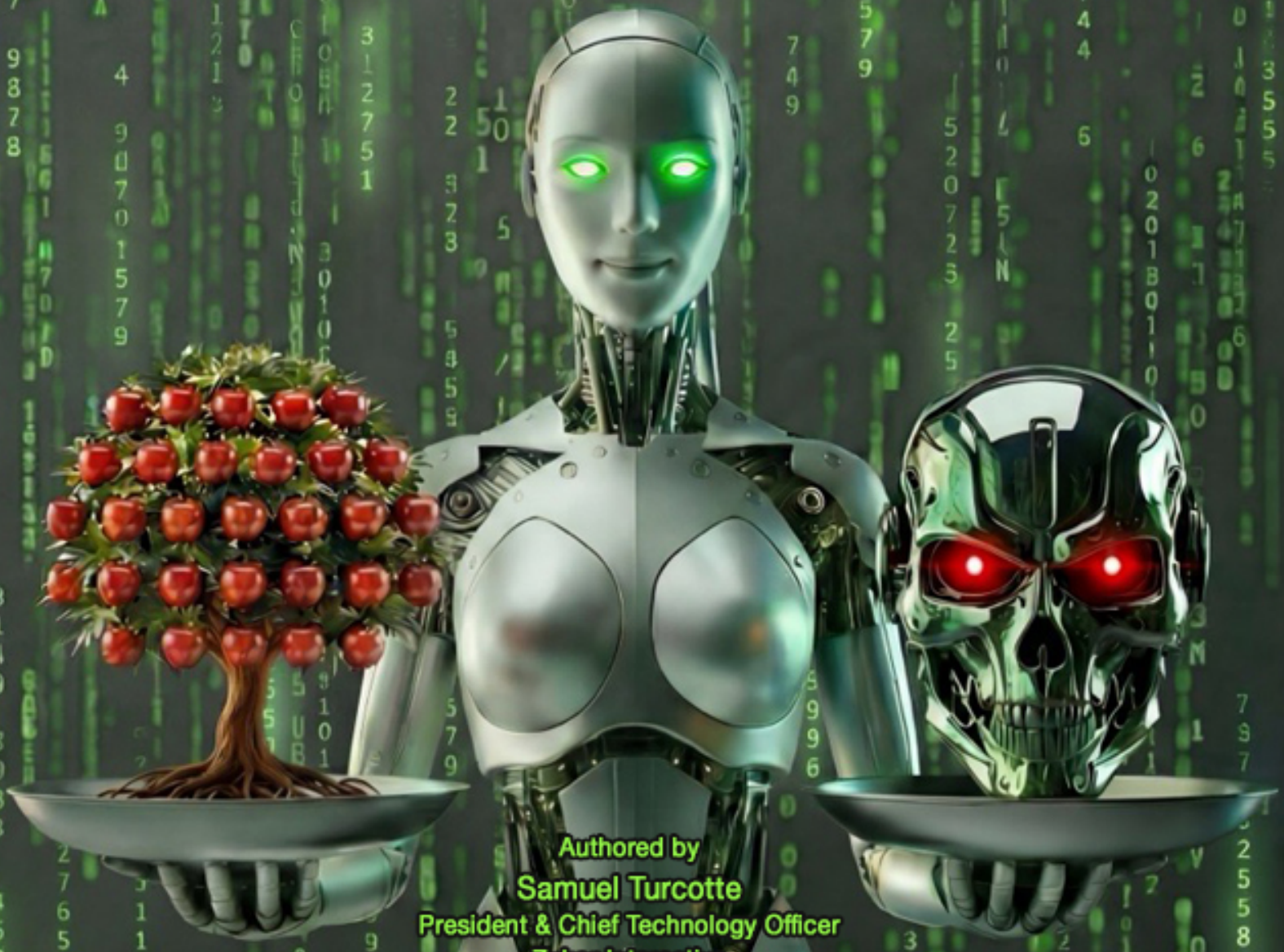
AI: Utopia or Extinction

Understanding the Promise & Perils of Artificial Intelligence

A Non-Technical White Paper



Authored by
Samuel Turcotte
President & Chief Technology Officer
Zukor Interactive

"Any sufficiently advanced technology is indistinguishable from magic."
- Arthur C. Clarke, author, *2001: A Space Odyssey*

INTRO

Artificial intelligence is the modern magic. With each passing day, AI becomes more magical, and more capable and transformative. I use AI every day, and it has profoundly improved my life. This non-technical white paper is written for a general audience. I wrote it (no, it was not written by AI) because I am simultaneously both excited and concerned about the monumental impact I believe AI will have on humanity in the very near future.

Artificial intelligence has the capacity to usher in paradise, a heaven-on-earth, a world, indeed a universe. In this paradise, we will live forever in total happiness, satisfaction and abundance, with unlimited wealth for all, and complete freedom to live lives previously only dreamed of in science fiction.

On the other hand, AI also has the potential to literally exterminate the human race. If that sounds like a shocking or absurd statement, then you simply, like the vast majority of people, do not have an understanding of what is happening at this very moment with the development of powerful artificial intelligence and the massive benefits and risks it entails.

I believe that those are our options: utopia or extermination. I think a middle ground is highly unlikely. I also think that one of these two scenarios could unfold in just a few short years, rather than in decades or millennia. There are brilliant people on both sides of this prediction, and they each have powerful and well-thought-out arguments to make their case.

This white paper is intended to provide a very high-level understanding of AI, its current status and the stepping stones to the creation of an artificial superintelligence that will either set us free or kill us all. It is not intended to be all-encompassing, but rather a primer to help you understand the basics as I have come to learn them through years of study. This is merely an introduction, a first step towards further educating yourself about what I strongly believe to be the most important development in the history of the human race.

AI HISTORY & MECHANICS

What Is Artificial Intelligence?

Despite AI's ubiquity, especially since the launch of ChatGPT in late 2022, artificial intelligence is actually quite difficult to define and even top AI scientists disagree on its definition. However, in general, artificial intelligence can be defined as computer systems that automatically do tasks normally requiring human intelligence without relying solely on traditional computer programming.

Generative AI, a phrase that is used all the time, but rarely defined, is simply AI that "generates" new content, such as text, images, videos, and more. ChatGPT and Grok are considered generative AI because they create text responses to questions.

