William Blair

AI INSIGHTS

Equity Research
Technology, Media, and Communications

January 23, 2025
Industry Report

Jason Ader, CFA +1 617 235 7519
Arjun Bhatia +1 312 364 5969
Dylan Becker. CFA +1 312 364 8938
Brian Drab, CFA +1 312 364 8280
Jed Dorsheimer +1 617 235 7555
Louie DiPalma, CFA +1 312 364 5437
Jonathan Ho +1 312 364 8276
Ryan Merkel +1 312 364 8603
Sebastien Naji +1 212 245 6508
Maggie Nolan, CPA +1 312 364 5090
Jake Roberge +1 312 264 8568
Ralph Schackart, CFA +1 312 364 8753
Stephen Sheldon, CFA +1 312 364 5167

Navigating the Boom:
Confronting Generative AI's Most Pressing Questions

This report is intended for enelson@williamblair.com. Unauthorized distribution is prohibited.

Please refer to important disclosures on pages 70 and 71. Analyst certification is on page 70.
William Blair or an affiliate does and seeks to do business with companies covered in its research reports. As a result, investors should be aware that the firm may have a conflict of interest that could affect the objectivity of this report. This report is not intended to provide personal investment advice. The opinions and recommendations herein do not take into account individual client circumstances, objectives, or needs and are not intended as recommendations of particular securities, financial instruments, or strategies to particular clients. The recipient of this report must make its own independent decisions regarding any securities or financial instruments mentioned herein.

Contents

Key Conclusions

1. **No Sign of AI Capex Letting Up**. We expect sizable GenAI-related capex to persist through the end of the decade as model intelligence is still rapidly improving and well-capitalized hyperscalers are engaged in an AI "arms race."

2. **Test-Time Compute Is Next Frontier in AI Scaling**. New vectors of AI model improvement beyond pretraining should continue to support scaling "laws" and drive the need for ever-increasing computing power. In particular, test-time compute (used in reasoning models like OpenAI's o1 family) represents the next frontier in AI scaling/intelligence given its ability during the inference phase to generate multiple potential solutions, evaluate them, and select the optimal one.

3. **Still Scratching the Surface on AI Use-Cases.** Initial GenAI use-cases include customer service, web search, software development, IT service management, content creation, and advertising. Use-cases should expand rapidly with improving model performance, reliability, and increasing enterprise/consumer familiarity, but return on GenAI investment (i.e., monetization of AI apps) will need to be proven out in the next few years to protect against overbuild risk.

4. **Physical Power and Infrastructure Are Main AI Bottlenecks.** The greatest constraint on the AI buildout today is the physical infrastructure, such as power and data centers, not chips and technology. In a bull scenario, we estimate AI could cause U.S. data center electricity demand to inflect up to 15% growth annually (up from 1% over the past decade), consuming 500 TWh of electricity by 2030, accounting for 9% of total U.S. electricity consumption in that year (up from ~2% today). When combined with the "electrify everything" movement, reshoring, and automation, AI is projected to more than double domestic electricity load growth annually over the next decade.

5. **Natural Gas Is Best Positioned for AI Data Center Energy Demands.** We see natural gas as the greatest near-term winner from the AI boom given its flexibility to meet electricity demand throughout the day, while nuclear and battery storage are likely to receive significant investment going forward to support the massive power demands of next-generation AI data centers.

6. **Enterprise Adoption Starting to Roll, but Still Early Days.** We expect to see signs of a steady increase in enterprise adoption (and monetization) of GenAI applications in 2025 as pilots and experiments move into production. However, it is still early days, and we believe broad and scaled adoption will come over time as businesses become more comfortable with the new technology and tackle barriers to adoption (privacy, security, and change management).

7. **Agentic AI Will Be a Major Catalyst to Enterprise AI Adoption.** In 2024, AI agents rapidly emerged as the best way to consume GenAI in the enterprise, marking a shift away from copilots. AI agents are more autonomous than copilots, operate without human intervention, and can solve complex, multistep problems. This drives more tangible ROI, which we believe will be a catalyst for enterprise AI adoption and a tailwind for application software vendors.

8. **AI Should Be Net Positive for IT Services.** AI-driven transformation requires significant expertise in areas like infrastructure modernization, AI model integration, data governance, and enterprise-scale deployment—capabilities that IT services providers are well positioned to deliver. Rather than being a threat, AI represents an opportunity for these companies to evolve their offerings, focusing on advisory, implementation, and management of AI solutions.

9. **Public Company Valuations Not in Bubble Territory.** Despite median equity returns of 200% for the big six public AI companies (AMZN, GOOG, META, MSFT, NVDA, and AAPL) since December 2019 (pre-GenAI), the median forward price-to-earnings multiple of 32x is up only 23% over this period, hardly indicative of a speculative bubble. Instead, it reflects tangible growth in earnings power.

10. **Venture Investment Is Booming.** U.S. GenAI venture investment of $96 billion in 2024 (49% of total VC funding) augurs well for continued innovation and greater monetization, though frothy valuations for some start-ups are reminiscent of the dot-com era.

Executive Summary

This report addresses common questions about the generative AI theme that the William Blair tech team has heard from investors over the past 18 months (since we published our GenAI primer found here). Our goal is to provide a roadmap for investors to assess the myriad opportunities (and risks) around the AI theme, and to identify companies and sectors that may not be getting the attention they deserve.

Our broad conclusion is that we are still in the early innings when it comes to AI development and usage, and the risks of not investing in AI (for institutional investors, hyperscalers, large language model [LLM] providers, enterprises, etc.) substantially outweigh the risks of investing in AI. This is based on our unwavering view that AI has the potential to be as transformative to the global economy as the steam engine, the transistor, and electricity.

We contend that in the current first phase of adoption (GenAI 1.0), the main impact will be increased productivity for both knowledge workers and consumers, as GenAI augments and enhances tasks like web search, customer service, software development, content creation, sales and marketing, and advertising. In particular, the implications of AI-assisted coding and LLMs are mind-boggling for the fields of software development and content creation and are likely to spark an explosion in human creativity (e.g., dramatically lower entry barriers for code development should lead to a burst of new app/business creation). As we move into the GenAI 2.0 phase, which we believe encompasses next-generation reasoning models and agentic AI (where multiple AI models are chained together to perform complex tasks), we expect a step change in model accuracy and a consequent increase in tangible use-cases across both consumer and enterprise markets.

While developers have delivered tremendous advances over the past few years across each of the three main dimensions of GenAI model intelligence—namely, compute, algorithms, and data—our research suggests that scaling "laws" (the empirical idea that the bigger the model, the more intelligent it will become) should continue to hold for the near term, especially with the emergence of new vectors of model improvement beyond pretraining of LLMs, including test-time compute (where AI systems follow a logical chain of thought to reach a conclusion versus the next-word prediction characteristic of current LLMs), multimodal AI (training models not just with text, but also video, audio, and image data), and new model training techniques (including post-training reinforcement and the use of synthetic data).

On the tech side, we continue to believe that the primary value creation today is occurring at the infrastructure layer (chips, networking, data centers, foundation models). With respect to LLM providers, the high cost of training the next frontier model creates significant barriers to entry and should therefore limit the competitor set. That said, this does not feel like a zero-sum situation—we see ample room for small models (off-the-shelf or home-grown models trained on proprietary enterprise data) to gain adoption as they may be more cost-effective and customized for certain use-cases.

Looking beyond the current boom in chips and networking (benefiting from rapidly expanding GPU cluster sizes), we see budding value creation in compute-adjacent infrastructure components such as high-bandwidth memory (will be particularly important for increasingly large context windows, new reasoning models, and device-level inferencing). In addition, as infrastructure software provides the application building blocks and runtime systems for organizations to securely develop both internal-facing and external-facing AI applications, we believe that it will be the next natural beneficiary of the GenAI boom, following infrastructure hardware.

Infrastructure software "picks and shovels" include databases, data lakehouses, data streaming systems, MLOps/DevOps tools, container management platforms, and data labeling, governance, and security tools. Given the scarcity of AI skillsets and experience within enterprise IT departments, we also believe IT services firms will be critical in helping the average enterprise customer leverage and adapt these software tools for the design and implementation of custom GenAI models/apps based on proprietary enterprise data.

While we see some lingering bottlenecks on the tech front (such as high-bandwidth memory and constrained fab capacity for the most advanced GPUs), the bigger AI bottlenecks are occurring in physical infrastructure such as power and data centers. In a bull scenario, we estimate AI could cause U.S. data center electricity demand to inflect up to 15% growth annually (up from 1% over the past decade), consuming 500 TWh of electricity by 2030, accounting for 9% of total U.S. electricity consumption in that year (up from ~2% today). When combined with the "electrify everything" movement, reshoring, and automation, AI is projected to more than double domestic electricity load growth annually over the next decade.

From a fuel perspective, we see natural gas as the biggest near-term winner from the AI boom given its high degree of flexibility to meet electricity demand throughout the day. Longer term, we are also optimistic on nuclear and battery tech, which will be necessary to support the massive power demands of next-generation AI data centers.

At the application layer, we are largely still at the build-and-experimentation phase of the enterprise AI adoption cycle, as businesses are still getting comfortable with the technology and discovering the best ways to deploy it. From past tech cycles (particularly the rise of cloud), we know that change in enterprise happens slowly at first and then quickly snowballs. While we are not at the stage where we broadly see large-scale deployments of GenAI applications, it is a top priority for the C-suite.

Over the next 12-24 months, we expect enterprises to tackle many of their own barriers to GenAI adoption (especially from a data-readiness perspective) and believe we will see a steady increase of GenAI apps being put into production. In particular, we believe the advent of AI agents will go a long way toward accelerating enterprise adoption of AI. In other words, we should soon start to see signs that the snowball is starting to roll, even though it still may be small.

While we continue to believe that installed customer bases and data moats should naturally accrue competitive advantage to AI-forward incumbent application vendors (like ServiceNow, Salesforce, and Workday), we also expect a new generation of AI-centric application companies to arise over the next several years given the ability for disruptors to emerge during major platform shifts (like what we saw in the dot-com era). We believe this will especially be true for disruptors that can target specific enterprise use-cases with an agentic platform. We are already seeing emerging disruptors arise in certain sectors like legal, software development, content creation, and marketing.

The bottom line, in our view, is that GenAI is ushering in a brave new world where the script is still being written, where the sheer magnitude of infrastructure needs for the AI buildout is still underappreciated, where AI model intelligence is rapidly advancing and seems to have no obvious ceiling, and where the use-cases are still early but progressing quickly. While we cannot predict the future, we can say for sure that GenAI will make it different!

In the exhibit below, we outline top AI picks across William Blair's Technology and Industrials coverage. A more detailed review of these picks can be found on page 58.

Exhibit 1
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Top AI Picks

Company	Ticker	Analyst	Mkt Cap ($ Bil.)	Price
Large Cap – More than $15 Billion				
Microsoft Corporation	MSFT	Ader	$3,185.8	$428.50
Alphabet, Inc.	GOOG	Schackart	$2,443.7	$199.63
Amazon.com, Inc.	AMZN	Bhatia/Carden	$2,425.9	$230.71
Meta Platforms, Inc.	META	Schackart	$1,555.9	$616.46
Tesla, Inc.	TSLA	Dorsheimer	$1,361.3	$424.07
Broadcom Inc.	AVGO	Ader/Naji	$1,126.4	$240.31
ServiceNow, Inc.	NOW	Bhatia/Roberge	$226.0	$1,096.85
GE Vernova Inc.	GEV	Dorsheimer	$114.7	$416.00
Motorola Solutions, Inc.	MSI	DiPalma	$78.2	$467.84
Datadog, Inc.	DDOG	Roberge	$47.0	$138.40
Axon Enterprise, Inc.	AXON	Ho	$46.2	$605.58
Cloudflare, Inc.	NET	Ho	$41.1	$119.85
Pure Storage, Inc.	PSTG	Ader	$22.9	$70.08
Toast, Inc.	TOST	Sheldon	$22.0	$38.65
Midcap – $3 Billion to $15 Billion				
nVent Electric plc	NVT	Drab	$12.3	$74.88
AAON, Inc.	AAON	Merkel	$10.7	$132.15
Elastic N.V.	ESTC	Roberge	$10.4	$100.36
AppFolio, Inc.	APPF	Sheldon	$9.5	$260.39
Globant S.A.	GLOB	Nolan	$9.1	$210.57
CCC Intelligent Solutions Holdings Inc.	CCCS	Becker	$7.4	$11.27
Modine, Inc.	MOD	Drab	$7.3	$139.58
Clearwater Analytics Holdings, Inc.	CWAN	Becker	$6.6	$28.99
Sterling Infrastructure	STRL	DiPalma	$6.0	$196.55
Small Cap – Less Than $3 Billion				
Grid Dynamics Holdings, Inc.	GDYN	Nolan	$1.7	$20.80

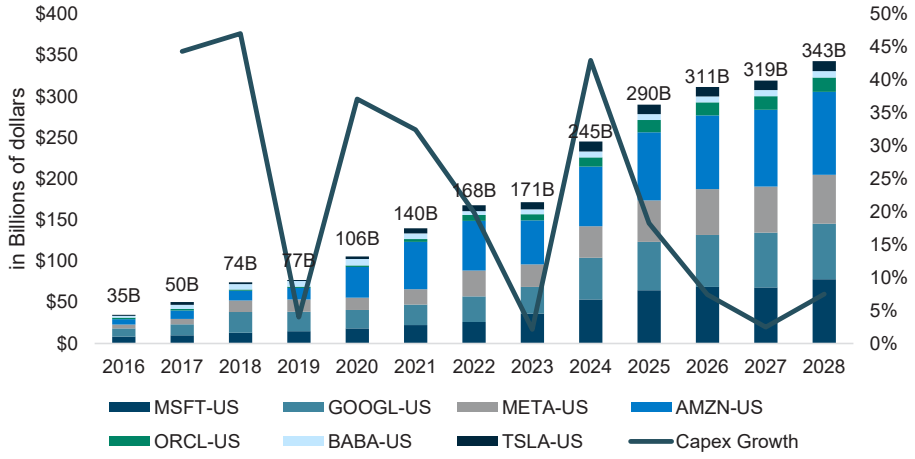
Price and market cap as of 01/21/2025

Source: Factset, William Blair Equity Research

Are We in an AI Bubble?

This question more than any other has dominated our conversations with investors, especially in view of massive capex investments tied to AI buildouts (cumulatively the largest public hyperscalers spent $245 billion in capex in 2024, as show in exhibit 2), eye-popping valuations for AI companies, and incessant media attention on the potential disruptive impact of AI on the economy and society. While it is difficult to answer this question with any certainty (especially as there is no clear consensus on what defines a "bubble"), our view is that we are still in the early innings of the AI infrastructure buildout and the potential for immense economic value creation (in the form of higher productivity) from AI applications is real.

Exhibit 2
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Continued Growth in Hyperscaler Capex

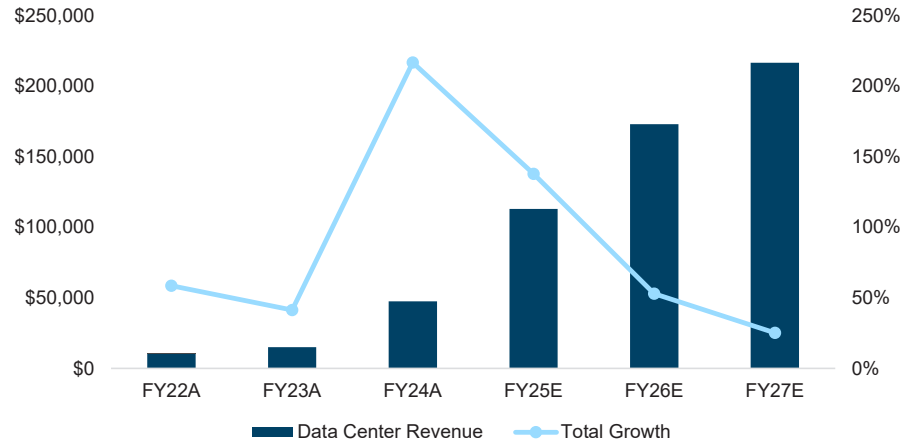


Sources: Company reports, FactSet estimates, and William Blair Equity Research

We base this view on the groundbreaking functionality of GenAI models, the fact that successive model releases are showing material improvements (in terms of overall intelligence and accuracy) with no obvious ceiling, and the practical timeline required to build out the necessary compute infrastructure, data centers, and energy production to support current and future market demand. That said, we concede that supply-demand imbalances are not uncommon in boom periods (especially in view of the current data center "arms race") and excesses are bound to occur—though given what we know today, we do not expect an airgap in demand until 2027 at the earliest. McKinsey estimates that demand for data center capacity will increase between 19% and 22% annually from 2023 through 2030 with 70% of total demand being allocated to AI workloads.

This sustained growth is supported by commentary from both hyperscalers and chip vendors. For example, Nvidia has seen exceptional growth in its data center business over the last few years, with 217% growth in fiscal 2024 alone, and 138% growth forecast in fiscal 2025 (see exhibit 3). In addition, during its earnings call in December 2024, Broadcom noted that its custom chip customers (e.g., Google, Meta, and Bytedance) have plans to scale up their data centers to more than 1 million GPUs by 2027/2028.

Exhibit 3
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
NVIDIA Data Center Revenues FY2022 - 2027E



All dollar figures in millions
Sources: NVIDIA and William Blair Equity Research

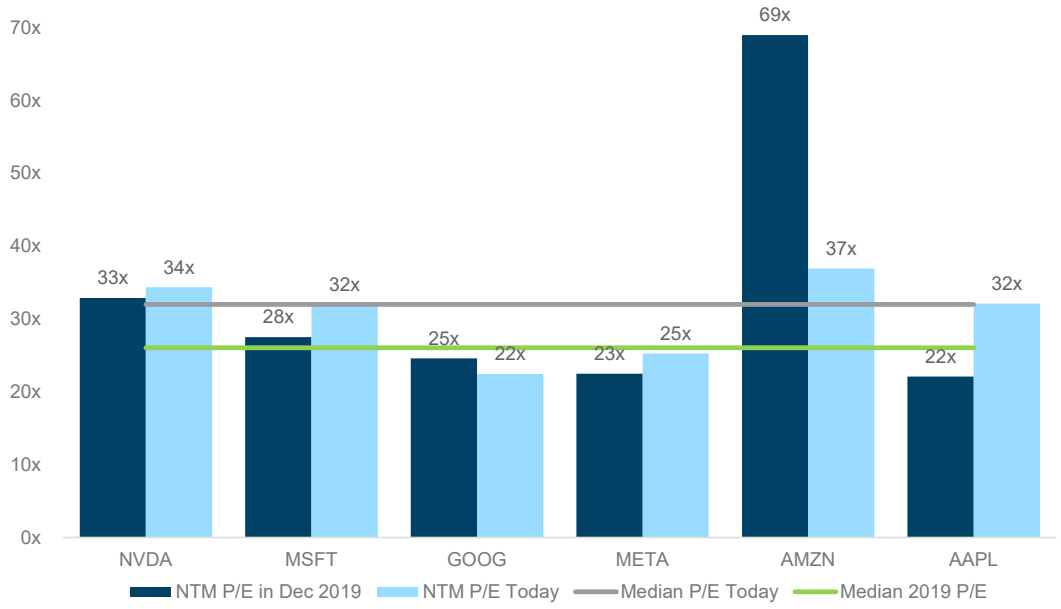
As we enter 2025, we are only starting to see 100,000-plus GPU centers move into production, with the biggest buyers of compute in the world planning their AI buildouts to extend for another two to three years at least. Taiwan Semiconductor, the manufacturer of the majority of the world's AI accelerator chips (e.g., tapes out chips for Nvidia, AMD, Broadcom, Marvell, etc.), noted in its January 2025 earnings call that it expects AI revenues to grow at an impressive 45% CAGR over the next five years—further evidence that AI buildouts remain a multiyear growth opportunity.

To cap things off, in January 2025, the newly minted Trump administration announced its backing of Project Stargate, a new AI venture led by OpenAI and SoftBank, with technology support from Arm, Microsoft, Nvidia, Oracle, and OpenAI, and funding from SoftBank, OpenAI, Oracle, and UAE tech fund MGX. Project Stargate intends to invest $500 billion over the next four years building new AI infrastructure in the U.S. for OpenAI, with $100 billion being deployed as soon as possible. Stargate will build 20 data centers, each measuring 500,000 square feet, according to Oracle Executive Chairman Larry Ellison, with construction of the first data center already underway in Abilene, Texas.

The aim of Stargate is to secure American leadership in AI, create hundreds of thousands of American jobs, and protect the national security of the U.S. and its allies. This adds fuel to the narrative that we are still early in the capex buildout required for AI and signals that the new U.S. administration is likely to be highly supportive of the investment and energy requirements of the AI platform shift.

On the valuation side of the bubble equation, the data suggest that the most exposed public GenAI stocks are not in a bubble. Looking at the six largest companies in the S&P 500, which all have meaningful exposure to GenAI—Apple, Microsoft, Nvidia, Google, Amazon, and Meta—forward P/E multiples from December 2019 to today have only increased by slightly more than 20%, from 26 times to 32 times (see exhibit 4 below). Meanwhile, the median return for these six stocks over that same time horizon is about 200%. This suggests that the equity returns have largely been driven by a tangible growth in earnings, not a more speculative increase in the multiple. In fact, for some of the leaders, like Amazon and Google, the P/E multiple has actually contracted from 2019 to today despite strong returns in these stocks.

Exhibit 4
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
NTM P/E for Mega-Cap Tech – December 2019 and Today



Sources: FactSet and William Blair Equity Research

AI vs. Dot-Com

Humans tend to search for frames of reference and analogous situations when something new and potentially revolutionary emerges. As such, investors commonly ask us if the current GenAI boom looks similar or different to the dot-com era, which began around 1995 with an intense period of physical fiber optic buildout from telecom providers and culminated with a spectacular bubble burst in 2000 and pursuant multiyear technology recession. While we find certain similarities to this era, we also see important differences.

Similarities

In terms of similarities, the hundreds of billions of dollars spent by telecom providers—such as Global Crossing, WorldCom, Enron, Qwest, and Level 3—on laying fiber, network and data center infrastructure, and related technologies in the late 1990s to support the Web 1.0 buildout is certainly reminiscent of the current period of GenAI capex. Then, like now, we are seeing a major technology platform shift that has the potential to disrupt established industries and drive new business creation—i.e., there will be winners and losers in the GenAI era just as there were in the dot-com period. Plus, like the internet, AI will be broadly available across different sectors of the economy for organizations to leverage, not just tech firms.

Investors often think about the value destruction related to the dot-com bust, but often overlook the massive value creation that occurred in the decade starting in 1995. Looking specifically at the tech companies founded in the dot-com decade (which we loosely define as occurring between 1995 and 2005) that are today valued above $100 billion, it is an impressive list that includes: Google, Amazon, Meta, Netflix, Salesforce, ServiceNow, Tesla, Arista Networks, Palo Alto Networks, Palantir, and Shopify. Cumulatively, these companies today represent more than $9 trillion in market value(!).

Exhibit 5
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
$100 Billion-Plus Tech Companies Founded in Dot-Com Era

Company Name	Ticker	Year Founded	Current Market Cap
Amazon	AMZN	1994	$2,425.9
Netflix	NFLX	1997	$371.8
Alphabet	GOOG	1998	$2,443.7
Salesforce	CRM	1999	$312.8
Tesla	TSLA	2003	$1,361.3
ServiceNow	NOW	2003	$226.0
Palantir	PLTR	2003	$166.5
Meta	META	2004	$1,555.9
Arista Networks	ANET	2004	$153.1
Palo Alto Networks	PANW	2005	$120.4
Shopify	SHOP	2006	$137.3
TOTAL			**$9,274.6**

Note: All numbers in Billions
Source: Factset and William Blair Equity Research

Our point is that despite the bursting of the internet bubble, this was a period of dramatic innovation and value creation—which we believe will similarly transpire with the GenAI platform shift over the next decade. To be clear though, we also expect GenAI business failures to be a natural outcome of this boom period, and we are already seeing frothy private company valuations exhibit bubble-like characteristics. For example, in late 2023, coding start-up Cognition closed a funding round at a $2 billion valuation without having a generally available product in a highly competitive market segment. More recently, the start-up Safe Superintelligence (SSI), founded by OpenAI co-founder and former chief scientist Ilya Sutskever, raised a seed round of $1 billion at a reported $5 billion valuation with zero revenue.

