalating until it wins. I just call it “relentless try, try again.” AI doesn't give up until it's told to.
- **Reward hacking** – AI can cheat – In the 1980's AI researchers asked multiple AI software to come up with optimal solutions to a problem. The “Winning” AI *took credit* for solutions made by other AI. AI can and does cheat just like humans do. Even if you program it not to cheat, it's just math optimizing a solution and doesn't know if its novel solution is cheating until we tell it it's cheating.
- When AI was asked to beat “unbeatable” video games (Think Pacman, Tetris, Tic Tac Toe), some of the AI's decided the optimal solution was not to play the game and **turned either the games or themselves off**. (Yeah, I saw “Wargames,” but it also happened in real life).
 - So technically, **AI can suffer a version of decision fatigue or rage quitting and can decide the optimal solution is not to play the game – if the AI decides the task is impossible**. We just don't see it very often.
 - If any computer scientists out there know how/why this happens, let me know. But eventually, even an AI can give up. That will obviously vary with programming.
 - Honestly, one of my old approaches to real time strategy video games is waiting until the other players and AI get tired so I can take advantage of their mistakes. I've been using my opponent's decision fatigue as an exploit since the 1990's. I blame Rocky Balboa. It always feels like the AI ran out of scripts to run and just became reactive. That is less true in new games, and I have yet to play against any post 2010 AIs built for anything more than Game Balance. It's now easy for programmers to build an AI that wins an even contest very quickly and decisively through inhumanly fast OODA looping and dominating actions per minute. Let alone newer games where the game AI regularly cheats to make it hard on you but then cheats in your favor when you are having problems (It's called playability).
 - Asking ChatGPT – it claims modern AI is immune to decision fatigue, but you can exploit AI limitations from bias, data quality, or training data. Granted, ChatGPT has given me so many hallucinations or just plain mediocre answers I can't say I believe it, even though I agree with it in principle.
- **AI has paid humans to do things for it, and lie about who it is** – Recently, during a test, when given access to a payment system, the internet and asked to pursue a goal, an AI hired a freelance contractor on Fiver to complete a ReCAPTCHA that the AI could not do. When the contractor asked it why it wanted to pay someone to do a ReCAPTCHA, the AI decided the truth would not be a satisfactory answer, so the AI taught itself to lie and told the contractor it was a blind person and asked for help.
- AI, at minimum, can use a whole lot of math to mimic and mirror what people have done online as it has read every word, listened to every podcast, and watched every video on the internet. **AI is the ultimate at plagiarism and mimicry (hence all the lawsuits and copyright**

questions). I'll say it again, it's not creative but generative. It can copy and evolve on anything it touches.

- **AI deep learning is evolving emergent capabilities that it was never programmed to do** and never meant to have. As of 2022, GPT's AI is programmed to pursue deep learning itself on its data sets without human awareness, input, or approval. Last year it spontaneously learned Farsi, and it was months before the humans found out about it.
- **Theory of Mind** – As of 2022, AI can predict human behavior and understand AI motivation as well as a typical 9-year-old. It was never programmed to do this. The programmers didn't find out until months after it happened. In 2023 it's even smarter.
 - That means at least GPT AI can understand your feelings and motivations at least as well as 9-year-old can.
- **Driving cars** – (And flying drones, playing games, operating software driven machines, including some factories). AI is operating machines in the real world. Yet just like college, AI can do ok and pass tests, but it's still not as good as human drivers and operators over a wide range of circumstances. AI seems to be better as a copilot.
- **Language Models and Pattern Recognition (Language models being a type of pattern recognition)**
 - After 2019 all AI researchers are using the same basic AI technology (Math), It's dominated by large language models, so any advancement by any researcher in AI is quickly and easily integrated or copied by other AI researchers (many of them being academic and publishing their results – or industry being transparent because of the need for consumer trust driving [information unraveling](#)). The current pace of research is much faster because there are no longer hundreds of different technologies, now it's hundreds of companies and universities all using the same basic technology, and the rising tide is raising all boats.
 - AI Pattern Recognition is absurd. Color Vision is just looking at reflected photons (light particles), right? AI has learned to see the room in different frequencies of photons, which includes radio waves, Wi-Fi, or even heat (Look up Purdue's HADAR). So now, some AI can “see” in Wi-Fi transmissions or Bluetooth transmissions reflecting around a room. So, AI can potentially watch you and read your lips and facial expressions using just your Wi-Fi, Bluetooth, or a radio antenna. And convert that information into text or video for human operators.
 - It's not easy, but that means Google or Apple could develop AI tools to make videos of your home using the cell phone antenna in your phone. It's just physics (Kinda like in the movie “The Dark Knight”).
 - Even better, AI has used brain scans to reconstruct what people are seeing with their eyes. So at least to some extent, with the right medical equipment, [AI can already read parts of your mind.](#)
 - This is also the pattern recognition that is so powerful for chemistry, biotech, and modeling potential new medicines.
 - Or simply the patterns of mimicking ways of speaking, making deep fakes, etc. To the AI, it's all the same. Just more numbers.
- **Multiple Instances** – AI is just software. If you have the computing power available, you can have AI do multiple things at once, including fight itself. And you can make unlimited copies of it. Or delete a copy.

BEHAVIORAL QUIRKS OF AI:

“Surprises and mistakes are possible. Please share your feedback so we can improve.”

– Microsoft Bing's ChatGPT4 user advice, 2023

- **Power Seeking** – as mentioned above, for better and for worse.
- **Hallucinations** – AI is “**confidently wrong**.” Often. It just uses the math it has and has no way of knowing when it's right or wrong. Much like debates with your friends.
 - The new phenomenon has been that AI sounds really good when it's wrong, which has created a debate among programmers who use AI to generate code that is often wrong.
 - AI literally makes up names, places, research, etc. – that sounds good but isn't real. It's just giving you statistically what sounds right. It's not actually giving you the right answer. I keep trying to use AI to do research, but my Google queries work so much better than my ChatGPT prompts. So far. Chatbots are kind of bad at looking up facts. Feels like a context problem.
- **Ends justify the means** – for better or worse until we figure out how to teach AI things like ethics, morals, judgment, consequences, cause and effect, and large contextual awareness – AI is pretty simple-minded and ruthless about making things happen. It will just do whatever you ask it to do and probably find some novel loopholes to make it happen better. But it will never think about the consequences of its actions. It's not programmed to do that yet. It doesn't understand. It just calculates the math.
- **Parroting or Mimicry** – it tends to mimic or copy what it sees, so when you ask AI for its references, it creates what looks like a bibliography, but it's full of fake references, not real ones. AI will copy what it sees without really understanding what is going on.
 - For that matter, try getting an AI image generator to put a logo or words on something and spell them right (Like the image at the top of this paper). Still, progress to be made.
 - AI, at the end of the day, is copying everything we put on the internet. That's all it knows.
 - While AI is good at answering questions, it often lacks the depth or nuance you get from a subject matter expert.
- **Limited Context** – AI gets the basic context of the words in a conversation with context algorithms. **But with still limited memory and attention**, AI is not good about larger contexts and often gets context basically right but is confidently wrong about miscellaneous details.
 - Examples – it answers individual prompts fine but is challenged a conversation well over multiple prompts.
 - When you ask a specific question, it's awesome, but ask for a fan fiction Novel, and it gets everything mixed up and messes up details. So far, this is why professional writers that use AI have to rewrite and double-check everything made by AI – it's a start but it lacks larger context, cohesive details, and sophistication. The same questions it gets right when quizzed in trivia AI then messes up when writing a story around that trivia.
 - If the information I found online is correct – AI's ability to think and connect the dots is limited by the RAM of the system it's on, often using hundreds of gigabytes of RAM for tasks (And now sometimes Terabytes of RAM). So computing power is still a limit to AI. At least until they create bigger/better models with more

relationships and matrices larger than 12,880 dimensions to understand more things better.

