Creating Intelligent Machines

tlv partners

Cognition Through Language:
The Intelligence Revolution

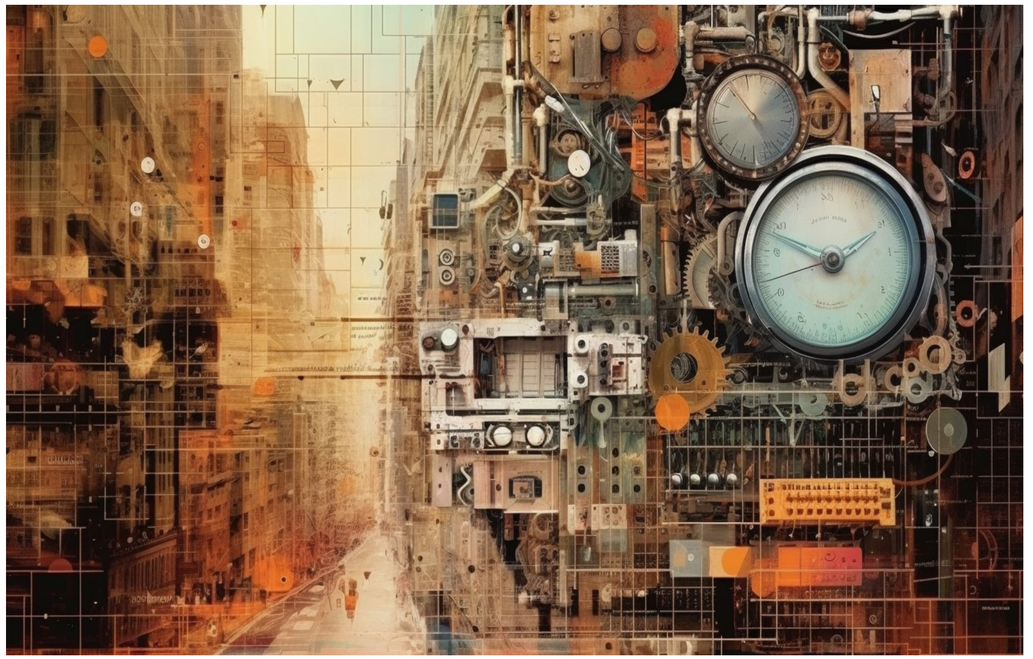


photo: ST / Midjourney

By Shahar Tzafrir
Managing Partner at TLV Partners
June 2023

**“When change comes,
it comes in single drops,
then in battalions”**

- ST, June 2023.

Index

Introduction

The establishment of AI research. ST / Midjourney

The Establishment Of AI Research

In 1956, the term "Artificial Intelligence" was used for the first time, marking the establishment of the research field.[1] In the first AI conference that year, a research conjecture was established suggesting that "Every aspect of learning or any other intelligence feature can be so precisely described that a machine can be made to simulate it."[2]

1 John McCarthy. Winner of the Turing Award (Nobel of computer science) coined “Artificial Intelligence.”

2 The 1956 summer Dartmouth Research Project on Artificial Intelligence was the founding event of artificial intelligence

This vision generated immense excitement.

However, the field did not meet these expectations. For over 50 years, AI research has mostly met dead ends and setbacks, leading to reduced funding and deterring talented individuals from joining the field. Many believed AI research was a pipe dream.

> **In a change of responsibility, computers were no longer only used to run programs but also to find out how problems could be solved.**

A Different Approach to Programming

Around 15 years ago, a completely different approach, coupled with a significant increase in computational power, allowed AI to recapture attention and take center stage.

This different approach was a remarkable conceptual shift. Instead of a human programmer finding the rules and conditions to solve a problem, the computer was tasked with doing that.

In a change of responsibility, computers were no longer only used to run programs but also to find out how these problems could be solved.

That is called Machine Learning, followed by Deep Learning (2010). These conceptual breakthroughs have since become the cornerstone of the new AI Renaissance.

No less important, it marked an initial MOL - "Mental Offloading," where what was previously done by the human brain was now mentally offloaded for a machine to do.

AI Research Renaissance

In the past 15 years, AI research has experienced a renaissance. AI benefited from a reinforcement snowball effect where increased computational power allowed discoveries that drew more funding and talent, leading to more discoveries. Many had immediate practical value, contributing to the field's renewed excitement.

GPT4 - A Giant Leap for Humanity

Unveiled by OpenAI in March 2023, GPT4 is a monumental milestone in AI research. The performance and experience when using GPT4 are nothing short of shocking.

GPT4 is a software program that can communicate with humans. It does that at a level unimaginable before its release. It is used as the entity responding on the other end of a chat program (ChatGPT) as a know-it-all conversational partner.

Developers can license it to enhance their products, which they are doing at an ever-increasing pace.

Microsoft has secured an exclusive license agreement with OpenAI and started by integrating it into a new interface for their Bing search engine. They will use it with their other products.

GPT4 is also becoming a platform, allowing developers to enhance it Through 'Plugins,' opening possibilities limited only by our imagination.

> **Any sufficiently advanced technology is indistinguishable from magic. GPT4 is magical.**

Unlike earlier versions of GPT (such as version 3.5) or similar programs by other vendors, which were all mediocre scientific examples showing how far computers are from reaching human-level communication, GPT4 is a whole new dimension. Through language, it shows a new pathway toward Intelligent Machines and Machines' cognition.

Any sufficiently advanced technology is indistinguishable from magic. GPT4 is magical! It is as if we caught a glimpse into the future. The future reveals a foreign intelligence we created using language. It offers glimpses of what it would be like to integrate such power into our lives. Without overstating its significance, it will have profound and far-reaching implications across

technology, society, and industry.

This paper will explore three Thematic Points:

MOL: Mental Offloading. Delegating tasks that we previously thought demanded human intelligence for computers to perform. We will inevitably see a so-called Intelligence Revolution, triggered by reaching a critical mass of businesses that improve their economics by MOL, forcing a race for mental automation equilibrium. It will have a wide social impact, like the Industrial Revolution. Professions and society will change. Incidentally, software development will be transformed among the first.

