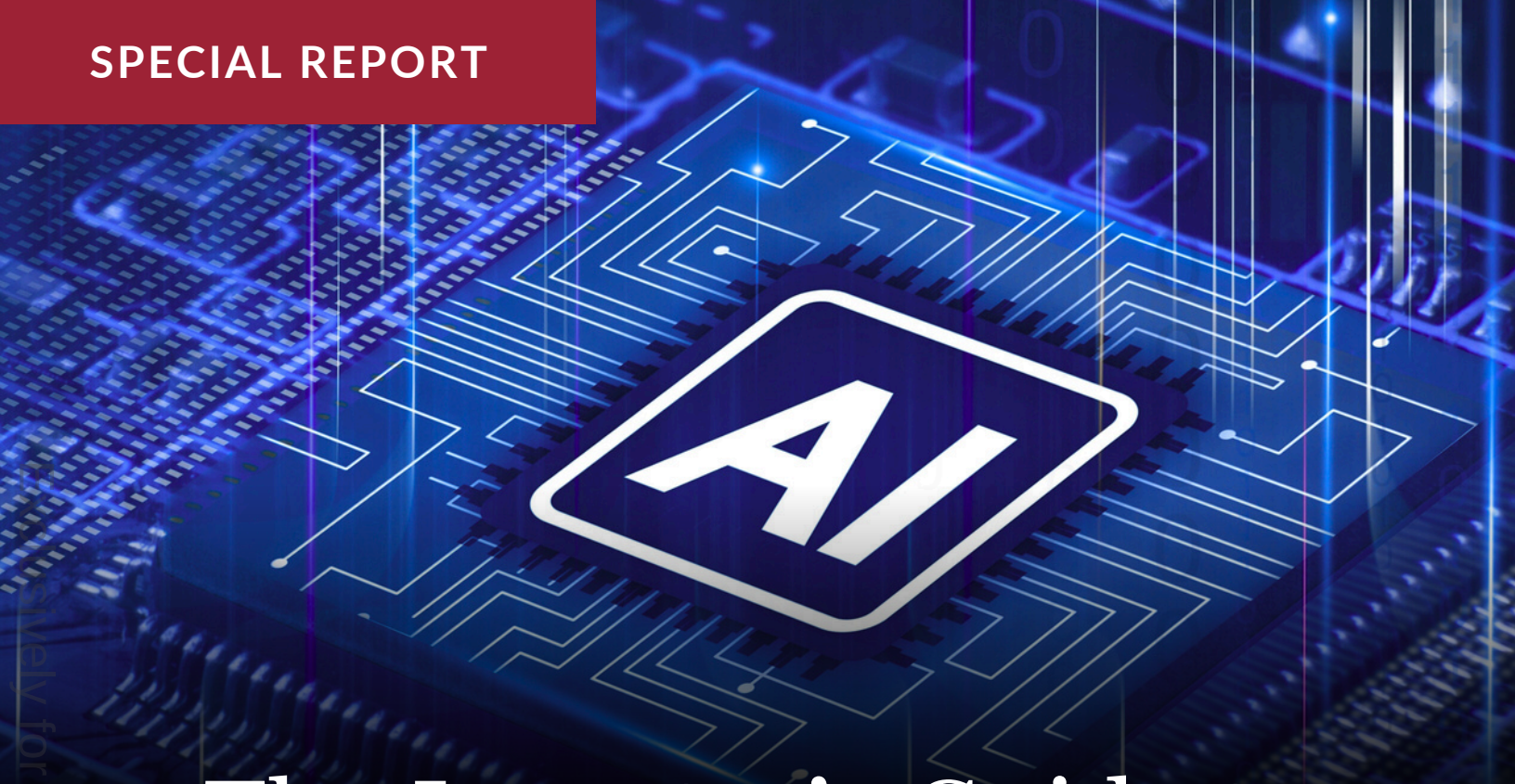
SPECIAL REPORT



The Investor's Guide to Artificial Intelligence

How to Profit from the Next Chapter in This "Fourth Industrial Revolution"

Rosenberg Research

John Smolinski, External Research Consultant
David Rosenberg, Founder & President
July 14th, 2025

+1 (416) 583-5566
information@rosenbergresearch.com
rosenbergresearch.com

Table of Contents

Executive Summary

In this Special Report, we identify and wade through the far-reaching impacts that Artificial Intelligence (A.I.) will exert across the overall economy, a wide swath of industries, and various asset classes. As well, we discuss investing opportunities and potential risks. An appendix on A.I. models, technologies, and infrastructure is included at the end for interested readers.

There are, indeed, many risks to contemplate. But so long as they can be identified, they won't end up falling in the Rumsfeldian bucket of "unknown unknowns" — rather, they become calculated risks that can be priced out before any investment decisions are made. Risks must always be calculated, not avoided — especially in the realm of generative A.I. with all its complexities and ramifications across the markets, economy, and society at large. This is otherwise known as risk management, and that is a key aspect addressed in this report. Knowledge, after all, is power, and this report serves as a guide for making educated and well-informed decisions. Yet, as extensive as it is, the topics covered are by no means exhaustive, and we welcome any thoughts or suggestions from our readership regarding further development of the storyline.

Overall, it appears to be full steam ahead for investors with long time horizons. With "The Next Generation" being the central thesis, how apropos to quote the great Captain Picard: "Engage!" Those investors who missed out on the first chapter of this A.I. story will be comforted to know that many more chapters are on the way, and they can look forward to more significant opportunities ahead. The multiplier effects from the steam engine, the cotton gin, the railway, electricity, automobiles, and the internet lasted many years and spawned numerous profitable investments.

Introduction

There has been much debate as to the degree that Artificial Intelligence (A.I.) will affect the global economy and society in general. The term "*The Fourth Industrial Revolution*" was the title of a prescient book written in 2016 by Klaus Schwab, the founder of the World Economic Forum. At that time, A.I. was just emerging and Mr. Schwab envisioned meaningful changes to the *status quo* resulting from new technologies such as A.I., autonomous driving, IoT (Internet of Things), and quantum computing.

It is becoming readily apparent that A.I. has much potential to alter every industry and company in a meaningful way. It is opening up new opportunities for scientific discovery, such as developing new medicines to cure many diseases. It will spur innovation in other fields, enable productivity gains, and boost the global economy. Manufacturing efficiencies will be achieved that help solve a growing problem of our time: a general shortfall in the skill set and productivity required of workforces to cope with the Schumpeterian "creative destruction" pervasive in today's world. A.I. can also bolster many businesses seeking to make bold investment decisions that capitalize on emerging technologies. It truly is an exciting time. But there are also some very worrisome aspects of A.I., such as having a broadly disruptive impact on employment levels or being used by bad actors to harm humanity via cyberattacks and bioterrorism.

In this paper, we highlight the impact that A.I. will have on a wide range of industries such as data centers, electric-power generation, healthcare, and manufacturing, to name a few. We also discuss the effect it could have on various asset classes, including global equity markets, infrastructure/alternative assets, private equity, and private credit. And we will look at the economic consequences, such as job creation/destruction and the overall impact on GDP growth.

We conclude by discussing the opportunities for individual companies to benefit from the buildout of the A.I. infrastructure over the years to come. It is critical that business leaders, employees, politicians, and the everyday person understand the opportunities and risks that lie ahead. CEOs and business owners will have an especially difficult task assessing how to adapt to an up-and-coming and highly disruptive technology. Much strategic planning and key decision-making will be needed from senior executives and their board of directors. The key decisions include whether to invest significant amounts of new capital in their businesses and how to raise it in order to take advantage of the new opportunities and remain competitive.

Overall, there will be many exciting new technological advances and products that benefit people all over the world. Unfortunately, there will also be many hiccups/setbacks ahead. With the foregoing as a backdrop, let's analyze A.I. and how it is reshaping the world.

Industry Impacts

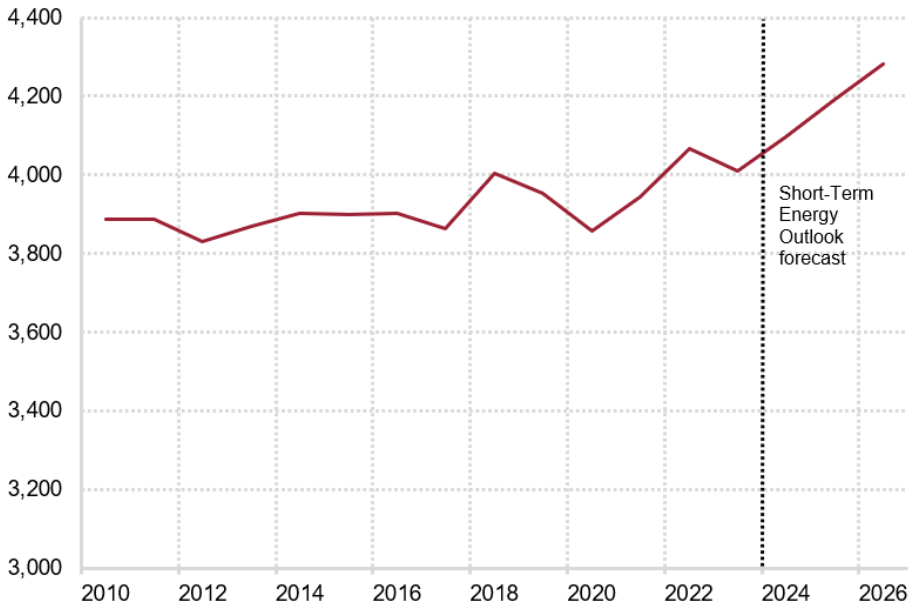
Power Generation (U.S. Market)

According to the International Energy Agency (IEA), 80% of the increased electricity demand due to data centers is expected to come from the U.S. and China, with the U.S. being the largest contributor. For this segment, we will focus entirely on the U.S. electricity market.

U.S. electricity demand grew at +2.0% (average annual growth rate) from 1990-2005 as per the Energy Information Administration (EIA), then plateaued for the 2005-2020 time period (mainly due to efficiency gains), and once again has trended upward up since 2020. The renewed growth is being led by increased industrial demand, a growing population, and rising demand from data centers.