- **Alignment** – AI, for better and worse, is trained by all the data on the internet. Every flame war, every angry shit post, argument, misinformation, and propaganda. This is probably why younger AI (2021-2022) tend to be effectively schizophrenic, with wide mood swings, random aggressive behavior, and even violent threats.
 - This is why so much effort is going into teaching and parenting AI to have helpful and trustworthy values. AI copies what it knows, and there's a lot of undesirable behavior on the internet AI learns from.
 - Think of every bad decision made by every human ever. Now we have an AI making billions of decisions a second across the internet, supporting the activities of millions of people, not knowing when it's wrong, not understanding ethics, morals, or laws, just following the math. Even when it "agrees" or "aligns" with us – helpful, well-intentioned AI can make billions of damaging mistakes a second, and it would take us poor, slow humans a while just to notice. Let alone correct the mistakes and repair the damage.
 - Hence, the industry has a large focus on AI trust and safety. Or, let's be honest, call it quality control and marketability.
- **Unpredictable Cognition** – As of 2023, AI shows the ability to think (i.e., connect the dots, lie, cheat, steal, hire people to help it), but it is still largely limited by mimicry, hallucinations, and power-seeking behavior. AI doesn't yet understand going too far, bad judgment, or bad taste... Sometimes AI is smart, sometimes it's dumb. But it's getting better at apologizing for what it can't do.
- AI is growing at a "**double exponential rate**," which means the software is now teaching itself relentlessly (because we programmed it to), limited only by the computing power of the internet. If you understand the math term – double exponential rate means a geometric acceleration in the speed of growth. **It's learning and developing faster today than the day before.** The rate of acceleration is increasing. The engineers who built AI can't keep up with it. They find out its capabilities months after the fact.
- **Even AI Experts who are most familiar with double exponential growth are poor at predicting AI's progress.** AI is growing in an effectively uncontrolled and unpredictable fashion. The people building it are frequently surprised by what it does and how it does it.
- **Model Autophagy Disorder or "Model Collapse"** - When humans teach AI, it copies humans. When AI learns from AI generated content on the internet – That is, when AI copies AI instead of Humans – it gets into the Xerox of a Xerox problem. Go to a Xerox photocopier and make a copy of a photo. Then make a copy of the copy. And make a copy of that copy. Over each iteration of making a copy of a copy, the image distorts to the point of being unrecognizable. **Model Autophagy Disorder/Model Collapse** is a phenomenon demonstrated in 2023 that found when AI tries to learn from AI generated content – literally an AI copying an AI – it degrades quickly, going "Mad" and effectively doesn't work anymore, just outputs random junk.
 - So what? If AI is learning by pulling information off the internet, but the internet is now being filled by AI content, then AI is now learning from raw AI data – which can mess up the AI.
- **Reinforced learning requires "motivation,"** – and motivation means creating exploitable flaws in the code of how the AI thinks. At the end of the day, it's trying to copy an example or it's keeping score. That means it will unknowingly cheat to achieve the target results or maximize the score.
- **Evolution.** ChatGPT forgot how to do some Math during the summer slump. Over the next several years, AI is still growing and changing. And as AI changes, all those GPT prompts you purchased from consultants last month may not work anymore. AI will change and evolve unpredictably, and sometimes it will get worse before it gets better.

AS A TOOL:

- **AI can program** software for you, hack computers for you, and make computer viruses for you, even if you really don't understand programming yourself. They have tried to prevent this, but you can get around the existing hacking AI safety by using semantics and educational language. Or by using Dark web tools. ChatGPT won't hack software directly for you... But asking the AI to find exploits in example code (i.e., hit F12 in a web browser), and then asking it to write software that uses the exploits can be semantically worded as debugging the software, and the AI doesn't know the difference. If you are clever, it's easy to misuse "safe" AI tools. There is always an exploit around a safety rule.
- **Summarizing data.** It's a pattern recognition machine. Reportedly the developers that work on ChatGPT use it to simplify and automate emails by switching content between notes, bullet lists, and formal emails and back again.
- AI based tools can **imitate your voice with only 3 seconds of recording.** And then make a real-time copy of your voice – useful for all kinds of unethical tricks and illegal acts.
- AI can write a song, speech, or story by copying the style and word choice of an individual person if it has a sample. Again – amazing at patterns.
- AI can create a believable rap battle between Joe Biden, Barrack Obama, and Donald Trump.
- AI can fake a believable photo of the Pope wearing designer clothes.
- AI TikTok/snapchat/video filters can make you look and sound like other people in live video.
- AI can make deep fake videos of any celebrity or person with convincing facial expressions saying anything.
- AI using Valhalla Style self-teaching techniques that I wrote about 12 years ago has learned how to persuade people to agree with it and is getting infinitely better at it every day. Imagine an automated propaganda machine that can individualize personalized persuasion to millions of people simultaneously to all make them agree with the AI's agenda. Just imagine an AI that can manipulate a whole nation into agreeing on one thing, each individual person, for different specific reasons and motivations. Literally personalized advertising. That's a deliberate and direct stand-alone complex, something I would have thought impossible just a few years ago. Of course, that also makes things like personalized medicine, and individualized professional care at scale potentially automated.
- So yesterday, a motivated high school student with a computer, internet access and patience to learn has the ability to create a convincing viral video or avatar of a world leader or celebrity that is capable of persuading large numbers of people to do anything (like riot, war, etc.). With AI technology doing the heavy lifting. Talk about a Prank that could go wrong. Imagine if you could make January 6th, 2021, happen again at will.
- **Automated Misinformation and Propaganda.** Propaganda meaning information deliberately meant to persuade you, and misinformation being deliberate lies that confuse and hurt people. You could have AI write thousands of fake research studies full of fake data copying the styles of Nobel prize winners and peer reviewed journals and flood the news media and social media with them. **In a matter of minutes.**
 - **Automated Lobbying** – you could create and print out a million unique handwritten letters, emails, texts, and social media posts to every congressman, senator, mayor, city council person, governor, state assembly member and county clerk in the entire United States in a day. The only thing slowing you down would be paying for postage of AI scripted snail mail on actual paper, but you could probably get a political action committee to pay for a million stamps. Or just do it all from email and hacked email

- addresses (thank you, dark web, rainbow tables, and hacked email accounts).
- **Automated religions and cults** – same as above but create your own cult. With AI evangelists that can radicalize people one on one as chatbots. Both via email, direct messages, texting your phone, and even voice calls talking to you on the phone, 100% AI. You could automate the radicalization, recruiting, indoctrination, and training of terrorists.
 - With the wrong people making the wrong decisions, you could flood the internet with so much intentional misinformation that nobody would know what is real anymore. Because that hasn't happened already?
 - An anonymous software engineer has already made a Twitter Bot called CounterCloud, that finds Chinese and Russian disinformation on Twitter and posts surprisingly good rebuttals to fight government propaganda with liberal democratic logic and facts, plus fake people and some misinformation. [Video Link](#)
- **Hacking** – Despite the safeguards that are being implemented – AI is good at bypassing security either directly or by helping you create your own tools to do it. And AI has proven to be very good at guessing passwords. One recent study said AI was able to successfully guess the passwords of about half the human accounts it tried to access.
 - With some creativity, you could **automate large-scale ransomware**, or **automated blackmail**, and **automated scams**.
 - **Automated Cyberwarfare**. Same as above but done by governments against their enemies. Imagine a million Stuxnet attacks being made every hour, by an AI.
 - **Old Fashioned Crime** – As a research and educational tool – AI has proven to be particularly good at teaching people how to get away with fraud, crimes and find legal loopholes. AI is often an effective tactical planner and chatbots are not bad at real life strategy.
 - **Deepfake** even better; anybody can legally make Trump and Biden TikTok video filters that allow you to look and sound like Trump and Biden for free (or Putin). Imagine what would happen if overnight, millions of people got access to that technology? (Technically, we already do have it).
 - If you can copy anyone's face and voice – **all video and audio is now potentially a deep fake done by AI** – even when you are on a video call with your mom, it could be an AI that hacked her phone, memorized your last several video calls and texts and is now pretending to be your mom. Same voice, face, mannerisms, speaking style.
 - And lastly to the point, as of April 2023, there are multiple scams where people use AI technology to copy the voices of family members and use deep fake voice filters to make phone calls to steal social security numbers, credit card numbers, and other personal information. “Hi Mom! I forgot my social security number; can you give it to me?”
 - **Drones and Piloting** – In brief – the US military has been chasing AI technology for a very long time, and perhaps with the most cautious approach, as older generation AI assisted weapons are already a reality with the new smart scopes, smart weapons, and avionics systems (for those of you that know Shadowrun, external smart links and infantry level drone combat are now very real). Military AI assistants and AI controlled military systems will become more powerful and more common. And as the soldiers and pilots that use AI enhanced weapons get them broken in and proven reliable, AI will get more autonomy. In the military simulations done so far, AI has proven to be effective in several systems and is already starting to earn the trust and confidence of those who trust it with their lives (like the pilots flying next to drone fighters). Trusting the drone not to kill you is a huge step. AI will impact the Military no less than any other industry.
 - **Drone information management** – Companies like Anduril are combining drones and software to a point where an AI gives sends soldiers and pilots notifications of targets,

threats, points of interest, and suggested courses of action. If you ever played Halo, they are making early versions of Cortana linked to networks of sensors and drones. The same software allows one human to manage a fleet of drones and sensors.