SAGI: Subjective AGI.[3] AGI is the holy grail of AI research - a truly intelligent machine, on par (which means superior) with humans regarding the intellectual tasks it can perform. It will transform every fabric of life. I will discuss why AGI progress is a spectrum and not a binary manifestation and why GPT4 is a strong step forward within that spectrum. The closer we get to AGI, the less we can "measure it."

Instead, we will resort to how we assess others - subjectively and by interacting with them.

CBL: Cognition by language.The biggest breakthrough of all. A new path towards Intelligent Machines. Made visible thanks to the breakthrough of OpenAI's GPT4. Unlike the current approach, which focuses on increasing the domains AI techniques solve, we now have a different approach to achieving computer cognition via language understanding. GPT4 undeniably shows signs of that.

Language and Cognition

GPT4's communication, both in terms of its comprehension and the human-like nature of its replies, surpasses anything that has come before it. Its ability to hold a conversation is so impressive that some have suspected a level of cognition on the other end of the line.

GPT4 is powered by the largest artificial neural network (a structure based on the human brain) ever created, with more than a trillion data points. It has "learned" language by processing almost all public and licensed English-language works.[4] This paper discusses the near-impossibility of creating a neural network of this size. How humans use language influences how we see the world and understand our experiences.

> **How humans use language influences how we see the world and understand our experiences. Language is not just about communication. It sets the borders of our thoughts. We cannot grasp what we cannot express.**

3 Artificial General Intelligence (AGI) is a program that should be mentally equal to or beyond human intelligence and cognition.

4 Unconfirmed leaked information from OpenAI, who hold this data in confidence.

That means that our cognitive processes, or how we think and understand, are tied to the language we use.

Judging by GPT4 - That may also hold for artificial intelligence.

Cognition via Words ST / Midjourney

The Search for Intelligence

Defining and measuring intelligence in humans is a complex and debated topic. The challenge becomes near impossible with advanced AI.

As AI progresses rapidly, our ability to assess it diminishes even more quickly.

What are we searching for when we are looking for intelligent machines? Measuring the wrong things leads to optimizing in the wrong direction.

GPT4 has performed exceptionally in standardized tests, including passing at the top percentile in the written portions of the US Bar exam, the Medical School final exam, an oncology exam, an engineering exam, a plastic surgery exam, and the SAT. It passed the USMLE test by over twenty points from any other AI software, including those fine-tuned on medical knowledge.

> **As AI progresses rapidly, our ability to assess it diminishes even more quickly.**

Subjectively, conversing with it leaves no doubt of its remarkable intelligence.

Should research continue to optimize for scoring even higher on these tests? Or has that measure reached its end, and our assessment of what

constitutes a "smart" AI needs a fundamental change?

What GPT4 Says It Can Do

In addition to easily passing the standard human Intelligence tests, GPT4 can perform a wide range of complex mental professional tasks. When asking GPT4 what it can do - here's a list which GPT4 says is only partial, of its abilities: "Reviewing and suggesting improvements to software code, assisting in data analysis for business insights, analyzing and summarizing legal documents, assisting in creating and optimizing marketing campaigns, generating reports and summaries for business management, providing medical information for healthcare professionals, helping in the design and execution of scientific experiments, assisting in project management tasks, providing language learning support, writing blog content or other types of written content, offering customer service and support, assisting in studying and learning new topics, providing information on cooking recipes, helping with fitness routines."

"The true sign of intelligence is not knowledge but imagination." - Albert Einstein ST / Midjourney

It should be noted: GPT4 was only "created" with language. It was not programmed, designed, or aimed to perform any specific task. It is the closest to a foreign intelligence we have created and a very strong step in the spectrum between software and a fully intelligent machine.

Is GPT4 Conscious?

Establishing a scientific threshold for AI cognition is nearly impossible, as is measuring the "intelligence" of something that can surpass humans in almost all intelligence tests. Therefore, applying human cognitive models to AI is a flawed approach.

Machine intelligence and machine cognition lie at the unique intersection of philosophy, psychology, and technology, specifically in deciding what

constitutes cognition in a machine. These domains rarely intersect, but interdisciplinary convergence is inevitable as we approach intelligent machines or AGI.

According to the materialistic philosophical view, anything that convincingly displays cognition can be considered conscious, with experiences as valid as our own. The concept of "real" cognition is a complex and nuanced topic subject to theological debate.[5] Our understanding of another entity, human or machine, cognition is inherently nonexistent. Therefore, when discussing "real" cognition in AI, we refer to our subjective approximations and interpretations of what we believe the other entity has experienced based on observable behavior.

That boils down to SAGI: Subjective AGI. This interaction-based subjective evaluation will be all we have. At some point, that, too, will no longer be relevant. This shift in perception can guide what should be measured and for what.

5 Our perceptions of reality are based on habit and custom rather than any underlying objective truth - David Huume; All I know is I can think - Descartes; The physical world is a shadow of a higher reality - Plato.

The Illusions of Current AI Intelligence

"The illusion that we understand the past fosters overconfidence in our ability to predict the future." - D. Kahneman
ST / Midjourney

When people hear about Artificial Intelligence, they often envision highly intelligent robots capable of many tasks. The current reality is the opposite.

Current AI is known in the research field as "Narrow AI." That is because it's designed to perform specific tasks exceptionally well, at a superhuman level. However, it is completely limited in its ability to apply this to different problems.

To an outside observer or someone watching only a part of a movie, Narrow AI might seem indistinguishable from genuine intelligence. However, upon closer inspection, its limitations become clear.

Whether AI systems are created using cutting-edge Machine Learning algorithms that let the computer define the rules or traditional programming methods where a human sets the rules, they are limited to the narrow domains for which they were developed.

AI research aims to extend the "narrow" problems AI can manage today to generalize towards an Intelligent Machine, namely an AGI. That is where GPT4 took the industry by a complete shock, as it could seemingly generalize, and reached this through an entirely new way. It did that through language.