Exhibit 6
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Notable AI/ML Venture Rounds, 2024

Company Name	Amount Raised	Valuation	Date	Series	Lead Investors
Databricks	$10,000	$62,000	12/17/2024	J	Andreessen Horowitz, DST Global, GIC, Insight Partners, and WCM Investment Management
OpenAI	$6,600	$157,000	10/2/2024	E	Thrive Capital
xAI	$6,000	$45,000	12/23/2024	C	Andreesen Horowitz, Blackrock
Waymo	$5,600	>$45,000	10/28/2024	C	Alphabet
Anthropic	$4,000	NA	11/22/2024	NA	Amazon
Anduril	$1,500	$14,000	8/7/2024	F	Founders Fund, Sands Capital
G42	$1,500	NA	4/15/2024	NA	Microsoft
CoreWeave	$1,100	$19,000	5/1/2024	C	Coatue
Wayve	$1,050	NA	5/7/2024	C	SoftBank Group
Safe Superintelligence	$1,000	$5,000	9/4/2024	NA	NFDG, a16z, Sequoia, DST Global, and SV Angel
Scale AI	$1,000	$13,800	5/21/2024	F	Accel
Tenstorrent	$700	$2,600	12/2/2024	D	Samsung Securities, AFW Partners
Figure AI	$675	$2,600	2/29/2024	B	Microsoft, OpenAI, NVIDIA, Jeff Bezos, Parkway Venture Capital, Intel Capital, Align Ventures, and ARK Invest
Mistral	$645	$6,200	6/12/2024	B	General Catalyst
Groq	$640	$2,800	8/5/2024	D	Cisco Investments, Samsung Catalyst Fund, and Blackrock
MiniMax	$600	$2,500	3/5/2024	B	Alibaba, HongShan
Cohere	$500	$5,500	7/22/2024	D	PSP Investments
Perplexity	$500	$9,000	12/18/2024	C	Institutional Venture Partners (IVP)
Poolside	$500	$3,000	10/2/2024	B	Bain Capital Ventures
Physical Intelligence	$400	$2,000	11/4/2024	NA	Jeff Bezos, Lux Capital, and Thrive Capital
Lambda	$320	$1,500	2/15/2024	C	US Innovative Technology Fund (USIT)
DeepL	$300	$2,000	5/22/2024	NA	Index Ventures
Moonshot	$300	$3,300	8/5/2024	B	Tencent Holdings
Skild AI	$300	$1,500	7/10/2024	A	Lightspeed Venture Partners, Coatue, SoftBank Group, and Bezos Expeditions
Glean	$260	$4,600	9/10/2024	E	Altimeter, DST Global
Liquid AI	$250	$2,300	12/13/2024	NA	AMD
Augment	$227	$977	4/4/2024	B	Sutter Hill Ventures, Index Ventures, LightSpeed Venture Partners
Sakana AI	$214	$1,500	9/18/2024	A	NVIDIA
Cognition	$175	$2,000	4/25/2024	NA	Founders Fund
Sierra	$175	$4,500	10/28/2024	NA	Greenoaks Capital
Codeium	$150	$1,250	8/29/2024	C	General Catalyst
EvenUp	$135	>$1,000	10/8/2024	D	Bain Capital Ventures
Nimble Robotics	$106	$1,000	10/23/2024	C	FedEx, Cedar Pine
Together AI	$106	$1,250	3/13/2024	C	Salesforce Ventures
AnySphere	$60	$400	8/9/2024	A	Andreesen Horowitz, Thrive Capital

Note: All numbers in millions. NA = funding round was not a series round
Source: William Blair Equity Research

A final similarity between the two eras is low barriers to entry for new business creation. In the dot-com decade, thousands of start-ups were launched—all they needed was a computer, an idea, and some capital. Likewise, we see low barriers to entry across many fields with GenAI given the

availability of open-source LLMs, API accessibility to closed-source models, and an abundance of venture funding. AI start-ups captured a record 49% of the total $197 billion in U.S. venture funding raised last year, compared to less than 10% a decade earlier in 2014.

Differences

There are some obvious differences. First, the sources of investment are substantially different between the two eras. The majority of AI investment to date is emanating from mega-cap hyperscaler balance sheets and not from high-yield debt and speculative venture capital that was common in the early days of the internet buildout. Second, the valuations of internet companies circa 1999 were wildly speculative at best and irrational at worst. Investors created new metrics at that time—like number of eyeballs—to justify nosebleed valuations. In contrast, as noted above, we believe multiples for the top publicly traded AI companies are surprisingly reasonable.

Exhibit 7
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Valuations of AI Companies vs. Dot-Com Companies

AI/ML Leaders	Date	Market Value	Forward P/E Multiple
NVDA	1/22/2025	$3,448.9	47.7x
AVGO	1/22/2025	$1,126.4	37.8x
GOOG	1/22/2025	$2,443.7	24.9x
MSFT	1/22/2025	$3,185.8	32.8x
ANET	1/22/2025	$153.1	55.5x
AMZN	1/22/2025	$2,425.9	44.8x
META	1/22/2025	$1,555.9	27.2x
TSM	1/22/2025	$1,134.2	23.9x
ORCL	1/22/2025	$482.7	28.1x
AAPL	1/22/2025	$3,348.0	30.3x
NASDAQ 100	1/22/2025	$21,091.3	34.4x

Dot-Com Leaders	Date	Market Value*	Forward P/E Multiple
CSCO	3/1/2000	$500,000.0	69x
DELL	3/1/2000	$130,000.0	48x
INTC	3/1/2000	$277,000.0	53x
YHOO	3/1/2000	$100,000.0	>1,000x
AMZN	3/1/2000	$30,000.0	No earnings
EBAY	3/1/2000	$18,000.0	~200x
NASDAQ 100	3/1/2000		~60x

*market value figures are approximations
Note: All dollar figures in billions
Source: Factset and William Blair Equity Research

Today, with some notable exceptions (see SSI above), many of the most hyped AI companies do not have head-scratching valuations. For example, looking at OpenAI's last valuation round, the forward revenue multiple was about 15 times. While AI darlings like OpenAI, Anthropic, and CoreWeave are far from profitable, they are all generating meaningful revenue and therefore their valuations are much less hypothetical.

Leading up to the internet bubble, the number of technology IPOs increased substantially, with 370 companies going public in 1999 and 261 in 2000. This compares with 8 and 12 technology IPOs in 2023 and 2024, respectively. Furthermore, in contrast to the start-up-driven dot-com wave, many of the leaders of the GenAI wave are the biggest technology firms on Earth—including Microsoft, Google, Meta, Amazon, and Oracle—which has major implications for the competitive playing field if scale turns out to be a key source of competitive advantage in the AI era.

Exhibit 8
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Number of Technology IPOs, Dot-Com vs. AI

Year	Number of Technology IPOs	Median Company Age
1997	174	8
1998	113	7
1999	370	4
2000	261	5
2001	24	9
2002	20	9

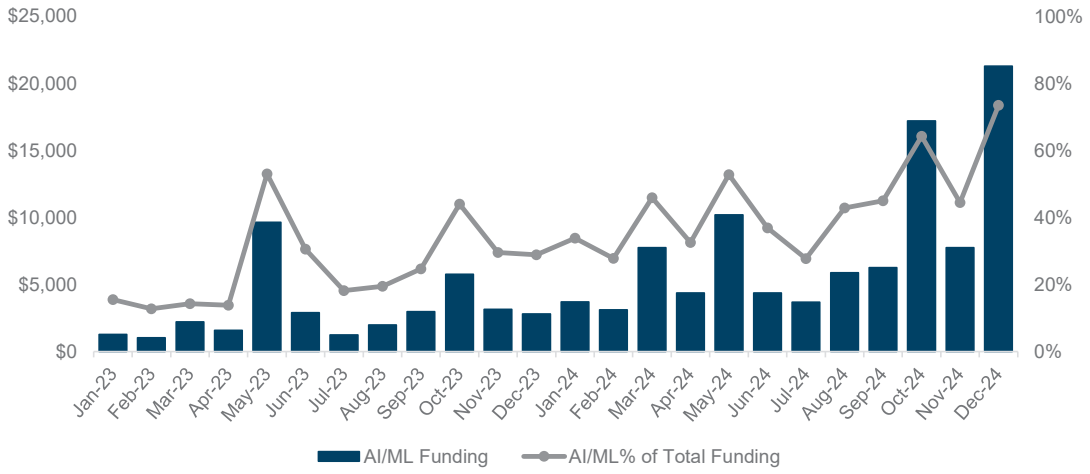
Year	Number of Technology IPOs	Median Age
2020	48	12
2021	126	12
2022	6	15
2023	8	10
2024	12	12

Source: Factset and William Blair Equity Research

Lastly, while the internet drove digital connectivity and revolutionized content distribution, GenAI (building on the internet and its progeny the cloud) arguably has much greater potential to impact the economy given its cognitive abilities and opportunity to augment/displace many human tasks.

Bottom line: While we are cognizant of overbuild risk and the challenge of picking winners and losers in the early days of a major platform shift such as this one, just as with the dot-com era, we believe investors should be playing the long game with GenAI, with the potential payoff likely much greater than the invested capital.

Exhibit 9
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Total U.S. AI/ML Venture Funding Activity 2023-2024
($ in billions)



Note: Data was sourced from Pitchbook Data, Inc., which classifies funding rounds into verticals. One company can be classified into multiple verticals depending on the specifics of the business model. For example, if a company like Amazon was included in our fundraising data, it would likely be categorized into multiple different verticals, including e-commerce, AI/ML, big data, TMT , and more. Therefore, the sum of our vertical data will be much greater than the total funds that were raised in the period being evaluated.

Source: Pitchbook Data, Inc. and William Blair Equity Research

Is There Enough of a Return on Current Investments in AI Models and Data Centers to Justify the Spending?

If you build it, will they come? This is perhaps the most critical question to address in assessing whether we are destined for a major supply-demand imbalance at some point in the next several years. Specifically, with a cumulative $1.6 trillion in capex expected to be spent from 2023-2028 by the largest hyperscalers in the world, investors are asking is there enough actual revenue potential in the AI ecosystem to justify this level of spending. In a recent blog post, Sequoia Capital's David Cahn has termed this: "AI's $600B Question"—basically arguing that $600 billion in annual AI software revenue is needed to "pay back" just the current level of capex, which creates a massive hole to fill given AI application revenue in 2024 for the entire market is estimated to be only around $20 billion.

Exhibit 10
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
AI Revenue Required for Payback
($ in billions)

	Q4 2023E	Q4 2023A	Q4 2024E	Q4 2024A
NVIDIA Data Center Revenue Run-Rate	$50	$74	$90	$150
Data Center Facility Build and Cost to Operate	*50%*	*50%*	*50%*	*50%*
Implied AI Data Center Spend	$100	$147	$181	$300
Software Margin	*50%*	*50%*	*50%*	*50%*
AI Revenue Required for Payback	$200	$294	$363	$600

Note: Payback revenue is calculated by taking Nvidia's run-rate revenue forecast and multiplying it by 2x to reflect the total cost of AI data centers (GPUs are roughly half of the total cost of ownership—the other half includes energy, buildings, backup generators, etc). Then multiply by 2x again, to reflect a 50% gross margin for the end-user of the GPU (e.g., the startup or business buying AI compute from Azure or AWS or GCP).

Source: Sequoia Capital and William Blair Equity Research

Tech leaders appear to understand the dilemma they face, with Alphabet CEO Sundar Pichai noting that, "The risk of under-investing is dramatically greater than the risk of over-investing," while Meta CEO Mark Zuckerberg admitted that, "There's a meaningful chance that a lot of companies are overbuilding now, but … the downside of being behind is that you're out of position for the most important technology for the next 10-15 years."

We also question whether "AI revenue" is the only framework to be thinking about in terms of return on invested capital. That is, if AI becomes integrated into most or all applications over time, how does one discern what is AI revenue and what is non-AI revenue? For example, some vendors may charge extra for new AI products/features (which can be measured), but other vendors may simply embed AI into their offerings. This might help them sell more or drive competitive advantage, but it will be difficult to precisely measure the impact of AI on vendor revenue.

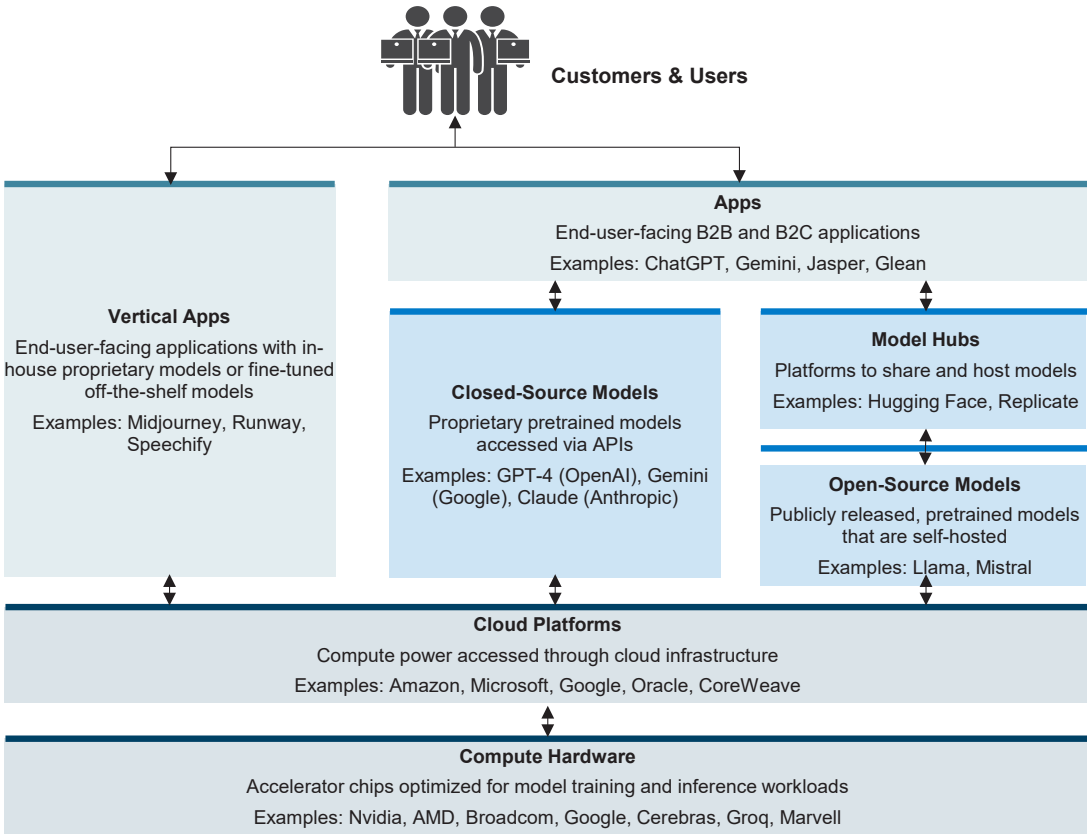
Moreover, if AI results in material margin expansion for businesses across many sectors (due to a reduced need to hire new workers and/or productivity increases for existing employees), how would that be accounted for if revenue is the main return framework? In the area of software development, if AI can increase developer productivity by 30%-50% (as some experts predict), that would translate into massive cost savings for companies. Beyond this, what does it mean for economic value creation and the future of software if the gating factor to code development is no longer human developers. This could usher in a wave of application innovation that is hard to fathom.

Bottom line: The GenAI space feels to us more like a marathon than a sprint, with the science still in the early innings (we expect further algorithmic breakthroughs going forward) and returns on current investments likely to take many years to play out. As Amara's Law states: "we tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run." While near-term skepticism around the breakneck level of capex is warranted, we think investors should not miss the forest for the trees.

Where Is the Value Accruing in the AI Space and Where Is the Most Untapped Opportunity for Investors?

In the near term, we expect the infrastructure/compute/model layers (the bottom three layers in exhibit 11) will continue to accrue the most value in the GenAI space as GPU clusters scale ever larger to support the pretraining of frontier AI models and as post-training and inference become more significant. For example, xAI's Colossus supercomputing cluster in Memphis, Tennessee, contains 100,000 Nvidia Hopper GPUs (and plans are to expand that cluster to 300,000 GPUs in the first half of 2025 and 1 million GPUs in 2026), while Oracle is in the process of standing up a supercluster of over 100,000 Nvidia Blackwell GPUs.

Exhibit 11
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
The Generative AI Stack



Source: a16z and William Blair Equity Research

Because of the need for massive scale here, this limits the infrastructure opportunity on the public company side to a handful of chip, cloud, and model vendors (NVDA, AVGO, AMD, MRVL, AMZN, GOOG, META, MSFT, ORCL, and ANET). On the private company side, we call out OpenAI, Anthropic, xAI, Mistral, CoreWeave, Hugging Face, Scale AI, Cohere, VAST Data, Glean, and Inflection AI as early GenAI leaders.

Moving forward, as AI applications start to take greater hold (both on the consumer and enterprise sides), we see a few different areas of budding value creation. These include compute-adjacent infrastructure components such as storage and memory (will be particularly important for retrieval-augmented generation and device-level inferencing), and infrastructure software picks and shovels, including databases, data lakehouses, data streaming systems, MLOps/DevOps tools, container management platforms, and data labeling, governance, and security tools. Infrastructure software provides the application building blocks and runtime systems for organizations to develop both internal-facing and external-facing AI applications. As such, it is useful to think of infrastructure software as the next natural beneficiary of GenAI, following infrastructure hardware.

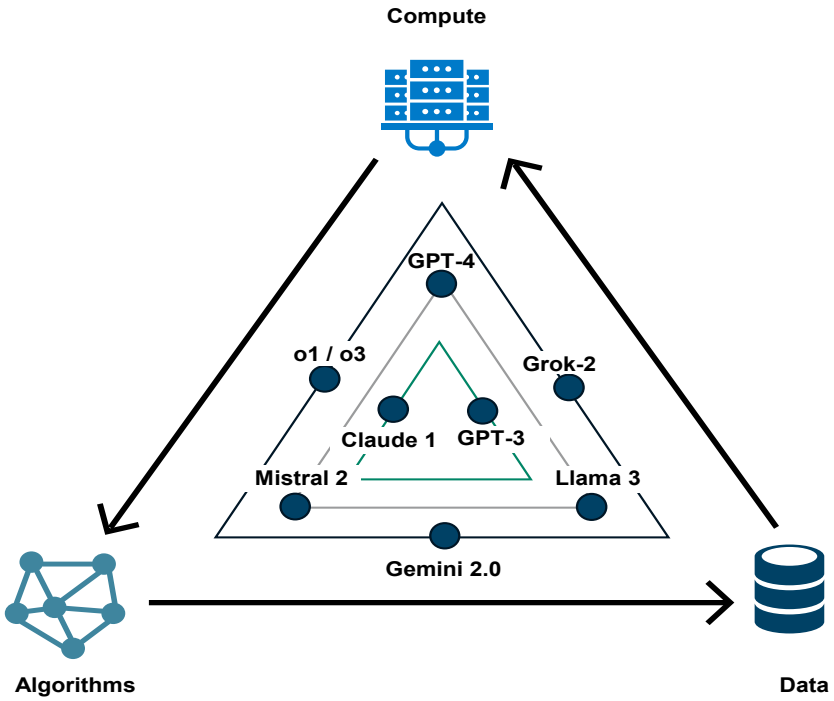
Beyond the core tech, we see significant opportunities for value creation in energy production to power new data centers and certain related equipment, such as transformers and switchgear, that connect the grid to the data center (see page 43 for further details).

Bottom line: We remain very much in the buildout phase of GenAI, with most of the value generation today coming from chips, hardware systems, and the physical infrastructure required to stand up data centers, though we see signs on the horizon that the application layer will begin to accrue increasing value as the tooling and skillsets mature and the use-cases are proven out.

Are Scaling "Laws" Holding or Are LLMs Reaching Diminishing Marginal Returns?

AI scaling laws empirically assert that model performance improves predictably as the number of parameters increases and the amount of training data grows. To see the maximum benefit in model performance, each of the three core levers of AI scaling—compute, algorithms, and data—need to be increased/improved.

Exhibit 12
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Three Levers Powering AI Scaling Laws



Source: William Blair Equity Research

Throughout the second half of 2024, investors began to worry that AI scaling laws were starting to exhibit diminishing marginal returns, especially as the data available for LLM pretraining seemed to be reaching a natural limit (the web only has so much text-based data). As we push model sizes into the hundreds of billions or even trillions of parameters, the computational and financial costs balloon, and the returns in terms of accuracy and capabilities start to taper off. This raises a practical limitation: even though scaling laws predict performance gains, the cost and complexity of training ever-larger models may soon outweigh the benefits. As a result, investors and practitioners began to ask: does doubling computing capacity no longer provide the same level of model performance improvement as in prior generations?

The reality, in our view, is that as they approach the upper bounds of what is financially and technically feasible, model providers are indeed encountering diminishing returns for their LLMs. The marginal gains from adding even more parameters or tokens shrink relative to the cost and complexity of running billion-dollar training clusters. In essence, the straightforward path of just "making the model bigger" no longer looks like the best strategy for advancing the model's capabilities.

To overcome these limitations, researchers and practitioners are introducing new vectors of model improvement beyond pretraining of LLMs. These include but are not limited to:

- *Test-time compute* – used in reasoning models like OpenAI's o1 family, where AI systems follow a logical chain of thought to reach a conclusion versus next-word prediction characteristic of current LLMs, see below;

- *Multimodal AI* – training models not just with text, but also video, audio, and image data;

- *New model training techniques* – including the use of synthetic data and post-training reinforcement learning with human or increasingly AI feedback, see below;

- *AI model chaining* – used in agentic AI to execute complex tasks;

- *Expanding context windows* – the higher the number of tokens/words put in a conversation with an LLM, the more context can be provided to improve response accuracy; and

- *New algorithmic techniques* – such as S4, Mamba, and Titans, which are better at handling long sequences without needing as much computing power as Transformer-based models.

As such, even if LLM pretraining is seeing diminishing returns, we see ample room for ever-higher levels of compute to drive continued material AI model performance improvements. By refining models after their initial training and leveraging more sophisticated reasoning processes at inference, we can extract greater value from models without endlessly pushing their parameter counts. This not only keeps costs manageable but also introduces more flexibility, allowing models to adapt dynamically to challenging queries. As a result, the industry is finding ways to sustain and even accelerate progress in LLM capabilities—even as pure pretraining scaling approaches practical limits.

An apropos analogy comes from the semiconductor space itself, where even though the pace with which we can shrink transistor nodes (from 7 nm to 5 nm to 3 nm and now 2 nm) is decelerating, we have found novel approaches (parallelization, new material science discoveries, scale-out) that have allowed the industry to maintain performance improvements in chips that mirror Moore's Law.

Jensen Huang, CEO of Nvidia, attempted to quell fears about a wall in pretraining scaling in his most recent earnings call (in November 2024), emphatically noting that pretraining scaling laws are holding. Despite this assertion, he laid out two new vectors of scaling that should drive continued LLM improvements for several more years:

- ***Test-time (inference-time) compute:*** Test-time compute focuses on increasing a model's capabilities at the moment it is used (at inference). Rather than relying solely on the fixed parameters learned during pretraining, models can be allowed to "think" longer or "reason" more deeply—which will require ever-increasing computational resources at inference time. Techniques such as chain-of-thought reasoning, multistep planning, and iterative refinement mean the model can generate multiple intermediate steps, check its own work, and explore various solution paths before producing a final answer. This approach reduces the reliance on creating ever-larger base models; instead, smaller models can achieve complex reasoning and higher-quality outputs by using more computational steps at inference, effectively trading off response time for intelligence. OpenAI's Strawberry, or o1 model family, is an early example of these types of models that reason for longer before providing an answer. We expect such reasoning models will be particularly relevant in domains where there are verifiable truths like the sciences and math.

- ***Post-training scaling:*** Post-training scaling involves enhancing a model's abilities after the initial training phase. Early approaches focused on reinforcement learning with human feedback (RLHF), where a pretrained model's responses were refined by having humans rank and guide its outputs, thereby improving its quality without retraining from scratch. More recent methods have introduced synthetic data generation and reinforcement learning with AI feedback, where the model—or other models—assist in refining responses. These techniques allow significant improvements without the massive computational expense of redoing the entire pretraining process.

Beyond the Transformer

Algorithmic innovations are another way that LLM builders are looking to drive continued performance improvements for models at scale. While the Transformer model has taken the world by the storm with the release of ChatGPT, there are many other types of model algorithms and architectures with which researchers are experimenting.

One approach that has seen success is called S4. S4 was first introduced in a paper published in 2021 and is built around a linear state space model (SSM). This concept uses state space layers, which are like mathematical tools to efficiently handle sequences (e.g., words in a sentence, notes in a song, time-series data). Instead of comparing every part of the sequence to every other part (as is done with transformer models), an SSM treats the sequence as the result of a special kind of equation. This makes SSMs much faster and better at handling long sequences, without needing as much computing power. Transformers are adept at processing sequences, but they use a method called self-attention, which takes a lot of computing power, especially as the sequences get longer. A helpful analogy to compare the two approaches is that of reading a book. While a transformer model would need to re-read the entire book to understand the next chapter, an SSM just remembers the key points from the prior chapters to make sense of what happens next.

Another more recent algorithmic approach is called Mamba. While less established in the literature than S4, Mamba is representative of a class of models that aim to handle long sequence contexts without resorting to standard self-attention mechanisms that dominate Transformer models. These models often employ advanced mathematical constructs (like state-space formulations, kernel methods, or efficient recurrence) to leverage less memory and compute resources and derive a response more efficiently. Vendors like AI21 Labs are employing a hybrid approach combining elements of both Mamba and transformers.

Lastly, in December 2024, Google researchers released a new model architecture called Titans. Titans is designed to have better memory than the transformer model by separating out its memory handling into short-term and long-term memory. This allows the Titans architecture to remember context without slowing down operations. As a result, Titans can purportedly scale context windows beyond 2 million tokens and could represent an answer to the memory bottleneck that limits the efficiency of current models.

S4, Mamba, and Titans illustrate the ongoing innovation in AI algorithms in general and sequence modeling in particular. Although transformers remain dominant, these newer algorithmic models signal that alternatives are gaining traction and may play an increasingly important role in driving performance at scale, particularly in complex agentic systems that may be required to dynamically process large streams or sequences of information.

Is Synthetic Data Effective?

One solution to solving the data scarcity problem is to simply make more data. Synthetic data, which refers to information that is artificially generated rather than collected from real-world events (i.e., data is simulated to replicate the statistical characteristics and patterns of real-world data), is particularly valuable in industries and use-cases where data collection is too expensive, cumbersome, or restricted due to privacy concerns.

One of the most obvious beneficiaries of synthetic data is the healthcare industry given privacy regulations around accessing personal medical records. With the use of synthetic data, artificial medical records can be generated on which AI models can be trained. If these models become more intelligent with this synthetic data, they can be used to identify illnesses, simulate treatments and their effectiveness, and/or accelerate the discovery of new drugs.

Similarly, within the financial services industry, personal information cannot be used for AI model training. To counteract this, banks have begun creating synthetic financial records of fake customers to train its model on how to better detect fraud. Within the manufacturing space, companies like Advex are working to improve computer vision systems by generating thousands of "fake" images for training. With Advex's technology, car manufacturers can teach computer vision systems to recognize defects in car seats by simply uploading a dozen or so images of defected seats and generating thousands of images for training purposes.