After More than a Decade of Little Change, U.S. Electricity Consumption Is Rising Again

United States: *U.S. electricity consumption*
(billion kilowatthours)



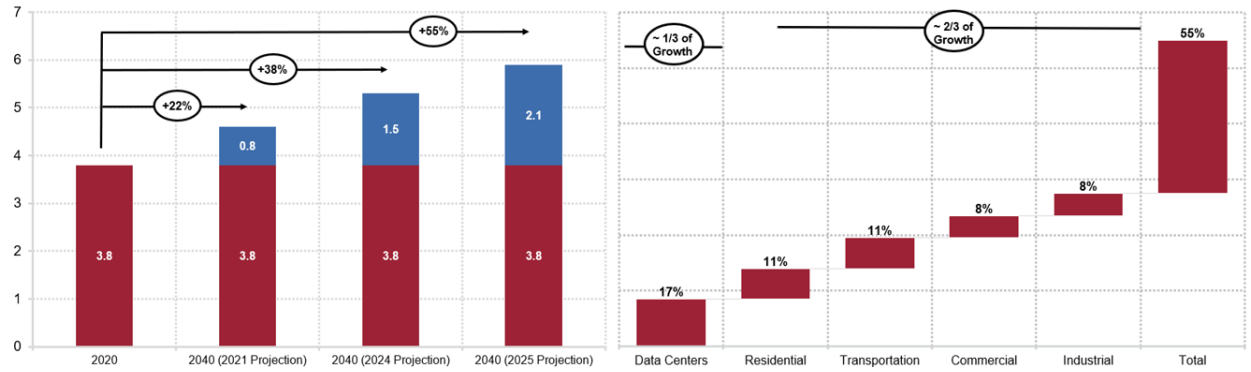
Source: EIA, Rosenberg Research

With the substantial growth in A.I. data center demand, U.S. electricity consumption is forecasted to increase by +44% from 2024 to 2040, which equates to a CAGR of +2.3%. In order to achieve this target, all forms of electricity generation will be needed, including aging coal plants that would otherwise be shut down. What is

most interesting about the chart below is how much the demand forecast has changed. The forecast issued in 2021 for terawatt hours (TWh) required in 2040 was 4.6TWh, but a forecast in 2025 now puts the requirement at 5.9TWh. That's a +1.3TWh increase in electricity usage (equivalent to five Californias).

Demand Forecast Growth Continues to Accelerate

United States
(left panel; red bars; historical U.S. power demand; thousands of terawatt hours)
(left panel; blue bars; incremental to 2020 U.S. power demand; thousands of terawatt hours)
(right panel; growth by sector in power demand; percent)



Source: NextEra Energy, EIA, Goldman Sachs, McKinsey, Rosenberg Research

In December 2024, the Lawrence Berkeley National Laboratory (LBNL) released a paper that examined the impact of data centers on U.S. electricity consumption. Data center demand currently accounts for 4.4% of electricity demand and is projected to grow to between +6.7% and +12.0% by 2028. Currently, there are six states that have over 10% of their electricity demand coming from data centers, with Virginia "Data Center Alley" leading the charge, at approximately 25%.

It will be a tall order to meet the growing electricity demand, and below we touch on the main forms of generation, namely renewables, natural-gas power, and nuclear.

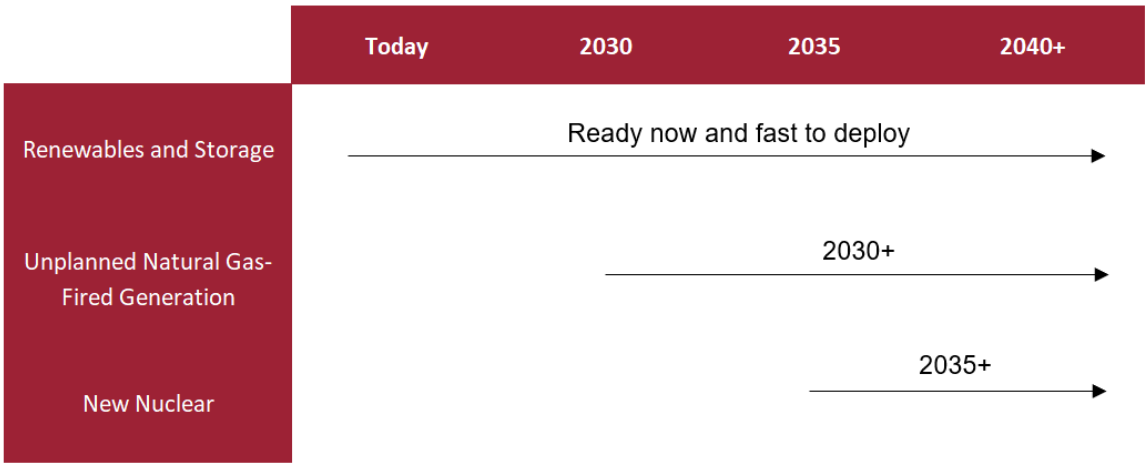
Solar (Renewables)

Renewables (solar and wind) accounted for 23% of U.S. electrical generation in 2024, with a large chunk of that being solar energy. The cost of providing solar power has declined by roughly 80% over the course of the last ten years. Solar has become the lowest-cost form of electrical generation, and is the easiest and quickest to deploy (under two years). Refer to the two slides from NextEra Energy, the largest publicly traded electrical company. Solar plus 4-hour battery storage systems (BESS) have become an important energy solution for many U.S. states.

However, there is considerable uncertainty surrounding Chinese tariffs, as over 80% of the world's polysilicon and solar panels come from China. Additionally, the Big, Beautiful Bill recent passed into law quickly phases out the 30% investment tax credit on solar projects. To top it off, there are restrictive FEOC (Foreign Entities of Concern) provisions in the bill that may prevent American power companies from buying cheap Chinese solar panels. This may result in a number of cancellations and/or deferrals of major solar projects.

Expected Deployment Timeline by Generation Type

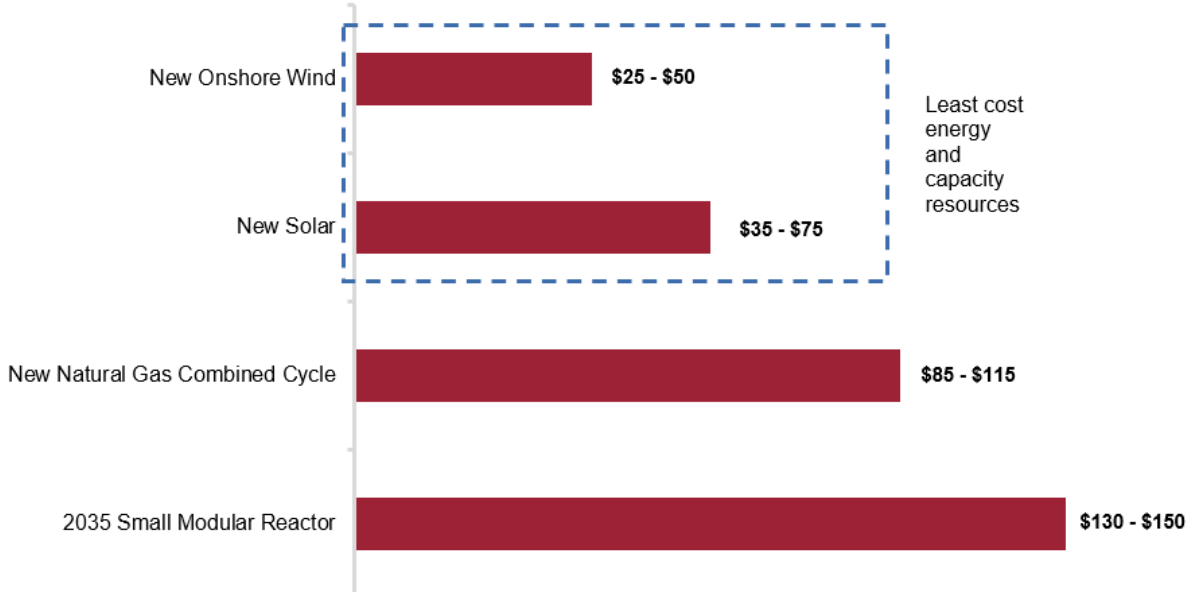
United States



Note: neither gas nor nuclear have supply chains running at scale — it is expected to take 5+ years to get natural gas plants built and running and nuclear is 10+ years away.
Source: NextEra Energy, Rosenberg Research

Estimated Costs of Firmed Generation Resources, 2030

United States
($ per MWh)



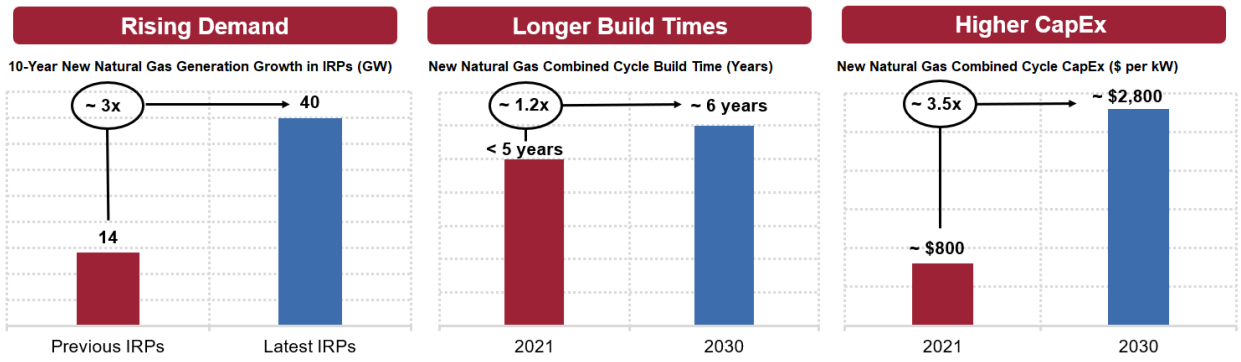
Note: renewables and storage are the most cost-effective energy and capacity solutions, and are ready now.
Source: NextEra Energy, Rosenberg Research

Natural Gas Power

Natural gas was the largest fuel source for electrical generation in 2024, with a 42% share. The U.S. has an abundance of very cheap natural gas, but unfortunately the lead times to build new natural gas plants have been stretched out to six years. Also, the cost has more than doubled since 2021. The culprit is that gas turbines are in very high demand, and costs have risen with rising steel prices, especially with high-performance metal alloys. The recent doubling of steel tariffs to 50% has not helped matters. Large gas turbine manufacturers like GE Vernova and Siemens Energy are running at full capacity and are cautious about expanding capacity as they have done so twice in the last fifteen years and have been burned. The chart below looks at both the increased costs and the extended build times.

Natural Gas Fixed Costs Are Rising

United States
(left panel; 10-year new natural gas generation growth in IRPs; gigawatts)
(middle panel; new natural gas combined cycle build time; years)
(right panel; new natural gas combines cycle capex; $ per kilowatt)



Note: natural gas-fired generation cannot meet demand in the near term and is a longer-term, more expensive solution.
Source: NextEra Energy, Rosenberg Research

Nuclear

Nuclear is a terrific solution to provide large amounts of stable baseload power for A.I. data centers, as we discussed in our recent report. Nuclear generation has long lead times, and the supply chain must significantly improve to bring on large amounts of new power generation. Despite the high capital costs of nuclear energy, it has an advantage as its plants have close to a 90% operating capacity versus 65% for the newer, most efficient gas plants.