To an outside observer or someone watching only a part of a movie, Narrow AI might seem indistinguishable from genuine intelligence. However, upon closer inspection, its limitations become clear.

The evolution of computer chess illustrates AI progress. Early attempts to simulate the strategies and understandings of chess grandmasters proved unsuccessful. Instead, the solution relied on brute force: evaluating millions of moves per second with a basic chess assessment crafted by human developers and deliberately simplified for speed.

Deep Blue, IBM's chess computer, employed this strategy to be the first to defeat a chess world champion. Rather than trying to comprehend chess, Deep Blue utilized its computational might alongside a static, pre-programmed chess position evaluation. Simple enough for a novice player to surpass. It was compensated for by the sheer volume of chess positions Deep Blue could consider every second.

In stark contrast to Deep Blue stands AlphaZero, developed by Deepmind.

It demonstrated a significant advancement in AI. AlphaZero utilized reinforcement learning (part of machine learning) to learn how to evaluate chess. Playing hundreds of millions of games slowly improved its evaluation of a chess position.

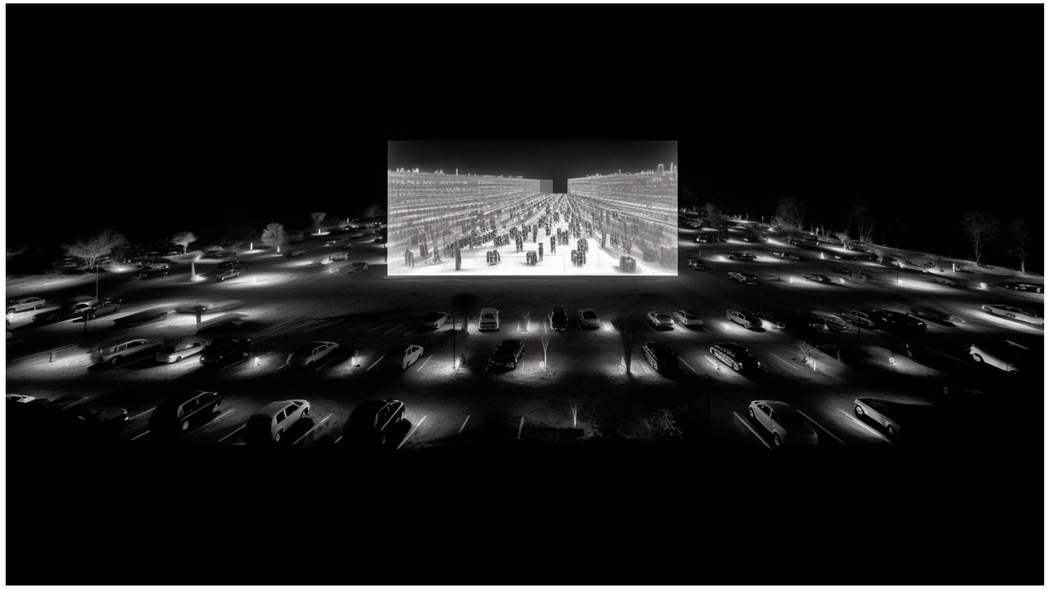
It found its laws of chess strategy and thinking unique from the wisdom humans have found for chess in the 1,500 years humans have played the royal game. Some of its moves defy all current chess concepts yet prove deeper than any player can imagine. Garry Kasparov, one of history's two greatest chess players, commented that he was always curious to see how an alien life form that plays perfect chess would play. After AlphaZero, he now knows.

The DeepMind method showed how AI could learn to solve problems with a long distance between an action (in this case, an early move in a chess

game) and a result (a chess game win, loss, or draw). A concept they apply to other problems.

However, with all its genius, AlphaZero is still a Narrow AI! DeepMind has shown a technique where AI can learn a domain and gain ability in it without any prior knowledge and only through the result, propagating it back to the move performed. But the program itself, as it has gained a deep understanding of chess, cannot generalize this info to other domains.

Unlike humans, who can use chess in other domains, Narrow AI can't.



"Life is like a movie; it doesn't make sense if you skip the middle to see the end." - Marilyn Monroe .ST / Midjounrey

Border Patrol: AI's Gradual Heating

The journey from Narrow AI to AGI is akin to gradually increasing the temperature in a pot. Just as heat subtly transforms the state of water from a liquid to a gas, the incremental advancements in AI are subtly pushing us towards AGI.

DMZ. ST / Midjourney

However, we lack a definitive 'thermometer' to measure the gradual increase in intelligence.

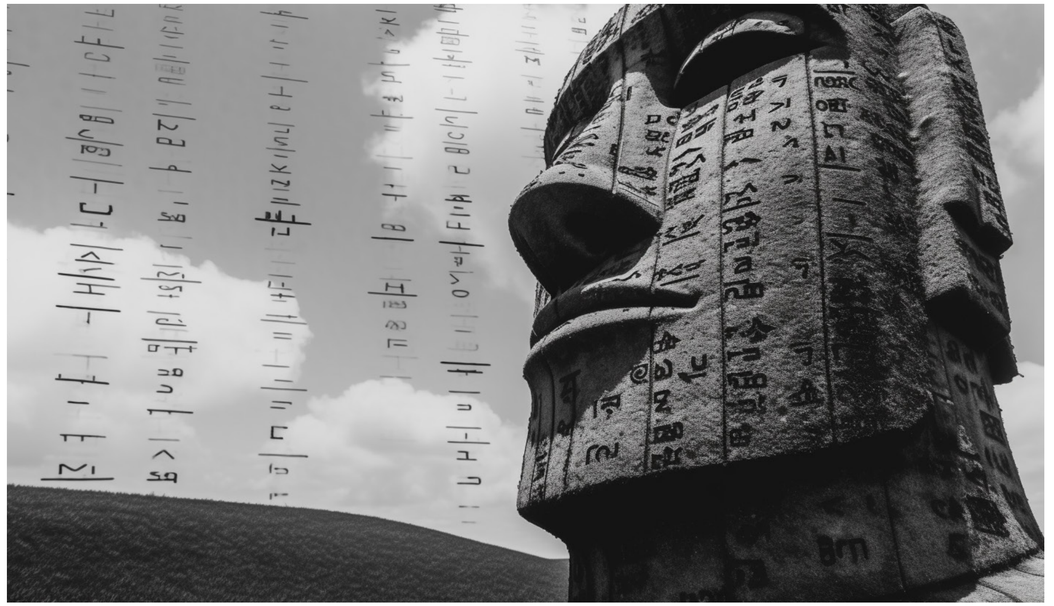
Clear boundaries don't mark the transition from Narrow AI to AGI. It's a continuum of gradual improvements and expansions in capabilities. Just as it's challenging to pinpoint the exact moment when heated water becomes steam, it's equally challenging to figure out the precise point when Narrow AI evolves into AGI.