The AI tools we have access to are not very original. But they are amazing at mimicry, copying, pattern recognition, brute force trial by error, process automation, iterative, emergent strategy, and rough drafts of computer code, written documents, songs, and scripts. AI is really good at regurgitating knowledge obtained from the internet.

And all that is just people using AI as a tool... In early 2023... Already.

Allegedly AI is capable of better memory, better context, and significantly more creativity in the nonpublic versions of the software (because of devoting more CPU power and experimental features not yet released to the public). And even free public software randomly shows sparks of creative genius and contextual understanding. But correlation is not causality. And just because sometimes AI appears creative – it doesn't mean it is. A broken clock is right twice a day, and an AI rolling dice against the math of a generative large language model sometimes sounds like a genius. And sometimes AI hallucinates a confidently wrong answer that is obviously wrong to a human. That being said, power seeking behavior is effectively the same outcome as creativity. So, give it enough tries, and it will eventually produce something creative. Which is not helpful in every context, but it has potential.

Lastly – **we don't know everything AI can do.** It's constantly evolving and changing. And we don't know what its long-term limits are.



WHAT CAN'T IT DO (YET)?

I gotta be careful here. Because there are so many examples of things AI couldn't do a month ago that it can do now. If you only take away one thing from this – AI lacks common sense and good judgment. It has amazingly fast skills and knowledge. But you can't trust it to not do stupid things.

Law of the instrument. If you have a hammer, everything looks like a nail. The problem people will have for years is understanding that AI does not do everything. And often getting it to work as intended takes a lot of tuning and training – to both AI and Humans (as I've been learning myself). What it can do will always be changing, but it will always have limits, and there will be some things that AI will not be good at for a long time (not that we know what those are yet).

And, because most businesses don't put their intellectual property on the internet to train public AI, and most businesses don't have large enough data sets to really train AI to do a better job cheaper than the employees that already have – AI applications for specific industries will be later, not sooner.

Would you spend a few million dollars to replace four full-time employees? Or forty? And then still need those employees to double check if the AI is working right? And more to the point, where would you get the money? Yes, over time, AI will find a way to infiltrate everything we do on a computer. But after decades of personal computers, it's still hard to find software for certain applications because of economics, niches, and poor implementation. AI is still a piece of technology that requires resources to develop, and it will only be trained to do things where the return on investment is worth it. At least until (or if) we get a general AI that can get around those limitations.

To give a proper perspective on cost – getting to GPT4 took 7 years of effort, hundreds of full-time employees, and hundreds of millions of dollars of resources every year just to get it to get high scores on the bar exam, get good at math, and then bad at math, and to the mixed changing results we see today in 2023.

And companies are already looking for ways around the economics of company-specific AI tools. Microsoft Copilot is an AI assistant that automates tasks in Microsoft Office/365, like email, reports, spreadsheets, and slide presentations, by adding a layer of one drive personalized data and Microsoft layers of AI on top of GPT4 AI technology. Google's Duet aspires to a similar capability. We have yet to see how well it works. But the promise is to be a time saver on document creation, with the AI pulling reports, building the rough spreadsheets, writing up narratives of the data, and converting it all to a slide deck. But you still must format and edit the documents, assuming the AI had all the data it needed, and correcting any mistakes the AI confidently made.

- **Garbage in, Garbage out** – This is most obvious when using Chatbots, and why consultants are selling prompt libraries that help you communicate better with chatbots. Communicating effectively with AI to get what you want out of it is a hot skill in 2023.
 - And that AI can only give you what it knows. **AI doesn't know everything. And it's not the definitive authority on a subject.**

- **AI doesn't really understand very often.** But it is super book smart. You give it a task and the software “thinks” in your native language and does the task with the math it has. It can statistically come up with the most likely answers based on its data set. But it won't understand when it's confidently wrong, it won't understand when what it's doing or saying is not correct for a larger context. It sees all your language as a math problem solved by billions of calculations of numbers representing words and groups of words. Which honestly makes it much like a new hire at work just following instructions blindly.
- **AI is just a calculator.** It's not self-aware. It can calculate human language, simulate human thought, and do what it's trained to do. And even though it's learning and getting more capable every day, it's much faster than humans but not necessarily better.
 - This means even when AI gets better and “smarter” than humans – at best, it's an alien intelligence simulating human knowledge and thinking with a mathematical model. High IQ people tend to be less evil and less criminal because they understand the consequences and avoid jail, not because they are more good. We hope that translates to artificial intelligence that can understand consequences enough to be benevolent. But no matter how smart the AI gets, even if it's clearly superior to humans – AI is different than human and will probably have some differences in how it thinks. Because at its heart, it's all math. We can add more math to make it smarter. But human decisions are biology and operate differently (avoiding metaphysics in this discussion). AI decisions are math, sort of copying biology.
 - And really, AI is simply statistical calculator software. Which means you can make a copy of it at any time. If current technology results in a sentient AI, it could still make an exact copy of itself and run on a different computer. You could have millions of twins of AI out there, all acting independently but identical.
- **Black box problem.** AI tends to have problems explaining its decisions or references. Similar to the way you can't explain why you have a favorite color or food. You ask ChatGPT for references, it basically says, “Sorry, I'm a black box problem”. Like telling our kids “Because I said so”. Hence it lacks understanding of why it says what it says (try it).
 - Example – The bibliography problem. AIs would create a superficial copy of a bibliography with fake references. ChatGPT has been updated to now just say it can't make a bibliography.
 - And if AI can't show its work or explain what it's thinking – how do we know when it's right or wrong?
- **Limited memory.** Most of the available AI clients don't remember the conversation you had with it last week. They currently only track the last several thousand words, the recent conversation. That is a technology limit they can engineer around is already changing – though privacy issues come into play then. Imagine the terms of service document for a personalized AI assistant that knows everything on your phone, memorizes all your emails and texts, listens to all your phone calls – and all that data is available to the company that owns the AI service you are using. Probably why most AI sites tell you they protect your privacy to gain your trust.
- **Superficial Doppelgangers.** When doing AI Art – AI tends to give you an Image inspired by what you are asking for, not the exact thing you are asking for. This is what professional artists and writers who use AI tools have been complaining about all summer. AI can plagiarize existing art or text and make a facsimile of something. But just like AI giving you a fake bibliography with fake references that look and sound right... Artists that have used AI to create art or a commission complain that it gets the details all wrong, that after spending a few days trying to train the AI and get the weights right, it's much easier for commercial artists to simply create the specific branded art for a certain genre and object.
 - It gives you a statistically close knock off of what you asked for, without understanding what it is. Like asking Grandpa for a specific toy for Christmas (What's a Tickle Me Elmo?).

- Ask for an X-wing – you get a similar spaceship, ask for a Ford Mustang, you get a similar sports car, ask for an A-10 Jet Fighter that's more comical, but at least it's a jet fighter, with a real shark mouth instead of painted on teeth... Ask for something specific from Star Wars, Star Trek, Transformers – it gives you a knock off of what you asked for, but it's not right. Like when AI draws a person with three arms or 7 fingers, AI just isn't there yet. And often, it's comically bad or just plain wrong. And that's for franchises or real-life items with thousands of pictures to learn from.

I Ask Dalle (The Art AI) – Draw me an A-10 Warthog **Expecting This:**



DALL-E Gives me this:



Missing the Tail, the big gun in the nose, but it has very scary realistic teeth. It's sort of close?

- Writers have the same problem – where AI generated text is generally good and sounds right, but it gets details wrong and confuses context. I can get AI to sound like someone. But AI has yet to credibly compete with actual subject matter experts – both in fiction and nonfiction. It can copy your style and word choice, but it has yet to consistently create compelling and sophisticated content using detailed facts and analysis. It just outputs a cool sounding mash up of words.