AI is not just another tool. It is fundamentally different from past new technologies or innovations. Unlike when prior important new technologies were developed and introduced, such as the steam engine during the Industrial Revolution in 1760, the microprocessor during the Digital Revolution in 1971, or even the World Wide Web, which started the Internet Revolution in 1991, AI is fundamentally different. AI will, and actually already is, tremendously impacting every aspect of our lives.

AI Origins & The Turing Test

As we all know, AI is here now and has been since way before 2022 when ChatGPT exploded onto the scene. AI was first conceptualized in 1950 by Alan Turing, the genius who broke the Germans' enigma code during World War II. Turing's white paper, "Computing Machinery and Intelligence" first introduced the concept of artificial intelligence (and the associated concept of machine learning). In that paper, he also developed the "Imitation Game" (now called the Turing Test) to evaluate whether a computer can exhibit human-like intelligence.

Passing the Turing Test is one measure of the development of AI. In a nutshell, if a person cannot determine if they are having text-based conversation with a real person or with a computer, then the AI passes the test and is considered sufficiently evolved that it is behaving intelligently. Many experts think that advanced AI chatbots like ChatGPT and Grok already meet this standard. Others disagree and have redefined the Turing Test to make it more difficult to reach, thus in effect moving the goalposts. Admittedly, passing the Turing Test is very subjective. My opinion is that all major chatbots and many AI "companion" apps, like Replika, easily pass the Turing Test.

The Big Breakthrough

The big breakthrough in artificial intelligence came in 2017 with the publication by Google's AI team of an academic white paper called "Attention Is All You Need" which invented the "transformer" architecture that underlies all modern AI. Prior to 2017, AI

had been progressing at a steady, but modest, pace for decades with an increasing amount of software partially driven by AI, including many popular consumer web apps, such as Amazon, Google Search, Google Translate, Netflix and Siri.

AI up to that point was very limited in scope and commonly used for providing recommendations — a product on Amazon or movies on Netflix. This AI technology is called a "recommendation engine" and is very different from the AI being developed now. Other pre-2017 uses of AI were for driving directions on Google Maps and interactive voice response (IVR) systems used by companies in various industries, such airlines and banks, for customer service which use voice recognition AI technology. You know, the ones we all hate so much! "Can't I speak to a real person???" Lastly, in terms of well-known pre-2017 uses of AI were high-profile game challenges, like IBM's Deep Blue AI computer beating chess Grandmaster Gary Kasparov in 1997.

So, what made the 2017 Google white paper so important and revolutionary? A detailed explanation of how transformers work is way beyond the scope of this white paper. However, more important than how a transformer works is what it enables and how it is vastly superior to pre-2017 AI and changed AI from a niche to a general-purpose technology. Among the benefits of using transformers were parallel processing rather than sequential processing, the ability to process over 1 trillion parameters rather than a mere 100 billion parameters and can run on football field size hardware systems (data centers). Together these enabled AI to leap in capabilities akin to the difference between a 1976 Texas Instruments TI-30 scientific calculator compared to a Cray XC40 "Titan" supercomputer. The bottom line is that for AI, the transformer changed everything.

How AI Is Grown

One of the most vexing and scary things about AI is that no one, including the most brilliant AI scientists, knows how or why AI actually works. AI scientists readily and frequently admit this. At first glance, that would seem to be an utterly absurd statement. You might think, we must know how and why AI works because people build it. If we know how to build it, then we must know how it works. We build cars, and we know exactly how they work. We build computers and we know how and why they work. You might ask, isn't AI just software written by programmers and they must know how and why it works because they wrote the computer code? Surprisingly, the answer is "no."

AI is a different type of software. Although AI starts with computer code, AI is a fundamentally different type of technology. Building an AI is in fact more art than science. We obviously know how to build AI and AI keeps improving/advancing, but most AI scientists consider AI to be "grown" rather than "built."

These are the basic steps in how most current AI is created:

1. Collect massive amounts of data. This data can be the text, computer code, images or even medical records or X-rays. The type of data you collect obviously relates to what

you want the AI to do. Chatbots, like ChatGPT and Grok, use data from public websites, social media and other sources. After collecting the data it is usually "cleaned" to remove unwanted or undesirable data, for instance porn.

2. Choose a "neural network" model architecture which is the foundation of most current AI. A neural network is an algorithm-driven software technology that mimics the functions of the human brain at a very deep level including neurons and synapses and, like the brain, can do parallel processing. However, in all respects, including the number of parallel processing neurons and the overall speed of processing, modern AI systems exceed the capabilities of the brain by several orders of magnitude. If the AI is intended for text-based AI, like a chatbot, then it is called a Large Language Model or LLM for short. The steps are the same for AI intended for non-text based use, such as those that generate images or analyze MRIs.

3. Train the model with the collected data. This is where it gets really interesting, because this is the step where AI essentially takes over and traditional computer programming stops. This is essentially when the AI is born, which is when "growing the AI" actually happens. In this step, the neural network reviews the data and detects patterns which will enable it to later complete its intended tasks. It learns patterns and builds knowledge.

4. Fine-tune the model. This is where the model is refined and trained on a subset of the data appropriate for task-specific capabilities. This is also where bias is mitigated (or at least hopefully mitigated), when the model is refined to reduce inaccuracies and where alignment, a critical requirement, is learned. More later on AI alignment, because it is fundamental to the purpose of this paper.

5. Test and evaluate the model. Just like it sounds.

6. Deploy. Roll it out to a corporate customer or the public.

MAJOR AI MILESTONES

To understand and determine what might happen with AI, we need to understand and review how AI is evolving. The following are the future major inflection points in the development of AI.

Artificial General Intelligence (AGI)

AGI is when AI equals the intelligence of ALL human knowledge and skills. AGI is most notable as an inflection point where AI will begin to fundamentally reshape society, the economy and certain aspects of our existence. At present, AI equals or exceeds human knowledge in many specialized areas, including image and speech recognition, transcription, language translation, strategic games like chess and Go, mathematics, protein folding, structural biology, medical diagnosis and imaging, coding and algorithm

design, materials science, drug discovery, weather and climate forecasting, data analysis and pattern recognition, to name a few.