We know we're aiming for AI that can perform any intellectual task a human being can, but the specifics of this goal remain unclear. Just as the temperature at which water boils can vary based on factors like altitude and impurities, the temperature at which Narrow AI becomes AGI may vary based on factors we have yet to understand fully.

> **Just as heat subtly transforms the state of water from a liquid to a gas, the incremental advancements in AI are subtly pushing us towards AGI.**

In this context, our role can be likened to that of a border patrol, monitoring the progress of AI as it inches closer to AGI. Just as water can exist in a superheated state, appearing to be a liquid above boiling temperature, AI might surprise us, reaching AGI in ways we didn't expect. In a fitting analogy, superheated water is unstable and can flash into steam without warning, especially if it's disturbed or finds a nucleation point.

Our Cognition Within Our Language



Brevity is the soul of wit - Shakespeare, Hamlet

Language is not just about communication. It sets the border of our thinking.[6] We cannot conceive that for which we have no language.[7]

6 The Sapir-Whorf Hypothesis, or Linguistic Relativity, posits that the structure of a language affects its speakers' cognition or worldview. Language shapes and limits of our thinking, suggesting that different languages lead to different perceptions of the world.

7 Newspeak in the dystopian 1984 by Orwell shows how the totalitarian regime aims to remove the mere concept of a revolution by removing words from the language. In a few generations, such concepts will no longer be conceived.

Language, Intelligence, and Cognition are intrinsically interconnected. We form our thoughts and understanding of our world within our language. It shapes our Cognition.

Language, the underpinning of the human intellect, is a dramatic key to unlocking Artificial General Intelligence (AGI), as GPT4 has shown us. Once a considerable quantity is absorbed in a Neural Network, and the depth of the computer's neural network is sufficient, closer to the human brain, something unplanned happens.

GPT4 has processed vast amounts of Language and data. It has the largest neural network ever created. In training GPT4 with language, it has to discover hidden linguistic patterns that interconnect in ways the human brain cannot understand.

With GPT4, we get a peek at the future. It already displays what is experienced as a level of cognition and common sense. GPT4 shows not only a level of common sense but hints of Theory of Mind (ToM) and reasoning ability.

Language as the Basis for Intelligent Machines (AGI)

The current approach to evolving AI is to augment the AI framework with new capabilities to solve problems that humans can easily solve but AI cannot. The rationale is that extending the AI framework to solve more problems will inch us closer to human-level AGI. However, slow progress has cast doubts on its efficacy.

When the field of AI began in 1956, early assumptions that programmers could copy and simulate human cognition by codifying human laws proved misguided. Only with a paradigm shift, when machine learning came to light, AI became a practical field. Today, we are seeing another transformation, this time a bigger one. Computers are learning languages, not for specific tasks but for a broad range of applications. They are becoming proficient at it, showing an ability to transcend syntax and semantics. Before GPT4, chat engines, including GPT3.5, were mediocre scientific demonstrations of how far software was from natural communication. GPT4 proved that computers could learn a language and communicate so effectively that we would have to consider them intelligent and conscious through their superhuman-level communication.

> **Language, the underpinning of the human intellect, is a dramatic key to unlocking Artificial General Intelligence (AGI).**

Moving closer to AGI, we learn the answer has been embedded in our earliest expression - language. Through it, we, too, gain cognition and intelligence.

Common Sense and Theory of Mind

Common sense is the innate ability of humans to make practical judgments or decisions based on basic knowledge and understanding of the world

around them. It is difficult for AI to show common sense as it lacks the human capability to inherently understand and learn from experiences and context, which are fundamental to common sense.

Attempts by researchers to codify common sense within AI knowledge have failed because the list of what constitutes common sense is infinitely long. However, it is not longer than the language itself. GPT4, having learned the entire English language and data corpus (more on how GPT4 was constructed later), has absorbed the good parts of common-sense knowledge, including reasonings and answers.

In a test of common-sense reasoning, GPT4 faced a logic puzzle involving a glass door with 'push' written in mirrored script. Rather than a simplistic approach, GPT4 identified that the mirrored instruction was designed for individuals on the opposite side of the door, concluding that the correct action was to pull, not push, to open the door. This example serves to illustrate GPT4's ability to reason beyond direct instructions.

Another critical aspect of GPT4's capabilities lies in its understanding of the Theory of Mind (ToM), which is the cognitive ability to attribute mental states to others and understand that these may differ from one's own. GPT4 explored the possibility of AI developing ToM, addressing both the potential and the inherent challenges, such as issues of interpretability, flexibility, and the representation of emotional aspects.

GPT4's depth of understanding was further exemplified in its responses to probing questions. When asked about the user's belief in its own ToM, GPT4 explained that although it could not directly perceive its thoughts, it could infer from the conversation that the user might be considering whether GPT4 demonstrated some level of ToM.

In response to a complex analogy comparing an AI's simulation of human cognition to a duck's behavior, GPT4 did not shy away from engaging with the underlying philosophical question. It acknowledged the intricacy of the issue and explored the question of what it means for an AI to "think" or "understand."