The power-intensive data center industry is on the hunt for reliable baseload power, and there have been some interesting announcements over the past twelve months. This entails both restarting mothballed plants and extending the life of existing facilities. Some examples:

- Microsoft, in conjunction with Constellation Energy, will restart nuclear reactor #1 at the Three Mile Island site that was shuttered in 2019. This will provide 835 megawatts (MW) of power starting in 2028 with a 20-year Power Purchase Agreement (PPA). For those curious, this is not Unit #2 that famously suffered a partial meltdown in 1979.

- In early June, Meta (the owner of Facebook and Instagram) announced that they had entered into a 20-year PPA with Constellation Energy to extend the life of the Clinton Clean Energy Center. The nuclear plant in Illinois was at risk of closing in 2027. This large facility is 1,090 MW in size and provides enough power for approximately 800,000 homes.

- In conjunction with this announcement, Meta provided an update on their RFP process to bring on an additional 1,000 MW to 4,000 MW of nuclear power in the U.S. They are in final discussions with a shortlist of partners to make this goal a reality.

- The most promising aspect of nuclear is Small Modular Reactors (SMRs). These smaller reactors are typically 300MW in size compared to a 1,000 MW full-scale reactor. The newer reactors are simpler in design and should be more cost-effective to build. However, investors are rightfully worried about potential cost overruns for nuclear power projects. The Vogtle facility in Waynesboro, Georgia, brought two new large nuclear reactors on stream in 2023-2024, providing over 2 gigawatts (GW) of power. They were the first reactors built in the U.S. in more than thirty years. The project was plagued by design changes, massive cost overruns, and construction delays. All told, the final cost totaled $35 billion versus an original estimate of $14 billion. Westinghouse Electric, the prime contractor, declared bankruptcy, and the business was eventually sold to Brookfield Business Partners.

- The nuclear industry is closely watching Ontario Power Generation's Darlington SMR project, just outside of Toronto. It will shortly start construction on the first of four 300MW SMRs. The estimated capital costs for 1,200 MW of power are approximately $15 billion (U.S.) and are slated to come onstream in the 2029-2030 time period.

The Electrical Grid and Transmission

Electrical transmission carries electrons from generation sources like nuclear and natural gas power plants to consumers in cities, manufacturing plants, and data centers. Electrical transmission is a significant bottleneck for bringing on new generating capacity, as the process can take eight to ten years, or even longer. Time-consuming environmental assessments, permitting, and landowner approvals are needed to bring new projects forward. In summary, no one wants high-voltage power lines near or running through their property.

The U.S. electrical grid is old and is struggling to handle the existing electrical generating capacity, let alone the new capacity coming on stream over the next five to ten years. Whether it's power lines, transformers, or circuit breakers, many of these key components are in the back-half of their fifty-year useful lives. Driven by power-hungry data centers, electricity demand is increasing once again, raising the topic of grid stability to the forefront. A prime example of grid stability occurred just this April when major blackouts affected Portugal, Spain, and Southern France. It's somewhat disconcerting that they still don't know what caused these blackouts.

The challenge for the transmission industry is the sheer size of A.I. data center projects. As referenced earlier, it is common to see 1,000 MW A.I. data center projects these days. By comparison, the largest aluminum smelter in the U.S. is the Century Aluminum Sebree plant in Kentucky, which is owned by Glencore. It has a capacity of 220,000 metric tonnes per year of aluminum and is estimated to consume 370 MW of power, roughly one-third of a full-scale A.I. data center. The American grid wasn't designed to bring this large amount of generating capacity online in such a short period of time.

With long approval times for new transmission projects, major utilities in the U.S. have focused their capital investment on replacing existing transmission lines and older equipment while using grid-enhancing technologies (GETs) to expand capacity. Although overall transmission capex is at record levels, little capital has been spent on new high-voltage power lines over the last seven or eight years.

In 2023, the Federal Energy Regulatory Commission (FERC) published a report identifying ten major corridors that are in need of new high-voltage transmission capacity to provide electricity to underserved markets. Each of these electrical corridors are significant in scale and require billions of dollars in capital.

Ideally, President Trump and his administration will speed up the process to bring new transmission capacity on stream. This is already beginning to happen with nuclear energy as in May, the administration announced that they would enact policies to speed up the regulatory approval process by the NRC for new nuclear plants.

In a nutshell, the power industry needs to bring on as much electrical power as realistically possible, and it must do so in a very timely manner. Transmission lines must be upgraded to increase capacity, and new high-voltage transmission lines are needed. If all this doesn't happen, a number of large data center projects are at risk of being delayed. In this scenario, the IEA estimates that up to 20% of data center projects could be impacted.

Health Care and Drug Discovery

From a humanitarian perspective, using A.I. to create new medicines to eradicate disease is the most meaningful opportunity that lies ahead. A.I. may help provide significantly better healthcare and meaningfully reduce costs across the healthcare system.

According to data from the Centers for Medicare & Medicaid Services (CMS), Americans spent $14.9 trillion on health care in 2023, a +7.5% increase over the prior year. That represented 17.6% of GDP and $14,570 per person, which is about 50% more than other G20 countries like France, Germany, and the U.K.

Government spending on healthcare amounted to $1.9 trillion or approximately 28% of overall expenditures. Reducing these outlays could help improve the fiscal deficit which has been a hot point for fixed-income investors. A.I. can improve efficiencies across all aspects of healthcare, whether it be reducing drug discovery costs, improving supply chains, better diagnostics, or reducing administrative costs.

Drug Discovery

In February 2025, during a *In Good Company* podcast interview, David Ricks, the CEO of Eli Lilly said it costs just under $3 billion to bring a new drug to the market. He noted that in 2015, it took Eli Lilly four to five years from the idea for a new drug to the first dose of a clinical trial, then a further ten years of clinical development to get final approval. So, just under a fifteen-year process, in total. Eli Lilly has done an excellent job by cutting this time in half, but it still takes them two-and-a-half years from idea to first dose and six years of clinical development time. A drug is given a twenty-year patent life, so that leaves the drug company with ten to twelve years to generate returns in the market.

The biggest cost for a drug company is failure. Thus, anything that can improve the probability of success for new drugs and reduce the overall time to market will be of great help. This is where A.I. can really help with the process.

AlphaFold

One of the most fascinating projects in health science is the AlphaFold project by Google that maps proteins. Proteins are the building blocks of life, and each has a unique complex 3D structure. Traditionally, scientists would target one protein, but it would take several years and hundreds of thousands of dollars. Using A.I., AlphaFold is able to predict these highly complex protein structures in minutes and has already identified over 200 million protein structures, nearly all proteins known to science. The protein database is currently being used by over 2 million scientists in 190 different countries. For their part, the two founders of the AlphaFold project won a Nobel Prize in 2024.

Below, we provide a couple of examples targeting certain proteins that have a good potential to cure human disease and how a company like Isomorphic Labs is using A.I. in the process.

Potential Treatment for Drug Resistant Antibiotics

Scientists are targeting the resistance mechanisms of bacterial proteins. If they are able to block this mechanism, it should allow existing antibiotics to keep working. A.I. will also aid in the development of new antibiotics by modeling interactions between proteins and antibacterial compounds.

Parkinson's Disease

Parkinson's disease occurs when a person's brain can no longer produce enough of the chemical dopamine. Up to 9% of people with the early onset of Parkinson's disease have a genetic mutation called Pink1, and it impacts approximately ten million people. This defect allows the build-up of mitochondria, which prevents nerve cells from producing dopamine. AlphaFold was used to predict the Pink1 protein structure, and will potentially create a cure in the upcoming years.

Isomorphic Labs

Isomorphic Labs is a company founded by Demis Hassabis, the CEO of DeepMind, who is now leading A.I. at Google. The company uses A.I. models to mathematically prove that certain molecular designs will work in a clinical trial. Remarkably, they are able to do this with a high degree of accuracy. This should help improve the overall chance of success in clinical trials and reduce the overall duration of the testing phase.

On average, the FDA approves approximately fifty drugs per year. An interesting question arises: if you have A.I. models with the ability to accurately predict clinical trial outcomes, can the FDA one day approve many more drugs in a far shorter timeframe?

Diagnostic Imaging

One area that is benefiting from A.I. is diagnostic imaging, more specifically ultrasound. A recent GE Healthcare presentation stated that over 80% of health care systems have reported a shortage of radiology therapists. This will only become more pronounced as experienced ultrasound clinicians retire or leave the profession.

A.I. has the opportunity to create meaningful efficiencies. Firstly, A.I. guided ultrasounds have superior accuracy and reduce exam times. This results in a more favorable outcome for clinicians and patients. Secondly, the interpretation of data can be far more accurate and reduce the time to analyze results. Finally, A.I. can be used to automate the report writing process, further reducing administrative burdens for doctors.

Software Development

Coding

Writing software is a critical aspect for every industry, including healthcare, finance, and industrial automation. Recently at Meta's Llama conference, Microsoft CEO Satya Nadella stated that A.I. is now on average writing about 20% to 30% of the company's software, and he expects this number to move closer to 50% in 2026. Some of their software programs are now almost entirely written by A.I.

A recent WSJ article highlighted that one of the biggest issues in the software industry is rewriting old, outdated code. Outdated code is cumbersome to update as it is written in legacy languages, like Cobol. Decades old code also poses significant cybersecurity risks and is becoming more difficult to maintain as there is a lack of skilled people versed in the old programming languages.

Morgan Stanley launched its own software program called Dev GenAI that is built on OpenAI's ChatGPT models. It ultimately translates legacy code into modern computer languages.

Cybersecurity

Cybersecurity is a mission-critical issue for companies and governments across the globe. The economic impact resulting from cybercrime already totals in the hundreds of billions of dollars each year. For enterprises, this can result in significant brand damage, business downtime, and lower productivity. Bad actors are using A.I. to create increasingly sophisticated attacks using very realistic phishing emails and deepfakes. Lee Klarich, the Chief Product Officer at Palo Alto Networks, provided some insights at a recent event titled, "*Hello Tomorrow: What's Next in A.I. Security*."

With bad actors now using A.I., there has been an increase in daily cyberattacks to 8.9 million per day versus 2.3 million last year, nearly a 3-fold increase. The sheer scale of attacks highlights the level of concern. More importantly, bad actors using A.I. have seen a 99% reduction in time for a successful breach, declining to a mere five hours from forty-four days in the prior year.

All is not lost, though, as Cybersecurity companies are using machine learning to prevent cybersecurity attacks from occurring. In a sense, it is becoming a battle of the machines.

Mr. Klarich posed an interesting question: "*Does A.I. benefit the attackers or the defenders more?*" His view differs from the prevalent industry perspective that it will benefit the defenders far more.

In order to combat the increasing level of sophisticated attacks, Palo Alto Networks created an A.I.-driven platform called Cortex XSIAM. It has over 10,000 different detection models, with 2,000 built by A.I. itself. This

compares to only 800 models for the nearest competitor. This allows for quicker detection and response time to cybersecurity threats.