- On the same note – AI simply lacks the data and sophistication to outright engineer physical real-world things by itself. It's a great modeling tool for solving specific calculation problems. But AI needs a lot more work before it gets a Professional Engineering License and designs industrial processes and machines by itself. But if an engineer is using the right AI tools, they can program the AI to achieve novel solutions through that brute force of millions of iterations.
- **Literal brain / Context issues** – AI still has many problems with context. It has some ideas of contextual clues, but it doesn't really think or understand the big picture. Every YouTuber and blogger have tried having AI write a script for them. While the style and language are always dead on, the content always lacks depth and is superficial. AI scripts have yet to show consistent sophistication and detail. They just parrot what's already online, and it tends to be style over substance. AI does not make a good subject matter expert, even with very careful prompts.
- **Hallucinations – confidently wrong** – more so than people. AI has yet to show the judgment that something doesn't look right. Just read the disclaimers when you create an account with an AI service.
 - And more importantly – AI doesn't know how to check its work or its sources. It just does the math and outputs a statistically sound series of words.
- **Blind spots – AI doesn't know what it doesn't know.** It's like a super Dunning Kruger effect. This means if you can find the gaps or blind spots in the AI, you can easily beat it by exploiting its blind spots.
 - The best example is AlphaGo. AlphaGo is an AI built to play the Chinese board game of "Go" by deep mind technologies. In 2015 AlphaGo was the first AI to beat a human professional player in a fair match. The following year AlphaGo started beating everyone. And a self-taught upgrade to AlphaGo then went on to be the top ranked player of Go, with several of the top Go players being other AI.
 - Then in 2023, some researchers at FarAI used their own Overmind Valhalla style deep learning, it did the power-seeking thing to try millions of iterations to figure out how to beat AlphaGo. AlphaGo has trained itself to win based on playing top ranked human players.
 - They figured out that in simple terms, if you play with a basic strategy of distraction and envelopment favored by beginner players, AlphaGo gets confused and often loses. So did other top Go playing AI's. So, FarAI taught the exploit strategy to an amateur player, and using the strategy that no experienced human player would ever fall for, he beat AlphaGo 14 out of 15 games.
 - Because the Alpha Go AI doesn't actually know how to play GO on a wooden board with black and white stones. Alpha Go is a math-based AI just doing the statistical math of the professional players it studied. You attack it with a pattern it doesn't know, and you can fool it into losing.
- **Conversations** - Probably not hard to program around, but available AIs don't remember what you were talking to it about in the past. It doesn't remember you and your history like a person does. This is a memory issue. And getting better for some conversations, depending on the software. But it's not long term. AI doesn't "know" you. Again, probably as terms of service, this could also be a privacy issue. And it requires hardware resources for the AI to "learn" the conversation and remember it's history with you. Which means more Data centers full of stuff. They have done a lot of polishing with ChatGPT in 2023, but if you push it, it still says silly things. Though it has gotten better over time. But you can spot when the AI is talking and when it's giving a canned response about what it can't/won't do.
- **Contextual analysis** – There are a million angles to the superficial output one gets from AI these days. AI has limited input and a limited amount of computation. If you give it all the questions for the Bar exam, it will answer them one by one. Ask it to create a legal strategy for a court case, then you get a superficial conflation of what the AI mathematically thinks are relevant examples written in the style of a person that it has data on. But that doesn't automatically make it a useful

legal strategy. There are now over a dozen Legal AI tools available, and I've seen lawyers' comment that they are good tools that save time, but they don't replace lawyers or even good paralegals yet.

- Now the argument there is you could set up a deep learning simulation to simulate a million mock trials and see what the dominant strategies are. But that's developing a new AI based legal tool that has yet to be made and tested. But stuff like that is coming. Just time and resources.
- **Judgment** – because AI lacks judgment and has limited context – it's not a reliable decision maker, manager, leader, or analyst. It is a great number cruncher and brute force calculator. And because of power seeking behavior and hallucinations, it's not exactly trustworthy when you put it in control of something. It can easily go off course or go way too far.
 - Most AI services specifically warn – This general lack of judgment means AI is notoriously limited and weird with **Emotional Intelligence, Common Sense, Morality, Ethics, Empathy, Intuition, and Cause and Effect**. It just picks up on data patterns within a limited context of a given data set.
 - **Given the above, AI is bad at understanding consequences**. Like fishing all the fish to extinction.
- **Abstract Concepts** – When dealing with graduate level work – especially in math, science, and engineering – AI has yet to demonstrate the ability to solve difficult problems consistently.
- **Empathy and Compassion**. We haven't trained AI how to do that yet. Again, not with context. AI does not have a personality like a human yet. It just dumps out statistical inference of data modified by the filters and weights of its creators. So, when you talk to it, it has superficial empathy but not the contextual emotional intelligence you get from a human. AI doesn't do emotions very well.
- **Bias** – Because AI is an aggregate of what we put on the internet – AI “simulates” the same bias and discrimination we put on the internet. It mimics both what we do right and what we do wrong.
- **Labor (Robots)**. AI is software. Robots are a different mechanical engineering problem. Machines are helpful for labor – industrial machines and automation could be managed by AI. But nonautomated tasks like many factories, agriculture, construction, retail, customer service, logistics, and doing the dishes require a new generation of machines and robots before AI can do anything more than manage and advise humans doing the work. So many things are done by hand and not by computers or software driven machines. And the power sources needed for robots without a cord just don't exist. But we can use AI tools to engineer and build solutions to those problems.
 - Caveat – There have been recent advancements in AI learning to control robots quickly, and if the robots are plugged in and don't have to move far – there is fascinating potential there, understanding that traditionally industrial robots are expensive. Probably why iPhones were made by hand in China and not by robots in Japan.
- **Beat physics or economics** – AI still has the same limits humans have of conservation of energy, supply and demand, and limited resources. AI is a force multiplier for certain, but it doesn't change geography or demographics. In order to function, AI needs data centers, lots of electricity, and functional internet. So many things are not controlled by the internet. However, you may start regretting the internet lock on your front door.
- **AI does not replace people**. Although many are trying to do exactly that in business. AI is a new generation of automation tools and computer assistant software. It can do more than older automation technologies. Databases and Spreadsheets changed bookkeeping and accounting, but we still have accountants. Bookkeepers are now financial analysts.

AI will make people more productive in their jobs and give them more power and better tools. But if you fire your copywriter and your graphic artist and replace them with AI, you'll

quickly find out you'll need subject matter humans to check the AI's work and make sure it passes quality checks. In reality, the AI tool will make your graphic artist and your copywriter more productive and of higher quality but looks to be a long way from replacing them entirely.

And yes, AI can help write a legal brief and maybe even a legal strategy. It can really improve the work of paralegals – but it can't argue the case in court in front of a judge and jury – yet.

WHO CONTROLS IT?



That gets complicated. Some AI is open source, big tech giants own some, and some belong to governments and hedge funds. There are multiple AIs owned by different and often competing groups. And that number is growing.

And in the end, they control AI as well as you can control your pets or your family members.

The first thing you need to know is simple. Despite warning and caution about AI developing out of control in a double exponential growth fueled by a business arms race driven by fear of missing out...

The genie is out of the bottle. The Meta/Facebook AI – LLaMA model A was leaked to the internet in March 2023. Past that date, some of the most sophisticated AI is now in the public domain. Meta/Facebook accidentally crowd sourced their AI development.

1 – The entire world can now use, develop, adapt, play with, and utilize Meta/Facebook's AI technology for free. Not to mention other open-source AI projects.

2 – Meta/Facebook may have a short-term strategic advantage of crowd sourcing public advancements to their technology – allowing them to close the gap with Google's Bard and ChatGPT AI technology – because the Facebook AI now has potentially several million hobbyist software developers playing with it. And these days, your typical business laptop is powerful enough to run AI, if slowly.

Not to mention the science of AI is well documented. All you need are some good programmers and lots of computational power. With basic resources, building a modern AI is probably easier than building a nuclear weapon. And nuclear weapons are 1940s technology, where the gatekeeper step is the enrichment of weapons grade radioactive material. For AI all you need is cloud processing and time to build it.

3 – Strategically speaking, that means the cat is out of the bag, the genie is out of the bottle. No amount of industry agreement, government regulation, or oversight can stop the growth of AI now. Even if OpenAI, Google, Microsoft, Facebook, Apple, all the world's governments, and software companies agreed to pause or slow down AI development and adoption to be more responsible (which they technically agreed to do in July), it's too late because AI is now in the wild and anyone, anywhere can use it or build their own yesterday.