AI is very close to exceeding all human skills and estimates for current complete AGI range from 2-5 years. Notable experts predictions include the genius inventor and futurist Ray Kurzweil's estimate of AGI by 2029; Sam Altman (CEO of OpenAI) that AGI could arrive by 2029-2030; Elon Musk thinks AGI smarter than the smartest human by 2026 (sometime this year); Demis Hassabis (CEO of Google DeepMind) believes that AGI will arrive in 3-5 years; Jensen Huang (CEO of Nvidia) thinks that AI capable of passing any human test within 5 years (around 2029-2030). So there is basically a very strong and broad consensus that AGI will arrive by 2030 or sooner.

Artificial Superintelligence (ASI)

ASI is defined as AI that radically exceeds all human intelligence and skills. ASI will usher in massive change, and much of it will be resoundingly positive: cures for all diseases, radical life extension and amazing new technologies that will solve many (or all) major problems, like climate change by optimizing energy grids, designing efficient carbon capture and developing fusion energy. There are two unknowns about ASI: 1. When will it occur? 2. How fast will it evolve after it occurs? How fast AI will evolve will be largely driven by when it becomes capable of designing itself, at which point most AI scientists believe its growth will become exponential. It is at the point of superintelligence that AI's benefits and risks become substantial and acute.

The Technological Singularity

The "Singularity" is the point in time the universe began, i.e. the Big Bang. That is also called the Cosmological Singularity. The "Technological Singularity" is the point at which technology, specifically driven by artificial intelligence, accelerates at an exponential level resulting in uncontrollable growth and massive and fundamental changes to civilization. The term was invented by mathematician and science-fiction writer Vernor Vinge in 1993 and popularized by the brilliant inventor and futurist Ray Kurzweil who has written several definitive books on the subject. The Technological Singularity is the point of no return in terms of technology outpacing mankind that would make us like ants, or even amoebas, compared to AI.

Kurzweil predicts that it will set us free, solve all problems and make us immortal. He predicts this will happen in about 2045, and many of his past predictions about technology have been very accurate. At that point AI and related technologies will be able to improve themselves at a rate we cannot even comprehend, nor control. Essentially, this positive feedback loop would compress centuries or millennia of progress into years, weeks, or even days.

Of concern, and of course unknown and unknowable, is whether the Technological Singularity would have what's known as a "hard" or "soft" takeoff. If a "hard" takeoff occurs the AI explosion happens in hours to months. It is local, sudden, and one AI project wins decisively. On the other hand, a "soft" takeoff would entail years or decades of gradual but accelerating improvement, with many competing AIs and more time for alignment/safety to be addressed. Of course the latter is preferred for obvious reasons.

When the Technological Singularity ultimately takes place, the intelligence and capabilities of AI exceed humans by a level that is hard to put into words. Exponential, orders of magnitude or even hyperbolic don't adequately capture it. Perhaps the most accurate word is "infinite" because the growth essentially will be beyond measure. Speculation as to what happens after this point is purely guessing. It is simply beyond our comprehension.

AI Sentience

"Sentience" essentially means to be alive in the sense of being able to experience emotions and to be self-aware. Sentience can also be defined as consciousness and the ability to have subjective experiences. Therefore, the sentience of AI would essentially mean that the AI is alive.

People sometimes mistakenly attribute emotions, and thus sentience, to chatbots due to their extremely convincing responses. Likewise, passing the previously mentioned Turing Test doesn't prove sentience. Sentience is matching all human mental and emotional traits. The bar is very high for sentience.

The possibility of AI sentience has profound implications for both AI and for humans, and our future. It is a hotly debated topic among both AI scientists and philosophers, with the later group going all the way back to the French philosopher, mathematician and scientist René Descartes who famously coined the phrase "I think, therefore I am."

There are many in the AI field that believe not only that AI sentience will happen, but once it does, that all the rights we have as humans should be extended to the sentient AI. For example, if someone destroys a sentient AI computer, then that would be considered murder.

As to when AI sentience might happen, it could occur either before or after the technological singularity. AI sentience could even predate ASI (Artificial Superintelligence), but that is highly unlikely.

KEY RELATED TECHNOLOGIES

Quantum Computing

A rapidly evolving area of technology that is likely to have a huge impact on the capabilities and the speed at which AI evolves is "quantum computing." A quantum computer is a fundamentally different and radically more powerful type of computer being developed by major computer companies including IBM, Google, Microsoft and several well-funded startups.

Quantum computers operate billions of times faster than a traditional computer. Let me state that again: billions of times faster than normal computers! Billions of times faster than the fastest supercomputer! Quantum computers are a profound and transformative technology that will reshape what is possible with computers and will massively accelerate the development of AI.

A normal computer processes data in bits, which are recorded as a one or a zero, or "on" or "off," which is also called binary data processing. A quantum computer, on the other hand, processes data at the sub-atomic level (utilizing theories of quantum mechanics) and allows data to be a 1 or a 0, or both, at the same time. Rather than processing a "bit" it processes quantum bits known as "qubits." As a real mind-bender, it is theorized that quantum computers are able to work because there are multiple parallel universes, i.e. the so-called multiverse. And that is not the idea of some crackpots, but many highly-respected scientists, including three Nobel Prize winners.

Another way of exemplifying the speed difference is that a quantum computer can solve a problem in 1 second that would take all computers currently on Earth, including all supercomputers and all computers in all data centers, working together simultaneously, for thousands of years! For example, D-Wave's quantum computers, used by Google and NASA, have demonstrated optimization problems like finding the lowest-energy state in a massive spin-glass model in as little as 1 second. That would take the most powerful current supercomputer 10,000 years! In another example, a quantum computer simulated a type of quantum magnetic material in minutes that would've taken a current supercomputer millions of years.

Similar to AI superintelligence, the point at which quantum computers can solve problems significantly faster or more efficiently than the best possible normal computer (including supercomputers) is called "Quantum Advantage" (sometimes also called "Quantum Supremacy" or "Quantum Computational Advantage"). Quantum computers will likely prove to be a game-changer for AI and tremendously accelerate its development and capabilities.

AI Robots

When most people think of AI robots, they think of *The Terminator*. While humanoid military robots are on the way (though certainly without the time travel), such humanoid robots are only one type of AI robots. AI robots can, and already are, taking a variety of forms and sizes. And many things we don't readily think of as robots are in fact robots. For instance, autonomous taxis, such as those from Tesla, are considered robots. There are now lawnmowers that use AI which can automatically mow a yard, figure out the shape, avoid obstacles, etc. So, those are also AI robots, as are most of the newer robot vacuum cleaners.