Nikesh Arora, the CEO of Palo Alto Networks believes the next phase for cybersecurity is creating A.I. cybersecurity agents to provide out-of-the box and bespoke solutions for customers. Their contribution currently under development is called Cortex Agent IX. Arora views it as a three to five year journey (likely longer) that will require a significant amount of development and trust to hand over important cybersecurity responsibilities to A.I. agents.

The $200 billion per year cybersecurity industry remains very fragmented. During a recent *In Good Company* podcast, Mr. Arora said Palo Alto Networks had a 1.5% market share seven years ago and was a top three player. With strong organic growth, coupled with numerous acquisitions, Palo Alto is now the largest cybersecurity firm with a market share between 4% and 5%. Mr. Arora thinks there is reason to believe there will be a 20% market-share player in the next five to ten years. Customers prefer consolidating vendors that have a platform approach with a full-suite of security products and capabilities. This allows customers to reduce costs and, more importantly, simplify their security infrastructure. With A.I.-powered ransomware attacks now being 100x faster, it is imperative to have real-time capabilities. This is especially important for critical infrastructure, like power and water plants, which have old computer systems that have been patched up with band-aid solutions over the years

.

Manufacturing

There have been numerous articles on A.I. successes for automation and improving manufacturing efficiencies. A prime example is how Audi used A.I. to automate spot-welding inspections. An individual car may have 5,000 spot welds and an assembly line may produce up to 300 cars per shift. Traditionally, Audi inspected welds using ultrasound sensors and employees randomly checked the vehicles on the production line. Audi developed an internal A.I. model, and with the use of sensors, was able to automate the process and increase the accuracy of spot-welding inspections.

The opportunity set for using A.I. to improve manufacturing processes is quite large over the long term. However, we would like to caution that this may take some time. In a recent Nvidia A.I. podcast titled *How Siemens is Bringing A.I. to Factory Floors*, some reservations about the timeline for rolling out A.I. solutions were expressed by Matthias Loskyll, head of industrial A.I. at Siemens Factory Automation (by way of background, Siemens is the global leader in providing industrial automation solutions to manufacturing companies).

Mr. Loskyll pointed out that 92% of the manufacturing companies lack skilled A.I. experts, which are currently rare and expensive resources. Training is going to be a big part of the ultimate solution, but this will take time. Trust also appears to be a big issue, as 40% of employees on the factory floor don't trust A.I. Making progress on this front is paramount since safety is such a big consideration for manufacturing plants. There have been some early successes, but in a recent survey, 70% to 80% of all A.I. projects failed to deliver adequate returns, and only 16% achieved their A.I. goals. Overall, the ultimate prize is huge, but it will take much training, a mindset shift, and time for the potential to be realized.

Digital Twins

Digital twins are virtual 3D replicas of physical assets such as factory floors, assembly lines, and office buildings. NASA pioneered the concept of the digital twins, dating back to the Apollo space program. It was famously used for the Apollo 13 mission that ran into trouble when the main engine and oxygen supply encountered problems. Space engineers were able to simulate the environment from 200,000 miles away and help the astronauts return safely to Earth.

The technology is obviously much improved since then. Using A.I. coupled with Internet of Things (IoT) sensors on factory equipment, 3D graphic computers can simulate a very realistic manufacturing environment. This can then be used for training new and existing employees, and very importantly, will help train future robots for when they eventually work on the factory floor. Digital twins can be used for predictive maintenance to monitor the performance of production lines and run various simulations, like replacing or adding new equipment to existing manufacturing facilities. All this will help improve efficiency and reduce maintenance downtime.

Robotics

Humanoid robotics is considered the next phase of the technological evolution and can be thought of as Physical A.I. Humanoid robots that are trained on A.I. and will have the potential to perform many everyday human tasks and reduce costs for a wide variety of industries.

Humanoid training is a complex undertaking as the movement of the human body involves numerous muscles and tendons, let alone brain power. The smooth movement of modern-day humans has been fine-tuned over of hundreds of thousands of years. Replicating this with humanoids has taken years of painstaking research to build lighter and ever more dexterous humanoids. Although significant improvements have been made, much more progress will be needed to enable humanoids to comfortably navigate human society.

It is fairly difficult to assess the scale of the opportunity and the ultimate timeline for mass adoption. The key here will be to meaningfully improve the capabilities of humanoids, all the while significantly driving down the cost curve.

There is a wide variability in industry views. Kudos to Adam Jonas and his team for providing some thoughtful insights as to the size of the humanoid opportunity and how to think of the potential timeline. Mr. Jonas is the head of Global Automobiles and Shared Mobility Research at Morgan Stanley. The following thoughts are summarized from the Morgan Stanley corporate website:

The Humanoid Market

Morgan Stanley estimates the humanoid robotics market at $5 trillion in 2050 based on 1 billion humanoids worldwide. The commercial and industrial market will be 90% or 920 million units of the total, while the household market will be 10% or 80 million units. China will be the largest market at 302 million units versus 77 million units for the U.S. The leadership by China may surprise many, but China aims to be the dominant leader in manufacturing, and robotics is a key enabler. Overall, the humanoid robotics market has the potential to be twice the size of the automotive market or even larger.

Price of Humanoids

Driving down the cost curve will be a key determinant in attaining mass adoption. In 2024, Morgan Stanley estimates the price of a humanoid robot being roughly $200k. By 2028, Mr. Jonas forecasts this cost will drop to $150k per unit and drop further to $50k by 2050. Significant improvements will be needed in the supply chain for this to occur.

Other Viewpoints

Within Tesla, there is a division that makes Optimus humanoid robots. If you visit a Tesla dealership mall location, chances are they have an Optimus robot in the front window. Optimus is 5'8" tall and weighs 150 pounds. Much larger is Elon Musk's outsized projections for Tesla's Optimus humanoids. He foresees production ramping up in 2026 and selling Tesla bots to external customers. He also envisions selling millions of units in 2029 at a price point of $20k to $30k. A characteristically bold forecast on the part of Mr. Musk, for sure. His anticipated price range is far below Morgan Stanley's forecast of $150k.

Marc Andreessen, a highly-regarded venture capitalist and tech visionary at the Reagan National Economic Forum in May said, "*The general-purpose robotics market thing, it's going to happen in the next decade and it's going to happen at giant scale.*" He went on to say, "*Robotics is going to be the biggest industry in the history of the planet.*"

Amazon (Industrial Robots)

Amazon is a Prime example (sorry for the pun) of how a leading company is greatly benefiting from A.I., and more specifically, robotics. Amazon's keen interest in robotics is two-fold. Firstly, to improve efficiencies in existing and new fulfillment centers, and secondly, to drive down the costs of the last-mile delivery to customers. In 2024, fulfillment costs at Amazon were $98.5 billion and they represented 18.6% of net sales excluding Amazon Web Services (AWS). Thus, even a -1.0% reduction in fulfillment costs results in significant cost savings. Amazon has always been obsessed with serving the customer, that is, offering a wide selection of products at the lowest prices and fastest delivery possible. Any efficiencies can be reinvested into the business to gain market share and further extend Amazon's market leadership.

Amazon already uses more than 750k robots across its facilities to sort, lift, and carry packages. The biggest opportunity is with Next-Gen fulfillment centers, like the one in Shreveport, LA. The facility is massive, 3 million square feet on five floors, which equates to over fifty football fields. It is powered by A.I., has 10x more robotics, and drives an estimated 25% improvement in productivity. Safety is improved and it provides 30% more high paying jobs in the area of maintenance and engineering. Overall, it's great news for Amazon and its customers, bad news for the competitors.

For the last-mile solution, there have been reports of Amazon constructing "Humanoid Park" in San Francisco. Amazon will develop its own A.I. software to train 3rd party robots. These robots will practice entering and exiting delivery vans and traversing different terrains at Humanoid Park. So don't be too surprised one day when you see a humanoid robot get out of a delivery van and deliver an Amazon package to your front door.

Neo Humanoid Robot (Household Market)

In 2025, 1X Technologies will begin selling the Neo Gamma humanoid robot. It was specifically designed for home use and has financial backing from OpenAI and Samsung. On average, people spend 2.3 hours per day doing various chores at home. Currently, NEO is proficient at vacuuming, able to do the laundry, getting better at folding, and capable of answering your door. The robot itself is 66 pounds and can lift up to 150 pounds. It will run semi-autonomously with tele-op, so privacy concerns are a key issue.

One should view this as a long expedition, as humanoids will need to train in the real world to become more proficient at simple tasks such as washing dishes or loading the dishwasher. Expect some broken dishes along the way as the humanoids learn. Safety and social acceptance are key issues for humanoid robotics to take hold. The safety issue is most critical, as most people would be reluctant to have a home robot around small children or pets. There are many issues to deal with, but the day of humanoids is coming. It will free up those 2.3 hours per day of chores and will be a big help in key areas like eldercare and companionship.

China

The Chinese government has identified robotics as a critical component of its "*Made in China 2025*" initiative. This initiative aims to make China the global leader in advanced manufacturing and robotics. China is currently the world's largest market for industrial robots, with over 50% of global installations.

China also aims for dominance in the supply chain of key humanoid robotic components such as actuators, sensors, motors, and controllers. They already have a 60% market share in this area. Increasing concerns about the reliance of the Chinese supply chain will surely mount as the humanoid market enters the mass adoption phase. This has already happened in renewable energy components such as polysilicon, solar panels, and EV batteries, where China dominates global production with a market share of over 80%.

Defense

With advancements in technology such as A.I. and drone effectiveness, the nature of warfare is rapidly changing, as we discussed in our recent Special Report. Military prowess has traditionally been defined by sheer size and capability. The U.S. was seen as the dominant military power, but this advantage has eroded as China has invested heavily in defense over the last couple of decades. The U.S. has tended to focus on larger, more expensive systems and the importance of maintaining air superiority with the F-35 fighter jet program. With first-hand observations of the Russian-Ukraine war, many military experts are now rethinking modern-day warfare.

In early June, Ukraine launched a surprise attack using 117 small drones that were smuggled into Russia. The Ukrainian attack used real-time visual guidance with onboard cameras, resulting in a high degree of accuracy in hitting vulnerable points on the aircraft. Each drone cost less than $1,000 and was able to hit forty-two strategic bombers with nuclear capabilities. It is estimated that this drone attack cost roughly $1 million, yet it inflicted $7 billion in damage. This asymmetric warfare can be thought of as a combination of the Trojan Horse fable, coupled with David and Goliath.

China and Taiwan

The potential Chinese attack on Taiwan is at the forefront in the minds of many geopolitical experts. Taiwan is critical for the global economy as it manufactures over 90% of the most-advanced semiconductor chips. They go into everyday uses such as computers, automobiles, smartphones, robotics, and A.I. data centers. You can see why reshoring critical manufacturing has become such an important topic for governments, especially for the Trump administration.