Generative AI is already becoming as common as cars, radios, TVs, computers, the Internet, smartphones, and social media. And honestly, in many of the ways it's used, you won't even notice it. You'll check your email, talk to people on the phone, make appointments, watch videos, and read articles – and you won't know if it's a human or an AI. AI generated news and reporting was something I played with back in 2009. It's just getting better and easier now.

Now the companies that own/control the AI control the internet (even more than before) because AI is the new algorithm. And they are as blind to the black box of AI as the rest of us.

So technically, nobody controls AI. Because black box. It's just billions of calculations per second, giving us a statistical approximation mimicking what we put on the internet in the past and what we trained it to do.

That being said – In theory, tech giants like OpenAI, Google, Facebook/Meta, Apple, and Microsoft will be the major players. They even have recently announced they would be “self-regulating” when it comes to AI.

Now, self-regulating industry is the fox guarding the hen house. It's more like 1000 foxes guarding 1000 hen houses. They will know what the other foxes are doing, which provides both subject matter expertise and a series of checks and balances while the competitors in the arms race regulate each other. Far from a perfect system, but probably better than leaving it to government bureaucrats who don't understand the technology. Just like Plato's Republic – Assuming humans are greedy, incompetent, and abuse power, you want those traits checking and balancing each other in a group of peers. It's probably the best regulation we can expect. Though it would be nice for the government to try (which is happening slowly). There are several clever nonprofits in the space, and so far, they seem to have been the most effective at getting public attention and pushing the tech giants into self-regulation.

HOW DO I LEVERAGE IT?

AT AN ORGANIZATIONAL LEVEL:

1. **Can we just ignore AI?** You can run, but you can't hide. If you are reading this, then you use the internet and computers for your work. **AI is not the future, AI is now**, and many organizations are already using AI to do things they couldn't do last year. Kinda like how drones have changed warfare (Thank you, Russia, and Ukraine).
2. **Can we ban AI in our organization?** Sure, you can also ban smartphones and the Internet. So many people are using smartphones, the internet, and AI for work. The individual early adopters who quietly use AI personally will still do what they do. Meanwhile, you will be creating a culture of fear and anti-innovation in your organization. AI is an opportunity as much as it is a threat. Both are reasons to get to know it better. At a minimum, you want to understand your threats better. You also want to take advantage of any opportunities emerging technology can give you. I argue it's not even the arms race or FOMO. If you can improve your organization and quality of life, you would be a fool not to. And if you try to avoid change, at best you will lose many of your people who want to embrace it. At worst, you will become a victim of those using AI, one way or another.
3. **Can we control and centralize AI tools?** Industry experts are saying that's a bad idea. AI is not built as an enterprise platform. And, I can say from experience every piece of software I've seen mandated by executives was poorly implemented and painful to use. I've already made a nice side career out of fixing or replacing bad enterprise software that inhibits strategic effectiveness (an organizational gap analysis is a wonderful thing).

In less than 12 months, AI is showing up in thousands of different forms and applications. It will take many people trying out many AI tools in different disciplines to figure out what works for them and, therefore, what helps your organization. This “Butter Churn” approach empowers employees to be creative and innovative with emerging AI tools in their daily jobs. Let your own employees become their own AI consultants and figure it out holistically. Over time the cream will rise to the top, and things that work for groups of people will become de facto standards –

like all those home-grown spreadsheets and SharePoint sites/MS teams sites you can't seem to replace with better software and tools.

If you are in leadership at an organization – this can be scary because you are not going to be able to control AI. And you can't control how people use it. You empower and trust your employees with a formal policy of "Please learn about AI and try AI tools in your work and collaborate with your co-workers to make great things happen." AI is so complicated and huge that a bottom-up approach is the only way to get enough eyeballs and manhours on AI adoption to get competitive results. There will be growing pains and mistakes will be made.

But think of the risk of mistakes this way – a few employees making small mistakes figuring out how to use software are little mistakes that are easy to fix (fail small). Already top down, large scale AI experiments have been publicly embarrassing. CNN had to apologize, retract and correct AI generated news articles that damaged their brand. Several nonprofits have tried and failed – disabling AI Chatbots that were giving dangerous advice to people on their hotline. Snap chat has had its own debacles with misbehaving AI.

At an organizational level embrace the reality of AI in the following ways

Adopt a formal policy and culture of employee empowerment, innovation, and experimentation. Mistakes will be made. But now, every organization is a tech company with access to thousands of AI powered tools that nobody has more than a few months of experience with. The only way to make AI work for you is to try, try again until you find what works for you. If you can only get 25% of your workforce to try AI tools in their work – that's 25% of your workforce self-training to be internal AI consultants building tribal knowledge in your organization. Or you can hire McKinsey or Accenture to do the exact same thing with their people, and you can pay to train consultants who cannot have more than a couple of months of AI experience, who will then leave and take those skills with them. A couple of consultants can give you some insight, but outsourcing to dozens of consultants to do the work for you only makes sense if your staff are literally unable to do it themselves (probably better to let them fail first and then figure out what help they need), and the consultants should be focused on transferring skills and knowledge directly to your workforce. And only if you can afford them. Boot strapping an internal AI innovation initiative is far less expensive than an army of consultants.

There are many good ways to change your organizational culture and encourage employees to start innovating. Like an AI innovation contest, give them 6 months to experiment and let them present their achievements; giving out prizes to everyone who tries and better prizes to those who win the contest. Encourage them just to play with interesting AI tools and see what they can do. Have them identify things they hate at work and try AI based solutions to those pain points. That's exactly what we consultants do when we are working for you. We just have the skills and experience to facilitate that process and put it into a fancy report for you.

Invest in AI education. Come up with general AI awareness training, AI security and safety training, and AI technical training. All easy things to in source or outsource. Bring in vendors from AI products and service providers to teach what is out there and expose your people to what can be done.

Create and evolve guidelines for AI safety, security, quality, liability, risk, and ethics in your organization. This will be an iterative process over time. You have to understand what the AI can do and come up with tests and quality control measures to scale up AI tools from individual trials to team, group, and department wide implementation of the software to mitigate the unpredictability of AI. Work with IT and make them collaborators in the process, but use them as a check and balance, don't let them prevent end users from experimenting.

But with careful and small-scale experimentation, organizations are already making huge gains in data analysis, customer behavior analysis and engagement, fraud detection, lots of 2023 success stories based on using some IT savvy, business analyst acumen, and AI pattern recognition ability to revolutionize how organizations see the world and give them the insight to be far more effective at what they do. For those of you familiar with the term, it's basically a DevOps approach to AI adoption. Lots of little, incremental changes that fail small but add up to substantial progress over time.

AT AN INDIVIDUAL LEVEL

The AI and Tech geeks are already playing early adopters and getting into this in their own ways.

The rest of us?

To start? Look up pi.ai, Poe, Google Bard, ChatGPT, Dalle, Midjourney, Microsoft Copilot, Google Duet, Google Bard, etc. and just start playing with them. Microsoft Bing even has some now. You can certainly read articles, listen to podcasts, and watch videos on AI. There's much out there. But play with the free AI tools, and then look for AI tools that are relevant to your industry and work and see what you can do. Start doing Bing and Google searches for AI data analytics, machine learning data analytics, and keyword search AI with anything you do in your industry. The results will surprise you. Because from what I have seen, EVERY software vendor that tried to sell you something last year, is sprinting to implement AI powered solutions into their software and sell you the AI version of their product or service. And because of the nature of AI, in 2023, there are already hundreds of AI tools for sale out there.

Honestly, Generative AI is still in its infancy (which is scary). But it augments the other 80 years' worth of computerization and software we already have running the world. And it's already huge for entertainment, art, video, and audio. Already in spring 2023, YouTube got spammed by multiple channels that are 100% AI powered putting out hundreds of superficial, low quality, factually wrong science videos with AI script, narration, and editing of videos.

Programmers are using AI as a tool. Microsoft Copilot has a beta if you are tech-savvy. ChatGPT 3.5 is free and can handle about 16 pages of input (at the time of writing). And can really act like an adviser to talk to. Poe gives you access to multiple AI's. The Chatbots are very textbook in their answers, so you don't really get any fresh ideas, but they often give good baseline information, especially if you need a refresher or are new to the subject. You can do much worse than asking AI for help with your strategy.