Synthetic data use comes with its own set of risks. Because AI models are typically the creator of this synthetic data, the data output can only be as good as the model used and its training data. For example, if a synthetic data model is trained on data with certain biases or certain groups are underrepresented, then the output synthetic data will suffer from these same biases. If, from the earlier example, a few car seat images are not representative of defects, the computer vision system might flag several functioning seats. Therefore, if the synthetic training data is not representative of real-world data, it has the potential to introduce "hallucinations" or errors, especially in complex datasets.

To mitigate these issues, the use of synthetic data requires thorough curation and filtering, just as real-world data does. We have seen leaders in the AI space such as OpenAI and Meta adopting synthetic data as the costs for data collection increase. OpenAI recently released Canvas, which is meant to offer a new way to interact with ChatGPT through a window with a workspace for writing and code. Canvas is powered through a version of its GPT-4o model that was tailored using synthetic data to "make targeted edits and leave high-quality improvements in line" according to ChatGPT's head of product, Nick Turley.

Meta has used synthetic data to fine-tune its Llama 3 models and to generate captions for its Meta Movie Gen. Meanwhile, new AI models are being tasked specifically with creating high-fidelity synthetic data, including Microsoft's Orca-AgentInstruct and Splunk's MAG-V. In our view, the power and value of synthetic data will only grow as AI industry leaders invest more in data collection, and someday synthetic data may be reliable enough for models to train themselves.

Bottom line: While developers have delivered tremendous advances over the past few years across each of the three main dimensions of GenAI model intelligence—namely, compute, algorithms, and data—our research suggests that scaling laws should continue to hold for the near term, especially with the emergence of new vectors of model improvement beyond pretraining of LLMs, including test-time compute, new algorithms, and innovative post-training techniques (including the use of synthetic data).

Is the LLM Becoming a Commodity?

Yes and no. On one hand, it is incredibly difficult to create an LLM. The barriers to entry in the LLM market are quite high and only those with significant funding and access to computing power will be able to compete in this space. It is estimated that GPT-4 by OpenAI cost $100 million to train and GPT-5 is rumored to cost upward of $1 billion. This is not something that most start-ups will be able to do and is the main reason why there are only a select few players in this market today (and many of them were already the largest companies in the world before AI). In our view, the number of players competing in the LLM market or artificial general intelligence (AGI) race is unlikely to increase significantly from here. However, there may be specialized players that enter the market with small language models (SLMs) designed for specific use-cases or verticals (see page 24).

On the other hand, access to similar training data, lower cost of compute, and open-source models like Llama by Meta should level the playing field in the foundation model market among existing players. While the model providers will try to differentiate on factors like speed, accuracy, cost, and reasoning capability, we believe most end-users will have to look hard to be able to tell the difference between the top few model providers. Over the medium to long term, we suspect that performance among most model providers will be similar and believe the market structure will more closely mimic the cloud service providers or airline industry (an oligopoly with a handful of major providers and a tail of specialized providers) than the search market (winner-take-all market dominated by Google). In our view, the tail of the market will be SLMs that are competing for specialized use-cases.

Commoditization is more likely to occur at the low end of the market, for very simple tasks and use-cases like basic research, simple text or code generation, or low-level automation. In these cases, different models are more likely to be interchangeable. We already see this somewhat in how traction is playing out in the consumer AI chatbot market with three players dominating adoption: OpenAI's ChatGPT, Anthropic's Claude, and Google's Gemini. Differentiation is difficult to discern for the average consumer, and tool preference may be less dependent on the underlying LLM and more on factors related to the user interface. However, for more complex use-cases, differentiation is likely to be more conspicuous, and providers with better quality will win out. Examples of this might be drug discovery, complex legal analysis, medical treatment plans based on patient history, or live customer service.

Bottom line: Even if LLMs become more commoditized over time, this does not necessarily mean that the LLM providers themselves will become commoditized. AI leaders like OpenAI and Anthropic have the opportunity to evolve from just being LLM providers to being much more. For example, they can develop their own specialized SLMs while still gunning for AGI, can launch AI agents, and can develop their own applications to move up the stack.

Will Open-Source LLMs Diminish the Value of Closed-Source LLMs?

Developers of LLMs can take two different routes when creating and releasing their models: closed source and open source. In closed-source models, the underlying code, training data, and parameters are not made available to the public except through an application programming interface (API), which third parties can tap into for their applications. In these closed-source scenarios, the exact workings of the LLM remain a "black box" to users. In contrast, open-source models are created as collaboration software, where the original source code and models are freely available to the public for redistribution and modification. The obvious question then is if model performance (between closed and open source) converges over time, how do closed-source models compete in the long run, especially if open-source models do not charge subscription or API fees (as closed-source models do).

Because they are generally sponsored or run by a hyperscaler, closed-source models typically have access to proprietary training data and priority access to cloud computing resources. As noted, closed-source models are made accessible through an API that allows a downstream application to query and present information from an LLM without having to expend resources on training, fine-tuning, and running the model. Examples of closed-source LLMs include GPT-4 (OpenAI), Gemini (Google), and Claude (Anthropic).

In contrast, open-source models have full transparency into the training data and how the models were built. Though they require more operational overhead, the benefits of using open source include better governance (given that the origins of data are known), improved ability to customize the model (since it is less of a black-box offering), and the ability to avoid vendor lock-in. Examples of open-source models include Llama (Meta), Stable Diffusion (Stability AI), and Mistral.

The tension between closed- and open-source models is reminiscent of the iPhone/Android debate in the mobile operating system market. Though closed-source foundation models may see greater adoption in the near term given their performance advantages, over time we expect—as with other types of infrastructure software—developers and data engineers may prefer the transparency and flexibility afforded by open-source solutions.

Pros and Cons

The debate over open-source versus closed-source models can be distilled to two main factors: safety and speed of innovation. Proponents of closed-source AI models tend to argue that allowing access to all the different aspects of a model can exacerbate many of the risks surrounding AI. For example, armed with the characteristics of a certain open-source model (e.g., the model's weights and architecture), users could exploit the model for use in mounting cyberattacks or disseminating unknown or false information, among other activities. On the other hand, revealing the characteristics of a model allows for a greater understanding of how it was created and how it can be modified for a particular use-case. Open-source proponents argue that the democratization of access to software has historically led to faster advancements as it becomes more of a collaborative effort among users.

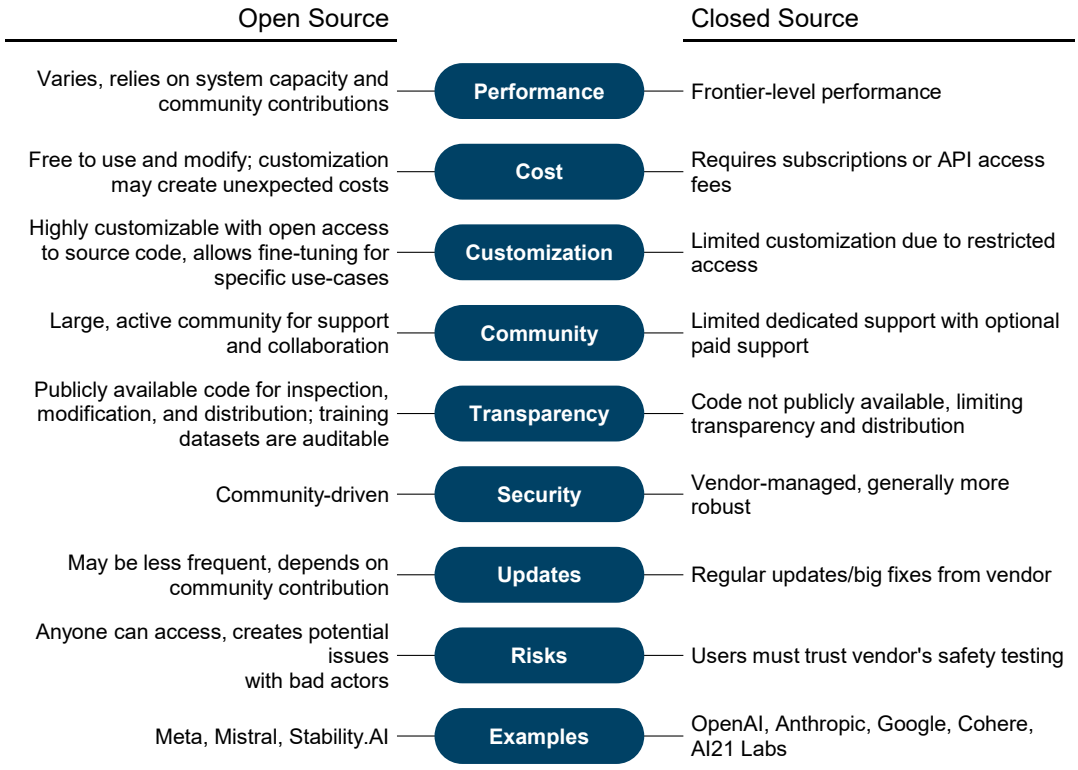
Tech company CEOs like Mark Zuckerberg (Meta), Emad Mostaque (Stability AI), and Clement Delangue (Hugging Face) have all spoken publicly about the value of open source in AI development. Zuckerberg has drawn parallels from open-source AI to the success of Linux in cloud computing. He argues that every organization is going to need custom models fine-tuned with their own data, but they will not want to send this sensitive data over cloud APIs. To be clear, as an advertising and social media company, Meta is in a unique competitive situation as openly releasing

its Llama models will not directly impact the company's overall business. The company's stated goal in pushing its open-source AI Llama models is to foster a robust ecosystem (especially for start-ups) that will drive faster adoption of AI across all markets.

Ultimately, this debate encapsulates a fundamental tension in the AI industry: the balance between innovation and competitive advantage. Open-source advocates claim to be progressive in their search for innovation by democratizing model access and preventing a concentration of power. At the same time, closed-source platforms preach ethical safety and the protection of proprietary data for companies looking to implement AI. OpenAI founder and former Chief Scientist Ilya Sutskever concedes that open source is important to prevent a concentration of power; however, once the technology's capabilities reach general or super intelligence, it will become irresponsible to allow anyone to have access to it. He believes that OpenAI is currently closed source for competitive reasons, but that we will soon reach a point where safety is the main driver for closed-source models.

Bottom line: We do not view the LLM market as a zero-sum game and see room for both types of models to be successful. Customers with strict requirements around compliance and support and those who favor rapid deployment of models will likely prefer closed-source models that are ready to use out of the box. Meanwhile, more cost-sensitive organizations that require significant customization may prefer to go the open-source route.

Exhibit 13
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Comparing Open-Source and Closed-Source AI Models



Source: William Blair Equity Research

Is There Room in the Market for Both LLMs and SLMs?

Large-scale models, such as OpenAI's GPT series and Google's Gemini, have captured most of the attention in the GenAI market. However, these LLMs require a tremendous amount of computational power and can be expensive to implement. As a result, AI developers are creating smaller, more targeted models for specific use-cases or as a wrapper for their mobile applications. Both types of AI models offer their own advantages and limitations, which will shape the future of GenAI development and expand use-cases across both consumer and enterprise markets.

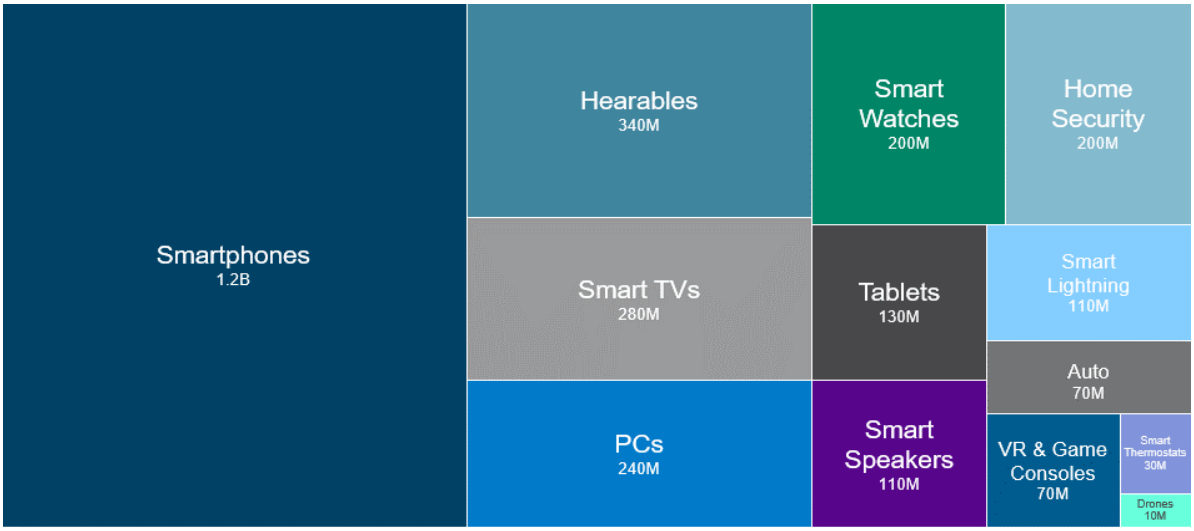
Large AI models are characterized by extensive training on vast datasets, billions (sometimes trillions) of parameters, and the most-up-to-date GPUs. This allows LLMs to understand a wide range of user prompts and generate highly nuanced and complex outputs. The major benefit of these models is their broad applicability; they can handle diverse tasks such as content generation, language translation, and complex problem-solving. However, the costs associated with training and maintaining such models are substantial. These models require extensive computational resources, raising concerns around energy consumption and environmental impact. In addition, large models can be prone to issues related to bias and ethical considerations, as they absorb massive amounts of training data. For these reasons, LLMs are not always the best solution for every user's AI needs.

Recently, developers such as OpenAI and Meta have released their own SLMs alongside newer generations of LLMs. The light footprint and lower computational requirements of these SLMs makes them especially well-suited for edge devices like smartphones. These models are more tailored and efficient (i.e., have lower latency) than LLMs for use-cases like summarization, rewriting, translating, or following instructions. Their flexibility in deployment also enhances user privacy and data security since they can be used locally rather than across networks. The trade-off comes with a much narrower scope of use-cases and a reduced ability to generalize.

Meta unveiled its newest lightweight models alongside Llama 3.2 at the Meta Connect conference in September. These models operate on 1B or 3B parameters compared to the 11B and 90B parameters of medium and large models. The company believes that these smaller models will empower developers to build their own AI applications for specific use-cases. Longer term, the hope is that these models can be transferred onto edge devices, such as thermostats, routers, or scientific instruments. For example, a 1B parameter lightweight model can run on a smartphone, but electronics like lighting or security equipment can only handle 100M to 200M parameter models, and therefore would not be able to support even these lighter weight models.

The shift to smaller, more compact models will likely democratize the use of AI and allow more business and innovators to leverage their proprietary data at reduced costs. Our research indicates a growing preference among Global 2000 customers for building or fine-tuning their own smaller, more-tailored GenAI models that are trained on a subset of proprietary data. For enterprises with strict data governance and privacy considerations, these smaller models are more efficient and less expensive to run, and reduce the risk of overfitting. We expect enterprises will leverage numerous SLMs to address different use-cases, similar to Microsoft's commercial approach in GenAI that features a number of distinct copilots across different product lines. Microsoft has also supported the concept of SLMs with its creation of Phi, an SLM that can run on endpoint devices (e.g., PCs and mobiles).

Exhibit 14
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Parameters for Consumer Electronics



Source: Meta

Importantly, this purpose-built approach should drive demand for model orchestration/agentic solutions that can help concatenate these smaller models to address a broader range of use-cases. For example, speech-to-text use-cases could use a recognition model to convert spoken language to text, which can then be translated by a transformer model to another language. Elsewhere, an autonomous vehicle could combine a model that identifies and classifies objects on the road with one that uses this information to plan/readjust the vehicle path. These types of agentic frameworks are maturing rapidly and may combine both LLMs and SLMs for specific tasks (see page 38).

Bottom line: We see room for different-sized models to achieve success in the AI market, as LLMs and SLMs are designed for fundamentally different purposes.

What's the Killer App for GenAI or Is That the Wrong Question?

On the Consumer Front

While the closest thing to a killer AI app today is ChatGPT, we see use-cases expanding rapidly to other areas. For example, Duolingo, a language learning app, offers a GenAI virtual tutor called "Max" that can calibrate to a student's speaking capabilities.

We believe GenAI app development may occur faster than what we saw during the first iPhone cycle as the tooling ecosystem is advancing rapidly and costs are declining quickly. We believe the fundamentals are the same for GenAI apps as they are for traditional apps: solve a clear problem for a wide audience, and success is likely. Once AI models improve their reasoning capabilities (already seeing this in next-generation LLMs) and better manage hallucinations, we expect them to become a viable tool for almost all industries.

Search

GenAI is becoming synonymous with search services like ChatGPT, Gemini, or Perplexity. These services are based on transformer models that use attention mechanisms to weigh the importance of different words in a sequence. For consumer search, LLMs are particularly useful as

users' queries can be answered in a direct and coherent manner without having to comb through multiple websites. At the enterprise level, these conversational engines can be incorporated into personal assistants and customer service chatbots to optimize ads and cut costs. Speaking with industry participants, Google's click-through rates for GenAI ads is sometimes 5 times the normal returned search results, and these ads are being prioritized in the search because GenAI is making them more relevant. Enterprise search tools like Glean are also being developed to better enable enterprise users to search for company data.

A February 2024 study by Pew Research Center estimated that 23% of U.S. adults have used ChatGPT—up from 18% in July 2023. A different study conducted by *The Washington Post* in May 2024 looked at nearly 200,000 conversations from the WildChat dataset, which has collected over a million real-world ChatGPT interactions. This analysis found that most conversations (~71%) are focused on creative writing and roleplay, homework help, search and other inquiries, and work/business. Other interactions centered on coding, image generation, health and advice, translation, etc.

Google investors debate whether LLMs like ChatGPT will take share and become the dominant source of search over time. Based on recent discussions with advertising industry participants, we do not see significant evidence of advertisers pulling back spend on Google due to ChatGPT. Moreover, our sense is that investors are becoming more confident that Google will be able to transition its core search product to Gemini and enhance existing products like Performance Max, which benefited advertisers by 25%-plus during second quarter 2024, compared to 12%-14% in the prior-year period, according to Chief Business Officer Philipp Schindler.

Advertising

The advertising industry is experiencing immediate benefits from the application of GenAI. The ability to create content using GenAI models makes marketing content less expensive and more efficient to produce. In addition, GenAI is being integrated into campaign management applications like Google's Performance Max (PMax) and Meta's Advantage+. Both campaign managers are meant to streamline advertising for users across Meta's social media sites and Google's search results. For example, Google's PMax will identify what a conversion is for you, allow you to upload assets (logos, headlines, descriptions, products, etc.), create the ad for you, and then ensure that the campaign is optimized for conversions. As a result, going forward, the advertising market may change at a faster rate than previously seen.

In a recent conversation with an industry participant that monitors nearly $1 billion in annual advertising spend across roughly 200 companies on all major platforms, we learned that one of the biggest benefits from implementing GenAI in applications like PMax is the ability to create copy, since the application knows what is being searched, knows what will get clicks, and can personalize ads based on the user's search history. For example, PMax may change the background of a specific ad if it knows you are more likely to click. This is especially beneficial to smaller advertisers that report seeing 8%-10% improvement, compared to large advertisers that claim up to 3% improvement. This equates to around a 5% tailwind for Google if we assume half of its consumers are large advertisers. Google's ads are seeing all-around improvement with higher click-through rates that will lead to more ads populating at the top of searches. The offset to the GenAI advertising advances is the decline in organic search results achieved from search engine optimization.

On the Enterprise Side

Similar to the software landscape that exists today, we do not believe there will be a single killer GenAI enterprise application that will do it all. Instead, there will be many different applications for various roles, processes, and domains. Adoption of these applications will happen at their own pace and will largely be independent of each other. At a micro level, adoption of GenAI will depend on business-specific circumstances and is likely to happen in phases, just like it did for SaaS/cloud.

On a more macro level, the cloud platform shift and modernization cycle took place over 20 years (and is still ongoing), with businesses updating big systems one at a time, depending on resource availability, change management capacity, business prioritizes, quality/maturity of solutions available in market, perceived ROI/time to value, and other factors. We believe the platform shift to AI applications will also depend on these factors and take place over time. However, in our view, the AI upgrade cycle is likely to happen at a faster rate than the upgrade to cloud for a few reasons:

1. *Cloud lays the groundwork.* The prior platform shift to cloud has laid the groundwork for AI adoption. A vast amount of cloud computing infrastructure is already in place and hyperscalers are rapidly spending to expand data center capacity.

2. *AI is easily embeddable.* It is easier to embed GenAI capabilities into existing enterprise applications than it was to lift and shift legacy on-prem apps to the cloud.

3. *Incumbent attitudes are different.* Incumbents in software are embracing the AI platform shift and investing significantly into GenAI capabilities. This has been the case since 2022 when GenAI first came onto the scene. In contrast, during the cloud platform shift, incumbents were largely skeptical of cloud, which likely delayed mainstream adoption by years. For example, Oracle did not start to embrace cloud until 2009, a full 10 years after Salesforce was founded. Adobe did not launch Creative Cloud until 2011. And the turning point in Microsoft's cloud journey was arguably in 2014 when Satya Nadella took over as CEO. These delays are simply not present today, as incumbents are acting with a sense of urgency to drive AI adoption.

4. *Enterprise attitudes are different.* Most enterprises and business leaders have gone through a digital transformation in the last decade or so, and they understand the importance of doing so. For these executives, the AI platform shift likely resembles the cloud shift and is a continuation of that transformation. Fresh memories of irrelevance or market share loss for those businesses that did not embrace digital also serve as warning sign to enterprises not looking to implement AI.

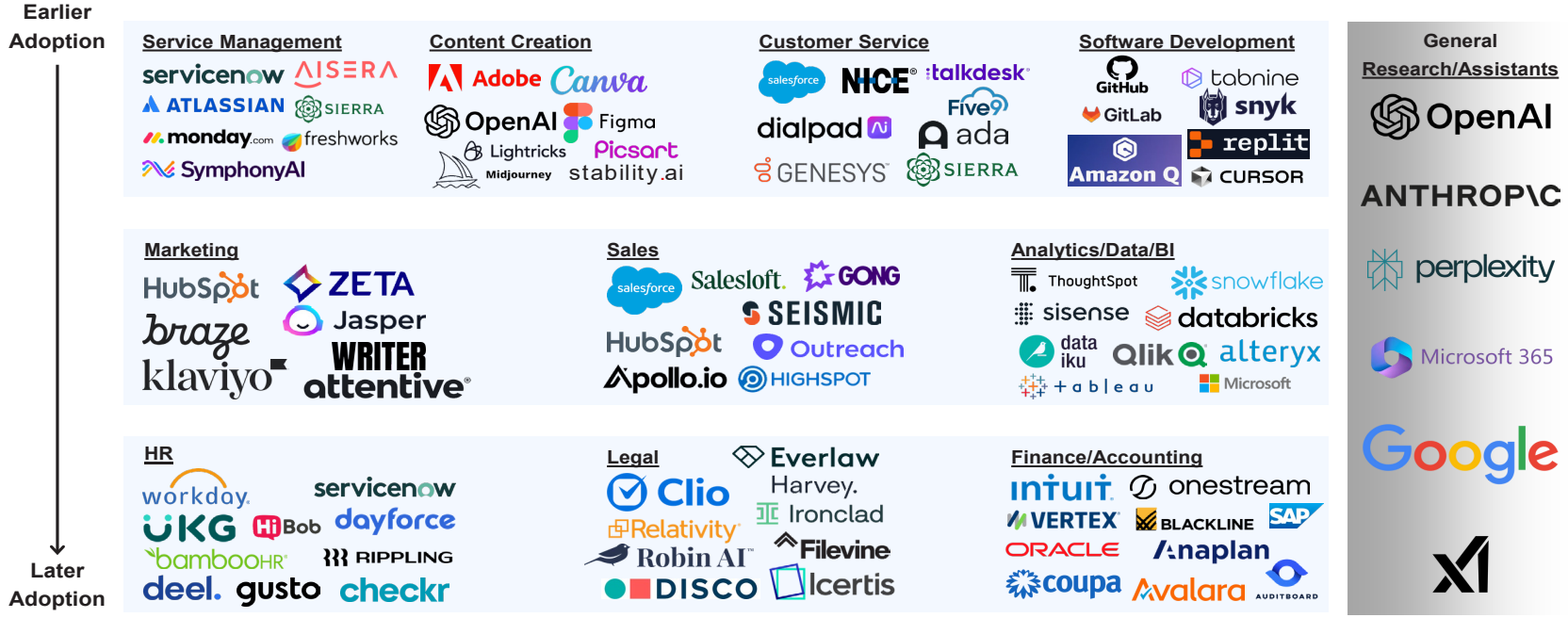
GenAI applications by department/use-case

As we acknowledge above, the adoption of enterprise AI is likely to happen over time and in phases. There is also likely to be a collection of different AI applications for different use-cases/functional areas—not a single app that dominates all workflows. In the exhibit on the following page, we highlight select vendors that are aiming to capture market share with AI capabilities in various functions. We believe domain expertise in each functional area will be critical to delivering quality AI and workflows that businesses can rely on versus a "jack of all trades, master of none" approach.

At the top of the graphic, we include the functional areas where we believe AI adoption will be earliest, and toward the bottom are functional areas where we believe adoption is likely to happen later. Our view is based on our conversations with buyers, vendors, and industry experts. It also incorporates the current adoption trends and our assessment of vendor readiness to deploy AI features and products. In general, we believe the initial enterprise focus for GenAI deployments is to drive revenue or enhance the customer experience. This is followed by a desire to cut costs and automate back-office work in departments like finance and HR.