In April of 2025, there was a fascinating TED Talk with Palmer Luckey, founder of a defense technology company called Anduril Industries. Mr. Luckey founded Oculus, the virtual reality headset company, which was sold to Meta for $2 billion in 2014. In his TED Talk, he discusses how the large U.S. defense companies have slowed their pace of innovation, while China has built the largest army and navy in the world. He is on the record saying that defense systems need to be smarter and "*A.I. is the only possible way we can keep up with China's numerical advantage.*" A.I. will become critical for flying autonomous aircraft and will be used to design better weapon systems while speeding up the manufacturing of weapons. With A.I. offering much faster response times to attacks, it should ultimately be a strong deterrent and prevent war from happening in the first place.

Financial Services

There are many aspects of how A.I. will impact financial services companies such as traditional banking, credit card issuers, investment banking, and insurance.

Goldman Sachs has been using A.I. for years to streamline their operations and financial reporting. For example, David Solomon, CEO of Goldman, said that ten years ago, drafting an S1 IPO prospectus took a team of five to six people two weeks, including one very senior person. Now with A.I., they can write 95% of the prospectus in a few minutes. Goldman spends about $5 billion per year on technology, but would like to spend $7.5 billion, except that level of spending would reduce the company's ROIC. Goldman has 46k employees, with 11k of them being engineers and programmers. Huge cost savings could be attained if they could simply reduce software coding costs by 30%.

Mr. Solomon is in a unique position as he has the opportunity to meet many global CEOs who are important to the capital markets. He said all the CEOs he met stressed the importance of using A.I. to make their companies more efficient and productive. He thinks it's still early in the cycle, but that it is accelerating fast. Overall, he is super optimistic and excited about A.I., but also views it as very expensive and thinks costs will be driven down over time.

The financial services sector is very much a people business, and there are numerous opportunities to improve overall customer service with the use of A.I. powered agentic chatbots. In insurance, the claim handling process could be sped up with better analysis and A.I.-aided reporting. Improved fraud detection should benefit property and casualty insurers and credit card companies. In traditional consumer banking, streamlining back offices, improving regulatory compliance, better risk control, and superior customer service are just a few of the initiatives.

Overall, the A.I. opportunity set with financial services is large and diverse. The challenge here will be steering the ship of a very large organization that often has legacy computer systems. The proper handling of secure and sensitive personal information is a top priority.

Autonomous Driving

One area being revolutionized by A.I. is that of fully autonomous self-driving vehicles. It has taken years of development, but now they are beginning to roll out in large cities across North America. Although seeing a driverless car may at first seem disconcerting, the safety record is much improved, with over 75% fewer accident claims reported by insurance companies.

WAYMO (GOOGLE)

Waymo's roots date back to the Stanford University's entry in the DARPA Grand Challenge competition for autonomous vehicles in 2007. Larry Page and Sergey Brin, the co-founders of Google, attended the competition and decided to put a small team together in 2009 to pursue the opportunity. In 2015, the first completely autonomous ride on a public road took place, and Waymo was officially created as a company within Google in December 2016. Waymo One, the first ride-hailing public service, was launched in Phoenix in 2018. As of April 2025, Waymo is in four cities (L.A., San Francisco, Phoenix, and Austin), has grown to over 250k rides per week, and has surpassed 10 million cumulative rides.

Waymo is the world's most experienced driver with over 20 million miles of real world driving and over 20 billion miles of simulated driving. To drive the vehicles, Waymo uses complex A.I. trained computer models incorporating vast amounts of data from various sensors and highly detailed customized maps. The 6th generation Waymo Driver has the following specs:

Cameras

- Thirteen cameras (29 cameras in 5th gen.) provide a 360-degree view around the vehicle, removing all blind spots. The cameras work in low light conditions, and the long-range cameras can see up to 500 metres. This is important for seeing stop signs, traffic lights, and pedestrians.

Lidar

- Four lidar sensors that provide a detailed 3D picture of the surrounding vehicle, no matter what time of day. It does this by emitting millions of laser pulses and measuring the time for the light to bounce back.

Radar

- Six radar sensors that are effective in all weather conditions.

Onboard A.I. Computer

- Waymo Driver takes all the complex data from the sensors and uses A.I. to determine the best route all the while monitoring surroundings. It even uses A.I. to predict how other drivers will react on the roadway.

Although, Waymo is only currently available in four cities, there are aggressive plans to quickly expand to other cities like Atlanta, Washington D.C., Miami, Las Vegas, and San Diego in 2025 and 2026.

TESLA ROBOTAXIS

Much like the Optimus Robot, Mr. Musk has grand visions for the future of Robotaxis. The first pilot program launched in June 2025 with service in Austin, Texas, using ten to twenty vehicles (Model 3 and Model S). The Tesla autos will likely have nine cameras versus thirteen cameras for Waymo and have no lidar or radar sensors. Tesla will rely much more heavily on machine learning to drive the vehicles. Safety considerations will be key here, especially in difficult driving conditions like rain and darkness.

The Austin pilot has received considerable attention, but a few words of caution. The service will be geofenced, limiting it to certain parts of Austin. It will also avoid complex intersections and use remote tele-operators if a Tesla vehicle should run into problems.

Tesla's Robotaxi (also known as Cybercab) does have two big advantages. The first one being cost, Mr. Musk's goal is to drive down total costs to under $30k per vehicle by 2027. This low cost would allow a speedy introduction to major markets, which is a multi-billion dollar opportunity. For Waymo, on the other hand, the 5th generation costs were well above $100k, prompting Mr. Musk to call it "WayMore." However, for the recently released 6th generation version, the cost is likely to come in under $100k. Secondly, Tesla has vast troves of driving data as there are over 2 million of Tesla's full self-driving (FSD) vehicles on the road with over 15 billion miles driven.

Overall, it will be extremely interesting to watch the battle unfold between two industry titans.

Economic Impacts

As mentioned in the various industry segments, the opportunities for meaningful productivity gains are numerous and diverse. This will undoubtedly have many positive effects on U.S. economic growth and the global economy in general. The build-out of the A.I. infrastructure will require massive amounts of capital totaling trillions of dollars. A recent Dell'Oro study projects capex on data centers will exceed $1 trillion by 2029. We are already seeing the positive impact on U.S. GDP in 2025Q1, via IT and construction spending that led some economists to say it added one whole percentage point to the GDP growth rate. But let's not get carried away with our expectations as the growth rate in 2025Q1 may prove to be too torrid a pace for the entire year.

Labor Market Impact

A.I.'s impact on the job market is one of the most hotly debated topics, and opinions on this are wide-ranging. In a late May Axios interview, Anthropic CEO Dario Amodei stated that A.I. could wipe out half of all entry-level white-collar jobs and spike unemployment to between 10% and 20% in the next one to five years. However, in the same week, Nvidia CEO Jensen Huang stated in an interview with CNBC's Jim Cramer that A.I. is a generator of jobs, and that there will be a shortage of thirty to fifty million workers by the end of the decade. A.I. will help create new jobs and grow overall GDP. Sam Altman, CEO of OpenAI, has a more measured approach on the topic: in an interview, while touring the Stargate data center, he declared A.I. "*will take a bunch of jobs away and create a bunch of new ones. That is what technology does. The thing the world is not ready for, is the pace of change.*"

The NYT published an article in May titled ***For Some Recent Graduates, the A.I. Job Apocalypse May Already Be Here***. It highlighted the challenge of recent graduates finding jobs with an unusually high unemployment rate of 5.8%, with a high concentration in the finance and computer science sectors.

Financial Markets Implications

Equity Markets

If global investors want exposure to the leading Technology companies, they are going to have to come to the U.S. market. That is investing in the likes of Microsoft, Amazon, Alphabet (Google), Nvidia, Apple, Meta (Facebook), Tesla, and many other U.S. Technology companies. This has led Wall Street to coin the term "Magnificent 7" to describe the premier Tech companies mentioned above. With the strong performance of Technology stocks in recent years, the weighting of Technology within the S&P 500 has increased to over 30%. The forward P/E for Technology stocks is now 30x vs 22x for the overall market, that's roughly a 30% premium for the Tech group. This is not surprising, as the leading Tech companies are wonderful businesses. They have far superior growth prospects and terrific franchises, all the while generating substantial free cash flows. For example, take a look at Microsoft, a $3.7 trillion market-cap company. In fiscal 2025Q3, their most recent quarter, the company reported revenues of $70.1 billion, which grew +13% YoY, while net earnings grew by +18%. It is hard to find growth like that elsewhere in the market.

What may surprise investors most is that this trend may continue with P/E multiples for the best-of-class Technology stocks expanding further. One reflects back to the Nifty Fifty era of the late 60s to early 70s when a group of about fifty large-cap, high-growth stocks were the darlings of Wall Street. With the euphoria, the Nifty Fifty stocks traded at premium P/E multiples in excess of 100% of the overall market. A word of caution though, it ended badly for investors lingering too long in the group.

Infrastructure, Private Credit, and Private Equity

Infrastructure

There are two of important megatrends happening in the world right now, infrastructure spending and the reshoring of critical manufacturing. In order to build-out the infrastructure for artificial intelligence, a tremendous amount of capital is needed to construct facilities for power generation, transmission, and data centers. Governments around the world are financially strapped, utilities have limited capital budgets, and the project sizes are often too big for even the largest Tech companies. The Tech companies also lack expertise in developing large-scale power projects. In order to bring on all the electrical power generation needed, partnering with infrastructure companies with deep expertise will be a key part of the solution.

Total Infrastructure Investment Need

United States: *by sector between 2024 - 2040*
($ trillions)

Sector	Infrastructure Investment Need
Roads	25
Energy	21
Rail	8
Telecom	6
Water	4
Airport	2
Ports	2
Total	**68**

Source: BlackRock, Rosenberg Research

Layered on top of this, governments are deciding that critical manufacturing such as semiconductors should be done domestically, especially considering potential instability in various parts of the world. Access to critical materials, such as rare-earth magnets, is also becoming increasingly important. The mega-trend of reshoring and adjusting supply chains is a major endeavor and will unfold over the course of many years.

Another important theme is investors increasing their portfolio exposure to infrastructure. Infrastructure offers solid returns, inflation protection (with much less volatility), and low correlations with other asset

classes. Large institutions, such as pensions funds, sovereign wealth funds (SWFs), endowments, and life insurance companies have been increasing their exposure for years. The next big trend could very well be individual investors adding lower-risk infrastructure assets into their portfolios. On this note, we highly recommend people read the *2025 Annual Letter to Shareholders* by Larry Fink, the chairman of BlackRock. Mr. Fink discusses moving away from the traditional 60/40 asset mix of equities and bonds to a 50/30/20 mix, with the 20% being private assets (infrastructure, real estate, private credit, and equity). He also discusses the important topic of potential tokenization of individual assets such as infrastructure. A move into hard assets has been **one of our key structural investment themes**.