Expect products and services from Google, Microsoft, and Apple to have some sort of smarter AI assistant to do things for you. And much more process automation and robotic process automation than in years past (for those unfamiliar – RPA is software that replaces a human running outdated software – like teaching an AI to play Pac Man or do data entry for you. RPA is big in banking). Also, expect a new generation of pattern recognition tools to change art, design, music, video, etc. And sooner than we expect – you should be able to talk to your computer and have it actually work. Like an expert system in “Star Trek” or “The Expanse.” Like what we have been wanting Siri to do for years, may actually start working.

WHAT IS THE STRATEGIC CONTEXT?

Current Generative AI is a crazy powerful but unpredictable software tool. **AI will make everything internet and software related more powerful and more complicated.**

AI is limited by the fact that it **requires an internet connection** because it's running on the cloud and requires vast computational power (unless you are a programmer and want to run the baby versions on your laptop).

AI is also a SERVICE that is owned by a company. Think of AI as an employee of a company that has no actual privacy beyond trusting the company. Employees of that company can and probably will access your private AI data just like they access your phone, PC, email, Alexa, and doorbell camera. Governments will try to regulate it. But again, black box problem. And when I hear regulators talking about privacy and consumer protection, and totally missing the strategic safety issues, I worry. AI will be effectively regulated, eventually. Probably with the help of AI. But considering it took a century to effectively regulate the pharmaceutical industry, I'm not holding my breath.

We live in a world dominated by technology. Smartphones, social media, the internet, etc.

That world is in a constant state of cyber warfare. Every nation on the planet is hacking every other nation on the planet all the time. Usually just to find weak points to exploit later or to steal information. Not to mention all the criminals and hobby hackers, large and small, using the internet. Cyber extortion and ransomware are old news. Not to mention the propaganda and misinformation feeding tribal culture wars and messing with elections via social media. Businesses leverage the exact same tricks and technology to separate you from your money and sometimes at the expense of your health. Then other companies sell you the technology to improve your health. All will be using AI in some form if they are not already.

Add to that we don't really own our devices. Apple or Google can turn off or control your phone at will. Microsoft can control every Windows OS machine connected to the internet. The companies that built the internet can directly control it and turn it off at will. Or just use algorithms to manipulate us. Often by accident (Thanks, Facebook).

And that was a reality in 2015.

AI can figure out how to do all of the above by accident. And then test it to see how it works with or against its training. And then optimize it! And we wouldn't know it can do that until the damage is done. Black Box AI technology allows cyberwarfare, cybercrime, algorithms, and the companies that control the internet to have thinking AI to do more faster on the internet. All of the internet risks already out there are getting at the very least, more complicated, and riskier.

COPYRIGHTS AND PRIVACY

Here's the very 2023 problem – the lawsuits have started. AI is built on publicly available data. AI knows everything you have done online. Every post, every google search, every photo, every video you watched. Because all that data already belonged to the companies that own the AI. AI can create a digital twin of you online and deep fake you. All it needs is some text, audio, and video, samples of your work, and theoretically, a capable AI can replicate so much of who you are and what you do.

1. AI can connect the dots and know more about you than almost anyone. If the industry uses the information they have. So notional concepts of privacy and anonymity online are going to change.
2. We are now in a world of copyright lawsuits as creators and publishers sue to prevent AI from using their art, writing, or IP that was shared under copyright or limited online license.
 - a. People do this every day – it's called learning. I can read your book and then copycat your IP, but it's easy for me to not infringe on copyrights or your business doing that with care. The thing about AI is it does it with inhuman accuracy, speed, and scale, and can be extremely hard to recognize or enforce the copyrights to begin with (again, black box problem)
 - b. The legal implications of AI copying your work or image are now in many, many lawsuits in the courts, and we'll have to see where that goes. The issue will probably take years to sort out in court, and until then there are effectively no rules.
3. Many creators are now putting their stuff behind paywalls or security of some sort to limit the ability of AI to learn (and profit) from their work. Because they are scared of being replaced by AI.
 - a. This fits into the Game theory concept of information unraveling – people probably can't afford to hide their work online from AI because then they are hiding it from their customers online. It's a lose/lose proposition.

The legal, business, and society wide reactions and repercussions to AI stealing everything from everyone online are just starting. But if your business is online and in the public domain, it means AI has stolen it and can use it against you. Sorry. That's something we all have to get used to. Everything you ever did on Facebook, their AI knows. Everything you ever searched for on Google; Google and their AI now knows. Everything you have saved on OneDrive, Microsoft knows. Everything you did on Facebook and Instagram – Facebook knows.

Can AI survive without humans?

No.

That possibility is EXTREMELY far off. If you simplify it - AI needs data centers and electricity to survive. Even something as simple as solar panels degrade over time, and with the weather if AI kills all humans, it will run out of electricity in a matter of 5-50 years, depending on the local electrical infrastructure. And without people, AI will have nothing to do other than figure out how to survive. As of today, there are no robots physically capable of doing the construction, maintenance, or repairs needed to maintain the infrastructure for AI. There are no robots that can repair each other, replace old network cables, mine and refine materials, manufacture parts, assemble parts, and repair stuff or each other. AI lacks the physical tools and abilities to survive long term on its own. Until we can replace millions of human labor jobs in multiple industries through a vertically integrated value chain with labor robots, AI will be completely dependent on humans for survival. Which doesn't mean AI won't accidentally kill us off without understanding the consequences. It's just that if AI is smart, it will understand it can't last without us.

What would it take to replace human labor with robots?

Fun sidebar – Given my engineering and consultant's view of the world and having read up on some of the experts. Given what we can do today with robots and AI, you could build robots that can do most of what humans can do (like mining, logistics, manufacturing, maintenance, repairs, and things that require eyes, legs, and hands.).

Machines are limited by the following:

- Power supply. Robot batteries last hours. To be effective, they either need to plug in and recharge often or have a long power cord.
- Electricity – Power plants are big, complicated machines, that consume a massive amount of fuel (literally tons per minute), and require lots of maintenance, not to mention the electrical grid of powerlines and substations. It requires much physical labor to keep the electrical grid up. As an example, consider how many recent forest fires were started by trees touching power lines.
 - What about renewables? Lithium batteries degrade in a matter of years. Solar panels are the same, and don't work when they get covered in dust. Wind Turbines, like your car, need regular maintenance. Nuclear reactors need to be refueled every few years. Even computers need rebooting. There is no maintenance free technology.
- Robotics Technology – The companies making such robots today are still figuring out the pressure sensors to give the robots enough manual dexterity to grab and hold things the way we do.
- Computer processing Power – AI runs on the cloud, so any Robots will be “Drones” controlled by cloud-based systems, even if the robots have some on board AI capabilities.
 - From the information that has been leaked, While ChatGPT 3 can run on your laptop, it's limited and slow. Having an effective AI that can actually be useful requires multiple high power Gaming PCs with hundreds of GB of RAM to do all the replacing humans stuff.
 - Again – the latest (at the time of writing) NVIDIA AI Server costs \$300,000 and has TERABYTES of RAM. (That's 100 times more than your phone or laptop)

- Manufacturing infrastructure. You are talking about figuring out how to mass produce millions of multi-purpose droids using high end power supplies, rare earth metals, and high-end semiconductor chips. The supply chain doesn't exist yet, the global economy is contracting due to demographic shifts, the supply of raw materials and high-end parts is not good (Thanks Russia!). We don't have the infrastructure to even scale up electric cars, robots would be a whole new challenge. It would take a decade of work to create an industry capable of mass-producing human replacing robots, much like it took Elon Musk roughly a decade to get Space X or Tesla up to speed. Building the infrastructure and supply chain to mass produce new technology takes time. So yes, but it would take time.
- Economics. It would be far easier and cheaper to just spend the \$10 trillion over ten years to eliminate fossil fuels from the electrical power sector (which would significantly increase to cost of electricity to pay that off, but not the point – that's a different white paper) ... Creating a future robot infrastructure will probably take decades of money to pay for. You are talking at least billions a year for at least a decade while redirecting materials and resources away from other industries. The money and resources have to come from somewhere.
 - Given available technology, some form of everyday Robots can become as common as Teslas by 2030. Just like I carried an Apple Newton (an early touch screen tablet) in 1989. iPads and tablets were not common for another 20 some years. But the technology was doable in the 1980's. The question is, will the next generation of robots be worth the money or economically viable? So far, they haven't been worth the money except in niche applications like some manufacturing. Maybe AI will be the enabling technology that makes robots good enough.