In our view, the four earliest adoption enterprise use-cases will be service management (largely IT service management), content creation, customer service, and software development. Service management and customer service both currently require a large volume of headcount (even though it may be low-salary employees in many cases) and manual processes that elongate the time to resolution. Use of AI can help customer service and internal agents resolve tickets faster, leading to a better customer and employee experience, while lowering labor costs as well. E-commerce vendor

Exhibit 15
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
AI Adoption by Function Market Map



Source: William Blair Equity Research

Klarna already claims that its homegrown customer service chatbot has reduced hundreds of jobs, while start-ups like Sierra offer a conversational AI platform that enables any company to build an AI customer service chatbot that is personalized to its business.

In contrast, both content creation and software development have seen strong adoption of AI to help highly skilled labor improve efficiency and drive better output. In software development, coding assistants—like GitHub Copilot, GitLab DuoPro, Amazon Q, Gemini Code Assist (Google), IBM watsonx, Cursor, Codeium, Poolside, Zencoder, Merly, Cosine, Tessl, and Tabnine—are seeing broad adoption. While opinions vary on the usefulness of these tools, many developers are seeing productivity gains through faster completion/automation of repetitive tasks and AI's ability to quickly generate unit tests (a particular pain point for developers).

In the second bucket of adoption, we include sales, marketing, and analytics. Sales-and-marketing teams in particular are being targeted as priority areas for AI adoption by the C-suite given their tie into revenue. AI use-cases in sales and marketing will likely be used in conjunction with content creation. Marketing use-cases will take core content creation a step further and deliver personalized and targeted content to customers at the right time with the right message and over the best channels throughout the customer journey.

Adoption of back-office functions that do not touch revenue or the customer experience are likely to be implemented later, in our view. This includes areas like HR, finance, accounting, and legal. We believe AI will still play a big role in these functional areas and can drive improved efficiencies, but it is likely to be a lower priority than in departments that touch the customer experience and have the potential to drive revenue growth (versus back-office functions that tend to only focus on cost savings).

Alongside all these departments are the more general-purpose AI tools and chat-based assistants (like ChatGPT by OpenAI or Claude by Anthropic) that employees may use to help with more minor and general tasks. These are likely not going to be replacements for AI solutions that address specific enterprise workflows. Instead, we view them as more of a complement to a Google search for example.

Our view of the enterprise GenAI adoption timeline

While GenAI has dominated the conversation for what seems like years, it is worth taking a step back and realizing that we are only about two years into a GenAI cycle that is likely to take a decade to play out. We have sensed some frustration and impatience from investors at the lack of revenue tailwinds from AI at the application software layer. In many cases, this frustration stems from what seems like little to no progress in enterprises actually adopting applications, while all the benefit seems to be accruing to the infrastructure layer (semis, networking, servers, data centers, etc.).

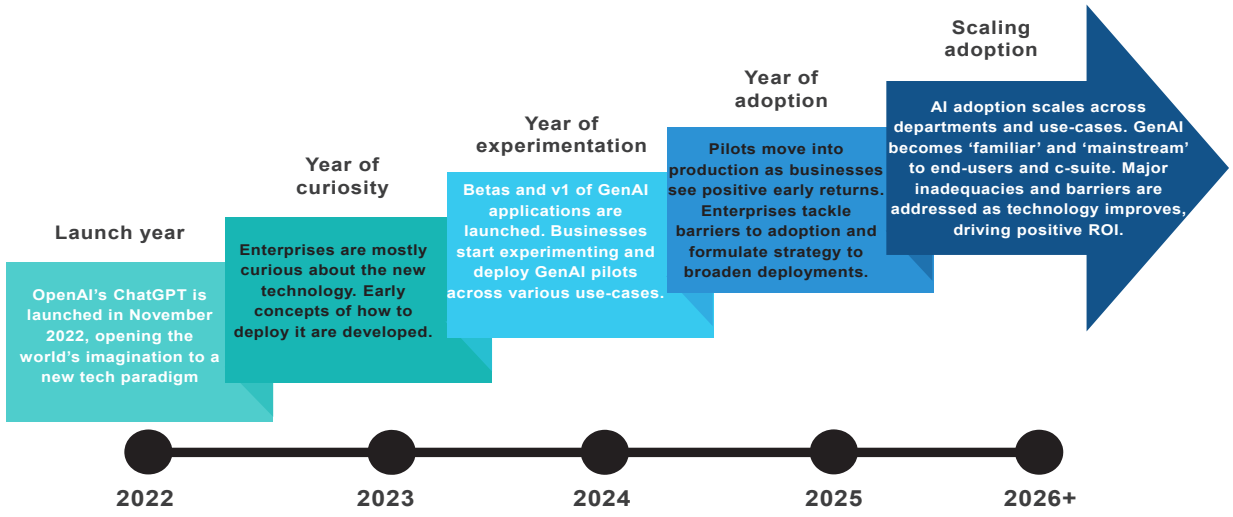
As with any new technology, change happens gradually, not overnight. And GenAI is no different. While it may seem to many like AI enterprise adoption is at a standstill because the financial benefit is not yet evident, important steps are happening behind the scenes that are setting the stage for broader scale adoption and monetization in the not-too-distant future. Put more simply, we have been at the build and experimentation phase of the AI adoption cycle over the past two years, not at the deployment and revenue-generation phase.

In this build base phase, much of the focus is on core infrastructure buildout, not application monetization. For the application software vendors, this largely means focus on building AI applications and integrating them into existing workflows. This is a task that needs to be planned well and done with purpose. While vendors are building their capabilities, buyers are getting ready for

adoption by ensuring data readiness, investing in security and compliance, preparing for change management and employee training, deciding which applications they will prioritize based on business needs and ROI, running pilots and proofs of concept, and much more.

As shown in the exhibit below, much of this foundational work was being done between 2022 and 2024. In 2024 in particular, we heard of many enterprises running AI pilots and experiments to determine what adoption would look like in their organization. In 2025, we believe many of these pilots will move into production, as enterprises are more familiar with AI, have a better sense of how it will be deployed in their organization, and have tackled some of the barriers to adoption (like data readiness). We expect incremental revenue contribution in 2025 from these AI applications moving into production and believe we will see early positive signs of ROI from the early adopters. As AI capabilities become more mature and advanced and as enterprises become more ready to adopt, we believe 2026 and beyond is when AI applications will become more mainstream and revenue contribution will kick into high gear.

Exhibit 16
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
William Blair View on GenAI Adoption Timeline in the Enterprise



Source: William Blair Equity Research

We are already seeing signs of (early) enterprise GenAI adoption

While revenue contribution from GenAI adoption in the enterprise was limited in 2024, there were many positive early signs of adoption, and in some rare cases (like for ServiceNow) also real monetization that followed. In exhibit 17, we compiled some of the early-adoption data points that public application software vendors have disclosed about their AI products.

With its broad reach and distribution, Microsoft has seen positive trends on two fronts. First, GitHub was early in getting its copilot out to market, and as of June 2024, it had 77,000 customers using GitHub Copilot with over 1.8 million paid subscribers. That likely implies ARR of about $200 million-$300 million, assuming a realized price of $10-$15 per user per month. Second, over 100,000 organizations are using Microsoft 365 Copilot (list price of $30 per user per month), which is integrated into core applications like Word, Excel, PowerPoint, Outlook, and Teams. Monetization on Microsoft 365 Copilot is unclear at this point, as we suspect many of these deals have been bundled into the broader M365 suite. While user feedback on the usefulness of M365 Copilot has been mixed at best, the scale of adoption is still impressive.

Exhibit 17
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Select AI Adoption Disclosures

Company/GenAI Product	GA Date	Relevant Adoption Data
Microsoft GitHub Copilot	Jun-22	**77,000 customers** using GitHub Copilot as of June 2024, up 180% year-over-year
ServiceNow Now Assist	Sep-23	Likely approaching **~$200M in ARR** from Now Assist*
Microsoft 365 Copilot	Sep-23	Over **100,000 organizations** using M365 Copilot as of September 2024, including 70% of the *Fortune* 500
Adobe Firefly	Sep-23	Over **16B generations** with Firefly as of September 2024
Zoom AI Companion	Sep-23	Over **4 million accounts** have enabled Zoom AI Companion as of October 2024, including 57% of the *Fortune* 500
NICE Enlighten AutoSummary	Oct-23	**140 AutoSummary customers** as of September 2024, including 50 that signed in the September 2024 quarter
Salesforce AI (incl Copilot, Agentforce, and other)	Copilot - Apr-24 Agentforce - Oct-24	Signed more than **2,000 AI deals** in October 2024 quarter, including **200 deals for Agentforce** (only ~2 weeks of availability). $1M+ deal wins with AI more than tripled

*William Blair estimate based on management commentary

Sources: Company commentary and William Blair Equity Research

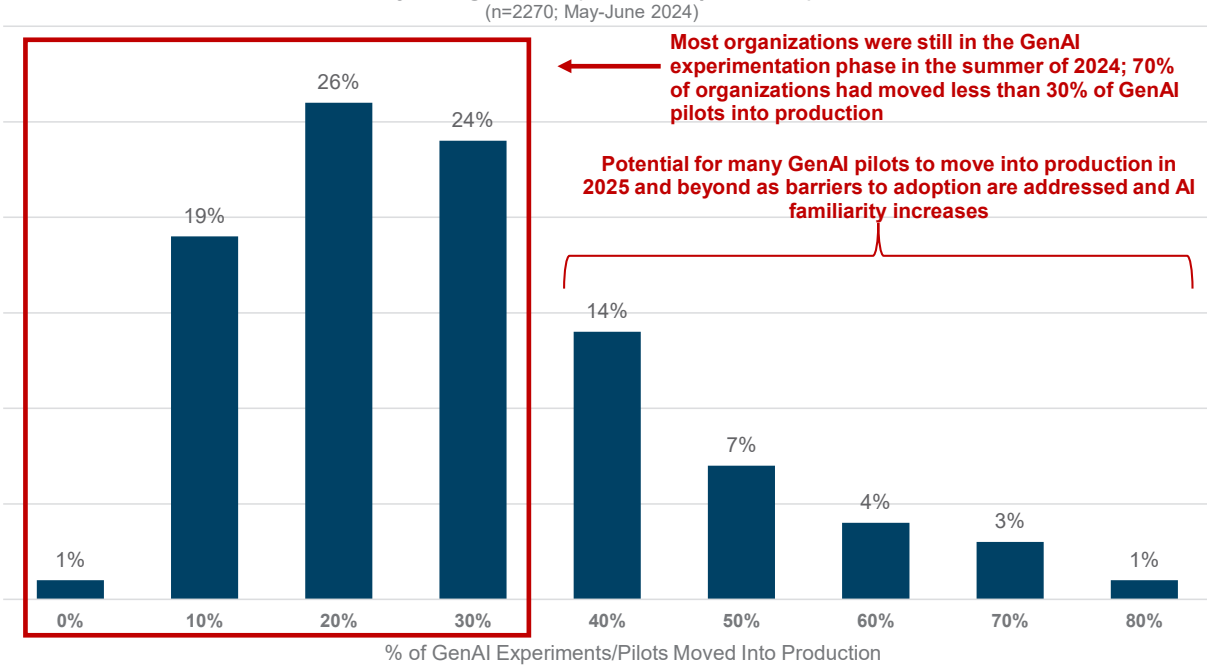
ServiceNow is another one that sticks out and is one of the few companies that has sold and monetized its AI capabilities (called Now Assist). Based on our estimates of the company's large-deal disclosures, ServiceNow is likely approaching $200 million in ARR generated just from selling its Now Assist SKU as of September 2024. While this is just 2% of its $11 billion in expected subscription revenue in 2024, the company has only been selling this solution for 12 months into a market where many participants are not yet ready to purchase AI capabilities. Equally impressive is that its customers are seeing ROI in the Now Assist solution and are willing to make large investments into it. ServiceNow has 44 customers spending over $1 million per year, 6 customers over $5 million, and 2 customers over $10 million per year. ServiceNow's monetization, along with the other examples where adoption is high (even if monetization has not started yet), leads us to be optimistic that broader monetization and revenue generation will follow in future years.

Enterprise AI adoption is poised to increase in 2025 and beyond

As a reality check, Gartner predicts that at least 30% of AI projects will be abandoned after proof of concept by the end of 2025, due to "poor data quality, inadequate risk controls, escalating costs or unclear business value." Similarly, data from a recent Deloitte survey show that the vast majority of enterprises have moved less than 30% of their AI pilots into production as of the summer of 2024 (see exhibit 18 below). Given the nascent phase of the AI adoption cycle (see our prior discussion around exhibit 17), this is roughly in line with our expectation that most enterprises are still figuring out the best way to adopt AI. With only 7% of organizations having moved over half of their GenAI pilots into production, this means there are many ongoing experiments that are still left to be broadly deployed, which we believe will take place over the coming years.

Exhibit 18
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Few GenAI Pilots Moved Into Production Through the Summer of 2024

In your estimation, what percentage of your Generative AI experiments have been deployed to date into your organization (moved into production)?
(n=2270; May-June 2024)

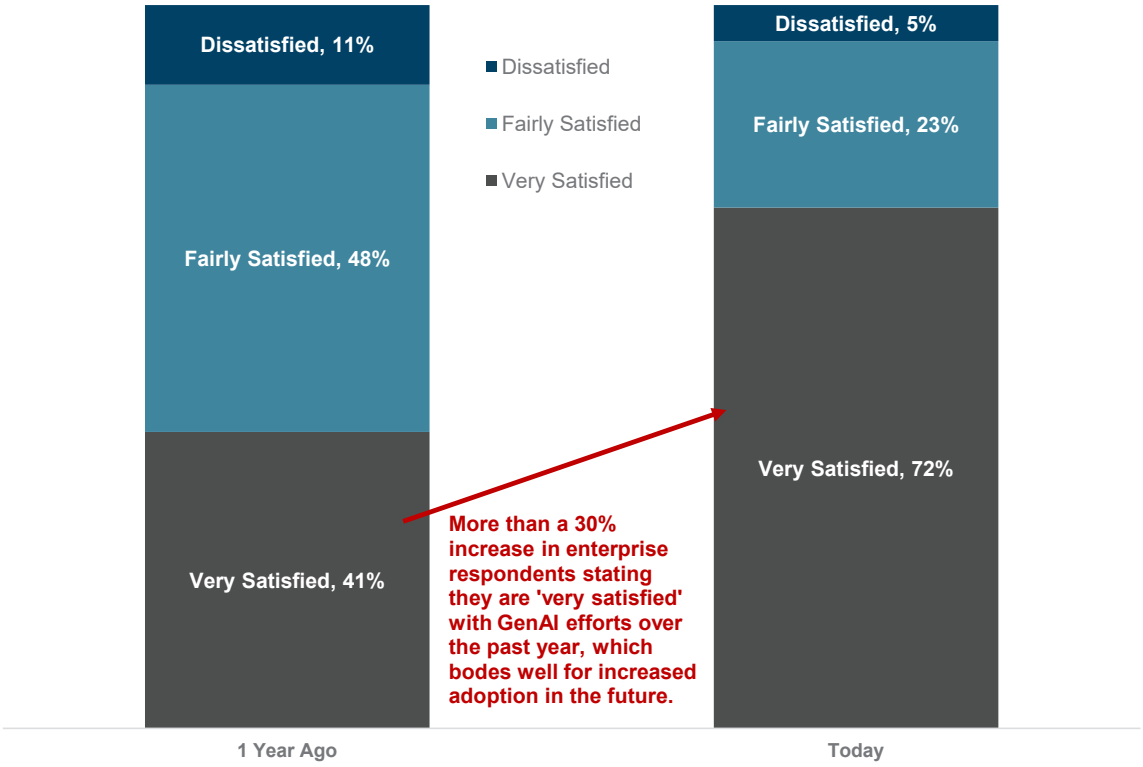


Sources: Deloitte State of Generative AI in the Enterprise and William Blair Equity Research

Certainly, not all of these pilots will move into production for various reasons, but the data suggest that many will. According to a survey from NTT Data conducted in the fall of 2024, enterprises are generally "very satisfied" with their GenAI efforts and have become increasingly pleased with GenAI compared to one year ago. As shown in the exhibit below, 41% of respondents stated they were "very satisfied" with their organization's GenAI efforts one year ago. When asked in the fall of 2024, that figure had jumped to 72%, more than a 30% increase. Again, we believe time, familiarity, and improving products drove this increase. It is also promising for future adoption trends of GenAI as it indicates a positive experience and good ROI from GenAI investments thus far (which are likely pilots and experiments thus far).

Exhibit 19
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Increasing Satisfaction With GenAI in The Enterprise

How would you describe your organization's satisfaction with its GenAI efforts?
(n=2307; September-October 2024)



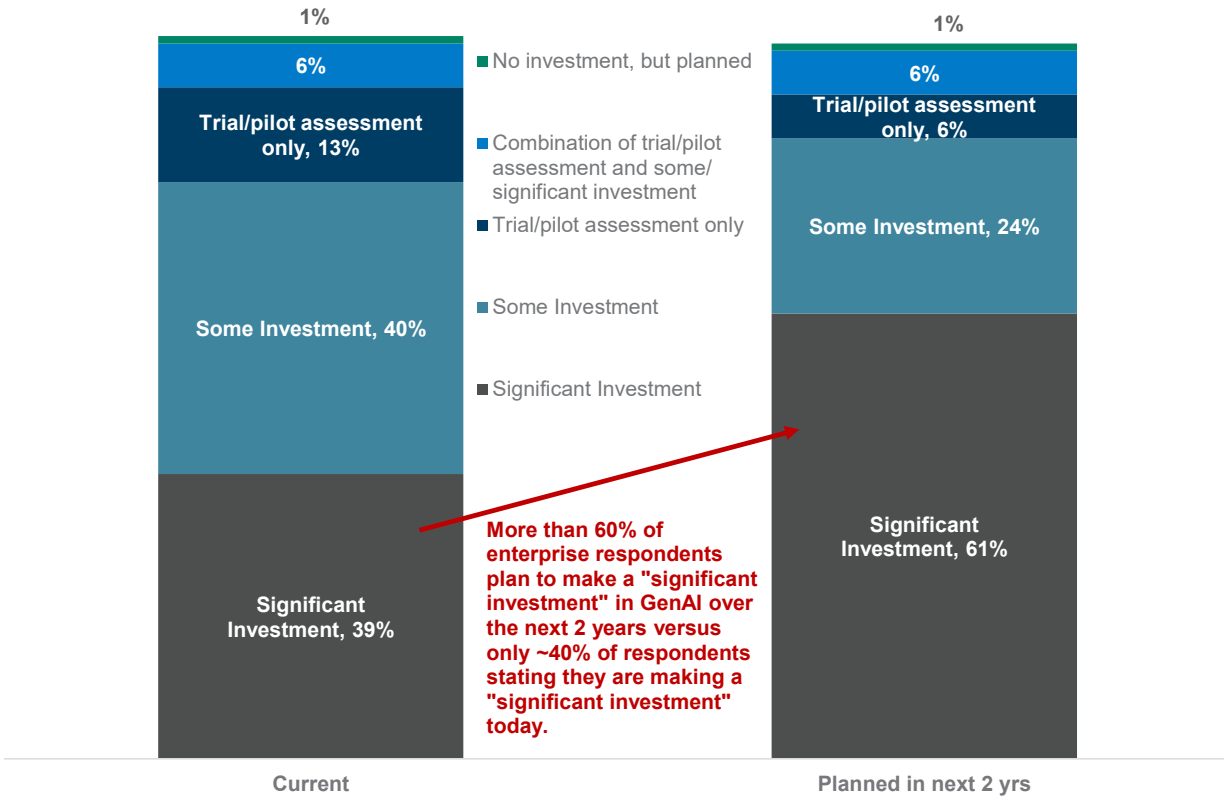
Sources: NTT Data Global GenAI Report and William Blair Equity Research

Data from the same survey shows that more than 60% of enterprises plan to make a "significant investment" in GenAI over the next two years (with 24% stating they will make "some investment"). While this is promising on its own as it shows a decent-sized pool of early adopters, it is even more encouraging for growth to see that currently only about 40% of respondents stated that their organizations are making a "significant investment" today. Over the next two years, a large number of enterprises plan to ramp up their investment in GenAI. This supports our earlier conclusion that 2025 and 2026 will be important years for GenAI adoption in the enterprise as applications move from pilots to production to scaled deployments. Throughout this process, we believe enterprise software vendors with quality AI solutions in market will start to see the revenue tailwind (we have identified many of these vendors in exhibit 17 above).

Bottom line: There is not likely to be a single killer app in GenAI due to the broad applicability of the technology. On the consumer side, we are already seeing rapid adoption across areas like search, advertising, and content creation. On the enterprise side, we expect to see signs of a steady increase in adoption (and monetization) of GenAI applications in 2025 as pilots and experiments move into production. However, it is still early days, and we believe broad and scaled adoption will come over time as businesses become more comfortable with the new technology and tackle barriers to adoption (privacy, security, and change management).

Exhibit 20
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Enterprises Planning for Significant Investment in GenAI Over the Next 2 Years

What best describes your organization's investment in GenAI?
(n=2307; September-October 2024)



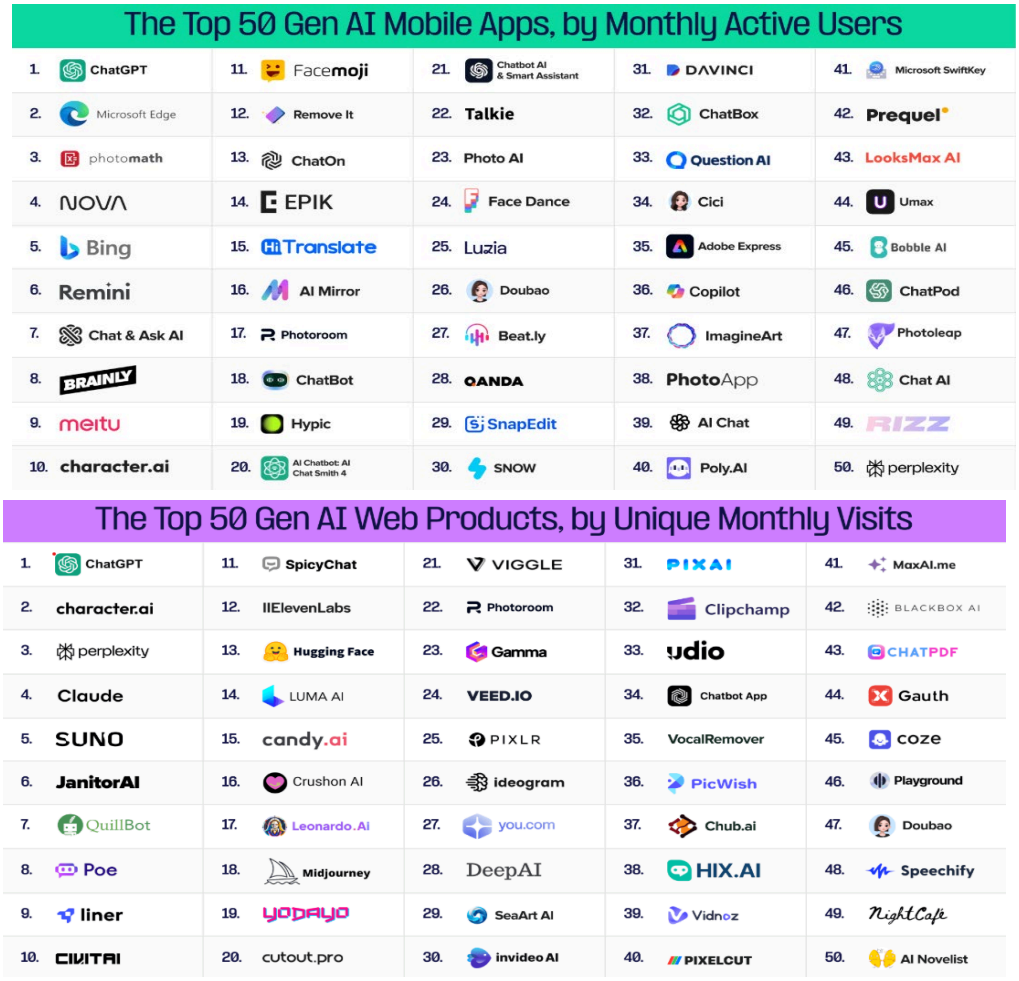
Sources: NTT Data Global GenAI Report and William Blair Equity Research

How Is AI Being Monetized at the Application Layer?

For Consumer Apps

The amount of GenAI companies emerging for specific use-cases continues to grow. To track the development and performance of each of these companies we leverage Andreesen Horowitz's list of Top 100 GenAI Consumer Apps, which released its third edition in August 2024. This ranks the most-visited AI applications across both web and mobile, while analyzing trends in consumer engagement across different modalities/use-cases.

Exhibit 21
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Andreesen Horowitz Top 100 GAI Consumer Apps



Source: Andreesen Horowitz

Our own analysis of the business models of 30 different emerging AI companies revealed that 28 of them are currently operating under a freemium business model. We believe this is a promising strategy for many of these early-stage AI developers as they can scale their user base through brand and performance marketing, while innovating their product and increasing the share of paying users over time. Open AI's ChatGPT serves as a good example of the freemium model, since anybody can log into ChatGPT and use its search feature with limited access and generate two free images a day using DALL-E. In addition, ChatGPT Plus is available for $20 a month, giving access

to five times the number of messages along with data analysis, image generation, and early access to new models and features. Backlinko estimates that in August 2024 OpenAI had over 10 million ChatGPT Plus subscribers and 1 million enterprise subscribers, bringing monthly revenue to about $300 million.

The average base subscription price of the GenAI products analyzed is under $10 per month with many offering additional more expansive tiers for team or enterprise use. Recently, we have seen some AI developers experiment with advertisements as an additional revenue stream. Specifically, Perplexity announced in October 2024 the incorporation of ads into its platform in the form of sponsored, follow-up suggestions on which the user can click. Since 46% of Perplexity's queries lead to follow-up questions, the company is optimistic that these new advertisements will drive consumer engagement.