Private Credit

Private credit is considered to be a relatively new asset class that has experienced substantial growth over the last fifteen years. With the huge amount of infrastructure required, the investment-grade private credit market should grow considerably, especially in the areas of power generation and digital infrastructure.

Private Equity

Even though private equity firms have many mature investments waiting to be IPO'd or sold, they have substantial "dry powder" for future investments. The whole A.I. trend may create numerous opportunities for private equity firms. Put yourself in the shoes of a CEO of an intermediate industrial company, wondering how you are going to navigate the world of A.I. with increased automation and robotics to achieve productivity gains. This may require substantial amounts of capital and the importance of strategic long-term thinking (please read the first Amazon shareholder letter by Jeff Bezos in 1997). This will be a difficult proposition for some publicly traded companies where the emphasis by Wall Street is often on short-term results. Certain companies may decide to sell their business to private equity firms who can provide capital, have deep industry knowledge, and can strategically plan over the long-term. This may entail selling the entire company or perhaps selling a major division (corporate carve-out).

Investing Opportunities

Below, we offer some comments on where exactly the leading companies intend to invest in the future to benefit from the mega-trend of Artificial Intelligence. These are companies that have superior business models, excellent growth prospects, are well-managed, and generate substantial amounts of free cash flow. We won't comment on valuation levels here, but to address this concern, one could consider adding to them during periods of market weakness.

Semiconductors and Equipment

NVIDIA (NVDA-Q)

The company is the industry leader in high-end graphics processing units (GPUs), like the Blackwell GB200 and GB300s. It's unfair to call Nvidia just a chip company because they build integrated systems that are being used in the buildout of large A.I. data centers. Mr. Huang, the CEO and founder of the company, is a true visionary with respect to high-performance computing and artificial intelligence. The risk to watch here is that gross margins of approximately 70% are likely not sustainable in the medium-term due to emerging competition.

BROADCOM (AVGO-Q)

The company is benefiting from the substantial growth in Application-Specific Integrated Circuits (ASICs) that are used to build data centers. While both ASICs and GPUs are considered to be accelerators, ASICs are designed for specific uses, allowing for better processing efficiency, lower costs, and reduced power consumption.

TAIWAN SEMICONDUCTOR (TSM-N)

TSMC is the global leader in contract semiconductor chip manufacturing and has a dominant market share in the most advanced chips (5nm, 3nm, 2nm) that are essential for artificial intelligence, smartphones, data centers, and automotive uses. TSMC just reported their May 2025 revenues, which grew +40% YoY.

ASML (ASML-Q)

The industry leader in semi-capital equipment with their unique Extreme Ultra-Violet (EUV) technology. EUV technology is critical in the manufacturing of high-end semiconductors, and ASML is considered to have an eight to ten year lead on the competition.

AMD (AMD-Q) — Honorable Mention

Honorable mention goes to AMD as the company has significant technical semi-chip capabilities and is very well managed with Lisa Su as the CEO. AMD has clearly fallen behind Nvidia in high-end GPUs, but has good opportunities to make inroads. As an aside, Lisu Su and Mr. Huang are first cousins.

Hyperscalers

Hyperscalers operate a massive global network of data centers, used for cloud computing and offering services such as storage, networking, and computer processing. For A.I. firms, hyperscaler infrastructure has become critical due to its ability to process huge amounts of complex data. The leaders here are Amazon Web Services, Microsoft Azure, and Google Cloud.

AMAZON (AMZN-Q)

Amazon is the industry leader in cloud computing with close to a 30% market share, and its annualized revenues from 2025Q1 surpassed $100 billion. Amazon also has A.I. exposure through their investment in Anthropic (Claude LLM) and uses A.I. for robotics in their Amazon fulfillment centers. Amazon is just rolling out Alexa Plus, an A.I. personal assistant that is purported to be a significant upgrade over the previous version.

MICROSOFT (MSFT-Q)

Microsoft has many different touch points with respect to A.I. The Microsoft Azure cloud computing platform is significant in scale (over 20% market share) but trails that of industry leading AWS. Microsoft has their own LLM model in Copilot and a large investment in OpenAI, the owner of ChatGPT. Satya Nadella, the CEO of Microsoft is one of the top global executives.

Other Leading Tech. Companies

ALPHABET (GOOG-Q)

Google is a diverse technology company with numerous exposures to A.I. Google Cloud is a number #3 player in cloud computing with roughly a 10% market share. They have an A.I. research division named DeepMind and their own LLM model called Gemini. Like Amazon, they also have an investment in Anthropic. Google has exposure to self-driving autos through Waymo. The risk here is that A.I. will cannibalize Google's traditional search business. My view is that this concern is already reflected in the share price and one should look at the sum-of-the-parts valuation for the company.

META (META-Q)

Meta is a large technology company with a significant advertising platform. Over 95% of the company's revenues are derived from advertising. Meta is investing heavily in A.I. as it is integrated into key platforms like Facebook, Instagram, and WhatsApp. Meta also has their own LLM called Llama and an A.I. personal mobile device in the form of Meta Ray-Ban AR glasses.

Noticeable Absentees

APPLE (APPL-Q)

Apple did not make the list as their A.I. initiatives have underwhelmed. Siri A.I. has been a total flop, and they are badly lagging in developing their own LLM. We wouldn't be surprised if Apple fills a strategic gap by acquiring Anthropic or Perplexity. Apple manufactures great consumer products and generates a significant amount of free cash flow. One can never count Apple out, but it makes you wonder if they have lost their way.

TESLA (TSLA-Q)

Tesla has tremendous potential with Robotaxis and the Optimus Robot. Expectations are exceedingly high though, and it appears Tesla has lost focus on its core business of electric vehicles (EVs). Tesla was once the world leader in EVs and has ceded that to BYD. The Chinese auto maker has cheaper cars, a wide selection with much improved design and performance. But one can never underestimate Mr. Musk and his ability to motivate his teams to drive manufacturing improvements and develop new innovative products.

Alternatives and Infrastructure

Brookfield Asset Management (BAM-N)

Brookfield is a leading investor in alternative assets with over $1 trillion in assets under administration (AUM), and it has plans to double that over the next five years. The company is one of the world's largest investors in the A.I. value chain with investments in key areas such as power generation (has a leading renewables business), logistics, and semi manufacturing. For example, Brookfield recently announced that they will invest $10 billion to build a large new data center in Sweden. Brookfield is well-managed under the leadership of CEO Bruce Flatt.

Other Companies to Consider

In addition to Brookfield, one should consider other large asset managers that have strong private credit businesses with exposure to A.I. infrastructure projects. This would include Apollo Global Management (APO-N), BlackRock (BLK-N), and Blackstone (BX-N). They are terrific businesses with excellent management teams.

Within the power infrastructure sector, GE Vernova (GEV-N) and Siemens Energy (ENR-FF) are helping to electrify the world. Both are top three players in gas turbines manufacturing and have lucrative services businesses. Siemens also has an attractive high-growth grid technologies business with revenues growing +34% YoY. The biggest blemish for both companies has been the offshore wind business where they have lost billions of dollars due to bad contracts, quality, and warranty issues.

Electrical Power Companies

There are numerous electric power utilities to choose from in both the U.S. and Canadian markets. These companies often have large regulated rate bases and have stable earnings growth in the +4% to +7% range with a dividend yield of 2% to 3% or more. For long-term investors, this sector can provide close to 10% returns.

Examples of companies: NextEra Energy (NEE-N) is North America's largest utility, and Constellation Energy (CEG-N) is the largest U.S. nuclear power producer.

Assessing The Risks of A.I.

Given the high stakes, a further discussion on the risks of A.I. is warranted. There is a possibility that A.I. could be used by bad actors to create incredible harm to human society. Anyone can drum up nightmare scenarios like using A.I. to create dangerous biological weapons and A.I.-enabled cybersecurity attacks to shut down critical infrastructure, i.e. Zero Day scenarios. The jobs impact could be disastrous if unemployment levels reach 10%+ around the world. Or one could go and envision full-blown Sci-Fi scenarios with robots taking over the world and computers refusing to be shut down. Maybe it's time to rewatch *2001: A Space Odyssey*, which was made in 1968, and think about Hal the computer.

There are many questions to ponder, such as:

- Should A.I. have some form of regulation?
- Should the pace of A.I. be slowed down so people can have a better understanding of the risks?
- What happens if electric grids become strained and experience blackouts due to power-hungry data centers?
- Should A.I. companies have an independent safety board, and how to ensure that the board remains independent?
- Should one person be allowed to control an LLM, like Mr. Musk has control over xAI and Grok?
- Advertising surely is coming to LLMs; if you have a highly influential companion chatbot, how independent will the advertising really be?
- What to do if A.I. decimates the job market?

In our coming sequel to this in-depth report, we will address these complex issues head on.

Appendix: Introduction to A.I. Technologies

What Is A.I?

A.I. is the development of computers, systems, or machines that have the ability to perform high-level tasks that normally require human intelligence. This involves learning, reasoning, and ultimately decision-making. At its core, an LLM like ChatGPT is essentially a neural network that uses advanced algorithms to analyze vast amounts of data to create text, generate images, and most recently produce video content. A.I. can be used to create models/simulations for important areas such as manufacturing processes, mapping detailed 3D proteins in biotech, and forecasting climate changes. A.I. models are becoming larger, more complex, and increasingly computing-intensive with the advent of powerful advanced semiconductor chips.

Key Points in A.I. History

AlexNet (2012)

AlexNet was an A.I. project at the University of Toronto. It was overseen by Professor Geoffrey Hinton and two students named Alex Krizhevsky and Ilya Sutskever. The U of T team entered an A.I. competition called ImageNet whereby the models were tasked with properly sorting 1.2 million different images into various categories. The U of T team astonished the A.I. world by handily winning the competition with an error rate of 15.3% vs the previous record of 26%. The breakthrough here was using two Nvidia GPUs (Graphics Processing Units) in parallel instead of CPUs for training their model. This was the "aha moment" for the A.I. industry and Nvidia itself. At the time of writing, Nvidia has become the most valuable company in the world, ranked by market cap.

Professor Hinton is often referred to as the Godfather of A.I. and won the Turing award in 2018 which is considered to be the Nobel Prize of computing. Ilya Sutskever was a co-founder of OpenAI along with Mr. Altman and Mr. Musk and became the company's first Chief Technology Officer.

Google Transformer (2017)

Google released a paper detailing their new Transformer architecture, which was an important step in natural language processing versus the traditional method of sequential learning. Think of it like reading a book where humans start on page 1, sentence 1, and read it sequentially. The Transformer architecture enabled A.I. models to read books all at once and focus on key words regardless of its position. It allows A.I. models to quickly formulate key thoughts on a wide variety of topics. This results in far superior analysis and much quicker responses.