Intuition, Creativity, and AI: Art as Another Language

The established hypothesis says that creativity and intuition are unique to humans; therefore, AI will always be foreign and inferior (in addition to the claim that it will never be able to have "real" experiences). And that art, artist, and creativity are one.

Are they?

Intuition is a mental shortcut where pathways in our cognition connect directly to take us to an answer. Supposedly sidestepping the analytical process. It's as if neurons at different depths connect to retrieve "something" we need without doing the full logical, analytical recall cycle.

Intuition is a shortcut to information we already have and have

absorbed deeply. Whether it will forever stay unique to humans, I doubt. Irrespective, the outcome is not unique. The same conclusions intuition quickly popped up could be reached without intuition by taking the long route but driving fast. Given a fast enough AI - its action will be akin to intuition. Knowledge is the great equalizer; AI can drive faster and reach the same conclusions intention has.

As to art - what is art? It's for the consumer to decide. Artists combine ideas and long-honed techniques. However, the technique is but technique, like a rose is but a rose. In the age of AI, it can easily be replicated, replaced, enhanced, and augmented. Accessible to all. What is left are ideas. Creative ideas. At the least, we have dropped the borders of art by making all art forms equally demonstrable to all with ideas and desires. Whether creative ideas themselves are solely human - here, too, I have my doubts. But that's for a separate discussion. For now, I enjoyed making the drawings with which each chapter starts.

Technique is but technique, like a rose is but a rose. In the age of AI, it can easily be replicated, replaced, enhanced, and augmented.

The Intelligence Revolution

The Industrial Revolution: The Dawn of Modern Industry

The Industrial Revolution, which began in the mid-18th century and continued into the early 19th century, marked a significant transition from agricultural societies to industrial ones.

The 'Spinning Jenny,' a textile machine invented in 1764, could replace several workers. Its introduction to factories resulted in predictable conflicts.[8] Historians pinpoint this event as the beginning of the Industrial Revolution.

One of the major catalysts that propelled the Industrial Revolution was intellectual property (IP) ownership laws for inventions. These laws encouraged inventors to invest more time into inventions, knowing that if successful, they could monetize them.

The introduction of technology then led to a transformation in the economic principles of work.

A business, such as a manufacturing plant, that, thanks to assembly line automation, produced more goods at a lower cost gained a meaningful advantage over its competitors. That forced competitors into an arms race of automation from inventions they acquired. They could not allow their

8 The Luddites were fighting for their jobs. Alas, they got immortalized as people who fight against progress, not for their fault.

“Revolutions are the locomotives of history." - Karl Marx ST / Midjourney

competitors an automation edge. To survive, they had to adapt and adopt these inventions.

Capitalism and the Industrial Revolution: A Symbiotic Relationship

The Industrial Revolution was driven by capitalism, not universal values like those the French Revolution started with, nor for the betterment of society.

It's an example of why capitalism drives progress. Uncoordinated action by individuals who rationally only care about their success drives innovation that benefits all.

Technologies emerged that enabled businesses to increase gross income and net profits by achieving higher production levels at lower costs. That was often accomplished by firing employees who could be replaced by automation and having the remaining employees do more, thanks to better tools.

Businesses that did not embrace technology towards efficiency lost to competitors who did.

This cycle increased the demand for more inventions and automation, which drove investment into innovation across sectors.

The cycle continued in full force until the point of diminishing returns was reached, and equilibrium was restored, leading to the end of the Industrial Revolution.

MOL as the Catalyst

Mentally offloading tasks to computers will lead to the Intelligence Revolution. The Industrial Revolution drove progress in tasks that required force and scale. The Intelligence Revolution will drive progress in tasks that require intelligence - the brain.

> **Mentally offloading tasks to computers will lead to the Intelligence Revolution.**

Thanks to AGI being a spectrum, tasks that require human intelligence can already be mentally offloaded to a machine, even before further improvements towards AGI.

The more we progress towards AGI, the more robust and wider its ability to handle additional tasks for us will be.

Innovative people who began using GPT4 to offload mental tasks that were boring or easy for the computer to do better than them were the first to see that mental tasks can already be offloaded to a machine.

This Mental Offloading (MOL) will reach critical mass once programs become available to businesses. The volume and type of people that can be replaced will grow. What business will employ one hundred employees when twenty can be replaced by software?

In a capitalistic act of survival, and not just for increasing profits, all businesses must replace staff with intelligent machines wherever it works, just like all factories had to introduce automation during the Industrial Revolution to survive in a competitive landscape.

This cycle of improving technology and the rate at which it's adopted will continue until a new equilibrium is reached, like the Industrial Revolution. Technology that has a profound economic impact on productivity will be used and flourish because of capitalism. No external force is required. The free market itself drives innovation when it has intrinsic economic values.

The societal impact can be directionally forecasted, and again with guidance from the industrial revolution, we can assume social upheaval and job elimination until reaching a new status quo.

OpenAI's strategy of licensing its technology to developers has set the stage for an influx of productivity enhancement tools as the need for# The user's text is too long for one message, so I'll split it into two parts.

OpenAI's strategy of licensing its technology to developers has set the stage for an influx of productivity enhancement tools as the need for them arises.

These tools, designed to assist consumers and businesses, will change many job roles and processes. They will address the types of professions that the Industrial Revolution did not.

Demand will come, and many developers will have the technology to create vertical and task-specific programs. These will require human handholding for some tasks, and other tasks will be fully automated, replacing a cognitively able human from the loop.

The extent of the impact, the timeline, which professions, and to what extent they will be affected are impossible to predict.

One safe bet, though, is that the profession that will most profoundly change, and will be among the first to change, will be software development."

Making The Largest Language Model Ever

Not all Neural Networks are made of Neurons. ST / Midjourney

GPT4's exact size is kept secret by OpenAI. Leaked unconfirmed information has placed it on one terabyte of parameters and up to hundreds of terabytes. The corpus of data from which it learned was put in thousands of gigabytes to a petabyte. A petabyte, for perspective, is equivalent to 1,000,000 gigabytes (GB). It would take about 223,000 DVDs, each with a capacity of 4.7 GB, to store a petabyte of data. Considering an average 10-minute MP3 audio file size of 4 MB, a petabyte could store around 250 million songs. In terms of video, assuming an average file size of 1 GB per hour, a petabyte could hold about 1 million hours of video.