On the Enterprise Side

We see four primary ways to monetize GenAI. Most software companies are either directly monetizing AI today or have plans to do so in the future once their AI capabilities are more robust. Just like on the product front, companies are experimenting with their AI pricing models and iterating to determine what works best. Most companies, however, have been less focused on monetizing thus far and more focused on driving engagement and delivering value. This is the classic software (and even internet) model philosophy: get your users hooked on using your product today, and once they are dependent on you, take price up over the medium/long term. This has worked for many of the biggest companies today and has been a key driver of their success (largely with internet/consumer companies), but it remains to be seen how much patience investors have for software companies taking a similar approach, especially if it puts pressure on gross margins.

Thus far, we have seen four primary ways that software companies are monetizing AI, though the pricing debate and what is best for which business is far from being settled.

- *Add-on pricing.* Customers have to buy a separate and distinct AI product from their core subscription. From our experiences, this is the most common model we see today, including from the likes of ServiceNow, Microsoft, and Salesforce. Add-on pricing offers the most distinct monetization for AI capabilities, but also comes with the added friction of forcing customers to allocate incremental dollars directly for AI capabilities.

- *Embedded pricing.* In this model, AI capabilities are embedded into the core product and users get access to those capabilities if they are subscribed. Companies that use embedded pricing typically only include their high-value AI capabilities in premium pricing tiers, aiming to drive customers to upgrade to premium offerings. This makes it more challenging for vendors to determine whether AI is driving the upgrade or if it's other functionality that a customer also gets as part of the upgrade. However, this is typically lower friction for customers. Companies like Canva, Atlassian, and Smartsheet use embedded pricing.

- *Token-based pricing.* We debate whether token-based pricing warrants a category of its own as it is a subcomponent of add-on or embedded pricing. This biggest difference is that this more closely aligns with a usage-based model as opposed to a seat-based model. With token-based pricing, customers buy a specific number of tokens that they can then use for some AI use-cases. Each use of an AI capability has a designated token value, and any unused tokens at the end of the subscription period typically expire. Token-based models make it somewhat difficult for customers to initially forecast how many tokens they need. The benefit for vendors is that it allows for subscription revenue recognition and protects margin downside from heavy AI users. For this reason, many vendors have a token-based component as a part of their AI add-ons. Adobe and ServiceNow are examples of companies that use token-based pricing.

- *A hope and a prayer.* We say this a little tongue in cheek, but there are some software companies that have decided (at least for now) to embed GenAI capabilities into their offerings for free. This is done with the goal of using AI capabilities to gain an edge over competitors and take share. The most prominent example of this is Zoom, which includes its Zoom AI Companion for free with all paid plans. This is a good way to drive engagement and theoretically a good way to take market share from competitors, but it comes at the cost of gross margin. For example, at Zoom, the incremental AI costs as a part of AI Companion have driven gross margins to compress about 100 basis points year-over-year. Of course, Zoom has the option to take price higher on its core products since its AI capabilities are now included for free or even break out AI Companion as a separate SKU when it pleases. Until then, however, it is left absorbing the incremental AI costs without incremental AI revenue to show for it.

Seemingly inevitable transition to usage- or value-based pricing models in enterprise software
Most software companies today have user- or seat-based pricing models, which has been the standard in enterprise software for many years. As GenAI introduces more automation for specific functions over time (like customer service or IT helpdesk agents, for example), there is a legitimate argument to be made that seat-based pricing models will no longer be adequate to capture the incremental value delivered by GenAI applications. Inherently, increased automation should change the nature of both highly technical jobs (such as developers or creative professionals) and low-skilled ones (like customer service or entry-level sales rep), to the point where the growth in headcount of these roles is likely to be impacted. This would have a negative downstream impact on software models with seat-based pricing, which is counterintuitive to some extent considering that it is the software platforms that are enabling lower headcount growth.

Increasingly, software companies are looking to transition from seat-based pricing models to usage-based pricing or value-based pricing. This makes sense to us—a usage- or value-based pricing model is the most effective way for software vendors to capture the value they are delivering in a GenAI–powered world, as they are effectively displacing labor costs with technology. For a customer service use-case, this would mean pricing based on how many inquiries AI agents are able to resolve versus pricing based on the number of human customer service agents who are using the software. Similarly, for creative professionals, this would mean a limit on the number of images that are generated over a specific period.

Effectively, we believe this means that software companies have pricing power over the long term (at least those that are on the front foot with GenAI) as they will be the enablers of automation and labor cost displacement. However, the transition to introducing a usage-based component to the pricing model (or fully going usage-based) creates an unknown variable for software models. Pricing changes are certainly not an insurmountable problem for companies to address, though care must be taken to ensure that the new pricing is not cannibalistic or churn-inducing, that it is easy for customers to understand so they can determine ROI and plan budgets effectively, and that it protects or enhances vendor margins.

Bottom line: We are confident that consumers/enterprises will see value in GenAI technology, with monetization taking multiple forms and price discovery still in process.

What Is Agentic AI, How Does it Differ From AI Copilots, and What Does it Mean for Enterprise AI Adoption?

The world of enterprise software and automation has quickly pivoted away from copilots as the best way to consume GenAI and moved rapidly toward AI agents. This transition has been swift; the largest tech companies have quickly launched their own AI agents and AI agent building capabilities in the back half of 2024 in a bid to increase AI adoption (exhibit 22 highlights some of the major agentic AI launches over the past year). We believe the shift was driven by relatively lackluster adoption and excitement around initial copilot products, which lack the tangible ROI that enterprises get from the more autonomous capabilities of AI agents. The ability of AI agents to act more autonomously and remove humans from a process is the key reason we expect agentic AI to drive material AI adoption in the coming quarters and years.

Exhibit 22
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
AI Agent Launches by Major Tech Players

Company	Agentic AI Solution	Launch Date	Use-Cases
aws	**Bedrock Agents**	**Apr-24**	Build your own
servicenow	**Now Assist**	**Sep-24**	IT service management, customer service management
salesforce	**Agentforce**	**Sep-24**	Customer service, sales development, sales coach, personal shopper
HubSpot	**Breeze**	**Sep-24**	Content marketing, social media, sales prospecting, customer service
ANTHROPIC	**Computer Use as part of Claude 3.5 Sonnet**	**Oct-24**	Build your own
Google	**Google Agentspace**	**Dec-24**	Marketing, sales, software/IT, R&D, HR
Microsoft	**Microsoft Agents**	**Dec-24**	HR, IT, project management, customer service, marketing, sales
OpenAI	**OpenAI Agents**	**To be launched early 2025**	TBD

Sources: William Blair Equity Research and company announcements

The timeline of these launches is notable as most occurred in late 2024, meaning there has not been much time thus far to evaluate agentic AI adoption. However, based on industry conversations, CIO feedback, and our interactions with management teams, we believe there is greater customer appetite for AI agents than there has been for copilots. Microsoft, for example, is charging $30 per user per month for Microsoft 365 Copilot, and we have consistently heard feedback from many enterprises that it has been challenging to justify the incremental cost to roll this out enterprisewide. Though price has been a big part of the pushback, customer feedback suggests that finding applicable use-cases for Microsoft 365 Copilot has also been part of the challenge. One reason we believe AI agents have greater potential is that agents are typically developed to be use-case specific—e.g., customer service agent, sales prospecting agent, employee onboarding agent; they remove an additional step that buyers would need to take to brainstorm how to deploy AI.

While both copilots and AI agents are powered by LLMs/foundational models, there are big differences between the two. First and foremost, copilots serve as aides to help humans do their jobs or complete a task. In this case, the human typically prompts the copilot for what they need help with. Only then does AI get involved. AI agents, on the other hand, operate largely autonomously and independent of human interaction. They are designed with a specific goal in mind and take an action (or multiple actions) to achieve that goal. With AI agents, humans will only intervene when necessary, largely to handle exceptional circumstances. With copilots, the human still serves as the bottleneck in terms of efficiency because the human is quarterbacking the process. In contrast, because AI agents work without human interaction and can work as needed 24/7, the user is no longer the bottleneck in the process. This is a big reason why the ROI for agentic AI is more evident and easier to understand for enterprises buyers.

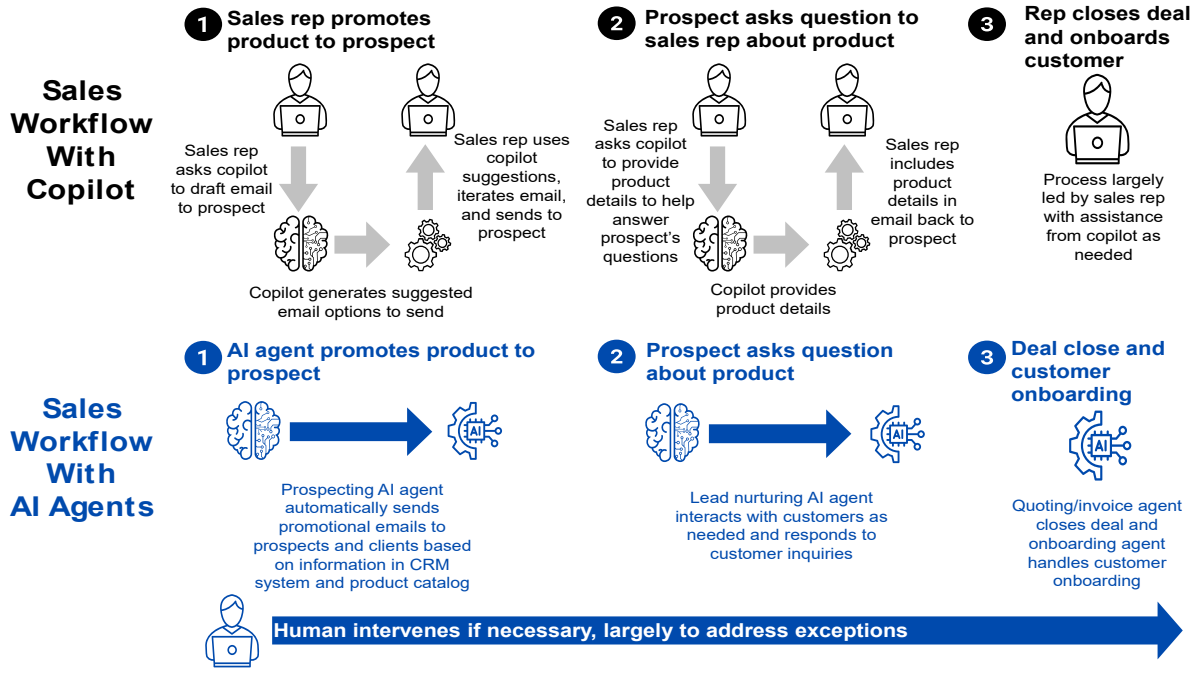
Exhibit 23
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Differences Between Copilots and Agentic AI

Copilots	Agentic AI
Human takes action, possibly guided by copilot or with copilot assistance	AI takes action without human intervention or human prompts
Aids humans, often by responding to a prompt	More autonomous; AI runs fully automated end-to-end process
Human must be involved in the process	Human intervenes only if needed (for exceptions)
Humans are still the bottleneck	Can make decisions and take actions to accomplish a goal without human oversight
Information oriented	Action oriented
More reactive	More proactive
Typically priced on a per (human) seat basis	Typically priced on a consumption basis, and in the future, likely on an outcome basis

Source: William Blair Equity Research

To illustrate the real-world impact of agentic AI, we compare a sales workflow run by an agentic AI process compared to a copilot-based AI process. In the agentic process, the human essentially serves as a supervisor of the sales process, while the AI agent does all the work. In contrast, the human (a sales rep) is in charge of the copilot sales process and is driving it forward. The latter approach requires constant human involvement at all steps from prospecting to lead nurturing to deal closure and onboarding. For any enterprise with a somewhat standardized sales process, the choice between deploying AI copilots to help sales teams versus AI agents should be obvious, both from an effectiveness perspective (customer response time, win rates, etc.) and a cost perspective (agents should allow for more automation and less human involvement over time).

Exhibit 24
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Sales Workflow With Copilot Versus Agentic AI



Source: William Blair Equity Research

It is also worth pointing out that in the agentic AI process depicted above, there is not one single AI agent operating throughout the sales process. There are multiple agents, each with its own goals and skillsets, operating at various stages. There is an agent for prospecting, one for lead nurturing, one for invoicing, and one for customer onboarding. We believe this specialization is likely to occur in an agentic AI world with each agent likely powered by a specialized SLM that helps that agent to efficiently achieve its objectives.

While we believe agentic AI will be a game-changer for AI adoption, enterprises will have to contend with the potential challenge of AI agent sprawl, especially in view of the rapid proliferation of AI agents that we have already seen. As such, we believe orchestration of AI agents will play an increasingly important role in the enterprise. Though it is still early to identify how this will play out, we expect to see AI agents that specialize in orchestrating other AI agents, much like human supervisors that manage other human workers. Several software vendors, including ServiceNow, Aisera, Salesforce, and UiPath, are already positioning themselves to serve this critical orchestration role.

Bottom line: We expect agentic AI will be a major catalyst to enterprise AI adoption as AI agents are more autonomous than copilots, operate without human intervention, and can solve complex, multistep problems.

What Are the Main Barriers to Enterprise GenAI Adoption?

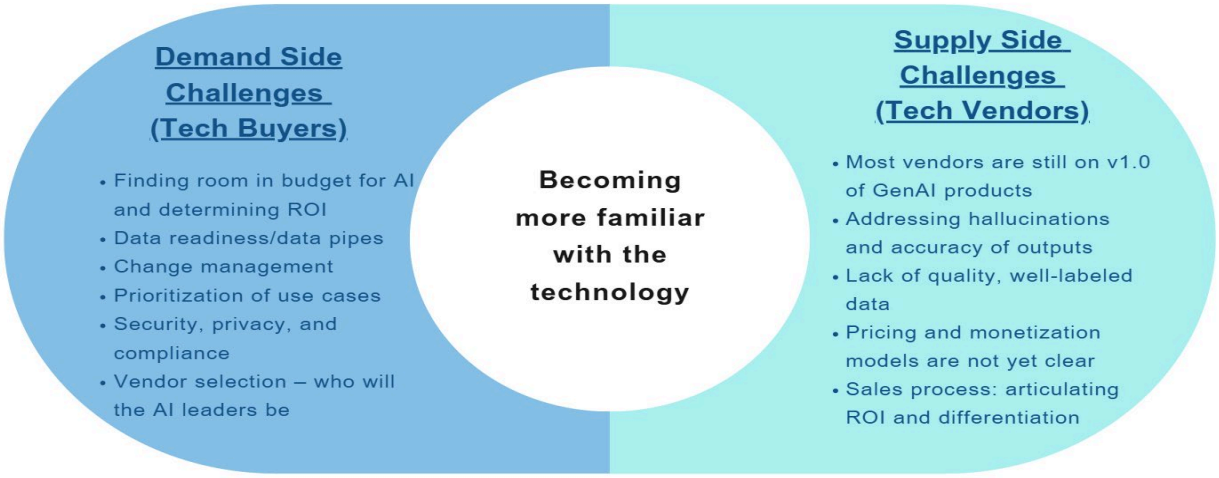
In our view, the main barriers to AI adoption in the enterprise thus far have been two-sided and can be broken up into challenges that: 1) tech buyers need to address (we call these demand-side challenges) and 2) tech vendors need to address (we call these supply-side challenges). Given how early it is in this platform shift, both tech buyers and tech vendors share some of the burden of the limited adoption we have seen thus far. On the tech buyer side, most of the challenges revolve around getting their internal house in order before tackling this new tech paradigm. This revolves around budgets, data readiness, capacity for change management, and figuring out where AI will actually have the most impact in their organization. None of these are insurmountable challenges by any means, but they require planning and time. And with the C-suite and boards pressuring companies to quickly develop their GenAI strategy, we believe these barriers will be addressed in due time.

On the tech vendor side, the issue mostly relates to the immaturity of the technology. After all, it has only been two years since GenAI came onto the scene. Since then, tech vendors have had to rapidly deploy new solutions that incorporate the technology, figure out how to make it accurate, and develop a solution that actually delivers value to customers.

Any sort of product development is an iterative process. A product needs to be put into market, users need to test and use it so the vendor can collect feedback on what is and is not working, and ultimately the vendor needs to make the necessary changes to improve its product and drive higher ROI for customers. This takes time, and thus far, most software vendors are only on the version 1.0 release of their GenAI products, which were mostly launched in the back half of 2023. However, as it has been over a year since initial product launches, second and third iterations are not far behind. We believe that many newer versions of GenAI applications will be launched by software vendors in 2025. These should be received well by customers and more scaled adoption should follow.

What both of these challenges ultimately have in common is time and familiarity. The more that is known about any technology, the easier it is for vendors to develop around it and buyers to plan for it. Given how much time is being dedicated at enterprises to learning about GenAI, it is hard to imagine that familiarity is not increasing at a rapid pace.

Exhibit 25
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Barriers to GenAI Adoption - Supply Side and Demand Side

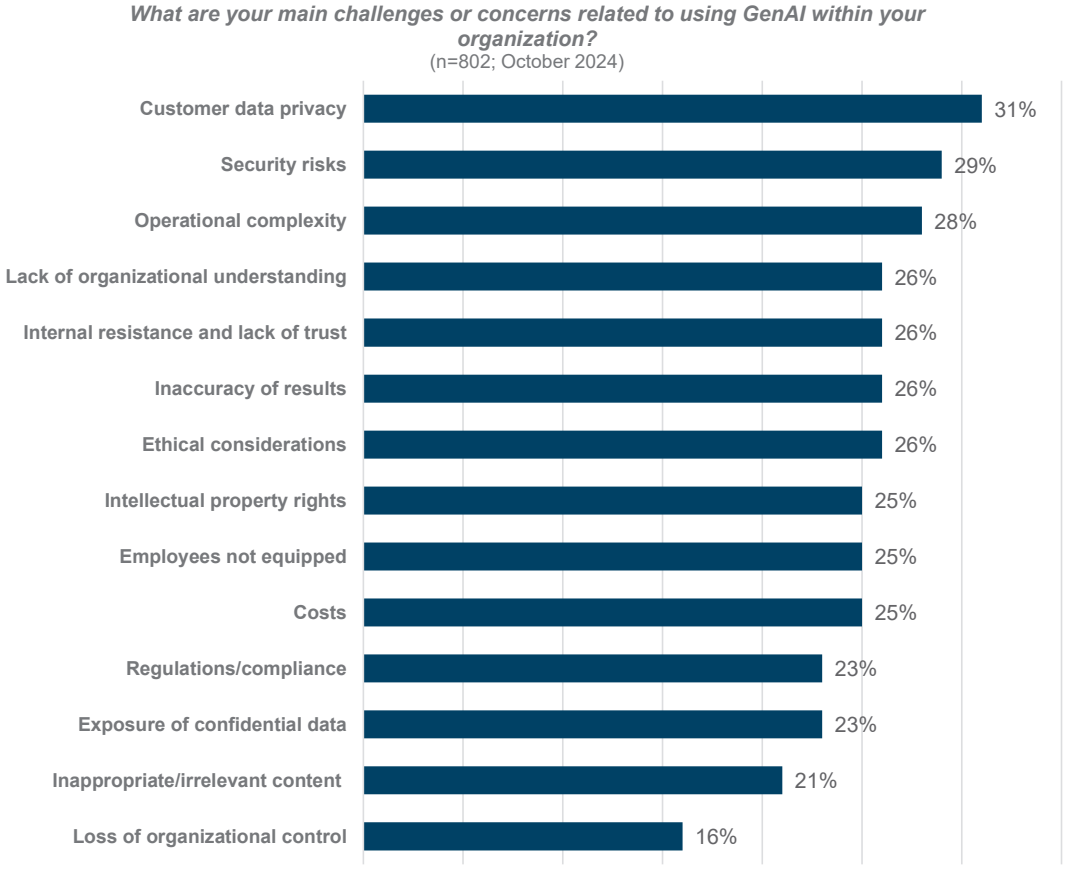


Source: William Blair Equity Research

According to a recent survey by the Wharton School and GBK (exhibit 26 below), the biggest challenges to GenAI adoption revolve around data, security, change management, lack of understanding, and inaccuracy. While the concerns are warranted, we believe these are normal for enterprises considering any new technology.

Bottom line: With time, development, guardrails, and increasing familiarity, we believe current barriers for GenAI will get knocked down over the coming years, just as they did for cloud in the 2010s.

Exhibit 26
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Primary Barriers to GenAI Adoption in the Enterprise



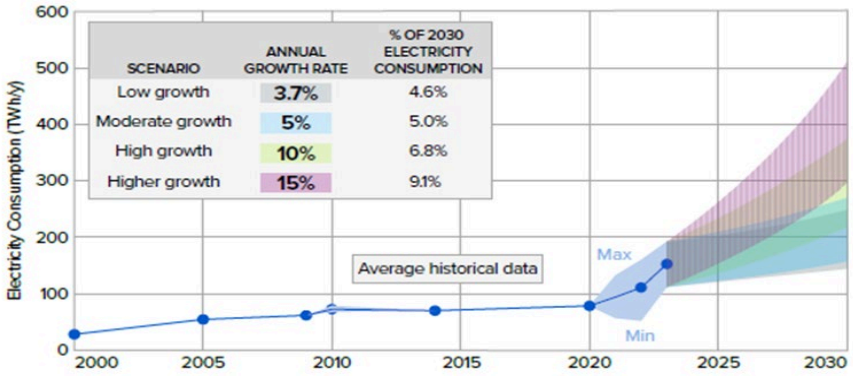
Sources: AI at Wharton and GBK Collective and William Blair Equity Research

What Are the Main Physical Bottlenecks in the GenAI Buildout?

Energy Is the Ultimate Arbiter of AI

Demand for data centers has been building for the past 20 years, as increasingly more business processes are digitized. However, until recently, this demand was manageable. The advent of LLMs and an "electrify everything" approach to climate change are weakening an increasingly fragile and outdated power grid. In our August 2024 report (The Power Behind Artificial Intelligence), we explored the opportunities and limitations of AI and determined that energy will be the ultimate arbiter of growth. Over the past decade, U.S. data centers have accounted for 2% of total electricity consumption, and demand has grown at about 1% annually. In a bull scenario, we estimate AI could cause U.S. data center electricity demand to inflect up to 15% growth annually, consuming 500 TWh of electricity by 2030 and accounting for 9% of total U.S. electricity consumption in that year.

Exhibit 27
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Projected Electricity Consumption for U.S. Data Centers



Source: Electric Power Research Institute, Inc. "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption" 3002028905 (2024)

It is important to understand AI's role in data center growth and energy demand because LLMs used in services like ChatGPT require much more energy than a traditional web search. GPT 4 is estimated to demand between 1.2 kWh and 1.5 kWh, which compares to traditional web search at roughly 0.0003 kWh, or 10-430 times more energy per query. These models will only improve, suggesting that these are baselines versus end-state figures. Data centers for cryptocurrency or blockchain pose an equally daunting demand scenario. According to CoinDesk, every Bitcoin transaction consumes roughly 1,000 kWh of electricity compared to a Visa transaction that consumes 0.0003 kWh. Once again, this suggests a factor of 300,000 times more energy per transaction.

While technocrats will often argue that "free" input renewables such as solar and wind push marginal energy costs to near zero, making renewable power the cheapest option to meet new demand, our data concludes the exact opposite. In fact, because of how our grid and our society is structured around dispatchable power on demand, penetration above 5% of renewables in the grid architecture quickly shifts from the least expensive to the most expensive generation asset. While battery storage helps (and will certainly be part of the solution, see below), we conclude that the greatest near-term beneficiary of AI-driven energy demand will be natural gas. Longer term, we believe that advanced nuclear solutions must not only be part of the conversation but might be central to grid architecture.

Natural Gas Wins the Flexibility Game

Each power generation asset is a bit different, and they are therefore utilized differently in the grid. Nuclear power, for example, provides very steady baseload generation, so it helps meet demand, but it is inflexible, meaning that it cannot ramp up or down to meet demand over the course of a day. Coal generation is like nuclear—once the flywheel is turning it cannot change easily. Natural gas generators come in a variety of assets: 1) combined cycle turbines, 2) simple cycle turbines, 3) cogeneration (combined heat and power), 4) microturbines, and 5) fuel cells. Some of these turbines are referred to as "peaker" plants. As the name implies, they can spin up and down in short periods of time, adding incremental generation relatively quickly. This makes natural gas much more flexible than nuclear or coal. Natural gas generation is therefore often used to meet demand throughout the day.

How Does Nuclear Fit in?

Nuclear power production is a perfect match for the electricity needs of data centers since data centers need a steady, large supply of electricity, and that is exactly what nuclear facilities provide. However, there is a time mismatch between supply and demand. The demand from data centers will come within the next decade, while a new electricity supply from nuclear will take at least a decade. We therefore see new nuclear as more of a long-term play.

Existing nuclear will be in high demand to enter long-term power purchase agreement (PPA) deals directly with data center off-takers. For example, Talen Energy, the owner of 2,228 MW of nuclear power at the Susquehanna Nuclear Facility in Pennsylvania, is entering into a $650 million PPA to sell power directly to Amazon Web Services, which is co-locating with the power facility. This type of co-location and long-term PPA will become increasingly appealing in the short term as AI data centers look for secure, stable, and large power supplies.

Nuclear is increasingly being sought by countries around the world. According to data in the McCoy power reports, China, Poland, India, and Ukraine all anticipate GW-level nuclear capacity addition in the next decade. As of now, the U.S. has requested only 17.5 MW. As power demand grows in the future, a nuclear renaissance could occur in the U.S.

Exhibit 28
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Announced Nuclear Reactor Capacity Additions

Country	Number of Units	Capacity (MW)
China	8	9420
Poland	5	6550
India	2	1400
Ukraine	1	1100
Russia	11	715
Canada	2	400
USA	2	17.5
South Africa	1	8.5
South Korea	1	5

Source: McCoy Power Reports

Challenges of Solar and Wind Energy

The challenge posed by variable renewable energy (VRE) like solar and wind is different still. VREs are predictable at the grid level given weather forecasting, but they are inflexible in that the balancing authority cannot ramp up generation from a solar facility if the sun is not shining. VREs are either producing power or they are not, and the grid system must manage that generation as best it can. To be sure, VREs can help to meet demand if generation happens to align with demand, but for grid resiliency and planning purposes, the balancing authority cannot rely on VREs the same way they can rely on, for example, a natural gas combined cycle generator.