ChatGPT (2022)

ChatGPT was released to the public in November 2022. It started out as a chatbot model and has developed into a highly complex reasoning model over time. In a recent CNBC interview, Mr. Huang (CEO and founder of Nvidia) stated that reasoning models require 100 to 1,000 times the level of computing than previous models.

The growth of ChatpGPT has been nothing short of phenomenal. It reached 100 million weekly active users in late 2023, hit 200 million in August 2024, and 500 million in March 2025, as the capabilities (image and video generation) of the model improved. Some observers think the current totals are between 800 million and one billion users. A SimilarWeb study measuring the Internet traffic of desktop users highlighted that in February of 2025, ChatGPT was the 5th most used website in the world. ChatGPT had significant traffic growth in the month, while all others in the Top Ten suffered traffic declines.

Artificial General Intelligence (AGI)

Currently, the most recent reiteration of LLMs are considered to be very good reasoning models capable of many complex tasks. The next step is when these models achieve artificial general intelligence (AGI), the level of overall intelligence that surpasses that of human beings. It is a very subjective question, but signifies an important milestone for A.I. models. In a recent interview on the Big Technology podcast, the question of whether AGI can be achieved by 2030 was asked to both former Google CEO Eric Schmidt and Demis Hassabis, the CEO of Google DeepMind. Eric Schmidt said yes, just before 2030, while Mr. Hassabis felt it would happen just beyond 2030. Either way, that milestone will be reached relatively soon.

Key Enabling Technologies (SEMIS And EUV)

Before further delving into A.I. data centers, we first want to touch on three amazing companies with very unique technologies that enabled A.I. to get to where it is today. The first company is Nvidia, which designs high-end GPUs and high-performance computing systems, essentially the infrastructure of today's A.I. models. The second company is Taiwan Semiconductor (TSMC), which manufactures over 90% of the world's most advanced semiconductor chips for Nvidia, Apple, Qualcomm, and other companies. You can understand why TSMC has raised significant geopolitical concerns on the topic of reshoring key manufacturing capabilities back to places like the United States. Lastly, there is ASML, which manufactures semi-cap equipment by using EUV technology, which enables the production of the most intricate and powerful semiconductor chips.

NVIDIA (GPUs)

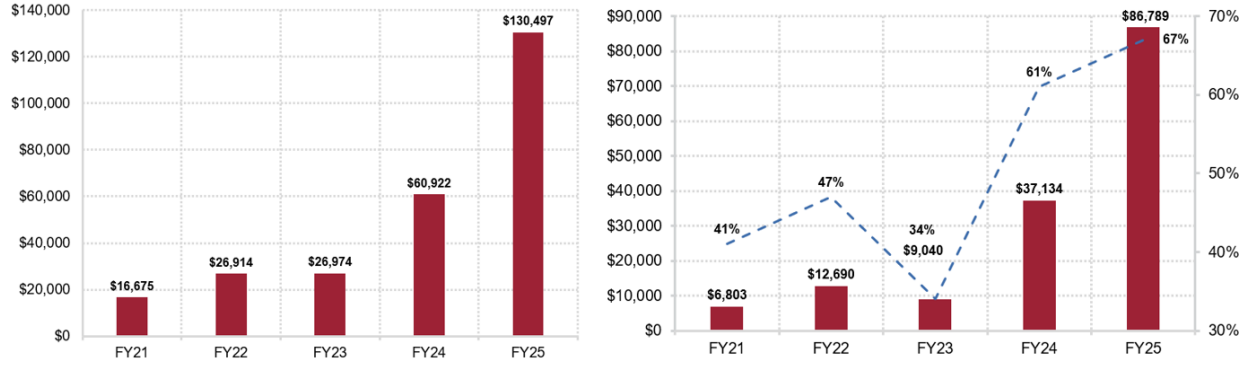
Nvidia was founded in 1993 and, over many years, grew into the market leader, selling GPUs to the gaming, automotive, and data center industries. Data centers have been around for years but have increased in importance with the significant growth of cloud computing. Rather than companies building out and maintaining their own computing infrastructure, they outsourced this aspect to companies such as Amazon Web Services.

Nvidia's fundamentals dramatically changed with A.I. companies and their large language models taking hold. A.I. data centers need to be much larger and have significantly more computing power compared to their cloud computing counterparts. In essence, A.I. data centers need the most advanced GPUs, such as Nvidia's Blackwell GB200 and GB300 chips.

In the chart below from a recent Nvidia corporate presentation, you can clearly see how dramatically the company's fortunes have changed. In fiscal year 2021, the company generated $16.6 billion in revenue, but that number grew to $130.4 billion a mere four years later. Meanwhile, their gross margins expanded from 41% to 67% during that time frame. The result was that net profits exploded from $6.8 billion in 2021 to $86.8 billion last year. Equally as important, the company generated $60 billion in free cash flow in fiscal year 2025. The Nvidia growth story has truly been remarkable!

Nvidia's Financial Performance

United States
(left panel; Nvidia revenue; $ millions)
(right panel; Nvidia operating income; $ millions; LHS)
(right panel; Nvidia operating margin; percent; RHS)



Source: Nvidia, Rosenberg Research

TAIWAN SEMICONDUCTOR (Semi Manufacturing)

TSMC is arguably one of the most important companies in the world. They are the world's largest contract manufacturer of semiconductors with approximately a 60% market share. Companies like Apple and Nvidia design their own chips and TSMC exclusively manufactures them. Critically, the company holds a dominant market share in manufacturing the most advanced semi chips, the latest being the 2 nanometer (nm) design. They have over 20 semiconductor manufacturing facilities (fabs) in various parts of the world, but are highly concentrated in Taiwan.

With governments realizing the importance of having a secure supply of advanced semiconductors, TSMC has embarked on an aggressive capital program to build fabs in the U.S. TSMC initially announced in April 2024 that they would build three fabs in Arizona for $65 billion, with a $6.6 billion subsidy from the CHIPs and Science Act. The first fab has already started production, while the second is currently under construction. In March 2025, TSMC announced an additional $100-billion investment program in the U.S. for 3 more fabs, 2 packaging facilities, and a major R&D center.

ASML (Semi-Capital Equipment)

In the small town of Veldhoven, Netherlands, resides ASML, also one of the most important companies in the world. They manufacture semiconductor equipment for customers like TSMC. ASML has developed and commercialized EUV technology that is used to etch intricate patterns on high-purity silicon wafers, which have an ultra-thin coating of a photo-sensitive polymer. The NXE3600 EUV system is roughly the size of a school bus and costs between $150 to $200 million. Some of the most advanced units like the EXE5200 can cost upwards of $400 million.

The technology is truly remarkable. ASML zaps a tiny droplet of molten tin with a specialized laser three different times at 50,000 pulses per second. The tin droplet is vaporized, which creates an extremely hot plasma at roughly 500,000 degrees Celsius. The hot plasma emits light with a wavelength of 13.5 nanometers. The EUV light is then concentrated and reflected using ultra-smooth Zeiss mirrors, which are among the smoothest objects humans have ever made. To put that 13.5nm wavelength into perspective, imagine the width of a human hair being the length of a basketball court. Then imagine a grain of sand on that basketball court, that is 13.5 nm.

The 13.5nm EUV technology is then used to etch advanced 4nm semiconductor chips like Nvidia's Blackwell GB200s. ASML is the only company in the world that is capable of producing EUV systems on a commercial scale and is considered to have an 8 to 10 year lead on the competition. Watch China though, as there have been recent articles on Huawei developing a semi chip using their own EUV technology.

To give you a sense of where the technology is going, TSMC will be tapping out 1.6 nm chips for its customers in 2025.

A.I. Models (LLMs)

There are currently dozens of major commercial LLMs like ChatGPT. We will focus on the largest LLMs and provide some snippets. As you can see below, almost all the big tech companies are involved as they cannot afford to miss out this powerful trend.

OpenAI (ChapGPT)

ChatGPT is the largest and most used large language model in the world. It is owned by OpenAI and the last funding round in March 2025 valued the private company at $300 billion. To put that into perspective, it would put OpenAI in the top 50 of the largest market-cap companies in the S&P 500. It has received significant funding from both Microsoft and Softbank.

Google (Gemini)

Gemini is Google's A.I. model and it is integrated into services such as Google search and Gmail. Google DeepMind is being praised for its new generative A.I. video model called VEO3. VEO3 will make it easy and inexpensive to create realistic generative video ads that likely will disrupt the advertising industry.

Grok (xAI — Elon Musk)

Grok is estimated to have a much smaller user base compared to ChatGPT, but provides very nuanced responses on a wide variety of topics.

Microsoft (Copilot)

Microsoft has its own LLM called Copilot and exposure through their investment in OpenAI.

Anthropic (Claude)

Anthropic's LLM called Claude has received significant financial backing from both Google and Amazon. The CEO, Mr. Amodei, has been one of the most outspoken A.I. executives about safety and the potential disruptive effects that A.I. could have on the global economy.

Meta (Llama)

Llama is Meta's LLM that is integrated into Facebook and Instagram. Meta has delayed Llama 4 as insufficient improvements were being made. This has led A.I. industry observers to seriously question scaling laws, or the ability for the industry to create new and improved models.

Training

This is the most intensive part of developing LLMs. The data surrounding this tend to be a closely guarded secret by the A.I. companies. However, most analysts assess that models like ChatGPT are trained for several months on very large GPU clusters. Think of training like going to graduate school, but instead of focusing on one subject matter, the A.I. model will learn about all subjects in all of history. Once trained, models will periodically undergo additional training to improve overall capabilities.

Inference

Inference is far less computing intensive than training. The A.I. model will answer queries from users by generating nuanced text, and when paired with an image generating model, it can create a wide variety of images such as charts, cartoons, and avatars. ChatGPT, with a user base of 500 million, still requires a very large amount of computing power. The reasoning models of today are able to provide much insight and detail on a wide variety of topics.

A.I. Devices

Currently, most people access consumer-based A.I. services on the desktop or on their smartphone devices. It is very easy to download an app for ChapGPT, Gemini, Copilot, or Grok, and in minutes you can start using the service. As mentioned earlier, ChatGPT has become wildly popular. The service is very easy to use, with new features added all the time, like image and video generation.

New form factors hold the promise to make the A.I. user experience much better. AR glasses like Meta Ray-Bans with a built-in camera and speaker are gaining in popularity. You can go to the local mall and try them out at the Ray-Ban store. There was a recent article in TechCrunch that hinted Apple will come out with their own AR smart-glasses at some point in 2026. Hopefully this will end the trend of people walking with their heads tilted downwards looking at their smartphones.