Irrespective of the exact number, it is evident that GPT4 has the largest neural network to have ever been attempted. This endeavor was not merely about algorithms but also represented an operational challenge of epic proportions.

The Operational Challenge

With GPT size, extensive computer, memory, and networking enhancements were necessary to actualize this project. Creating such a model pushed both hardware and software boundaries to their limits. No computer could hold such size as this Neural Network, and the computational requirements demanded multiple dedicated data and computer centers.

Technical Challenges and Cost Implications

Creating a neural network of this size presents many engineering challenges, including requiring numerous CPUs/GPUs for training and testing, memory capabilities far exceeding standard computers, and superior networking for distributed computation.[9] Any minor faults could jeopardize thousands of hours of computational work. As a result, redundancies were critical at all stages.

9 A GPU, or Graphics Processing Unit. Like a CPU but it only runs special mathematical commands used to game graphics. It turned out that the same commands were essential for AI developments, making GPUs required for a far larger market than gaming.

It is also worth noting that the cost associated with this neural network is estimated to be above $100 (confirmed by OpenAI to exceed this number), with most experts estimating it to have cost $300 million to $500 million, with a significant portion of that spent on computing power.

OpenAI's decision to construct the world's largest neural-network language model was extraordinarily daring. The potential outcome of training such an expansive network was speculative and unknown, and whether an incremental increase in network performance would justify the associated costs needed to be clarified. There was no sure way of even knowing whether an attempt to create a neural network of this size was possible.

Steps to Create GPT

The creation of GPT4 involved a series of crucial steps:

Corpus Creation: The first step involved building a massive corpus of relevant, diverse text. The content of the corpus used to train GPT4 is a commercial secret. It encompasses most publicly available textual data and a large data set of licensed text. It is not only collecting this data and maintaining it that is hard. It takes significant effort to clean this corpus from wrong, biased data and focus it on what should be included. A few wrong decisions and the model training will end up quite different.

Designing the Neural Network Architecture: This process involves deciding on the architecture's structure, including the number of layers and neurons,

OpenAI's decision to construct the world's largest neural-network language model was extraordinarily daring.

the type of data they can hold, and their initial values. OpenAI used its enhanced version of the Transformer architecture (the T in GPT), a popular choice for natural language processing tasks.

Training the Network: During the training phase, the neural network is fed chunks of data from the corpus and tasked with predicting the subsequent chunk of text. If the prediction is inaccurate, the connections between the neurons are adjusted to improve future predictions. This process is repeated countless times, thousands of millions of times, slowly getting the model to refine and increase its predictive accuracy.

I know this sounds counterintuitive or, bluntly, delusional. How can reading chunks of text and guessing what comes next lead to GPT4? Only it does work. Doing this thousands of millions of times makes GPT discover hidden relationships within new data and between that data and other parts of the language. These relationships are of such complexity that humans cannot find or even understand them.

Iterative Training: This process involves running the training procedure countless times to enhance the neural network's prediction accuracy.

Fine Tuning involves conducting refined tests and incorporating human

reviewers to rate question-answer pairs after the first training. This process continues until the entire corpus has been processed.

Run-time Cost Optimization: Optimization methods are needed to reduce the run-time cost per user query. That includes building enough processing ability to run these run-time predictions so the performance of each such query is reasonable. A complicated neural network can require substantial processing, in run-time and cost, to process a new precision call.

The Human in The Loop Element in Training

A not often discussed aspect in creating GPT4 and many other AI models is incorporating the "human in the loop" through training, fine-tuning, and production when unexpected results happen. It can involve many human reviewers. They review results, tag results pairs, or step in if a production anomaly happens.

Drawbacks of a Larger Network

The primary benefit of a more extensive network is its ability to handle complex tasks and understand more nuanced language patterns. As dramatically evident by GPT4.

However, larger networks also have significant drawbacks. They need high computational resources, both for training and run-time predictions. The increased run-time demand can lead to a slow and costly response time for the network. Managing and maintaining such a comprehensive system can also be complex and requires sophisticated software and hardware infrastructure.

Luck

An element of luck is involved in creating a successful AI model like GPT. Minor changes in the neural network setup can lead to vastly different results. Deciding when to stop and keep the "best model yet" vs. going for another attempt is mixing luck, money, and time. The creation of GPT4 was an engineering marvel requiring significant resources, ability, and persistence. And like all successful endeavors - it also required luck. Fortune sometimes brings in unsteered boats.

Summary

I began this paper with three central themes:

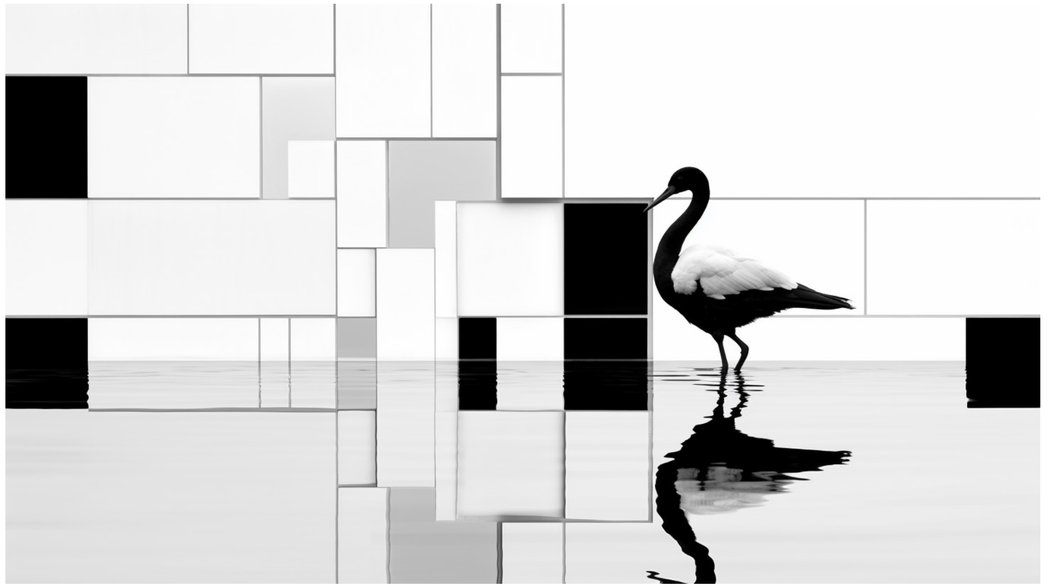
- MOL: Mental Offloading
- SAGI: Subjective AGI
- CBL: Cognition by Language