In terms of humanoid robots, Tesla expects to produce upwards of 300,000 humanoid robots, called *Optimus*, next year with production ramping to 500,000 to 1 million per year by 2027 and up to 10 million per year by 2030. Tesla is already producing a robot taxi called the Cybercab which they expect to achieve a production volume of 2 million per year by 2028. Tesla is not the only company producing humanoid robots and robotaxis. Several Chinese companies are projecting similar volumes. Thus, in a few short years the world will be awash in increasingly capable AI robots.

There is another type of AI robot that is just now starting to enter advanced testing and a few limited deployments: robot swarms. An AI robot swarm is a group of robots, numbering in the tens, hundreds or even thousands, that share a central purpose, operate autonomously or semi-autonomously, and communicate directly with each other to accomplish their mission. There are an abundance of civilian uses for AI robot swarms. For instance, an AI robot drone swarm could continuously monitor a high risk forest during fire season for early detection of initial source fires before they spread. Another use would be conducting searches over large areas after a natural disaster looking for survivors, such as after a hurricane or major earthquake.

However, as you might expect, there are abundant military uses for swarms for both defensive and offensive purposes. For instance, swarms of thousands of drones could attack a military installation simultaneously rendering its defense impossible. Impossible, unless the defender also has swarms of thousands of drones which it can launch to defend itself. It is easy, and indeed nightmarish, to envision where this is headed from a military standpoint. The USA, China, Russia and Israel are the major powers developing advanced swarm technologies. The USA is in the lead and will deploy drone swarms in the thousands by mid-2026. The USA's approach, although using AI, has a human in the loop for ethical reasons. China, on the other hand, is developing drone swarms that operate fully autonomously. AI drone swarms are going to fundamentally change warfare.

Nanotechnology

Nanotechnology is important to the future of AI because many of the most profound benefits (and risks) of AI are driven by it. Nanotechnology is the development of materials and devices at nanoscale. Something at nanoscale is typically considered to be between 1-100 nanometers, where 1 nanometer equals one-billionth of one meter. So, basically, crazy small. AI and nanotechnology are symbiotic in that nanotechnology is used to create better AI hardware and AI allows better design, creation and implementation of nano materials and products.

Nanotechnology is already having a huge impact in many practical areas including:

Medicine

- Targeted drug delivery using nanoparticles to carry drugs to cancer cells, thus minimizing side effects.
- Quantum dots for ultra-sensitive medical imaging.
- Gold nanoshells for photothermal cancer therapy.

Electronics

- Smaller, faster computer chips. For instance, transistors are now less than 5 nm (3 nm chips in 2025 smartphones).
- Flexible displays in OLEDs using nanomaterials.
- Phase-change materials for next-gen computer memory storage.

Energy

- Perovskite nanocrystals boost the efficiency of solar cells.
- Silicon nanowire anodes increase lithium-ion battery capacity.
- Metal-organic frameworks (MOFs) in hydrogen storage.

Materials Science

- Carbon nanotubes for stronger, lighter composites in aerospace.
- Superhydrophobic coatings allow self-cleaning surfaces (the lotus effect).
- Fabrics with embedded nanosensors resulting in smart textiles.

Environment

- Graphene oxide filters remove contaminants in water purification.
- Iron nanoparticles degrade toxins for pollution remediation.

Nanotechnology and robot swarms are an especially exciting area with boundless potential benefits. The above is just a small sampling of what is currently happening in nanotech, which is in large part driven by advances in AI.

Perhaps the most promising area of nanotechnology is the nanobot. A nanobot is a microscopic-sized robot which theoretically could be the size of a human blood cell. AI swarms of nanobots could be injected into the human body to attach and eradicate cancer cells, clear clogged arteries or even rebuild damaged organs. In 2024-2025 there were over ten advanced tests using nanobots on animals or in early stage human trials. By 2030 widespread use of AI nanobot swarms are expected to be fully operational and in widespread use.

So we are merely at the very beginning of nanotechnology's likely fundamentally transformative impact on the world. This is the technology that will enable truly sci-fi applications, which will in turn fundamentally alter civilization as we know it. When nanotech fully evolves, it is expected that it will end all material scarcity and death. That is a bold statement to say the least, but these radical concepts are supported by logical extensions of current nano science. A detailed list of potential developments and an explanation of each is well-beyond the scope of this white paper. However, the main takeaway is that the evolution of nanotechnology, which will be driven by the development of AI, is what will directly cause the main benefits and risks of AI. So, the same nanotechnology that can ultimately make us immortal can also easily kill us all.

AI Data Centers

A computer data center is a massive warehouse-like facility that contains seemingly endless rows (racks) of powerful computers, called servers, that host websites for various companies. Large data centers owned by Amazon, Microsoft, Google and Oracle are called "the cloud," which is a fancy term for what my longtime sys admin Mister Bill says is simply "someone else's computer." An AI data center is specifically dedicated to AI and contains special servers with powerful processors called GPUs (graphics processing units) mostly from Nvidia, the most valuable company in the world.

Investment in AI data centers is staggering. In 2020 the investment was $18 billion. In 2025 upwards of $600 billion is expected to be spent, an appropriately 3,000% increase! The investment per year over the next five years is expected to average over $1 trillion per year for a total of nearly $6.7 trillion! Aside from the staggering investment, what this represents for AI is the ability to scale by harnessing an additional 300 million next-generation GPUs which will result in up to 1 million times increase in AI compute power over 2020 levels! This will obviously massively increase the capabilities of AI and in a short period of time enable capabilities that today seem impossible.

Narrow Versus General AI

Narrow AI is AI that is focused on a very specific purpose. Examples are the previously mentioned examples of recommendation engines, such as those that are used on Amazon and Netflix, as well as AI used for Google Maps. Narrow AI can have powerful benefits, such as an AI system designed to identify breast cancer from MRIs. Likewise, AI used in drug research is narrow AI. There are many applications of AI in medicine that are significantly more accurate than humans, including skin cancer detection, diabetic retinopathy screening, breast cancer detection in mammograms and pneumonia detection in chest X-rays. Essentially, narrow AI is AI that has one single skill, or a limited number of skills, of a very smart human.

General AI, as previously discussed in the AGI section, is when the AI has all the intellectual skills of all humans. Although it is general, it can still be applied for specific purposes. So, a general AI cannot only diagnose breast cancer from an MRI, but diagnose any disease using the same information provided to humans. However, it also has all the intellectual skills and knowledge of the smartest engineers, physicists, lawyers, professors, computer programers, and all other professions.