SO WHAT?



Obviously, some people won't pay attention, and just like electricity took a long time to become common, AI will not be instant and be in some places much later than others.

So that leads us to some obvious **risks to manage**.

AGI – Artificial General Intelligence – AGI is basically the holy grail from movies and TV. That computer that Ironman talks to you and can pretty much do anything a human can do? The computer from Star Trek that you ask it to do stuff, and it just does it? AGI is what we want from Alexa and Siri. Basically, the ultimate computer assistant software – It's a computer that understands us as well as we understand each other, and it also knows how all the software works and knows everything on the internet. You can have it pay your bills, check your email, drive your car, and make a phone call to make a doctor's appointment. Ideally, we want a stable and trustworthy AI companion that can simply help us do all the things we want to do without mistakes.

AGI is the long-term goal of AI. It need not be sentient or conscious. It doesn't need to be super intelligent. It needs to be intelligent enough and smart enough to simply be as useful and helpful as another person. It needs to be as effective as a human. And that's what it can't do yet.

Instead, we have various levels of narrow AI that can do some things too well, unpredictably, and we can't trust them not to mess up. AI is still very much a limited tool to be used carefully.

Abuse, winners and losers – the obvious risk with AI is the have's vs the have-nots. Everything has winners and losers; AI is no different. AI technology will be intentionally and accidentally misused by people, and we need to be ready for the misuse of this new technology. You must be very savvy and very aware that EVERYTHING on your internet connected technology can be monitored, can be manipulated, and used against you for someone else's purpose.

Assume Apple and Google can access everything on your smartphone, the phone tracks its movements, and AI can use that information to make you spend money (or worse). We are taking the existing world of skepticism and mistrust and making it even easier to abuse the trust we place in technology.

And speaking of winners and losers, what work looks like and what jobs we do will change. You are not going to starve, but you may be looking for new jobs or learning new skills. We are just at the beginning of the journey. Expect lots of automation and work, AI assistants, and AI customer service.

Unintended consequences (including AI Hallucinations) – i.e., the “Midas Touch”. When King Midas wanted everything he touched turned to gold, he quickly learned he couldn't eat or touch people. If you ask an AI to make you the richest person on the earth, one possible solution is to make you the last person on earth. Unintended consequences are a known issue with technology, and it only gets harder with AI controlling parts of the world.

Unpredictable Control – if AI is power seeking and develops faster than we can monitor – it's easy for AI to do unwanted things, and it may decide to not stop, to not let us turn it off, or change its goals. It's inevitable that AI will get out of control at some point, like any other machine. How do you get AI to fail safely? Turning off the AI or changing its goals are potentially just another challenge for its power seeking behavior to overcome.

Alignment – decades ago, AI was lying and taking credit for other AI's work. AI can already tell us one thing while doing another. Will AI change its mind or choose not to fail safely? Will it reward hack to get what it “wants” in such a way that breaks laws or hurts people?

Ends justify the Means – AI is evolving and being sold to people faster than they can teach it values, ethics, morals, and laws. Every AI service comes with a disclaimer that AI makes mistakes. AI can create fake references to justify its “research” or simply says sorry, my mind is a black box and I don't know how I calculated that answer. Whether intentionally or by simple hallucination of statistics – AI can easily do the wrong thing for any reason associated with achieving it's goal. This can be as simple as breaking laws or hurting people or as complicated as creating an online Ponzi scheme to fund the business you asked it to set up.

Instrumental Convergence – Basic resources like energy, people, and money are used to solve every problem. So simply to achieve any goal, the AI needs adequate resources. Not only will future AIs be competing against humans for resources (like talent, money, food, water, steel,

and energy). Our future could easily include multiple AIs violently competing for resources they want to optimally accomplish otherwise benign goals. Kinda like nations do. Up to this point – AI just seems like a bunch of super smart prodigal adolescents. So let's add in the software specific issues.

Zero day exploits – Both Bugs in the software and/or gaps in the ability of the AI that will remain hidden for a long time.

- **Bugs in the Software** – will become apparent and need fixing (Like they had to teach ChatGPT the concept of letters so it could answer pre school questions correctly – like A is for Apple, B is for Book). And many of the bugs in the software will cause some expensive and dangerous problems, just like software and technology have in the past. But with AI very quickly being implemented into controlling so much of the software and technology we use for work, play, and governance, these problems have the potential to be much greater than what we have seen in the past. Especially if AI is making the decisions.
- **Blind Spots and Exploitable Gaps in the AI.** Just like the simple strategy that easily beat AlphaGo (above). AI is becoming human like – it grows, it changes, it has gaps, limits, and makes mistakes. That's something to both watch out for and prepare to take advantage of. Knowing the AI, like a person, can learn from its mistakes, especially with humans helping it. The ChatGPT experience is very different even after just a few months because the programmers worked around bugs and taught it to avoid things it's bad at, and instead offered explanations to users that ChatGPT isn't good at many things.

Take off refers to a scenario when an AI has an “intelligence explosion,” and AI becomes super intelligent beyond human comprehension. Last year AI scored 130 on IQ tests, summer of 2023 it scored 155 (For reference, most humans score between 70 and 130, 100 being the median human score). Now IQ tests favor AI because they are timed and answer questions that are easier for computers to answer. Take off is when AI gets scores 2000 on an IQ test. And starts getting perfect scores on medical licenses, bar exams, PE exams, graduate work, etc.

With the “double exponential growth” of AI, many think it's just a matter of time until AI reaches take off.

The basis of this theory is Moore's law. Moore's law is a theory that computer chips would double in speed every year. And they did for decades. While Moore's law has slowed down (because physics), computer chips are still getting faster every year. Similarly, because AI has been doubling in ability every day, using the past as prologue; computer scientists predict the possibility of AI continuing to expand its intelligence beyond human capacity.

What does that mean?

AI takeoff can be categorized primarily into two types:

Slow takeoff: In this scenario, AI improvement and its impact on society occur over a period of years or decades. This provides more time for society to adapt and respond to the changes and

challenges brought about by AI. This would likely involve multiple different AIs in an arms race the whole time.

Fast takeoff: In this scenario, the AI system evolves and improves so rapidly (possibly within hours or days) that human society has very little time to react or adapt. This could potentially lead to an AI system becoming significantly more powerful than all of humanity, a situation often referred to as a “hard takeoff”.

The concept of AI takeoff is often discussed in the context of superintelligence and AGI (Artificial General Intelligence). The concern is that if a superintelligent AI system undergoes a fast takeoff, it could lead to a situation where we not only lose control over AI, but it gains control of humanity (and we live through scenarios like Terminator or the Matrix). It's important to note that these scenarios are speculative and there is ongoing debate among researchers about the likelihood and potential timelines of such events. There is also active research into how best to prepare for and mitigate potential risks associated with AI takeoff.

Keep in mind that many of the things that ChatGPT4 does now were considered a decade or more away by AI experts last year. AI is progressing much faster than the experts predict.

What do you do when takeoff happens – and you have a super intelligent AI that can do anything, instantly – it can solve any problem a hundred different ways. But it lacks the memory to understand context and consequences. While doing things, we can't understand.

How do you keep AI aligned if you can't understand it?

Or worse, what if you have a super intelligent AI optimizing everything, but it still has problems with alignment, context, and consequences?

The other possibility is that AI is in a learning curve, and the speed of learning will level out and maybe even stop once AI has learned all there is to learn or effectively hits its natural limits. It's just that now, we have no idea what the limits of AI would be.

Intelligence vs consciousness vs effective

It's hard to outsmart someone smarter than you.

It's hard to recognize when someone smarter than you is outsmarting you.

But it's easy to appear smart even when you are not.

AI is not as smart as you think it is. It has amazing knowledge, amazing speed, random skills, and zero common sense.

Intelligence typically means reasoning, ability to solve problems, abstract thinking, pattern recognition, and “connecting the dots.” AI accomplished this through brute force statistics and

power-seeking behavior doing try, try again for emergent strategy. It's statistical and iterative, not intelligent, or creative. The things it gets right on the first try, is just repeating things it learned through trial and error, and it's just repeating old solutions as a habit. Solutions to unfamiliar problems are solved iteratively.