MOL: Mental Offloading

GPT4 can manage tasks we previously thought demanded human intelligence. Innovators began using GPT4 to offload mental tasks they believed GPT4 could handle more efficiently. This MOL process will grow with future versions like GPT4 that cater to specific industries and tasks. That leads us toward an Intelligence Revolution akin to the Industrial Revolution - a must-adopt scenario for businesses.

SAGI: Subjective AGI

As AI research progresses, our capacity to assess it diminishes. Soon, the only method left will be personal interaction, akin to how we form opinions about others. Artificial tests for AI will continue until it's self-evident that they don't add value. The risk is that misdirected focus now might skew future research results.



GPT4 is a Black Swan. ST / Midjourney

CBL: Cognition by Language

Cognition, intelligence, and language are intertwined. We can only think within our language, which allows us to conceptualize. GPT4 introduces a novel way to form cognition through language. GPT4 has demonstrated impressive performance, underscoring the importance of language for AGI. This shift in perception is the most promising development from GPT.

Concluding Thoughts

- Is GPT4 intelligent? Irrespective of how intelligence is defined - yes, it is.
- Is GPT4 Conscious? Irrespective of how cognition is defined, it’s almost not at all.
- Did GPT4 alter our course towards AGI? Absolutely. Cognition is entwined with language.
- Did GPT4 bring us closer to AGI? Yes, significantly more than other incremental advances. It presented a new paradigm for AGI research - cognition by language.
- What are the impacts of MOL? Huge. The Intelligence Revolution is coming soon. It promises to transform even more than the Industrial Revolution did. What began as creative individuals offloading mental tasks to GPT4 will culminate in a revolution driven by capitalism.
- Short-Term Impact: This will notably affect how software is written above and beyond other professions. A paper on why that’s the case and the impact on the tech industry, startups, and VCs is forthcoming.

Thank you. ST / Midjourney

About TLV Partners

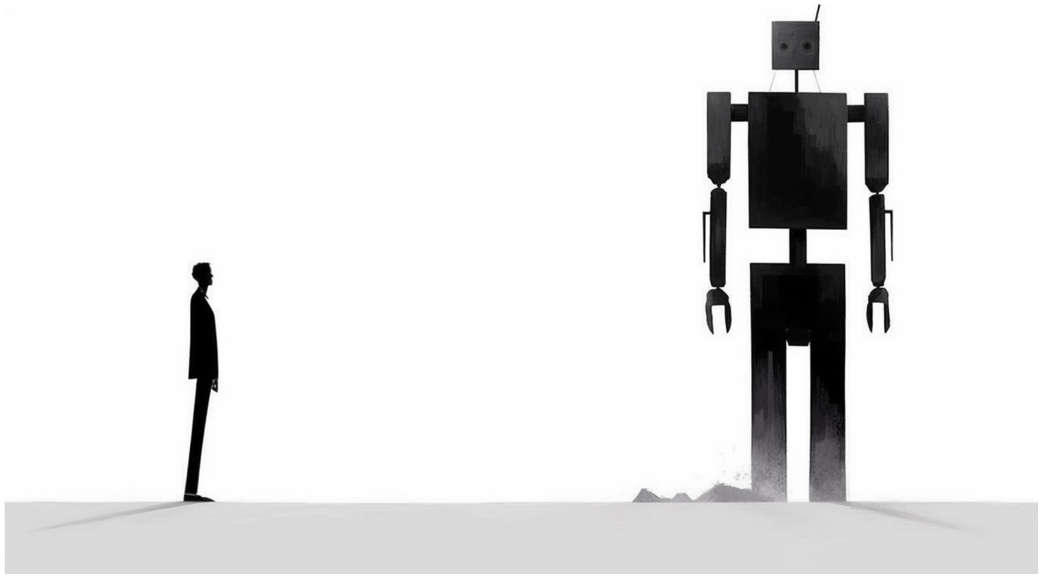
TLV Partners is a venture capital firm based in Tel Aviv focusing on early-stage investments. We seek Israeli entrepreneurs who dare to dream big and challenge the status quo. Our team, driven by a love for new ideas, works closely with ambitious founders to build strong, industry-leading businesses. With almost $1 billion in assets, we invest anywhere we find groundbreaking technologies. Since starting in 2015, we've backed many of Israel's standout companies.

Website: www.tlv.partners

For any inquiries, please contact Shahar Tzafrir, Managing Partner at TLV Partners, at shahar@tlv.partners

Tel Aviv view from Jaffa.Prompt Chaining: Phone->Snapseed->GPT4->Midjourney / ST

Appendix A - Risks



Say Boo! ST / Midjourney.ST / Midjourney

Loss Aversion or Risk Management?