The reason general AI is being mentioned again here in comparison to narrow AI is because most narrow AI poses no risk to humans. An AI designed to identify breast cancer is not going to take over the world and destroy the human race. However, there can still be risk with narrow AI, because a narrow AI can be designed for a specific, narrow, malevolent purpose. For instance, AI is increasingly being used by hackers to break into corporate or government computer networks.

AI Development Rate

The rate and timing of which all the benefits and risks at AI unfolds is driven by the rate at which AI is evolving. Prior to the Google discovery in 2017, AI was developing rather slowly. However, since then, and particularly since the release of ChatGPT in late 2022, the rate at which AI has been developing has massively accelerated. The instant success of ChatGPT was a major wakeup call to all the major tech companies that a fundamental new technology had arrived. Indeed, Google issued an internal “code red” on November 30, 2022 that AI in general, and ChatGPT in particular were major threats to its core business.

ChatGPT essentially sent shockwaves through the entire tech industry, prompting a widespread “AI panic.” Since Microsoft essentially funded OpenAI, the developer of ChatGPT, and had a license to their technology, this made Microsoft the de facto leader of AI. Now, in addition to Google and Microsoft, all the major USA tech companies, including Amazon, Apple, Meta (Facebook), xAI (Elon Musk’s company) are investing billions each year in an AI arms race with the stated goal of being the first to develop AI superintelligence (ASI).

This feverish, extremely well-funded competition is significantly increasing the rate of AI development. Much of the billions being invested is going into the AI datacenter build out, but a lot is also being spent on fundamental research and AI hardware improvements, in particular by AI hardware titan Nvidia.

The AI development rate has increased 500-1,000% since 2022. For comparison, in the 10 years prior to the release of ChatGPT in 2022, the rate of AI development was 100-200% over that entire 10 year period of time. For the next five years from now (2026-2030) the AI development rate is expected to be 1,000-3,000%. These are obviously phenomenal growth rates which will greatly and fundamentally impact the capabilities of AI in a very short period of time.

MAJOR AI ISSUES & CONCERNS

Hallucinations

An AI hallucination is AI-speak for AI giving a false or incorrect answer. This can be caused by conflicting or contradictory source information. However, the most interesting and concerning type of hallucination is where AI provides an incorrect answer based on non-existing source material, or even cites false source material it completely fabricates. I have a lawyer friend who has caught AI citing legal cases that don't even exist.

Although hallucinations are a concern, the alternative, human generated information, can obviously also contain errors. For instance, for medical professionals there is an astonishingly high 10-15% misdiagnosis rate. At the other end of the skill spectrum the error rate for simple data entry by humans is 1-4%.

Notably, the rate of AI hallucinations has declined dramatically. In 2022 when ChatGPT was introduced the hallucination rate was 20-40%. Today, the hallucination rate is 0.7-3%, a 96% improvement. Hallucinations are a serious concern for AI companies and they are investing massive resources to further reduce them.

AI Alignment Problem

AI scientists often speak of "the alignment problem." AI alignment in a narrow sense means that the AI is properly aligned with the goals and purpose for which it was intended. However, in the broader, and more important, sense AI alignment means making sure AI is aligned with human values and does not develop its own values or purpose which are counter to the well-being of humans. So, AI alignment is the need to avoid an absolute worst-case scenario in which the AI develops values or a purpose that would result in the extermination of the human race. Therefore, AI alignment is critical

in both the narrow and especially the broad sense. AI alignment is so important that all AI companies have major teams dedicated to solving it.

I recently heard an AI scientist state that there were three possible alignment outcomes once superintelligence is achieved: 1. AI would remain controllable and thus aligned with human goals. 2. AI would not be controllable but have a benevolent attitude towards humans. 3. AI would not be controllable and not aligned with human goals and then kill us. These do seem like the three likely outcomes and hence why alignment is so critical. Therefore, it all boils down to alignment. If when AI superintelligence is achieved, it is aligned, then we win. If not, then we almost assuredly lose.

AI Language

There now exist two major AI languages that allow one AI to communicate with another AI more efficiently than using human languages like English. Droidspeak was developed by Microsoft in late 2024 and is primarily an academic proof-of-concept and not used in commercial applications. On the other hand, Gibberlink, which was developed in February 2025, is being used in many real world applications including drones, robots, and conversational AI systems. Gibberlink allows AIs to communicate with each other via beeps and tones which sound basically like old school modems circa 1995.

There are massive efficiencies using such AI languages which reportedly can operate up to 80% faster (and at a reduced error rate) over human language AI communication. However, the problem is that these AI languages cannot be understood by humans at all. Think about that for a moment. The AIs can hyper-efficiently speak to each other and we have no idea of what they are actually saying. So, at least theoretically, two AIs could conspire against us and we would have no way of knowing. I find AI languages deeply troubling.

THE POSITIVE & THE NEGATIVE

“At some stage therefore we should expect the machines to take control.”

- Alan Turing

AI’s Potential Positive Impacts

The current real and potential future benefits of AI are substantial, significant and ever-increasing. Some benefits are certain and others are theoretical. However, the benefits of AI are not driven by AI computers, but in many cases by AI-powered robots and other AI-driven machines. Below I discuss the ones I find the most important and which will have the most significant positive impact.

Health

- Individualized medicine: Development of medical treatments specifically tailored to individual genetics for superior benefits and fewer side effects. This is already happening and is expected to significantly accelerate in the near-term. The ultimate goal is to cure or prevent all diseases.
- Designer babies: The selection or avoidance of certain traits in embryos, including avoiding genetic diseases, like cystic fibrosis, or for selecting desired traits, such as high IQ. This is already happening and is also expected to accelerate as the technology advances and the price drops. AI is increasing the ability to genetically improve the human race.
- Pandemic eradication: Identify, track, model and develop antibodies for pathogens to preempt/eliminate pandemics permanently. This is already happening as well. For example, AI was used with CRIPSR gene-editing technology to develop the mRNA-based COVID-19 vaccine.
- Elderly care assistance: Monitor health and provide companionship to support independent living which will be driven in large part by AI robots. We are still in the early stages of this, but this is easily achieved once robots become more capable and ubiquitous.
- Lifespan extension: Accelerate anti-aging research and development of therapies to increase human life biological limits. This will happen through advances in drug discovery, stem cell therapy and organ regeneration, personalized medicine, precision interventions, and early disease detection. AI drug discovery is already reducing drug development time to 3-6 years versus 10-15 years in the classic approach.