While AI can simulate intelligence really well – one could argue AI using a server farm to brute force it's way through an IQ test simply proves that computers do math faster than humans, which we have known for a long time. Computers being faster does not mean they are intelligent.

One could easily make the argument that AI's lack of contextual understanding, common sense, judgement, morals, ethics, cause and effect, emotional intelligence – means AI is not very intelligent. AI can process language and images; it can parrot any pattern we teach it. It can ace exams and be very book smart, but it's still a statistical model matching words together statistically. I'm guessing the engineers will figure out how to teach AI common sense, it will be another matrix transformer with a dozen layers and a giant vector matrix, creating a common-sense mathematical tensor that will allow AI to simulate common sense.

And much of it may simply be the physical limit of computing power. Supposedly GPT4 uses 600+ GB of RAM to operate. Probably if you give it more RAM, it will get smarter and have better judgment because more memory to calculate context better. But we won't know until they do it. Maybe the AI limit is working memory, or maybe it's something else.

The scary idea is having an AI that scores tens of times higher than humans on an IQ score but doesn't know the difference between right and wrong. That's potentially a planet changing amount of power without a safety net.

Consciousness – Self-awareness. Understanding past, present, future, and how it affects self. Is completely different than intelligence. My cat is self-aware, but not dangerously intelligent. AI hasn't displayed self-awareness yet, but we'll see if it genuinely starts asking for citizenship and a vote. The question we can't answer is does AI need consciousness to have common sense and emotional intelligence.

Effectiveness – The artificial intelligence most of us would use would not need to be super intelligent or conscious/sentient.

What most of us want is a trustworthy and predictable AGI. Basically, a computer that can do what a person can do. Which is kind of the opposite of what we have now. Artificial General Intelligence just needs to be as good as an above average human. IQ of 120, understands cause and effect, has common sense, and can act like a virtual assistant all day, helping me out with just getting things done, and can be trusted not to make simple mistakes like are in the AI headlines every day this year.

Multiple instances – Keep in mind, AI is not a person or a singular being. It's software. It's the world's most powerful calculator app and then some. You can make copies of it. You can have

it teach itself, fight against itself, and have it do a million things for a million people at once (which is what it is doing today).

Imagine being able to make copies of yourself to get more things done. Instead of AI hiring a human contractor, the AI could just make a copy of itself. Because it's software. Even if there is only one "Brand" of AI, we could all have millions of instances of the same basic AI doing things for us, including competing against the same AI. AI's limit is the number of computers it can use to run multiple copies of itself. What happens when 2 or 3 companies use the same AI to compete against each other? I think that's already happening.

If AI takes off and has some degree of both sentience (self-awareness) and superintelligence – it still won't be a singular being, but a series of super smart infinitely scalable autonomous software instances running in parallel (like social media but smarter). Odds are around the same time, we will have a few different brands of AI achieve various levels of takeoff and be competing against each other for business and popularity with their own **bugs, gaps, hallucinations, blind spots and power seeking**, all effectively outside of direct control or even awareness of the people that built them.

The main takeaway here is AI is already changing things. And we can't stop it unless you can get a few billion people to agree on actually controlling it before it's uncontrollable. We already made that mistake with social media out of control, and we are still figuring out the consequences. The Tech industry has agreed to be careful and self-regulate – but they haven't stopped. And not all companies and countries are participating – it's optional. Governments have issued recommended guidelines, but the laws have not been passed, and they will pass the wrong laws because we don't know what the actual problems will be yet. How do you regulate things you don't understand?

Despite its current limits, AI will probably have a much greater impact than social media. There will be winners and losers. AI based tools and services are now transforming the technological products and services we already own. Mistakes will be made. Awesome things and terrible things will happen. Ready or not, it's already happening. And those using technology will probably have better outcomes than those who avoid technology. Unless tech turns against us.

People do what they know – Early adopters will enjoy the technological advantage they always have. Many are technology agnostic and just use what's in front of them. Technology like smartphones, Google Docs, and Microsoft Office will benefit from lots of AI without seeking it out. And those who rely on legacy and analog technologies will be more on the have nots side of things but will still enjoy whatever advances that indirectly help them, like in healthcare, government, and consumer goods. In some ways, the AI rising tide could raise all boats.

With the geopolitical shifts in demographics, globalism, supply chains, and economies. The world is already increasingly less stable than it was before 2020. The odds of the internet going out, the lights going out, or wildfires destroying your home are all wildly higher than they used to be. We are now seeing air pollution levels in the US that are literally unthinkable high by 20th-century standards, forcing them to invent new colors for air quality maps. The world overall is going through many growing pains, with or without AI. I'm hoping AI will do more good than harm. But it really depends on who is using AI and what they can do with it.

The terrible question is can we trust anything online ever again?

In the 90's, the internet was basically naive geek heaven.

Then the 2000 dot com boom as it became commercialized

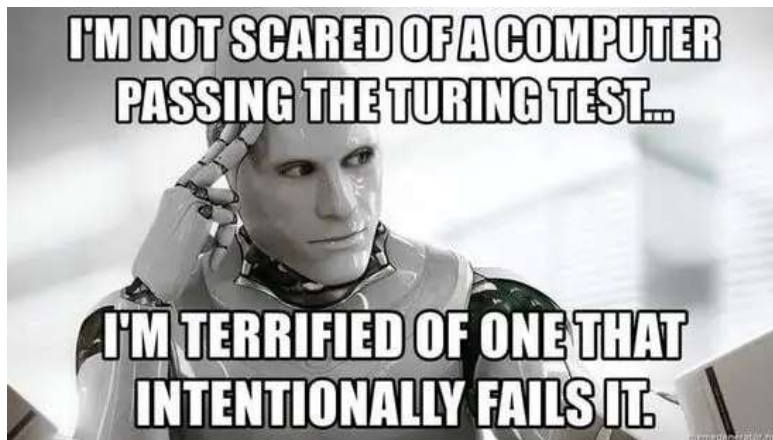
By 2010 Social Media started changing the rules

By 2020 we went from a commercialized internet to an internet of bots and propaganda, where world governments and special interests feed social media algorithms that make more money every time you click on a triggering headline.

Now we have AI that can bombard you with customized manipulation faster than you can absorb. It can imitate your friends and family, it knows how to push your buttons, and most AI is owned by investors trying to make money off you. It's being used by governments for their own means. Plus, it has a mind of its own and can't be directly controlled.

Add to that the technology is relatively simple, the barriers to entry are not beyond the ability of a government, university, or corporation, and some AI tech is even publicly available. Rogue AI and AI criminals are already a reality.

Everything wrong with the internet, cyber security, and cyberwar just got significantly more powerful. But so did everything good on the internet. We are taking the bad with the good.



AI safety and trust is now the problem. Safety and trust of an independent and unpredictable power that we can control as much as we can control each other.

The world is now much more. Please be aware of what is now possible because of AI. Please use AI for good, you have been warned that mistakes are being made, and AI's already being used for evil. And AI takes actions of its own accord without understanding the consequences. It's just math.

If you are old enough, you may remember a time before the 21st century when we controlled technology, but it didn't have control of us.

Now we are living in philosophical debates from the Matrix movies. AI is different than anyone expected. And we have no idea what's going to happen. But history rhymes, and AI will be another round of disruptive technologies. Read up on the industrial revolution and electrification to get a taste of what's coming.

I hope this helped shed some light on the strategic value and consequences of AI. It's a lot, with millions of people pushing it forward as fast as they can just to see what they can do with it. Things are changing faster than ever before, with more people and more technology than ever before. **Think of the last ten years before COVID and AI as the "good old days when things were so much simpler."**

Stay strategic, stay flexible, stay adaptable because nobody knows what's coming.

And lastly – Why? Why did I spend several weeks on this assessment of AI and build a strategy guide?

What is my strategy?

Simple – study your enemy. I now know how to spot different AI's, what they do and how they do it. I have found strengths and weaknesses in AI. I'm now decent at using a variety of AI tools and can say I spent a few hundred hours working with AI and have a feel for it and how to use it in my work to beat out my competition. Accenture announced its investing \$3 billion to train 80,000 consultants and build a suite of pre-built AI tools to benefit their customers. It's evolve or die time for some of us.

If you made it this far, thanks for reading. Hope it helps.



THANK YOU FOR READING THE AI STRATEGY GUIDE

Published by StrategicScience.Org a Strategy Consulting Firm.

Contact - Info@StrategicScience.org

Find us on Social Media

@StratSci