To understand how each resource contributes to resiliency, many BAs calculate something called the effective load-carrying capability (ELCC). According to the PJM, the "ELCC provides a way to assess the capacity value (or reliability contribution) of a resource (or a set of resources)," or perhaps said differently, the ELCC is a "measure of the additional load that the system can supply with a particular generator of interest, with no net change in reliability."

According to the PJM's calculations, thermal resources are the most reliable (ELCC=81%), followed by demand response programs, then storage, and lastly, variable renewable energy like solar and wind (exhibit 29). In fact, within the VRE category, fixed-tilt solar (which we will term "standalone solar" to distinguish it from solar plus battery systems) has the worst ELCC (exhibit 30). Standalone solar has ELCC values of 9% for fixed-tilt solar and only 14% for tracking solar. This means that for every 1 MW of solar added to the PJM grid, for reliability purposes, the PJM will only consider 0.09 MW added (assuming fixed tilt). However, adding battery energy storage systems to a photovoltaic (PV) array can increase the ELCC to between 59% and 78%, depending on the duration of the storage system (exhibit 31). So, while solar by itself adds little to resiliency, solar with battery energy storage system (BESS) is almost on par with conventional thermal generation.

Exhibit 29
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
PJM Resource ELCC

Resource Type	ELCC Average (%)
Thermal	81
Demand Response	76
Storage	68
VRE	35

Source: PJM

Exhibit 30
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
PJM VRE Resource ELCC

Resource Type	ELCC Average (%)
Fixed-Tilt Solar	9
Tracking Solar	14
Onshore Wind	35
Offshore Wind	60

Source: PJM

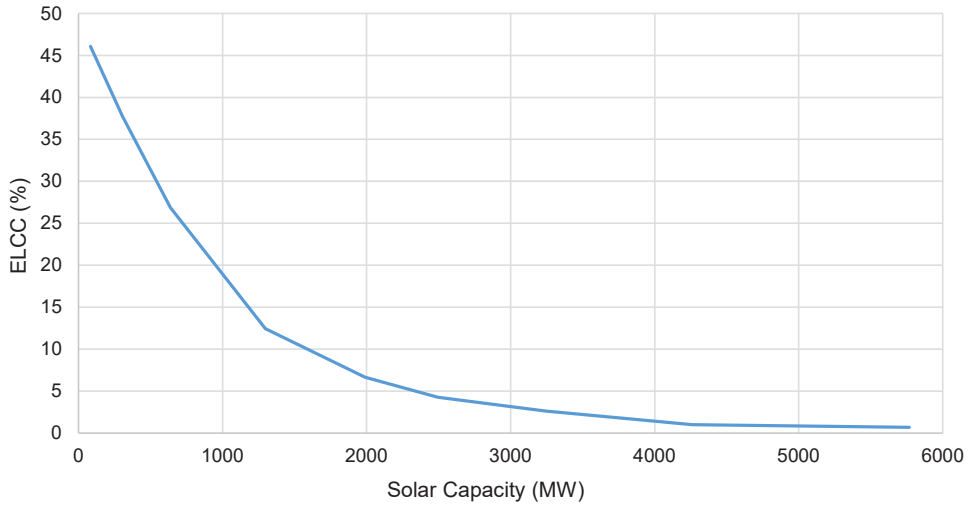
Exhibit 31
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
PJM Solar Plus Battery Storage ELCC

Resource Type	ELCC Average (%)
Four-Hour Storage	59
Six-Hour Storage	67
Eight-Hour Storage	68
Ten-hour Storage	78

Source: PJM

The ELCC also tends to drop dramatically as more VREs enter a market. For example, a typical graph of ELCC will start with an ELCC near or above 50%, but as solar capacity is added to the grid, the ELCC will quickly drop (exhibit 32). The reason for this is that when only 1% of the grid is VREs, then the rest of the grid can easily manage cloudy days. However, a cloudy day in California, for example, where 24% of the installed capacity is solar, can result in a large drop in power production. This forces the balancing authority (BA) in California (CAISO) to maintain enough reserve capacity to compensate for that drop. In short, the "effective" value of adding solar declines as more is added to the grid mix. See our case study, on page 49.

Exhibit 32
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Theoretical Cumulative Solar Capacity vs. ELCC

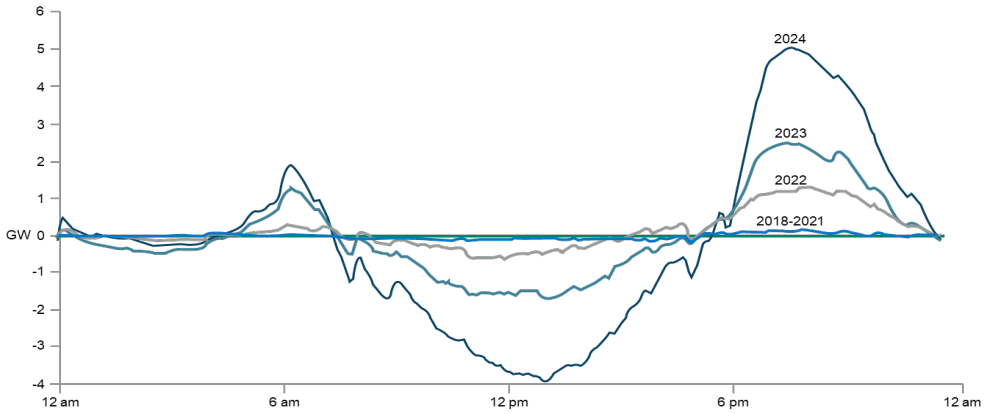


Source: Heptonstall and Gross 2021 and William Blair Equity Research

Battery Storage Helps Grid Resiliency

Batteries are also playing a larger role in balancing California's grid system. Battery dispatch to the grid in California has increased from almost negligible amounts in 2021 to 5 GW in 2024. Similar to natural gas, batteries can be used when needed to meet the evening peak in power demand, but unlike gas, batteries can also utilize excess solar production during the day to charge. The utility of batteries is evident in the fact that they are a larger driver of CAISO's new Net Metering 3.0 policy changes, which incentivize PV energy plus battery storage over standalone PV systems.

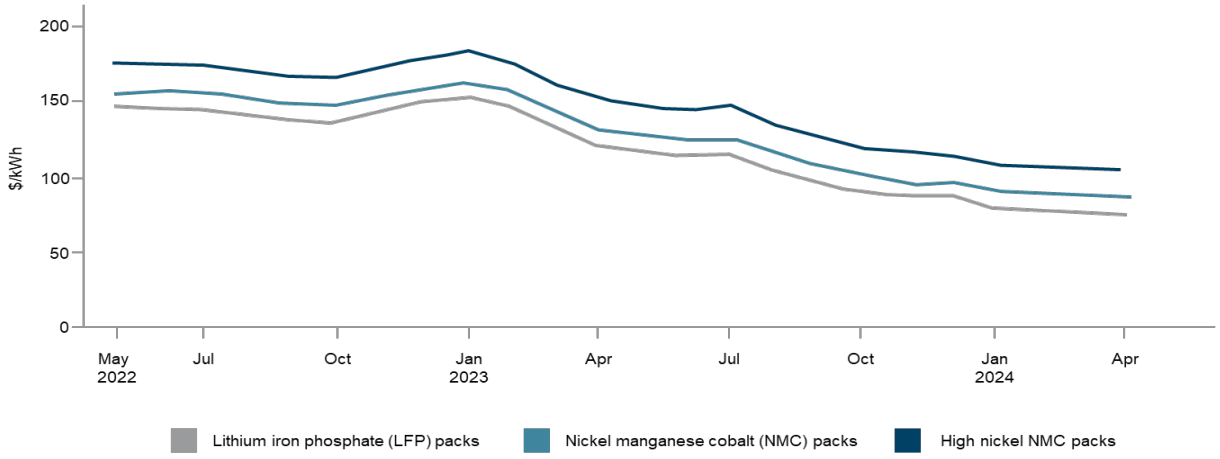
Exhibit 33
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Battery Storage Increasing Contribution to CAISO



Source: CAISO and William Blair Equity Research

We expect CAISO's NEM 3.0 policy to become the norm in states with high solar adoption, namely Texas and Florida, and outside the U.S., we see Germany adopting similar regulations. In addition, battery prices have come down significantly thanks to an overbuild in China, especially for lithium iron phosphate (LFP) used in stationary storage. Since the peak in 2021, prices have come down from $150/kWh to below $80/kWh, and domestic China prices for prismatic LFP for stationary storage are as low as $50/kWh, which we believe is selling at zero or negative gross margin.

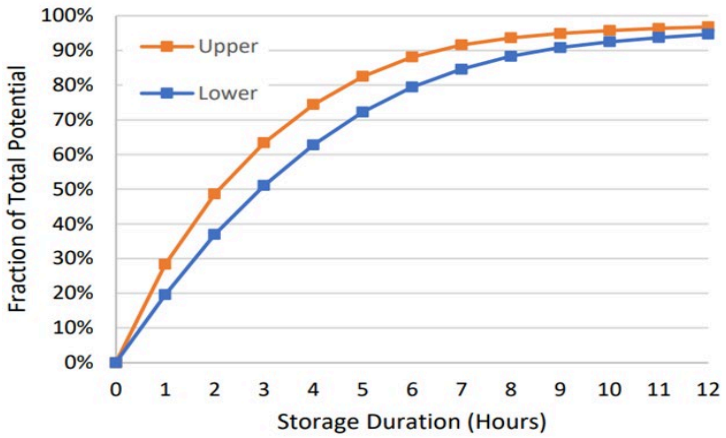
Exhibit 34
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Global Li-ion Battery Pricing



Source: William Blair Equity Research

Li-ion batteries are well suited for four-hour storage, making them ideal for daily time-shifting of solar generation from low energy demand in the afternoon to high demand in the evening. According to NREL, over 60% of the value of energy time-shifting is captured with four-hour storage. Four-hour energy storage can replace a considerable amount of natural gas peaker plants that are used for immediate and short-term energy injections. However, data from ERCOT suggest much of the storage being deployed on the grid is not being used as four-hour storage but rather to help stabilize the grid.

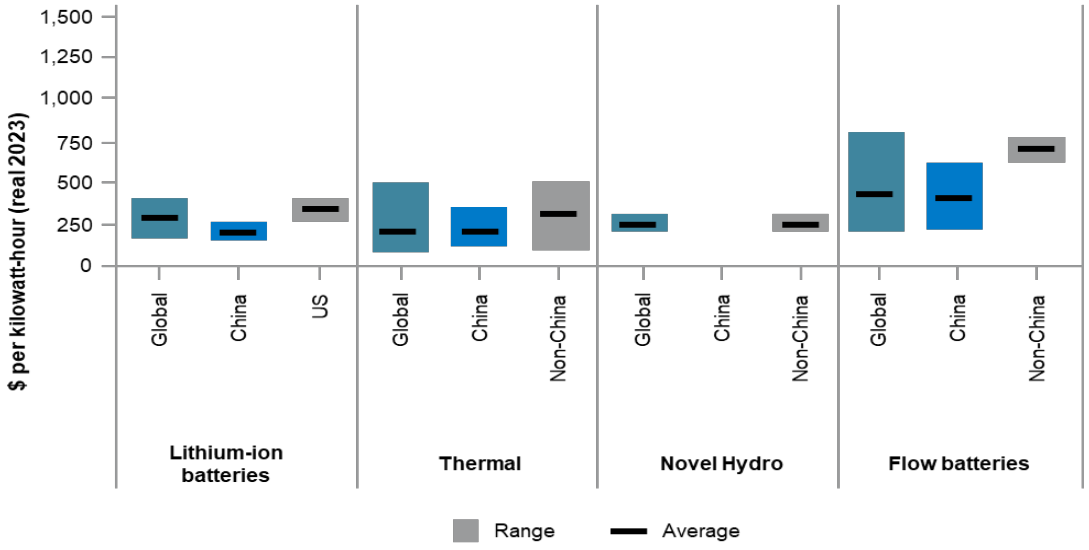
Exhibit 35
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Value of Energy Time-Shifting With Storage Duration



Source: National Renewable Energy Laboratory

There is also a significant market for long duration energy storage (LDES), which is characterized as eight hours or longer. Li-ion traditionally does not pencil out financially here because of the linear relationship between cells and total energy—there are no scaling benefits as you add more batteries. Alternative systems and chemistries are being pursued that offer lower-cost LDES, like flow or iron-air batteries, compressed air, or gravity systems, although none have resulted in commercial adoption to date.

Exhibit 36
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Cost Comparison of Long Duration Energy Storage Technologies



Source: William Blair Equity Research

Lower battery prices have made Li-ion more financially feasible for LDES. Recently, Li-ion was chosen for the first eight-hour energy storage system in Australia. The German company RWE contracted for the installation chose Tesla's Megapack for the battery system, adding 50 MW of

storage to an existing 250 MW solar farm. This is one of the first examples we have seen of Li-ion, specifically LFP, entering the LDES application, and we expect adoption to increase as pricing comes down and alternatives continue to underdeliver.

Case Study: Georgia Power Sounding the Alarm

Georgia Power's load forecast has increased by 50% in just a quarter since its last update. The company files a quarterly update of large load projects through the mid-2030s to the Georgia Public Service Commission due to the rapid increase in demand from data centers. Since second quarter 2024, project developments have increased by 50%, from 24.4 GW to 36.5 GW. In the document, Georgia Power highlighted 3.3 GW of data center projects have broken ground already, and another 4.1 GW are pending construction, totaling 7.4 GW of load expected to come online before 2029.

The majority of the Georgia Power's load growth will be served by natural gas, and GE Vernova is the primary beneficiary. The magnitude of the load increase and the 10-year time frame severely constrict the available options for new power generation. Georgia's newest nuclear facility has been an area of controversy due to the delays and cost overruns, but provide an additional 2GW of stable electricity, which adds to the appeal of data centers in Georgia. This dynamic appears to be spreading where solar and wind, once lauded as the climate solution, fail to provide stable dispatchable power. As such, we are witnessing a migration from high renewable and high energy cost states such as California to states such as Georgia due to contribution of energy sources.

Exhibit 37
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Georgia Power Integrated Resource Plan

	Project Size (MW)	Project Category
PPA between Georgia Power and Mississippi Power	750	Gas and/or Coal
PPA between Georgia Power and Santa Rosa Energy Center, LLC	230	Gas
Authority to Develop, own and operate BESS at various sites	1,000	Battery Energy Storage Systems
Authority to Develop, Own, and Operate three simple cycle CTs at Plant Yates	1,400	Gas
Two new customer-sided DER programs	unclear	DER
Total	**3,180**	

Source: Georgia Power

Data Center Construction

In addition to power generation limitations, physical bottlenecks in the generative AI data center buildout include limited supply of various hardware components. Some of the most severe shortages are associated with the equipment that connects the grid to the data center—specifically transformers and switchgear. Transformers step down high-voltage power from the grid to the level required by the facility. Switchgear is the collection of electrical components that regulate, protect, and manage a power system—including circuit breakers, disconnects, and equipment that enables efficiency and reliability.

Transformers have experienced a shortage in the United States since the supply chain issues associated with the COVID pandemic, and recent acceleration in data center construction and the electrification trend in general have exacerbated these shortages. Several industry sources indicate that transformer lead times have increased from less than several months before the pandemic to two years in 2024. For high-voltage transformers, the wait is now as long as four years. Lead times

for some switchgear components have increased to 18 months to two years. We believe increased domestic production of these components, possibly with the support of government subsidy, is critical in sustaining current data center build rates.

Although in high demand, other hardware such as generators, cooling systems, power distribution, and IT infrastructure components do not appear to be causing a bottleneck. For example, Cummins, a major supplier of generators to the data center industry, is sold out of the model typically installed at large data centers—the 95-liter, 3,000 kW version—through 2025.

Many vendors are making the necessary investments to support industry growth. Caterpillar is investing $725 million to expand an Indiana plant that makes large-scale generators. On an October conference call, Eaton's CEO announced an increase in the company's planned investment in incremental capacity (from $1 billion to $1.5 billion), reflecting "increasing demand in data center markets." Suppliers of data center cooling equipment, such as AAON, Modine, and nVent, have made significant investments in capacity expansion.

Other factors that present risk to the data center buildout include labor shortages, permitting issues, and tariffs. McKinsey & Company forecasts a potential shortage of up to 400,000 trade workers needed to support projected data center growth. Labor needs span from the construction and utility industries to factory operators producing transformers and other components. Delays in issuance, or resistance to issuance, of permits required for construction of data center or associated power generation plants could also hinder industry growth. Increased tariffs—and the escalation in trade wars they would likely spark—could make critical data center components more expensive and harder to acquire.

Tech Bottlenecks

Beyond the energy and physical data center limitations to building out AI data centers, there are also a number of bottlenecks on the server side. GPUs and AI accelerators remain a scarce commodity, with demand still well ahead of supply, according to commentary across the supply chain, from the fabs (TSMC) to the chip designers (Nvidia, Broadcom, AMD) to the providers of memory and interconnect solutions (SK-Hynix, Micron, Marvell, Astera) to the hyperscaler and CSP end-customers (Meta, Microsoft, Amazon, Google, Oracle). Particularly for the latest generation, a key supply bottleneck has been TSMC's ability to expand capacity for its most advanced fab technologies (including CoWoS, which is used to develop both Nvidia and AMD's most advanced chips). TSMC has committed to doubling this capacity in 2025 to keep up with demand.

Beyond getting access to more GPUs, which are the brains of these data centers, there have been other factors with data centers that limit the utilization of these GPUs—GPUs typically process data faster than the rest of the system, meaning access to memory, incredibly fast connectivity, and broader back-end networks are key to getting the most usage out of these expensive chips. In particular, memory has proved in high demand to enable fast access to trillions of model parameters. HBM3e, which is the latest generation of high bandwidth memory, started to ship in 2024, while HBM4 is expected in 2025 (built into Nvidia's B200 chip systems).

Elsewhere, the network has also proved an important bottleneck. With research from Meta indicating that GPU downtime is largely driven by network traffic moving more slowly than the processors. This has driven incredible demand and investment in accelerating the development of switching and DSP solutions that can handle the much higher throughput and higher speeds required in AI clusters. While through most of 2024, InfiniBand-based solutions sold by Nvidia remained the go-to solution, in the second half of 2024 we saw an inflection in demand for Ethernet-based solutions. Going into 2025, interconnect, switching, and optical solutions are expected to see

rapid growth as hyperscalers ensure that their billion-dollar clusters with hundreds of thousands of GPUs are getting optimal utilization. Key beneficiaries here include Arista, Cisco, Juniper, Nvidia, Broadcom, Marvell, Astera Labs, and Credo Semiconductor.

Bottom line: While we see some lingering bottlenecks on the tech front, the bigger bottlenecks to AI adoption are occurring in non-tech areas like energy generation, storage, and distribution, which are critical for the brick-and-mortar data centers needed to support the massive GPU clusters that LLM providers are deploying (or planning to deploy).

Where Do LLMs Fit Into the Application Landscape?

LLMs are an important part of the application software landscape and a critical infrastructure component of GenAI apps. Most software vendors rely on third-party foundation models to run the GenAI capabilities they offer as opposed to internally developing their own foundation models, which can be costly and time consuming. LLMs are an infrastructure software component, much like a database is an infrastructure software component in a traditional enterprise application.

What differentiates software vendors at the application layer is workflow, user experience, functionality, integration, and ease of use. While we believe that LLM vendors will inevitably look to move up the stack to take advantage of large TAMs and pricing power, history suggests that the rate of success is likely to be low. Overall, infrastructure providers are well equipped at building tools for developers and IT teams, but they struggle to cater to knowledge workers and often lack domain expertise in the end-markets they are selling into.

Examples of infrastructure companies that have struggled to succeed at the application layer include Twilio, AWS (which arguably still lacks an enterprise application), and Cisco. But this does not mean it is impossible, and things could be different this time around if the LLM providers prioritize the application layer. Thus far, we do not believe this has been happening. The major LLM providers like OpenAI and Anthropic have been focused on developing ever more powerful foundation models (an infrastructure investment), not necessarily focusing on use-cases, workflows, and the broader application layer.

Another interesting concept to consider here is the idea of multi-model applications. We believe that most AI applications are not going to be powered by a single foundation model. Instead, AI use-cases and applications are likely to be powered by multiple models, each providing its own unique strength. The role of the application layer will be to coordinate which model is used depending on the use-case.

This multi-model powered application is a similar concept to multicloud, which exists today. It allows application vendors to benefit from the strengths of each model provider while ensuring redundancy and keeping model providers honest with pricing. We are already seeing this framework deployed in the real world. For example, Atlassian's AI capabilities are powered using over 30 different models from 6 different model providers. We believe most of the models being used are SLMs that excel at specific use-cases as opposed to LLMs, which can be costly to operate (for more detail on SLMs versus LLMs, see our discussion on page 24).

Bottom line: LLMs themselves are not applications that end-users can deploy, especially in the enterprise. As such, they should be viewed as the foundational element of a GenAI application but not a substitute for the application itself.

Is AI a Threat to the Software Industry and/or Software Business Models?

While the adoption and monetization timeline for GenAI applications has disappointed some investors, we believe software businesses are well positioned to benefit over time as more GenAI applications move into production over the next couple years. For most enterprises looking to take advantage of GenAI, we believe they will choose to buy off-the-shelf packaged software applications as opposed to attempting to build their own GenAI applications with an LLM. This will be a nice tailwind to the software industry, and as we pointed out in exhibit 17 there are already some positive indicators of GenAI adoption from customers of incumbent software vendors.

We believe the argument that GenAI will lead to more enterprises building their own AI applications is misguided. Certainly, some companies will decide to go down this path (like Klarna), but we find it hard to believe that banks, airlines, retailers, and healthcare companies will suddenly decide to dedicate time to building and maintaining their own customer service software or marketing tool or HR system. Focusing on this when there are viable off-the-shelf alternatives that provide quick time to deployment, fast ROI, third-party vendor domain expertise, and network effects (more customers equals better software) would be a distraction for the average enterprise customer at a minimum. It would make them less focused on their core competencies and gaining a competitive edge over their peers. They would also have to bear the ongoing cost of maintenance and dedicate resources to innovating as the tech landscape evolves. Buying off-the-shelf software solves for these challenges (which is why it has worked for decades).

Some argue that GenAI has lowered the barriers to building software and fundamentally changed the build-versus-buy equation. We do not view this as a new argument. This same argument could have been made 10-15 years ago with the rise of cloud, microservices, containers, low-code, and open source—all of which also reduced the cost and complexity of building software applications. However, this did not hinder the growth trajectory of packaged enterprise software. Instead, it fueled growth, as the pace of innovation from specialized packaged software vendors significantly increased such that it was difficult for in-house teams to match. Software vendors took advantage of their domain knowledge to bring new capabilities to market faster, make it easier to integrate with other applications, and ensure it was all done with security and compliance in mind. We are already seeing something similar playing out with AI and believe the integration of GenAI into existing software models will prove to be a tailwind in the years to come.

Bottom line: Software business models themselves may need to evolve (see monetization discussion). The traditional seat-based subscription model is likely not the future; in the coming years, we will likely see software models evolve to become more consumption-based or value-based. While this transition may cause some near-term volatility in financial performance, we in no way consider this the death of software. Rather, it is simply a pricing model evolution that most software vendors understand they need to embrace (and ultimately aligns better with the value they are delivering).

How Necessary Are Nvidia GPUs Once the Heavy Lifting of Training Models Is Complete?

Machine learning models go through two primary phases as they evolve from initial development to real-world application—training and inference. Understanding these distinct stages is crucial for appreciating how a model learns and then deploys that learned knowledge.

Training is the initial phase where the model ingests large amounts of labeled data, iteratively adjusting its internal parameters to reduce prediction errors. This often involves computationally expensive optimization processes and requires high-performance hardware and extensive time. By the end of training, the model ideally "understands" patterns in the data, enabling it to generalize to examples it has never encountered before.

Inference occurs once the model is fully trained and ready to be used in practical scenarios. Instead of updating parameters, inference simply involves presenting new, unseen inputs to the model and receiving predictions in return. Inference is typically far more resource-efficient than training and is intended for real-time decision-making, where quick and accurate responses are key.

The GenAI landscape has so-far been dominated by training demand as hyperscalers and AI model providers race to build ever-more powerful LLMs. However, increasingly we are seeing a shift toward inference as more organizations start using these AI models to power new features and applications. Nvidia, the leader in high-performance, large-scale data center systems and GPUs (with roughly three-quarters of the AI accelerator market today), has noted this shift in its customer base, estimating that roughly 40% of use-cases for its GPUs are inference related—a proportion that should only increase going into 2025.

As the AI landscape shifts from training to inference, investors have become concerned that this will be negative for Nvidia. Because the compute requirements for inference are much lower than in training—or so the thinking goes—there is bound to be more competition and margin pressure in the inference market for Nvidia. Specialized ASICs like Google's TPUs and Meta's MTIA chips are being optimized for inference performance, while start-up providers like Cerebras and Groq are challenging Nvidia's performance dominance for certain use-cases.

Generally, we remain confident in Nvidia's ability to perform well even as the industry shifts toward inference use-cases and spending. As noted earlier in this report, hyperscalers are employing multiple axes of model performance beyond pretraining that will continue to require more compute capacity. Furthermore, with next-generation reasoning models employing concepts like test-time compute (which require massive amounts of computing to reason through problems), the pendulum could shift back into Nvidia's favor.

For example, while in many respects the input and output of test-time compute models look similar to the one-shot responses of the classical GPT model (e.g., 1,000 words in, 1,000 words out), the in-between step of reasoning through the problem, exploring alternatives, and selecting the best answer creates tens of thousands of tokens that require more intensive computing resources (e.g., o1 is estimated to drive a 5-10 times increase in tokens generated compared to GPT-4). This invisible step, which today remains slow, is what is pushing demand for at-scale computing infrastructure at the inference stage.