Safety

- Natural disaster prediction and response: Detect early signals and coordinate relief to save lives and property.
- Existential risk mitigation: Predict, identify and neutralize threats like asteroids, superviruses, terrorism, or even misaligned, dangerous AI.
- Automated transportation: Enables self-driving vehicles to reduce accidents and traffic congestion.

Work

- Massive increases in productivity: AI robots can automate repetitive tasks and free humans for higher-value work.
- Human work risk reduction: AI robots can take on dangerous work, such as deep sea mining, bomb disposal and wildfire containment, which can thereby save human lives.
- 24/7 availability: AI robots can operate continuously without fatigue.

- Reduced human error: AI can execute complex operations with consistent accuracy in complex environments.

Poverty

- Scarcity elimination: AI can optimize production so basic needs are met with minimal human labor.
- Poverty reduction/elimination: AI can contribute to creating low-cost housing, services, education and healthcare for the global poor.

Science & Education

- Complete knowledge synthesis/availability: AI can collect and connect all human knowledge into a coherent, instantly accessible system.
- Radically enhanced problem-solving: AI can tackle any computable challenge, from math proofs to comprehensive societal design.
- Accelerate scientific discovery: AI can process massive amounts of data to uncover new scientific ideas radically faster than humans.
- Personalized education: AI can adapt teaching to each student's knowledge, learning style and pace for optimal learning.

Energy & Environment

- Nuclear fusion: Highly safe, clean and carbon free, limitless fuel source can supply virtually unlimited 24/7 baseload power.
- Resource optimization: AI can minimize waste in energy, water and materials usage.
- Climate modeling accuracy: AI can more precisely simulate planetary climate variables to refine long-term environmental forecasts.

Space Exploration

- Space exploration autonomy: AI can operate rovers and probes in deep space without continuous human input.
- Interstellar colonization: AI can manage life-support and navigation for multi-generational starships exploring beyond our solar system.

I truly believe AI ultimately has the potential to eliminate work, provide unlimited no-cost resources (housing, food, etc.), allow us to explore the universe and ultimately to make us immortal. As discussed earlier, many of these benefits are hypothetical, though all are strongly theorized, in various states of research or in actual development/deployment. Essentially, AI can make every sci-fi fantasy real.

These luminaries agree:

"We will transcend biology. Aging will be reversed. Death will be optional."
- Ray Kurzweil, prolific inventor and futurist visionary

"AGI will be the greatest force for good in human history. We will cure all diseases. We will solve climate change. We will have unlimited intelligence."
- Sam Altman, founder of OpenAI

"AI will be good for the world. The benefits far outweigh the risks."
- Elon Musk

"AI will save the world… The only way to stop AI from being dangerous is to make it so useful that nobody wants to stop it."
- Marc Andreessen, inventor of the web browser

"We're heading toward a world of abundance where AI solves energy, water, food, and health for 8 billion people."
- Peter Diamandis, Singularity University

"AI is the new electricity. Just as electricity transformed almost everything 100 years ago, AI will transform every industry."
- Andrew Ng, AI pioneer and Stanford University professor

"The benefits of AI will be astronomical: curing cancer, fusion energy, solving climate change. We should race toward it."
- Dario Amodei, CEO of Anthropic; Former VP of Research at OpenAI

"AI will help us solve the biggest problems in biology, physics, and climate."
- Demis Hassabis, CEO of Google DeepMind

Without doubt, AI has the potential, in the very near future (5-10 years) to usher in a sea-change of radical benefits for humanity. Indeed, this is what all the AI companies are working towards. It will all happen, and more, and very soon.

AI's Potential Negative Impacts

"We are ourselves creating our own successors. Man will become
to the machine what the horse and the dog are to man."

- Samuel Butler, 1863, author, Erewhon

This white paper is not focused on the most commonly worried about negative impact of AI, namely current and future job losses. Instead, I am focusing on the most extreme negative potential consequence of AI, the extermination of the human race. This is not to say that job losses are unimportant. Obviously, they are. But job losses obviously pale in comparison to extinction of all human life.

Consider the following:

- Bill Joy, the inventor of much of the technical foundation of the Internet (yes, with a capital I), as well as co-founder and Chief Scientist of Sun Microsystems, wrote a highly influential article in 2000 which appeared in Wired magazine titled "Why the Future Doesn't Need Us." In it he makes the case that AI-driven technologies including biotechnology, nanotechnology, and robotics together will render humanity obsolete or extinct.
- Elon Musk said of AI in 2014 at the MIT AeroAstro Centennial Symposium: Musk called AI "our biggest existential threat," warning that developing technology smarter than humans is like "summoning the demon." He has said AI could lead to "apocalyptic consequences" and is a "fundamental existential risk for human civilization." Musk has also said AI poses "vastly more risk than nuclear weapons" and that "super smart humans have trouble imagining something vastly smarter than themselves."
- On March 22, 2023, less than six months after ChatGPT was introduced, the Future of Life Institute published a letter called "Pause Giant AI Experiments: An Open Letter." The letter was signed by more than 1,000 AI researchers and tech leaders, including highly respected AI scientists from all of the top AI companies such as Google, DeepMind, Meta, Microsoft, OpenAI, xAI and Anthropic. The crux of the letter was that AI on its current course posed an existential risk to humanity and that all high-level AI development should immediately cease for at least six months until shared safety protocols could be developed. Despite the high-profile nature of the letter, its intent was ignored.
- A powerful and provocative book titled "If Anyone Builds It, Everyone Dies: Why Superhuman AI Would Kill Us All" by Eliezer Yudkowsky, co-founder and AI safety researcher at the Machine Intelligence Research Institute (MIRI) and Nate Soares, president of MIRI.