In May, OpenAI announced an acquisition of io in a $6.5-billion all-stock transaction. Startup io is Sir Jony Ive's design company, and if you remember, he was Apple's chief designer for the iPhone and MacBook. During a nine-minute video, Jony hinted at an entirely new form factor that is currently under development. It will likely be a small screenless A.I. device that fits in your pocket, or you can put it on the table. It could also possibly be a necklace with a pendant or a clip-on button one could wear on their shirt. With having a long-term memory, it can be considered a companion device that you can take everywhere and get to know you over the years. Mr. Altman, OpenAI's CEO, has tested the product for over a month and was extremely impressed, while Jony Ive more or less called smartphones "legacy devices."

A.I. Infrastructure

The Stargate 1.2 Gigawatt A.I. Data Center

As mentioned earlier, A.I. data centers are physically much larger and a lot more computing-intensive compared to their cloud computing counterparts. They are typically described by power levels (i.e. 1.2 GW) rather than square footage, as power is such an important consideration. On average, a typical ChatGPT query uses approximately 0.34 watt-hours or roughly 30 to 100 times more electricity than a standard Google search. In other words, a ChatGPT query is equivalent to using a low-wattage lightbulb for a couple of minutes. It was common for A.I. data centers to be under 100 MW, but now they are being constructed at 1,000MW to 1GW or larger. To put 1GW into context, that is the power from one full-scale nuclear plant that could provide enough electrical power for 750,000 homes. It is commonly thought that a high-end 1 GW A.I. data center with associated electrical power could cost upwards of $40 billion to $50 billion However, in a presentation at Computex in May of 2025, Nvidia CEO Mr. Huang referenced a slide that showed the new 1.2 GW Stargate data center that would cost a mind-blowing $60 billion to $80 billion, with $40 billion to $50 billion of that being for computing. Nvidia's advanced Blackwell GPU chips would account for a significant percentage of the overall costs.

The Stargate Project is a massive $500-billion project that was announced at a press conference by President Trump in January 2025. It is a joint venture between OpenAI, Softbank, Oracle, and MGX (UAE-based investment firm). The first phase ($60-$100 billion) is currently under construction in Abilene, Texas. The 1,200 acres facility will initially house 8 data centers that are each 500,000 sq. ft. (8 ½ football fields) in size. The long-term plan is for the joint venture to build other data centers across the U.S. and perhaps internationally.

In Nvidia's 2026Q1 conference call, Mr. Huang said there are approximately 100 A.I. data centers under construction around the world, with lots more yet to be announced. It appears many countries will want to have their own data center infrastructure for security purposes. Expect many announcements of new data center projects, especially in Europe, in the upcoming months.

As an aside, the human brain is a remarkable creation and consumes an estimated 20 watts of power.

Next, we'll look at the basic components in a data center, to give you a better sense of the sheer scale and power-intensive nature of A.I. data centers.

Compute Rack

A compute rack, or rack, is essentially a cabinet that holds various compute trays (GPUs and CPUs), switch trays, and other networking equipment. In the visualization above, 8 compute racks are combined together to form a compute pod, or pod for short. In a 1.2 GW A.I. data center, there would be thousands of these power-hungry pods. Each compute tray holds 4 GPUs and 2 CPUs and there are 18 compute trays in a rack resulting in 72 GPUs per rack.

Let's do some very rough math:

A 1.2 GW A.I. data center like Stargate requires significant cooling to keep the chips at the optimal temperature of 50 to 70 degrees Celsius so it can run efficiently. For example, a GB200 GPU chip has a power rating of 1,200 watts, equivalent to 20 60-watt light bulbs. The most advanced chips give off quite a bit of heat and require liquid cooling vs previous generation chips that can be air cooled. If you assume 10% of the power goes toward cooling and another 10% for other networking equipment, transformers, lighting, and so on... that leaves 960 MW going towards computing and other IT equipment.

In addition to compute trays, there is storage and other networking equipment. Assume 85% of the 960 MW goes toward computing, with each rack rated at 130KW.

That results in 6,277 racks (cabinets) x 72 GPUs per rack = 451,944 GPUs.

Thus, there are roughly 450,000 GPUs in a full-scale modern A.I. data center. At an estimated $35,000 per GPU, that results in $16 billion for GPUs alone.

In reality, Nvidia has become a systems company selling entire racks that include GPUs, CPUs, switch trays, and software all linked together with more than 5,000 ultra-fast copper cables that ultimately create one giant supercomputer.

Full Line-Up of Publications

Macro Products

Early Morning with Dave

Provides a synopsis of the Key market-moving events and data that were released overnight and includes Dave's thoughts on how the day ahead will evolve from an investment perspective.

Thematic Analysis

As a part of our flagship "Breakfast with Dave" publication get access to Dave and his team's research on key thematic topics that are on top of everyone's minds.

Data Commentary

As a part of our flagship "Breakfast with Dave" publication get insight from Dave and his team on key macroeconomic data in the U.S., Canada and other notable regions with research that goes beyond the headlines.

The Rosie Macro Fund

The Rosie Macro fund tracks our highest conviction calls and is dedicated to the conservative investor, who is mindful of capital preservation as well as the preservation of cash flows.

Big Picture Investment Themes

As a part of our flagship "Breakfast with Dave" publication stay on top of the key Cyclical and Structural investment themes that are shaping the economy.

Investor Chartroom

Those who have seen Dave's slide presentations in person know that his material is unique and informative. The charts are forward-looking and each one has an actionable investment theme.

Special Reports

Dave and his team of economists and strategists publish in-depth reports that tackle important secular themes which will shape the future of investing.

Market Strategy Products

Strategizer

A monthly model-driven report, Strategizer is an effective risk-management tool, providing scorecards and expected future returns for global stock markets, equity sectors, bonds, commodities, and currencies.

What's Priced In?

A market strategy tool providing investors with actionable and tactical advice each month on which industry groups of the S&P 500 to overweight and underweight.

Market Insights

Weekly Fundamental and Tactical recommendations at the sector and industry level for equities.

Technical Analysis

This weekly publication consists of charts and analysis across various equity markets, sectors, fixed income, commodities, and currencies.

Video Content

Video Content

From Dave and his research team keeping you up-to-date on our latest thoughts.

Webcast with Dave

David is joined by industry and market thought-leaders, to share their views on the investment landscape and the economic outlook.

Contact Us

120 Adelaide Street West, Suite 2500 | Toronto, Ontario, M5H 1T1

+1 (877) 316-1322 (U.S. & Canada) | **+1 (416) 583-5566** (International)

information@rosenbergresearch.com

www.rosenbergresearch.com

Our People

Economists & Strategists

David Rosenberg
Founder & President

Marius Jongstra
Vice President, Market Strategy

Robert Embree, PhD
Vice President & Senior Economist

Mehmet Beceren
Senior Market Strategist

Paul Kim
Junior Economist

Nwal Anwar
Senior Analyst

Zackary Young
Junior Research Analyst

Advisory Board

Walter Murphy
Technical Analysis Consultant

Jeff Burchell
Senior Advisor

Ken Bercuson
External Economics & Policy Advisor

Edward Campbell
External Research Consultant

John Smolinski
External Research Consultant

Alp Erdogan
External Research Consultant

Operations

Jacob Rosenberg
Executive Vice President, Client Relations & Operations

Grainne McGlynn
Senior Vice-President, Institutional Research Sales

Rachael Zbar
Corporate Director

Dinesh Vel
Director of Technology

Alexandra Delle Donne
Executive Assistant

Vandhana Parthasarathy
Data Analyst, Sales Enablement

Josh Johnson
Client Relations Representative

Nik Veigli
Client Relations Representative

Ayush Anand
Marketing Manager

Alyson Taylor
Junior Marketing Associate

Important Disclosures

This report comments on the economy and capital markets. The opinions are those of the author as of the date of the report and are subject to change without notice. Rosenberg Research & Associates Inc. (Rosenberg Research) will not necessarily update the report in the event facts or opinions change.

This report is for discussion purposes only.

Although Rosenberg Research may provide commentary relating to investment approaches and opportunities to buy or sell securities, you should not construe such commentary as legal, tax, investment, financial or other advice. Nothing contained in any Rosenberg Research report or any other content on our web site constitutes a solicitation, recommendation, endorsement, or offer to buy or sell any securities or other financial instruments. Consult your own advisors if you require such advice.

This report should not be regarded by recipients as a substitute for the exercise of their own judgement and readers are encouraged to seek independent, third party research on any companies discussed or impacted by this report.

This report is not a research report under applicable U.S. rules and regulations.

Rosenberg Research does not guarantee the accuracy of any information in this report. This report may contain links to third-party websites. These links are provided as a convenience only. By providing these links, Rosenberg Research is not expressly or impliedly endorsing or approving these other sites, nor warranting or guaranteeing the content of the materials on such sites or the services or products that may be offered on such sites.

Rosenberg Research reports are distributed simultaneously to internal and client websites and other portals by email and are not publicly available.

Terms And Conditions of Use

Your receipts and use of this report are governed by the Terms and Conditions of Use which may be viewed at [www.rosenbergresearch.com]

THIS ROSENBERG RESEARCH REPORT IS NOT PUBLICLY AVAILABLE. UNAUTHORIZED USE OR DISCLOSURE IS PROHIBITED.

This report is prepared for the exclusive use of Rosenberg Research clients, subscribers to this report and other individuals who Rosenberg Research has determined should receive this report. This report may not be redistributed, retransmitted or disclosed, in whole or in part, or in any form or manner, without the express written consent of Rosenberg Research.

ANY USE OF THIS REPORT IS AT YOUR OWN RISK AND LIABILITY. NEITHER ROSENBERG RESEARCH, NOR ANY DIRECTOR, OFFICER, EMPLOYEE OR AGENT OF ROSENBERG RESEARCH, ACCEPTS ANY LIABILITY WHATSOEVER FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL, MORAL, INCIDENTAL, COLLATERAL OR SPECIAL DAMAGES OR LOSSES OF ANY KIND, INCLUDING, WITHOUT LIMITATION, THOSE DAMAGES ARISING FROM ANY DECISION MADE OR ACTION TAKEN BY YOU IN RELIANCE ON THE CONTENT OF THIS REPORT, OR THOSE DAMAGES RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER FROM THE USE OF OR INABILITY TO USE ANY CONTENT OR SOFTWARE OBTAINED FROM THIRD PARTIES REQUIRED TO OBTAIN ACCESS TO THE CONTENT, OR ANY OTHER CAUSE, EVEN IF ROSENBERG RESEARCH IS ADVISED OF THE POSSIBILITY OF SUCH DAMAGES OR LOSSES AND EVEN IF CAUSED BY ANY ACT, OMISSION OR NEGLIGENCE OF ROSENBERG RESEARCH OR ITS DIRECTORS, OFFICERS, EMPLOYEES OR AGENTS AND EVEN IF ANY OF THEM HAS BEEN APPRISED OF THE LIKELIHOOD OF SUCH DAMAGES OCCURRING.

If you have received this report in error, or no longer wish to receive this report, you may ask to have your contact information removed from our distribution list by emailing [subscriptions@rosenbergresearch.com]