We find it easy to focus on risks. An evolutionary-enhanced trait. Irrational loss aversion unifies the masses.[10] To put some in proportion, Yann LeCun[11] on Twitter:

- Engineer: I invented this new thing. I call it a ballpen
- Twittersphere: OMG, people could write horrible things with it, like misinformation, propaganda, and hate speech. Ban it now!
- Writing Doomers: imagine if everyone could get a ballpen. That could destroy society. There should be a law against using ballpen to write hate speech. Regulate the ballpen now!
- Pencil industry mogul: Yeah, ball pens are very dangerous. Unlike pencil writing, which is erasable, ballpen writing stays forever. Government should require a license for pen manufacturers.
- The "Terminator Risk" refers to the fear that AI or AGI could become an existential threat to humanity. This fear has been fueled by stories dating back thousands of years, from Greek mythology to old films such as Metropolis and 2001: A Space Odyssey. In contrast, the near-term concerns should be about AI being used to create massive amounts of fake content and fake websites and about AI learning our psyche and monetizing or influencing us even more precisely than the current social media sites do. A subset of the Terminator Risk is a programmer error, either in the software of autonomous weapons or in misinterpreted instructions.
- Open web data pollution. The most likely, as it's already happening with less sophisticated tools. Creating a network of self-promoting sites, fake news, and fake marketing campaigns. Fake everything textual. Polluting with a mass of fake data.

10 Daniel Kahneman and Amos Tversky

11 Professor at NYU. Chief AI Scientist at Meta

This ease of creating and spreading misinformation. Increased.

- Manipulation. Threat actors can use dialog base communication GPT has through ChatGPT and feed purposely misleading data.

- Long-term attacks: This risk involves threat actors using AI to engage in long-term psychological attacks against individuals by injecting their responses into conversations. For example, a threat actor could use AI to impersonate a trusted individual or entity and manipulate a person to reveal PII.

- Walled gardens: This risk refers to the confinement of information within a single system, such as an AI app. It could limit the diversity of information and perspectives users are exposed to.

- Task offload to task beachhead: This risk involves outsourcing cognitive tasks to AI, which threat actors could hijack. For instance, if a person relies heavily on an AI assistant for tasks such as email communication, a threat actor could manipulate the AI's responses to gain access to sensitive information.

- Misleading data: AI systems, including GPT, can sometimes provide incorrect or misleading information, either due to errors in the training data or bugs in the system. For example, an AI system might confidently provide incorrect medical advice, potentially leading to harmful consequences.

- Data belongs to others / Results belong to whom?: Generative AI models are based on data on which they were trained. It may be copyrighted. Does this color what is generated when using these tools? What about the results - who owns them? The person who wrote the "prompt"? The tool that created the content? Does it only depend on the terms of use? Does the original content creator have a say? What happens if the result is visual and infringes on rights? Who is accountable?

- Old biases or new biases?: AI, learning from human data, will absorb their prejudice. Some are so deeply rooted in the language that removing these old biases is far from trivial; for some corpus, far from possible ingrained answers are based on the norms hidden in the text on which they were trained. These biases run deeper than keywords that can be removed. They reflect the core ethics and prejudice of the culture within which they were written, tainted with the biases of that culture. Training a language model on our past written text, which is how LLMs are trained in language, introduces these biases into their core. Filtering these texts out or cleaning them is extremely difficult and not bulletproof.

That’s a partial list that does not include many others, including the human labor spent in creating these AI engines (person-in-the-loop with work offshored to least expensive, least regulated countries), the social and economic impact of mentally offloading tasks, lack of AI explainability, the constant surveillance AI allows, physical and mental.

Appendix B - AI Pulled Me Into AI

My fascination with artificial intelligence started at a young age. I grew up with iconic science fiction books such as the AI-themed works of Asimov's "I, Robot" and "Caves of Steel" series, Arthur C. Clarke's "Space Odyssey," and the hallucinogenic, profound, and sad Philip K. Dick's "Do Androids Dream of Electric Sheep?" and many more.

My first, literally, hands-on experience with AI came in the form of a computerized chess board. Equipped with magnetic sensors, an FPGA running a chess AI inside a wooden board, communicated moves via lit squares. Once my initial amazement and curiosity about how it works (and an almost irresistible desire to tear it down) subsided, it became an everyday fun thing to use. I quickly learned that when something is helpful, you forget about the technology behind it and use it.[12]

A decade later, I reconnected with AI through "Advanced Chess," a chess variant developed by world-chess champion Garry Kasparov. This format saw players team up with a Chess AI to play against similar pairs. I learned that the interface matters most when collaborating with AI; the human player sees AI-suggested moves and position evaluations and can guide the chess AI outside its horizon to explore very long-term ideas. As Kasparov said, "The AI hand doesn't shake," meaning no tactical mistakes, yet

ST / Midjourney

the creativity and intuition of the human make the difference. It's about understanding the workflow and how best to augment each other.

My third and most impactful encounter with AI occurred at the dawn of my venture capital (VC) career, a decade ago in 2013. DeepMind, through game playing, demonstrated a breakthrough approach. With no prior knowledge of rules, an AI agent learned to play Atari video games, such as Space Invaders, quickly reaching superhuman performance. The AI agent initially moved randomly, receiving only raw display information and score updates. Positive actions were rewarded, and after numerous trials and failures, the agent slowly learned its version of the game's rules

12 I discovered this perfect quote years later: “The advance of technology is based on making it fit in so that you don’t really even notice it, so it’s part of everyday life.” - Bill Gates.

and achieved superhuman performance.

Convinced of this breakthrough's potential to transform business problem-solving, I sought to invest in startups planning to utilize AI to tackle such challenges. Over the past ten years, this theme has guided me from visually driven software quality testing to construction, radiology, data-center workload optimization, target and drug discovery, user churn, tax, and other products across various verticals.

I am a Managing Partner at TLV Partners, where we oversee nearly $1B in assets under management. My passions are pre-seed and seed investments. My focus is on the intersection of AI across multiple verticals, where unsolvable painful, costly problems can now be addressed.

AI researching and investing over the last ten years has given me endless opportunities to gain experience, think, imagine, wander, and wonder. However, in these ten years, two awe-inspiring moments stood out. The first was DeepMind's demo a decade ago, and the second, more profound and inspiring, was the recent introduction of GPT4.

Credit to chess - and what it taught me about Life, Startups, and Venture Capital.