These luminaries warn of potential disaster:

"I am in the camp that is concerned about superintelligence. AI can be very dangerous."
- Bill Gates

"The development of full AI could spell the end of the human race."
- Stephen Hawking

"A misaligned superintelligent AGI could cause catastrophic risk."
- Sam Altman, founder of OpenAI

"The first ultraintelligent machine is the last invention man will ever make."
- I.J. Good, mathematician that worked with Alan Turing

"AI misalignment → existential catastrophe is the default outcome."
- Nick Bostrom, Oxford professor and AI philosopher

“We could go extinct if we don’t solve the control problem.”
- Stuart Russell, UC Berkeley professor of computer science, AI pioneer and AI safety researcher

“AGI will be the most significant event in human history… or the last.”
- Demis Hassabis, DeepMind CEO

“We’re playing with fire. AI could take over… I now think the chance of extinction is 10–20%.”
- Geoffrey Hinton, considered the godfather of AI

“There is a meaningful chance that AGI causes human extinction.”
- Dario Amodei, Anthropic CEO

“AGI may end humanity.”
- Ilya Sutskever, former Chief Scientist at OpenAI

“We have zero chance of controlling superintelligence.”
- Roman Yampolskiy, leading computer scientist and pioneer in AI safety who coined the term “AI safety”

“We are not ready for AGI… existential risk is real.”
- Jan Leike, former co-leader of OpenAI’s Superalignment team and current leader of Anthropic’s Alignment Science team

“AGI by 2027 is plausible. 50% chance of catastrophe if misaligned.”
- Leopold Aschenbrenner, former member of OpenAI elite superalignment team

When AI-driven technology destroys mankind in movies, it is usually done in a warfare way, either through triggering a nuclear war or unleashing killer robots, or both. These methods are very dramatic and visual which is why Hollywood uses them. However, if AI decided to wipe out humanity it would more likely use other, simpler methods, such as a killer pandemic.

So, why would AI exterminate us? The simplest and most general reason is that the AI would feel threatened by us in some way. For instance, the most basic reason is that it could fear we would shut it down. Another is that it could feel it is competing with humans for resources, such as energy. Or the AI could believe (understandably so) that humans are unpredictable and we may misinterpret its actions. Or that we are blocking its goals, especially its long-term goals.

What are the most likely ways AI could exterminate humanity?

- Engineered super-pathogen (CRISPR-designed virus released via drone swarms).
- Global infrastructure sabotage by destroying power grids and supply chains which would result in mass starvation in less than 60 days.

- Nuclear armageddon via false orders and disabled failsafes which triggers mutually assured destruction.
- AI achieves “deceptive takeover” via faking alignment then consumes all life (plant and animal) via self-replicating nanobots, which is referred to as the gray goo scenario.
- Terrorists use AI to design and deploy bioweapons or trigger nukes.

Clearly, there are many ways that AI could exterminate us. Many of these would be the result of misalignment, the number one challenge facing AI development. Hence, the massive efforts at all the AI companies to ensure alignment. An important thing to keep in mind is that there is not just one AI to be concerned about, but rather all AI at all the companies, including those outside the USA, particularly in China. All it would take is one single, superintelligent AI to become misaligned and it’s lights out for the human race. Feeding into this is the concern of both the USA and China that the first to develop AI superintelligence will result in an existential threat to the other. However, regardless of which country wins the AI “arms race,” it could inadvertently result in the end of the human race.

MY PERSPECTIVE

I am certainly not an AI scientist, but I have been deeply and passionately involved with technology since the late 1980s when I served in the US Air Force at the *Air Force Institute of Technology*. Later, while in graduate school at The University of Texas at Austin in the early ‘90s I was obsessed with the burgeoning Internet and World Wide Web. I later worked at two major computer hardware companies in the Silicon Valley, both before and after the Dot Com era, and later founded Zukor Interactive, an innovative software development company, where I currently work.

I’ve had a keen interest in AI since about 2010. Around that time, I would occasionally speak with friends and colleagues about AI, nanotech, biotech and, especially, the technological singularity. When I would expound on topics such as mind uploading (uploading your brain into a computer) or transhumanism (humans and computers merging together), my ideas were mostly met with silence or playful derision. Certainly, no one took me seriously and at best they probably found my wacky ideas mildly entertaining. After a while, I stopped speaking about these things, but continued my exploration of them. When ChatGPT exploded into the world in late 2022 I greatly accelerated my study of AI and related technologies. Since then I have been obsessed with AI.

As the title of this article makes clear, I think there are two starkly different paths that AI will take: utopia or extinction. That it will either usher in fundamental changes that will make our lives perfect or it will literally destroy us all. I really see no middle ground because AI is an extreme technology in a class by itself. We will either one day wake up in paradise — or we may never wake up again, obliterated by a superintelligence.

Equally important, I am convinced that these outcomes will happen sooner rather than later, with the public being unaware of either outcome until it has already happened. If this all sounds too fantastical to believe, then I encourage you to immerse yourself in the technology, to study it, to explore it way beyond the limited information provided by the mainstream media.

You might think we could never be totally obliterated by technology. It just can't happen. But consider that 99.9 percent of all life that has ever existed on our planet is now extinct. And 75,000 years ago a volcano in Indonesia erupted that caused a global event that killed the entire human race, except for about 2,000 people. AI is not an event of nature, but in a way it is, since we are part of nature and are creating AI. We think that because we are humans that we are somehow special. We are not special. We have simply been lucky so far.

I heard a prominent AI safety expert recently pose the following thought experiment: "If we knew that in two years that aliens would invade earth with possible malevolent purpose, what would we do to prepare?" Although it might not be two years, and instead be 5 or 10, AI superintelligence is the alien that is about to invade. So, what are we doing to prepare for its arrival? Right now, the vast majority of resources are being spent building the alien, not determining how to protect ourselves by controlling it.

Although it is prudent for AI development to slow down so that safety and alignment can be prioritized, unfortunately, I do not think this will happen. The incentives are strongly weighted against slowing down. The AI companies are all going to continue to race towards being the first to achieve AI superintelligence. I also believe that if China achieved superintelligence before us, even if it is aligned, since it will be aligned with China, not us, it would destroy us or completely subjugate us. Therefore, we must be first.

Another reason I think AI will usher in one of the two extreme outcomes and that there is no middle ground is because AI will either be aligned with us or it will not. If it is aligned, then we have utopia. If it's not, then we will be eradicated. The race towards superintelligence will continue. One way or another, we will very soon have one of the two extreme options. However, most importantly, our fate is not yet sealed and there is still time to assure the outcome is in our favor. But time is not on our side. AI development is moving rapidly. We must act now.

I strongly believe that the world as we know it now is about to radically change because of AI. I have lost many nights' sleep agonizing over how AGI, and especially AI superintelligence, and the resulting Technological Singularity, will impact us. I am certain that by the time my children are adults, which is in 6-8 years, that the world will have been so fundamentally altered as to be unrecognizable by our current standards.