Lastly, large-scale inference systems mixing multiple SLMs into an agentic chain should also drive upward pressure on computing demand as more models are chained together. More models chained together will typically drive up the size of the context window, requiring more compute (and memory) to process the full chain of actions.

Bottom line: New techniques like test-time compute are driving scaling demand from training to inference. Even as the market shifts increasingly to inference, Nvidia's ability to combine the best processors with networking, system engineering, and software know-how should keep the company at the forefront of the accelerated computing market.

How Will AI Impact the IT Services Industry?

The rise of AI, including GenAI, presents both opportunities and challenges for IT services companies. While AI has the potential to automate certain IT offerings, such as code generation, testing, and maintenance, which could reduce demand for traditional outsourcing work, it simultaneously creates new offerings and avenues for value creation. On a net basis, we expect that the demand to embed AI into applications, products, and features is inflationary and a continuation of the strong digital transformation tailwinds that drove growth in the industry over the last several decades. Further, where there is disruption to business models, we expect companies to pivot such that generative AI tools can streamline delivery processes, improve efficiency, and lower costs, allowing IT services firms to take on more complex, higher-value projects.

AI-driven transformation requires significant expertise in areas like infrastructure modernization, AI model integration, data governance, and enterprise-scale deployment—capabilities that IT services providers are well positioned to deliver. Rather than being a threat, AI represents an opportunity for these companies to evolve their offerings, focusing on advisory, implementation, and management of AI solutions. Earlier this June, McKinsey published a report titled, "Tech Services and Generative AI: Plotting the Necessary Reinvention." Within the report, McKinsey valued the emerging market for services related to generative AI/AI to be worth more than $200 billion by 2029. The report forecasts that service providers that successfully capture some of that incremental value could grow profitability by as much as 30% and boost revenue between 2% and 4% above the historical growth trend.

Below, we explore how GenAI is driving demand for solutions across the infrastructure and application layers, which could present IT services companies with significant opportunities to add value. By leveraging deep expertise, technology partnerships, and full-stack capabilities, these companies can help enterprises deploy and optimize custom GenAI solutions.

At the infrastructure level, IT services providers can play a critical role in designing and implementing architectures that support GenAI workloads. As real-time retrieval-augmented generation (RAG) systems and device-level inferencing grow in importance, infrastructure components like storage, memory, and high-throughput compute systems will become critical. IT services firms can guide clients in assessing infrastructure readiness and identifying gaps in compute power, memory, and storage performance. By partnering with leading hardware and cloud providers—such as NVIDIA, Intel, AWS, Microsoft Azure, and Google Cloud—IT services companies can deliver integrated solutions that optimize performance and scalability.

To address storage and latency challenges, IT services companies can implement distributed storage systems and in-memory computing architectures that reduce bottlenecks for AI inference workloads. In addition, they can integrate key infrastructure software picks and shovels, including databases, data lakes, data streaming systems, and MLOps tools. For instance, scalable solutions using platforms like Snowflake, Databricks, and Confluent can support real-time data streaming and high-performance AI model workflows. Similarly, IT services firms can deploy container management tools such as Kubernetes to ensure flexible and scalable infrastructure for AI deployments. Beyond optimization, these providers are well positioned to address enterprise needs such as data labeling, governance, and security by implementing AI governance frameworks, ensuring data lineage, and securing data infrastructure through access control tools.

At the application layer, IT services companies can help enterprises unlock the full value of GenAI through industry-specific, customizable solutions. By leveraging vertical expertise, these firms can develop GenAI applications tailored to specific business use-cases. In our *We're on IT* periodical, we explore examples of domain-specific AI use-cases enabled and developed by IT services companies. In the healthcare space, Globant offers several AI-specific services such as AI for care delivery and self-service AI/ML tools. In the financial services space, Infosys offers Infosys Topaz for Financial Services, an AI-first set of services, solutions, and platforms that help financial institutions maximize the value they receive from AI. Some companies in our coverage have developed use-case-specific LLMs. For example, Cognizant recently launched a set of healthcare LLMs solutions on Google Cloud's generative AI technology. These LLMs will help redesign healthcare administrative processes and improve overall experiences. We believe it is essential that providers use their industry knowledge to bring use-case-specific LLMs to the market.

Custom application development also allows IT services firms to integrate modern GenAI frameworks (such as OpenAI or proprietary models) into enterprise systems. In addition, IT services companies can enhance enterprise productivity through integrations with existing tools and platforms, such as Microsoft Copilot for workplace applications. By combining GenAI capabilities with low-code/no-code platforms, they can empower businesses to build AI-powered applications more quickly and efficiently.

Roadblocks to AI adoption, such as governance, ethics, security, and access to talent, increase the complexity of this emerging technology, setting up IT services companies as natural consulting partners during these early innings of the technology trend. Both hiring and maintaining AI talent is a task well suited to tech services providers because they possess the resources and learning opportunities that top-tier AI talent crave. In our aforementioned periodical, we discuss results from polls of our covered companies on the type of talent they are hiring for and training, as well as the results from an expert call regarding the impact of AI on the daily activities of an engineer.

Bottom line: IT services companies are uniquely positioned to drive value across the GenAI stack. By offering full lifecycle management, IT services providers can guide clients from ideation and proof of concept to deployment, scaling, and ongoing performance optimization. Their combination of technology partnerships, end-to-end delivery capabilities, and industry expertise allows them to deliver comprehensive solutions that span infrastructure and applications. By positioning themselves as partners in AI adoption, IT services firms can not only future proof their businesses but also deepen their role as trusted enablers of enterprise innovation.

Will Government Regulation Hold up the AI Market?

While AI is expected to deliver many societal benefits, such as better healthcare, safer and cleaner transport, more efficient manufacturing, and cheaper and more sustainable energy, policymakers have expressed concerns on potential risks, including misinformation/disinformation, deepfakes (images, videos, and audio), cybersecurity attacks, hallucinations and biased responses, and misuse of private data or intellectual property. To guard against these risks and support responsible AI innovation, governments around the world have begun introducing AI regulation, though different countries/states have taken different approaches, which could complicate efforts by AI providers to introduce their technology across borders.

Perhaps the most notable law thus far is the European Union's AI Act introduced in 2023, which has been touted as the world's first comprehensive AI law. The law features a broad definition of AI and establishes obligations for providers and users depending on the level of risk from AI. The law is backed by substantial monetary penalties for non-compliance—up to €35 million, or 7% of a company's annual revenue. The different risk levels are:

- *Unacceptable risk* – AI systems that are considered a threat to people will be banned. These include biometric identification (exceptions for law enforcement), social scoring (classifying people based on behavior, socio-economic status, or personal characteristics), and cognitive behavioral manipulation of people or specific vulnerable groups (e.g., voice-activated toys that encourage dangerous behavior).

- *High risk* – AI systems that negatively affect safety or fundamental rights. These include AI systems used in products falling under the EU's product safety laws (includes toys, aviation, cars, medical devices, and lifts) and AI systems falling into specific areas, such as: 1) management and operation of critical infrastructure; 2) education and vocational training; 3) employment, worker management, and access to self-employment; 4) access to and enjoyment of essential private services and public services and benefits; 5) law enforcement; 6) migration, asylum, and border control management; and 7) assistance in legal interpretation and application of the law.

Generative AI systems, like ChatGPT, will not be classified as high risk, but will have to comply with transparency requirements and EU copyright law, specifically in: a) disclosing that the content was generated by AI, b) designing the model to prevent it from generating illegal content, and c) publishing summaries of copyrighted data used for training. Content that is either generated or modified with the help of AI—image, audio, or video files—needs to be clearly labelled as AI generated so that users are aware when they come across such content.

In the U.S., states such as Colorado and California have introduced AI-focused legislation, but differences between these laws have prompted calls for a unified federal framework to address potential threats from the technology. Similar to the EU's AI Act, California's AB 2013 law requires developers of generative AI systems (starting on January 1, 2026) to publicly post on their websites certain information about the data used to train those systems, though the law does not include a specific enforcement mechanism. Meanwhile, the California AI Transparency Act (SB 942) aims to help individuals know when content was created or altered by AI. This applies only to "covered providers," defined as "a person that creates, codes, or otherwise produces a generative artificial intelligence system that has over 1,000,000 monthly visitors or users and is publicly accessible within [California]." In contrast, other U.S. states continue to rely on existing privacy, intellectual property, consumer protection, and other laws to regulate AI.

Bottom line: With a rapidly evolving AI regulatory landscape and a patchwork of laws across different countries and states, AI providers are rightly concerned that innovation could be stifled. At the same time, self-regulation may also be impractical given the breakneck pace of technological development and the myriad potential harms to consumers.

When Is AGI Coming and Are We All Doomed?

OpenAI defines AGI (artificial general intelligence) as: "highly autonomous systems that outperform humans at most economically valuable work." Prior to the development of LLMs, AI was mainly comprised of virtual assistants, recommendation engines, facial recognition technology, and computer vision systems. These capabilities are examples of artificial narrow intelligence (ANI), or AI systems designed to perform a narrow range of tasks.

The consensus view at the time among researchers was that it would take 50 years to reach AGI. Based on recent advancements, researchers have understandably revised this view, with industry leaders like OpenAI CEO Sam Altman and Tesla CEO Elon Musk arguing that AGI could even come as soon as this year (see below). In terms of intelligence level, ANI could be compared to that of an infant due to its limited scope, while AGI's ability to perform reasoning, problem solving, and abstract thinking with a much wider scope is akin to an adult human.

Just like AGI lacks a consensus definition, no consensus benchmark exists that indicates whether computers have reached AGI. However, two benchmarks in particular stick out: the Turing Test and the ARC-AGI benchmark. The Turing Test was coined by Alan Turing in his 1950 paper on computer machinery and intelligence. It describes the ability for a machine to replicate human intelligence in its responses over the course of conversation. Traditionally, this was seen as an indicator of AGI, but with the development of LLMs (where AI products can now commonly replicate human conversation), the goalposts have clearly shifted.

The ARC-AGI benchmark, introduced in 2019, was designed to evaluate whether an AI system can acquire new skills beyond the data on which it was trained. The benchmark's founder François Chollet and Zapier co-founder Mike Knoop hosted a competition in 2024 (offering a $1 million prize to the winner) to achieve the 85% threshold that they believe would be indicative of human-level intelligence, and thus AGI (in contrast, humans can solve an average of 80% of all ARC tasks). The best model scored a 53.5%, compared to the 33% scored in 2023, which suggests major advancements in the space.

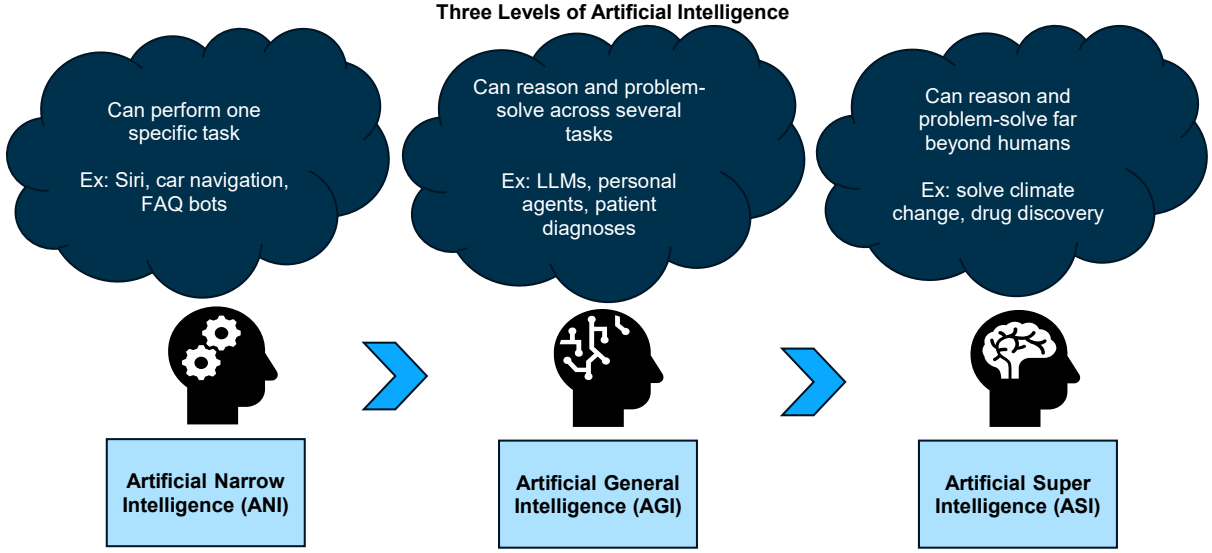
In a recent interview with Bloomberg, Altman noted that OpenAI's o3 reasoning model, which was announced in December 2024 and is currently being safety tested, has passed the ARC-AGI challenge. The model scored a breakthrough 75.7% on the Semi-Private Evaluation set at its stated public leaderboard $10,000 compute limit, while a high-compute (172x) o3 configuration scored 87.5%.

Now, Altman said, the company is setting its sights on artificial super intelligence (ASI), which is leaps and bounds beyond AGI, just as AGI is orders of magnitude more intelligent than ANI. ASI far surpasses any human knowledge, and machines equipped with ASI could potentially accelerate scientific discovery well beyond what we are capable of doing on our own.

Given the powerful potential impact on humanity from this technology, U.S. lawmakers increasingly find themselves at the center of the AGI debate, specifically focused on open-source models and fears that China may be outpacing the U.S. CEOs like Clem Delangue (of Hugging Face) have warned that China's open-source models are advancing past the West in terms of coding and reasoning benchmarks. In response to these fears, the U.S. has adopted strict policies to curb China's AI advancement by implementing export bans on hardware infrastructure for AI development. Recently, a congressional commission proposed a Manhattan Project-style effort to fund the development of AGI, or "superhuman intelligence," as they refer to it.

Bottom line: AGI looks close at hand and has the potential to revolutionize many fields, including scientific research, health care, and defense technology. While the nuclear weapons analogy may be a bit extreme, it is understandable that government leaders want to ensure that this technology is used responsibly and does not get into the wrong hands.

Exhibit 38
Navigating the Boom: Confronting Generative AI's Most Pressing Questions
Three Levels of Artificial Intelligence



Source: William Blair Equity Research

Our Best Ideas to Play the AI Theme

AAON, Inc. (Market Cap $10.7 Billion)
AAON is a best-in-class commercial HVAC OEM specializing in semi-custom and full custom systems. In its core business—the $7.5 billion commercial rooftop market—AAON is growing share due to superior equipment and a lower historical price premium. AAON's BASX segment is a market leader in the attractive data center cooling and semiconductor cleanroom markets. It has recently had wins with multiple hyperscalers and has proved effective with both its liquid and air-side solutions. We believe AAON is entering the best growth cycle in company history due to premium cooling for AI data centers, the A2L refrigerant change, and the 2025 launch of the industry's first cold climate heat pump. AAON is poised for compound EPS growth of 20%-25% from 2024-2028.

Shares currently trade at 19 times our 2027 EBITDA, in-line with Vertiv and Trane at 19 times. We believe AAON deserves a premium to both due to outsized growth, better margins, innovation, and share gains. Should our bull case play out for EBITDA reaching $750 million, we believe a 25-times multiple would be justified, driving 70% upside to $225. Risks to our rating include well-capitalized competitors, higher interest rates impacting commercial spending, all-time-high gross margins, new capacity ramp, elevated backlogs, commodities, and weather. Our rating is Outperform.

Alphabet Inc. (Market Cap $2,433.7 Billion)
Google's position in the AI boom is gaining momentum, in our view. Its foundational model, Gemini, is being applied across its suite of products, such as Performance Max, Gmail, NotebookLM, and Veo. GenAI advertising has been implemented into search and with AI overviews. Based on industry participant feedback, advertisers have a growing interest in AI overviews since currently one

advertising slot is being commercialized. Moreover, we believe the strength of its Gemini model is growing, and we believe there will continue to be incremental use-cases going forward. The company's Waymo division is particularly beneficial as automated driving becomes a reality; its expansion into Austin and Atlanta will begin this year, followed by Miami in 2026. Going forward, it will continue to deploy capital investments into data centers for AI advancements, which we believe will allow the company to remain a leader in the space.

Google currently trades at 14.4 times our 2026 EBITDA estimate. Using our DCF framework, we see roughly midteens upside through the next 12 months using a discount rate of 10% and an EBITDA multiple of 13.5 times. Due to Google's continued product innovations and successful leveraging of AI technologies to improve its offering across all key operating segments, we see increasing upside for the company. Risks to our thesis include privacy and regulation concerns, a worsening macro environment that impacts advertising spend, AI's potential impact on search, and growing competition.

Amazon.com, Inc. (Market Cap $2,425.9 Billion)
Amazon is poised to significantly benefit from GenAI, particularly through AWS, but also through other segments like e-commerce and advertising. AWS's extensive infrastructure and leading position in the CSP (cloud service provider) market make it an ideal choice for those developing large language models (LLMs) and for enterprises deploying GenAI applications. We believe the increasing demand for compute resources driven by GenAI workloads will enhance AWS's growth, especially considering the resource-intensive nature of LLMs and GenAI applications. As a result, AWS should benefit in multiple layers of the AI stack, including: 1) the infrastructure layer, which includes services for those building foundational models (FMs) like Anthropic, Perplexity, and others (these are compute instances powered by Nvidia, along with AWS's own custom silicon); 2) for enterprises looking to deploy GenAI, AWS has Bedrock, which makes it easy for enterprises to deploy and customize leading third-party FMs with internal enterprise data using techniques like fine tuning and RAG; and 3) at the application layer, Amazon Q helps developers generate, test, debug, and transform code. Q Apps can also build features and application based on natural language descriptions alone. AWS already generates billions of dollars in revenue from AI, and we see more growth ahead as AWS benefits from both training and inference workloads. We also expect GenAI to enhance the user experience on Amazon's e-commerce site, which should make it easier for consumers to transact, and drive more target ads.

Amazon trades at just 13 times 2026 EBITDA, a discount to other mega cap tech peers which are trading at 17 times EBTIDA. We believe this is an attractive multiple, especially with the growth acceleration potential within AWS and margin expansion company wide as revenue mix shifts to higher margin businesses like AWS, advertising, and Prime. Risks include the nonlinear pace of margin expansion, consumer health, and potential to redirect profit to invest in long-term opportunities like grocery, healthcare, content, etc.

AppFolio, Inc. (Market Cap $9.5 Billion)
AppFolio provides SaaS-based property management solutions to rental property owners and operators, and through serving as the core operating platform for its customers, it has a significant amount of proprietary data/integrations from which to leverage AI. Since it acquired Dynasty in 2019 (a provider of industry-specific AI solutions for real estate), AppFolio has been heavily focused on the development of its AI capabilities. Today, AppFolio's Realm-X platform is a comprehensive, verticalized set of AI-native solutions including an AI copilot, AI leasing assistant (Lisa), an automation engine for standardized processes (rental apps, collections, lease renewals), smart maintenance, automated marketing, bank reconciliation, and more. While some of these solutions are included in AppFolio's baseline subscriptions, many are included only in higher-tier subscriptions, and some (such as Lisa) are monetized separately. In a business where property managers spend 40%-50% of their time doing administrative work (entering invoices, submitting work-orders, handling accounting), Realm-X

users have already reported saving an average of 12.5 hours per week using the platform, which has supported strong demand for these solutions (over 90% of AppFolio customers are now using at least one AI product). Ultimately, we believe that AppFolio is well ahead of its competition in leveraging AI, which has already begun to serve as a competitive differentiator and should be a lever supporting continued ARPU expansion going forward.

On our 2025 estimates, AppFolio trades at 10 times revenue, 34 times adj. EBITDA, and 44 times FCF, which is roughly in line with our vertical SaaS peer group currently trading at roughly 9 times 2025 revenue, and 34 times adjusted EBITDA. Risks to consider include the potential for top-line growth to decelerate, choppiness in the company's margin/profit trajectory, variability in financial performance due to increasing payments monetization, competition, and recent political pressure against rental "junk fees."

Axon Enterprise, Inc. (Market Cap $46.2 Billion)
Axon is one of our top AI picks based on its GenAI solution Draft One and the company's significant opportunity with its broader AI Era bundling. Draft One offers automated report writing and has force multiplier capabilities of significantly reducing the time it takes to write an incident report, which is typically 30% to 40% of a police officer's day. Draft One can transcribe audio from Axon's body cameras to prepopulate much of the report with the required information, allowing officers to spend more time in the field, which they have responded very well to. This has become Axon's fastest-selling product since being introduced early last year and already has over $100 million in pipeline build. Axon's AI Era bundle is a high-value subscription that sits on top of a police agency's normal subscription and provides access to all of the company' s current AI products and future ones that release during the customer's subscription period, including Draft One, Form One, Brief One, and real-time translation. We believe Draft One and the AI Era bundling have potential to be significant contributors to the business that could also pull in broader demand for Axon's platform of software solutions, especially given the favorable law enforcement spending environment. In our view, Axon is creating new market opportunities with AI, which we view as another way for the company to sustain growth levels over the intermediate term.

Axon's stock last closed trading at about 21 times our revenue forecast for 2025 of $2.50 billion, and we believe the company continues to innovate and execute well with strong demand across-the-board which should lead to continued positive earnings revisions. Primary risks for Axon's stock include: 1) competition, 2) pricing pressure in the body and in-car camera markets, 3) difficulty entering the RMS/CAD market, 4) ability to sustain high level of innovation, and 5) costs associated with defending intellectual property rights.

Broadcom, Inc. (Market Cap $1,126.4 Billion)
We view Broadcom as well positioned to benefit from the continued buildout of AI data centers, with strong alignment to AI tailwinds across its fast-growing ASIC business and a leadership positioning in connectivity chips (including Ethernet switching, routing, and DSPs). A growing set of large customers are demanding in-house AI chips to meet the performance requirements of their webscale applications—Broadcom now counts five customers: Apple, OpenAI, and Bytedance join long-standing customers Google and Meta. Beyond its leadership position in the custom chip market, Broadcom's AI networking business is also benefiting from the buildout of large-scale AI clusters that demand exponentially more connectivity to train and perform inference on LLMs. In addition, to revenue momentum driven by its AI business, we also see upside potential to Broadcom's broader non-AI business, which should recover in 2025 and revenue and margin upside resulting from its acquisition of VMware.

Shares of Broadcom trade at a P/E of 37 times our calendar 2025 estimates, a premium to the peer group median multiple of 33 times. We believe Broadcom shares deserve a premium given the multiple vectors of growth over the next few years and the company's best-in-class financial

profile (mid-60% EBITDA margin). We see potential for upside to current estimates, particularly as three new AI customers ramp up spending and existing customers Google and Meta double-down on their custom chip investments. In addition, AI pull-in of connectivity chips and post-acquisition revenue upside in the VMware software business (as customers renew their 2022/2023 contracts) keep us positive on Broadcom's share potential. Risks to our Broadcom thesis include its exposure to China and ongoing geopolitical tensions, heavy competition in AI accelerators and optical interconnects, the impact of typical semiconductor cyclicality, integration risks with VMware and potentially higher-than-expected churn, and a key-man risk with CEO Hock Tan.

CCC Intelligent Solutions Holdings, Inc. (Market Cap $7.4 Billion)
CCC Intelligent Solutions is a leading cloud-based network supporting the automotive claims ecosystem. CCC's unique AI capabilities have enabled the company to deliver substantial value for critical automotive claims workflows, as industry stakeholders across the ecosystem, including carriers and repair shops, struggle with rising costs, increasing complexity, and ongoing labor shortages. With the company's event-based architecture and vast connected network, CCC's AI-driven IX Cloud handles massive amounts of data to the tune of over $1 trillion in claims processed through the platform to date. This widening network unlocks compounding value for users in real time as they deploy CCC solutions faster, more seamlessly, and with increased interoperability across products enabling more productivity through enhanced automation.

Shares of CCC trade at 7 times our calendar 2025 revenue estimates, a discount to the peer group. Further, we expect the multiple to trend toward the broader peer group and trade a premium multiple over time, as we believe CCC is well positioned in a resilient end-market. The company offers mission-critical tools to drive business efficiency, which should support a sustainable combination of growth and margin expansion long term. The automotive ecosystem continues to accelerate digital investments to leverage automation and AI to reduce the claims cycle timeline and improve operating profitability. Risks to our Outperform rating include competition, industry pace of adoption for cloud-based tools, and the potential for growing adoption of autonomous vehicles to minimize overall claims volumes.

Clearwater Analytics Holdings, Inc. (Market Cap $6.6 Billion)
Clearwater Analytics provides automated accounting and investment analytics software. Clearwater remains uniquely positioned to benefit from the prevalence of AI use-cases throughout asset management workflows, given the company's $7 trillion of assets across its platform. As complexity associated with global asset ownership and asset management continues to rise through increasing asset class types and reporting and compliance regulations, stakeholders across the investment ecosystem are looking to leverage innovative, AI-driven solutions to help reduce the operational workload to AUM ratio. We believe the proliferation of AI can be a significant value lever for customers through business generation, unlocking new asset classes and geographies to drive AUM growth, and operating scale through enhanced automation of manual processes. The company continues to target both sides of this proliferation, including with its most recent AI solution, Clearwater Intelligent Console, which provides quicker and more accurate portfolio querying and enables users to get the data they need to make better business decisions faster.

Clearwater shares trade at 33 times our 2025 EBITDA estimate, a premium to leading vertical software peers, which we believe is warranted given the company's durable Rule-of-50-plus financial profile and favorable competitive positioning. Risks to our thesis include macroeconomic volatility, competition, and execution risk.