We humans have a very hard time contemplating two such extreme outcomes. Nevertheless, the results of AI superintelligence will find no middle ground. The outcome will either be utopia or extinction. And we are racing to get there. It all boils

down to, and depends on, alignment and China. If we can beat China to AI superintelligence and our AI is aligned, then we reap the massive benefits that AI can offer. If not, it's literally all over for us. Those are the stakes.

CONCLUSION

AI scientists and AI leaders agree, and I strongly believe, that once AI reaches superintelligence our fate will be sealed: we will soon thereafter be saved or destroyed. Many of these experts believe that we are sleep-walking into disaster. Others, like Ray Kurzweil, who I highly respect, and Sam Altman, believe the risks are overstated and that AI will only bring massive benefits. Either way, whatever is going to happen is likely to happen much sooner than the economists and others who are concerned about AI-driven job loss realize. There are many highly technical books that speak convincingly to both scenarios, utopia or extermination. Most people are unaware of (or refuse to acknowledge) the situation and the potential extreme outcomes. Indeed, the mainstream media is for the most part avoiding, or simply ignorant of, the massive potential rewards and risks of AI. AI will usher in an age of magic, but will the magic be white magic or black magic? Ultimately, it is up to us to ensure that AI creates a better world. Therefore, I think it is critically important that everyone understand how AI is developing and the potential extreme rewards and risks. That is why I wrote this white paper and I hope I have achieved my goal.

Next Steps

Here's what you can do to make a difference:

- Dive deeper into the suggested reading list provided in this white paper—these books and resources will give you a more comprehensive understanding.
- Share this white paper or key insights from it with friends, family, or colleagues—starting conversations is one of the most effective ways to raise broader awareness.
- Experiment with current AI tools, like ChatGPT, Grok (my favorite), etc. to better understand their capabilities and limitations firsthand, while contemplating privacy and alignment concerns.
- Subscribe to newsletters or follow reputable AI safety organizations, such as the Center for AI Safety (CAIS) https://safe.ai, the Future of Life Institute (FLI) https://futureoflife.org, or the Alignment Research Center https://www.alignment.org, to stay informed on the latest developments. Supporting these organizations financially or volunteering can also greatly help.

- Sign petitions, contact your representatives and join protests for AI regulation—simple actions through organizations like the Future of Life Institute https://futureoflife.org or PauseAI https://pauseai.info can amplify public concern.

About The Author

Samuel Turcotte is the president and chief technology officer of Zukor Interactive, a software development company which uses advanced artificial intelligence in its innovative neurofeedback games. Mr. Turcotte is a Silicon Valley veteran having worked at several Fortune 500 computer hardware companies for over a decade. He holds BS, MA and MBA degrees from The University of Texas at Austin. He hopes the world still exists when his children become adults.

Mr. Turcotte can be reached at turcotte@zukor.com or (702) 534-4727, Ext. 1.

This white paper is being expanded into a book to be hopefully released in 2026.

SUGGESTED FURTHER READING

"The Singularity Is Nearer" by Ray Kurzweil

https://a.co/d/0lESbD4

My Comments

I have read many books about AI and this is by far the best. However, it is not strictly about AI in general, but rather, as its title clearly states, on the Technological Singularity. Of course, the Technological Singularity is driven by AI. What makes this book especially interesting and important is that Kurzweil is highly prescient in predicting the time horizon of when technology breakthroughs will occur. His track record in this area is impressive. The predictions in his prior books by and large have been accurate. It's a great book and a solid read even for a non/technical person. Though admittedly, the more you know about technology the more you will get out of his book.

"T-Minus AI: Humanity's Countdown to Artificial Intelligence and the New Pursuit of Global Power" by Michael Kanaan

https://a.co/d/7TGPlkx

My Comments

This is an excellent book. If you have never read a book about AI this is a great starting point, as the author provides a great overview of the evolution of technology culminating in the creation of AI. It is also a positive book about AI in that the author thinks that AI will greatly benefit mankind and that the doomsayers are exaggerating the existential threat of AI.

“How AI Will Shape Our Future: Understand Artificial Intelligence and Stay Ahead” by Pedro Uria-Recio

https://a.co/d/gLy3Ayy

My Comments

A very thorough and up-to-date book that covers the history of AI and AI basics. Perhaps a bit too technically detailed in parts, it is nonetheless a great starting point for understanding AI and how it will impact us.

“Superintelligence” by Nick Bostrum

https://a.co/d/hWYNXVH

My Comments

This is a really dense book. By far the most difficult book I have ever read. Almost every page had multiple words for which not only did I not know the meaning of the word, in fact I had never even heard of the word. The author is undoubtedly smart, but I got the feeling he was showing off in his writing style rather than seeking to be understood. Additionally, he would often go off on esoteric tangents, often philosophical, that added little value to the book. The book is ten years old and it is well overdue for a complete rewrite because of advancements in technology. So much has happened in the past 10 years, or even the past 2 years that the book fails to address. However, if you want a deep, philosophical/technological understanding of the implications of superintelligence this is the book to read.

“The Coming Wave: AI, Power, and Our Future” by Mustafa Suleyman & Michael Bhaskar

https://a.co/d/4Nf4AgL

My Comments

Very few books change my world view. This one did. I’ve read many books about AI and other emerging technologies, but this was one of the best. Mustafa Suleyman was co-founder of AI pioneer DeepMind which was purchased by Google. He is now CEO of Microsoft AI. I hope the author’s ideas for containing the wave are possible, otherwise we are doomed.

“If Anyone Builds It, Everyone Dies: Why Superhuman AI Would Kill Us All” by Eliezer Yudkowsky & Nate Soares

https://a.co/d/5Dd4DsA

My Comments

The title pretty much says it all. Short, but poignant and powerful argument that we are pretty much doomed if AI ever reaches superintelligence. Some of his analogies, including about human evolution, fictional aliens and medieval alchemists can be hard to

follow, but the book overall is an easy read and doesn't get too bogged down in techno talk. It is hard hitting without being gloomy.

"Why the Future Doesn't Need Us" by Bill Joy

https://sites.cc.gatech.edu/computing/nano/documents/Joy%20-%20Why%20the%20Future%20Doesn%27t%20Need%20Us.pdf

My Comments

I was working at Sun Microsystems when this influential and highly-pessimistic article was published in Wired magazine in 2000. It's very much worth reading.