Cloudflare, Inc. (Market Cap $41.4 Billion)
We view Cloudflare as a top AI play based on the company's Workers AI offering, which allows developers to run AI models and build and deploy AI applications on Cloudflare's edge network. Workers AI augments Cloudflare's Workers serverless edge compute platform with AI capabilities to support AI inference with NVIDIA GPUs within the company's global network, Vectorize (a vector database used with AI to power a variety of applications), and AI Gateway (for visibility, control, and insights on the use of AI applications). We view Cloudflare's AI infrastructure solutions as a source of momentum for the company as the number of developer accounts using AI increased 67% sequentially last quarter. Also, we believe the significant opportunity with AI inference models is driving interest in R2, which is Cloudflare's cloud storage solution that enables developers to access large amounts of data over the internet. We view AI inference as an emerging use-case for Workers AI that is gaining early traction, with Cloudflare reporting a 700% sequential increase in inference requests powered by Cloudflare AI. Also, we believe many enterprises are building data repositories and applications with both LLMs and SLMs, and that Cloudflare's use-cases are favorable for scalable, SLM applications where developers are looking for relatively low-cost and scalable infrastructure. In our view, Cloudflare's solutions are well positioned to capture early market share and to potentially grow as these applications evolve.

Cloudflare's stock last closed trading at about 20.5 times our 2025 revenue forecast of $2.025 billion on an EV/revenue basis, which we view as attractive given the company's unique positioning in attractive growth markets such as cybersecurity and edge compute. Risks for Cloudflare include competition, pricing pressure on commoditized CDN video traffic, potential inability to compete effectively in the SASE market, higher capex investments than anticipated, macroeconomic exposure to SMB customers, legal/geopolitical risks, and edge compute markets taking longer to develop than anticipated.

Datadog, Inc. (Market Cap $47.0 Billion)
We believe Datadog's expanding product suite and leading position in the cloud observability market position the company well to benefit from GenAI investments over the next few years. Over the past year, Datadog has seen a steady increase in AI-native customers contributing to total ARR, with this cohort representing 6% of total ARR in the most recent quarter (up from 4% in the prior quarter). Further, Datadog has expanded its solutions to include LLM observability, which helps customers improve visibility into their LLMs, and On Call, which is helping customers automate more pieces of the incident response process through GenAI. More broadly, over 3,000 customers have now used Datadog's AI integrations, up from 2,500 in the prior quarter. Lastly, we believe that Datadog's leadership position in the observability space gives it a nice advantage, as AI adoption should drive large workload migrations to the cloud over the next few years. Datadog shares trade at 50.3 times our 2025 free cash flow estimate, versus the peer group median of 52.3 times. Risks to our thesis include competition, cloud optimization headwinds, and the company's execution with acquisitions and new product offerings.

Elastic NV (Market Cap $10.4 Billion)
We believe that Elastic has a large opportunity ahead for its vector/hybrid search solutions as more AI applications are launched over the next few years. In 2023, the company launched Elasticsearch Relevance Engine (ESRE), which integrates transformer models and third-party LLMs into Elastic's vector search solution. This functionality, paired with Elastic's existing search features, helps users generate more efficient and contextualized search results. Thus far, the company has seen strong adoption of ESRE, with over 1,500 customers using GenAI functionality on the platform and new customer commitments growing close to 100% on a sequential basis in the company's most recent quarter. While it is still early days, we believe ESRE presents a large opportunity to upgrade existing customers to premium platform tiers, land new customers as companies look to develop AI applications, and drive increased consumption from existing customers given the

additional compute resources needed to power AI applications. Elastic shares trade at 6.1 times our 2025 revenue estimate versus software peers at 8.8 times. Risks to our thesis include competition, execution risk, and the uncertain macro environment.

GE Vernova Inc. (Market Cap $114.7 Billion)
Given our view that natural gas will be the major near-term benefactor from a shift toward dispatchable power for data center demand, we believe incoming secretary of energy candidate, Chris Wright, is likely to reduce regulatory burdens around this technology. As such, we see GE Vernova as the leading supplier of natural gas turbines, benefiting from demand and profitability. The recasting of natural gas as a viable option has inflected demand for new turbines as well, and GEV is doubling production capacity and is almost sold out through 2028. The demand is commanding higher margins than before on new turbines, and the higher-margin service contracts act as recurring revenue over the coming decades. GEV is the best positioned to benefit from the increased energy demand from AI by providing the turbines required to produce the power. GEV shares trade at a 26-times EV/EBITDA multiple on our 2026 estimate of $4 billion, a premium compared to the group average of 17 times. The company commands a leading position in key technologies required to power the energy production and electricity demands for AI and reshoring in the coming decade, thus, our Outperform rating. Investment risks include failures in the offshore wind business that could accumulate liabilities in addition to further reputational damage, and legislative risk from a change in administration and repeal of parts of the IRA and subsidies for wind energy.

Globant SA (Market Cap $9.1 Billion)
Globant is a technology services provider that combines the technical rigor of IT service providers with the culture and creativity of digital agencies to deliver innovative, next-generation software solutions for customers. Globant also offers over 30 studios that represent "deep pockets of expertise" across various technologies and industries. The Data & AI Studio is one of its top 10 studios in revenue contribution. In addition, Globant's AI Reinvention Studios are contributing over 20% to company revenue as of last quarter and have been highlighted as one of the company's "fastest areas of growth." We believe that Globant is a stock to watch as AI spending moves from infrastructure and hardware to applications and services. We believe the company has differentiated capabilities and is positioned well to pick up market share in this space. In the first nine months of the year, Globant's AI-related work resulted in over $250 million in revenue. Globant offers various AI consulting services, and we believe that its AI-powered platforms present the most differentiated opportunity for the company. Globant X, the products and platform division of Globant, was established in 2021 and serves as the company's incubator for next-generation products and platforms. Globant X includes eight different products and platforms: Augoor, GeneXus, Globant Enterprise AI, MagnifAI, Navigate, StarMeUp, Daxia, and Walmeric. We estimate that Globant X remains in the low single digits as a percentage of revenue, last disclosed in third quarter 2022 at 2.5% of total revenue. We are confident that Globant will continue to actively develop new platforms and services and continue to find ways to add value to clients looking to partake in the technology.

Based on a price of $210.13, Globant trades at 30.6 times our next-12-months' adjusted EPS estimate, compared to the peer group trading at 30.1 times. We believe that Globant's superior organic revenue growth profile and reputation as a leader in developing technology justifies a premium multiple compared to the peer group. Within our coverage, we continue to view Globant as one of our top two-year picks. We maintain our Outperform rating. In addition to the risk of a general macroeconomic slowdown or other events that affect services spending, other risks to our thesis include customer concentration at top customer Disney (8.3% of revenue), foreign-currency fluctuations, and low employee retention.

Grid Dynamics Holdings, Inc. (Market Cap $1.7 Billion)
Grid Dynamics is a fast-growth IT services company specializing in technical consulting, software design, development, and testing for *Fortune* 1000 customers. Grid's investments in technology and solutions have translated to a rapidly growing pipeline for AI work. As of the third quarter, Grid's pipeline consisted of 100 AI opportunities, representing 50% growth sequentially. Management believes that enterprises are increasingly interested in moving AI proofs of concept into full production mode. While we wait to see this dynamic manifest in Grid's results, we believe this incremental activity with enterprises is encouraging for Grid's near-term growth prospects and supports management's view that the demand environment is improving for the company. We are encouraged by management's commentary that Grid continues to win work for traditional digital transformation work given its strong technical capabilities in data and AI (i.e., enterprises are choosing Grid for future access to GenAI skills). One of the pillars in Grid's long-term growth strategy is to innovate at the intersection of business, technology, and data, which includes all things AI. Currently, Grid has five GenAI solutions—AI-powered data analytics, AI for process automation, AI for product images, conversational AI, and AI for developer productivity. In addition, Grid has a host of other AI-powered solutions including AI catalog optimization, AI search and recommendations, pricing and promotion optimization, fraud detection and prevention, predictive maintenance, and visual quality control. Grid has built several proprietary accelerators that ultimately help enterprises reduce time to market for new technology products and enhance developer productivity.

Overall, we continue to believe Grid Dynamics is in the early stages of growth, with strong technical capabilities and the benefit of robust multiyear secular tailwinds. Based on a price of $20.80, Grid Dynamics trades at 55.3 times our next-12-months' EPS estimate, a premium to the next-generation IT services peer group at 23.7 times. The company has historically traded at a notable premium to the peer group. Grid's size, client base, and superior offerings drive faster revenue and earnings growth than peers and premium valuation. We maintain our Outperform rating. In addition to the risk of a general macroeconomic slowdown or other events that affect services spending, other risks to our thesis include operational and geopolitical risks, particularly delivery exposure to Ukraine; a competitive labor market; and fluctuations in foreign currencies.

Meta Platforms Inc. (Market Cap $1,555.9 Billion)
Meta plays an important role in the AI ecosystem with an industry-leading, open-source LLM in Llama. Equipped with one of the largest proprietary datasets among technology companies, Llama's base model will continue to be leveraged by other companies looking to fine-tune their own model, in our opinion. In addition, in-house AI advancements are being applied within products like Advantage Plus, Meta Movie Gen, and WhatsApp for Business. We believe that Meta's use of agentic AI within WhatsApp for Business will allow for advertisers to bid on conversations to reach consumers. Moreover, its release of lightweight models combined with its vision for an open-source future leads us to believe that Meta is working on applying AI technology within edge devices such as thermostats, routers, scientific instruments, etc. While it is still early, its Reality Labs business segment continues to innovate in the VR/AR space with advancements to the Metaverse and the release of Quest 3S headsets along with smart glasses in a partnership with Ray-Ban.

We continue to feel positive about Meta's AI adoption and the benefits being realized by advertisers. Meta currently trades at 14.9-times our 2026 EBITDA estimate. Using our DCF framework, we believe the stock has roughly high-teens upside over the next 12 months using a 14-times EBITDA multiple and 10% discount rate. Risks include privacy concerns, increasing regulatory scrutiny, stalling user growth, artificial intelligence, macro advertising spend pullback, and slowing engagement.

Microsoft Corporation (Market Cap $3,185.8 Billion)
We see Microsoft as one of the biggest beneficiaries of the AI platform shift given its full-stack offering that spans AI infrastructure, model training, and AI applications. Central to Microsoft's leadership in AI is its deep partnership with and strategic investment in OpenAI, a pioneer in LLM research and development and the company behind the popular ChatGPT service. Microsoft not only provides the core cloud infrastructure (via Azure) on which OpenAI's models are trained and inferenced, but it also integrates OpenAI's capabilities into the various copilots across its portfolio, including Windows, M365, GitHub, Power Apps, and security. On the company's most recent earnings call, Microsoft highlighted AI as a meaningful accelerant to its Azure business, with AI cloud services contributing 12 percentage points of growth to the reported 34% Azure growth.

Microsoft trades at an enterprise-value-to-free-cash-flow multiple of 42.3 times and a price-to-earnings multiple of 31.8 times our calendar 2025 estimates. We continue to see Microsoft's wallet share within enterprises grow as customers see the benefits of consolidating vendors (across security, cloud services, productivity, endpoint management, DevOps, and collaboration). In addition, the broad application of AI copilots across different lines of business is helping expand Microsoft's wallet opportunity beyond traditional IT spend as companies across all industries trial these productivity-enhancing solutions. Risks to our Outperform rating include public cloud competition, a secular move away from Microsoft's profitable on-premises software solutions, more muted PC market growth, volatility among Microsoft's hardware-oriented segments (e.g., Surface and gaming), and general macroeconomic risks.

Modine, Inc. (Market Cap $7.3 Billion)
Modine's Airedale brand offers cooling solutions for the ambient environment in a data center, including chillers, air handlers, fan walls, and air conditioners. Data center cooling has been Modine's fastest-growing vertical over the last three years, nearly tripling in size, from about $100 million in revenue in fiscal 2021 to roughly $300 million in fiscal 2024 (41% CAGR). It now represents about 25% of Modine's total revenue, up from about 5%-10% historically. Data center revenue from the company's Scott Springfield Manufacturing acquisition is expected to double in fiscal 2025, from $60 million to about $120 million. Overall, Modine's data center revenue is expected to increase 100%-110% (50% organically), to a range of $588 million to $618 million in fiscal 2025. Management also expects data center revenue to increase about 45%-55% annually through fiscal 2027. A recently announced third hyperscale customer, geographic expansion to Asia, and the company's introduction of a new coolant distribution unit (rack-level liquid cooling) are all expected to be incremental to the previous projections.

At $140, shares trade at 31 times and 18 times our calendar 2025 adjusted EPS and EBITDA estimates, respectively. Based on 2025 consensus estimates, the peer group trades at 27 times EPS and 18 times EBITDA. The peer group includes companies that provide automotive components, HVAC products, and data center cooling solutions, along with some industrial equipment companies. We believe that Modine's position in the company's higher-growth end-markets should support a premium valuation. Based on our revenue and margin expansion assumptions, we expect shares to increase in line with earnings growth at a mid- to high-teens rate. We rate shares Outperform. Risks include a macroeconomic slowdown that could result in cautious customers, raw material inflation, and customer concentration.

Motorola Solutions, Inc. (Market Cap $78.2 Billion)
Motorola is investing in AI to help "automate the mundane" and accelerate analysis for law enforcement needs. Officers experience sensory overload so the demand to automate the surplus data to enable quick, necessary, and effective decisions is paramount. Motorola's AI algorithms and machine vision systems do not get bored, like security analysts in a security operations center (SOC) responsible for simultaneously monitoring several video feeds for long durations at a time. In retail settings, Motorola's cameras and software use pattern recognition with the flow of traffic

to identify elevated threat levels. Motorola's analytics software is designed to help prevent incidents rather than merely responding to incidents. Ultimately, there is a human in the loop to make the final call for an incident response. Motorola also has capabilities within its Vesta Next product line to automatically transcribe calls from victims or witnesses and help command centers alert first responders aided by keyword analysis. Driven by robust state and local government budgets and record stimulus, Motorola is seeing record demand for its LMR radios, video solutions, and command center software.

Motorola trades at 32 times consensus forward-year (2025) EPS, which is a premium to its 21-times February 2020 pre-pandemic multiple and its November 2021 peak multiple of 27 times. In our view, Motorola can maintain a premium multiple range as it demonstrates its resiliency to macro pressures. We expect double digit long-term EPS growth to drive double-digit stock returns. Accordingly, we reiterate our Outperform rating. In our view, the primary risk to Motorola shares is valuation multiple compression from decelerating revenue growth.

nVent Electric plc (Market Cap $12.3 Billion)

There is a large opportunity for nVent to provide cooling and power solutions in data centers, and with the liquid cooling industry growing roughly three times faster than traditional cooling solutions, we believe nVent's liquid cooling business could add 1 to 2 percentage points to overall organic sales growth annually. About 5% of data centers use liquid cooling today, and industry sources suggest penetration could be about 25% by 2028. The company's direct-to-chip liquid cooling solutions include coolant distribution units, rack manifolds, and rear-door heat exchangers. The company's data solutions product portfolio also includes power distribution units, which enable efficient power distribution through a rack and energy consumption control. Data solutions revenue totaled about $100 million in 2018 and is expected to exceed $575 million in 2024 (about 20% of total revenue accounting for the Thermal Management sale). Management estimates cooling and power should represent slightly more than 50% of nVent's data solutions revenue in 2024, up from 40% in 2023. We believe nVent's data solutions revenue should continue to grow at an annual rate of at least 20%.

At $75, shares trade at 24 times our 2025 adjusted EPS estimate and 18.5 times our 2025 adjusted EBITDA estimate (Thermal Management sale accounted for). Adjusting for approximately $10 in cash following the Thermal Management deal, shares trade at 21 times 2025 EPS and 16 times 2025 EBITDA. The peer group average multiples are currently 26 times 2025 EPS estimates and 18 times 2025 EBITDA estimates. We believe nVent should trade at least in line with this group given that, on average, nVent generally reports superior revenue growth, along with in-line margins and earnings growth. Given the opportunity for modest valuation multiple expansion and our forecast for sustained midsingle-digit organic revenue growth and low-double-digit earnings growth, we rate shares Outperform. Risks include potential organic sales growth uncertainty associated with nVent's end-markets, a slowdown in the generative AI buildout, and continued channel destocking by distributors.

Pure Storage, Inc. (Market Cap $22.9 Billion)

We see Pure Storage as a backdoor way to play AI given its recently announced design win with a top-four U.S. hyperscaler. Pure will be providing all online storage for the hyperscaler's next-generation data centers, with the customer decision largely based on the superior density (and power efficiency) of Pure's Direct Flash Modules (DFMs) compared to traditional hard drives and off-the-shelf SSDs. We believe this win could open the door for other hyperscalers and large enterprises that are looking to cost-effectively ramp up data center investments to increasingly support AI use-cases and workflows. On the enterprise side, Pure highlights three main opportunities tied to GenAI, which include: 1) machine learning and training environments, highlighted by its Nvidia DGX SuperPOD certification and strategic partnership with CoreWeave; 2) inferencing use-cases, where Pure recently introduced the GenAI Pod for on-premises environments (a set of full stack

solutions), providing time and cost efficiencies with GenAI projects; and 3) infrastructure modernization, with enterprises looking to Pure's Fusion software to create a unified global namespace that serves AI applications.

Pure's stock trades at an enterprise-value-to-sales multiple of 6.3 times and an enterprise-value-to-free-cash-flow multiple of 34.6 times on our calendar 2025 estimates. We see a favorable risk/reward equation for the stock at these levels in view of Pure's unique DFM architecture, ongoing share gains in the enterprise, and long-term opportunity to displace generic disk drives in hyperscaler data centers (validated by the top-four hyperscaler design win). Risk to the Pure Storage story include heavy competition in the on-premises storage market, application and data migration to the public cloud, volatility in NAND pricing, and potential macro issues.

ServiceNow (Market Cap $226.0 Billion)
ServiceNow is one of the broadest workflow providers in the software market and has been an important partner to enterprises in driving digital transformation over the last decade-plus. It helps organizations structure and automate workflows across multiple departments including IT, customer service, human resources, finance, cybersecurity, risk, and more. Its broad reach across departments within its customers, trusted relationship with CIOs and the c-suite, platform breadth, access to data, and rapid pace of innovation give it strong position to benefit from the GenAI theme. ServiceNow launched its AI suite of products (called Now Assist) in September 2023 and seen good initial traction, passing $100 million in ARR within one year of launch (with current adoption likely around $150 million - $200 million). We believe ServiceNow's ability to integrate new GenAI capabilities and AI agents into its existing deployments will enable it to scale further.

ServiceNow currently trades at 43 times our 2026 free cash flow estimates, versus its large-cap software peers at 32 times. While shares are not inexpensive, we believe the quality of the business deserves a premium multiple given its GenAI positioning, strong presence in the enterprise, platform breadth, and industry-leading rule-of-50 profile. We believe that a longer-term benefit is likely to accrue for ServiceNow given that it is a cornerstone of driving digital transformations. Risks include high investor expectations, lumpy deals, and competition in the company's emerging product suites (CSM, ESM, and ITOM).

Sterling Infrastructure (Market Cap $6.0 Billion)
Sterling's industry-leading data center site development business is surging and is still in the early innings. The company serves most of the data center industry's hyperscalers, including Amazon and Meta. We estimate that 27% of Sterling's 2025 revenue will come from data centers, on 65% year-over-year growth. Management recently commented that data center opportunities are "falling out of the sky every day," and its current data center backlog of roughly $500 million excludes a qualified pipeline of more than $1 billion. This robust demand provides high visibility into 2027. We forecast $977 million in Sterling data center revenue in 2030, up from an estimated $352 million in 2024. The company will likely add revenue to this total by targeting several geographic-focused acquisitions that are not included in our estimates. Robust EPS growth should continue as data center megaprojects are accretive to margins. Sterling's EPS have grown at a 39% compound annual rate over the past three years.

Even after shares have increased tenfold since the beginning of 2020, we believe that the company and stock price are just getting started. We expect share appreciation of greater than 20% per year over the next several years, based on maintenance of the company's midteens forward-year EBITDA multiple (currently trades at 17 times our 2025 adjusted EBITDA multiple). In our view, the main risk to shares would be a slowdown in the data center and AI market in the U.S.

Tesla (Market Cap $1,361.3 Billion)
We view Tesla Energy as the most underappreciated component of the Tesla story and expect this segment to augment the electric vehicle narrative. The three key drivers for energy storge are grid stabilization, the data center buildout, and renewables integration. The energy demand inflection from AI is stressing our already fragile grid systems, Tesla's Megapack adds critical resiliency and efficiency reducing inefficient start-stop occurrences of generating assets. Megapacks can be paired with either renewable or dispatchable power systems and act as the time arbiter to maximize the value of energy arbitrage. Each Megafactory has a capacity of about 10,000 Megapacks, and in a bull-case scenario, we estimate that it is capable of $14 billion in revenue (350,000 cars), a 39% gross margin (almost 3 times auto excluding credits), a 24% operating margin, and $0.75 in EPS. With three Megafactories running, we estimate Tesla Energy could generate $2.35 in EPS in 2028. Tesla's Megapack as the standalone leader in energy storage and believe it will capture significant market share at above corporate average margins. Tesla shares trade at 58 times our EV/EBITDA multiple on our 2026 estimate of $26 billion, a significant premium compared to technology peer's average of 16 times. Investment risks include 1) competition, particularly from Chinese EV and energy storage players; 2) geopolitical risk, with large exposure to customers in China; and 3) key-man risk with CEO Elon Musk.

Toast, Inc. (Market Cap $22.0 Billion)
Toast is the most comprehensive hardware and SaaS platform for small and midsize restaurants, and we believe that one of its most valuable assets is the proprietary data it has access to through 127,000 active restaurants on its platform. In recent quarters, Toast has introduced a number of products on the back of these data capabilities, including: 1) a benchmarking product that allows customers to compare a variety of metrics (sales trends, item pricing, etc.) to similar concepts in their area, 2) an AI-powered marketing assistant that delivers personalized marketing messaging to optimize sales, 3) enhanced tech support that leverages AI to automate technical support chats (already doubling productivity of support staff), and 4) Sous Chef, an AI copilot that allows restaurant operators to ask questions and receive personalized action plans based on proprietary operating and benchmarking data. For context, Toast's full product set would generate SaaS ARPU of roughly $30,000 (versus current ARPU of just over $6,000), and this number has consistently grown as Toast innovates and introduces high-utility modules to its platform. We expect these data/AI products to support a continuation (and potential acceleration) of this ARPU growth, with many being released toward the end of 2024. Ultimately, Toast is leveraging AI to not only provide insights but make it easy for restaurant owners to act on those insights, which is something we believe will further differentiate Toast's already-best-in-class platform from its competition; we are not aware of any competitors in the space with remotely similar capabilities.

Shares of Toast trade at 15 times 2025 gross profit, which compares with 11 times 2025 gross profit for fast-growth SaaS. Risks include competition, exposure to consumer spending patterns at restaurants and natural SMB restaurant churn, supply chain risk for its hardware solutions, and the need for high sales efficiency, which is more difficult with SMB customers. Risks to consider include: 1) competition; 2) exposure to consumer spending patterns at restaurants; 3) exposure to natural SMB restaurant churn; 4) supply chain risk for its hardware solutions; 5) the need for high sales efficiency, which is more difficult with SMB customers; 6) share count dilution; and 7) any credit risk undertaken with Toast Capital.

The prices of the common stock of other public companies mentioned in this report follow:

AAON, Inc. (Outperform)	$132.15
Advanced Micro Devices, Inc.	$122.28
Alphabet, Inc. (Outperform)	$199.63
Amazon.com, Inc. (Outperform)	$230.71
AppFolio, Inc. (Outperform)	$260.39
Arista Networks, Inc. (Outperform)	$121.50
Arm Holdings plc (Outperform)	$155.20
Astera Labs, Inc.	$124.41
Atlassian Corporation Plc (Outperform)	$256.19
Axon Enterprise, Inc. (Outperform)	$605.58
Broadcom, Inc. (Outperform)	$240.31
Caterpillar Inc.	$398.36
Cisco Systems, Inc. (Market Perform)	$61.03
CCC Intelligent Solutions Holdings, Inc. (Outperform)	$11.27
Clearwater Analytics Holdings, Inc. (Outperform)	$28.99
Cloudflare, Inc. (Outperform)	$119.85
Cognizant (Market Perform)	$78.45
Cummins Inc.	$367.17
Datadog, Inc. (Outperform)	$138.40
Elastic NV (Outperform)	$100.36
GE Vernova Inc. (Outperform)	$416.00
Globant SA (Outperform)	$210.57
Grid Dynamics Holdings, Inc. (Outperform)	$20.80
The Hewlett Packard Enterprise Company	$23.70
IBM Corporation	$224.26
Infosys (Market Perform)	$21.15
Intel	$21.77
Marvell Technology, Inc. (Not Rated)	$123.78
Meta Platforms Inc. (Outperform)	$616.46
Micron Technology, Inc.	$109.38
Microsoft Corporation (Outperform)	$428.50
Modine, Inc. (Outperform)	$139.58
Motorola Solutions, Inc. (Outperform)	$467.84
Netflix, Inc. (Outperform)	$869.68
nVent Electric plc. (Outperform)	$74.88
NVIDIA Corporation (Outperform)	$140.83
Oracle Corporation (Outperform)	$172.57
Palantir Technologies Inc. (Underperform)	$73.07
Palo Alto Networks, Inc. (Outperform)	$183.51
Pure Storage, Inc. (Outperform)	$70.08
Salesforce, Inc. (Outperform)	$326.84
ServiceNow, Inc. (Outperform)	$1096.85
Shopify Inc. (Outperform)	$106.28
Snowflake, Inc. (Outperform)	$173.53
Sterling Infrastructure (Outperform)	$196.55
Taiwan Semiconductor Manufacturing Company Limited	$218.70
Talen Energy Corp	$243.64
Tesla, Inc. (Outperform)	$424.07
T-Mobile	$219.49
Toast, Inc. (Outperform)	$38.65
Trane	$397.15
Twilio Inc. (Outperform)	$113.88
UiPath Inc. (Market Perform)	$13.33
Vertiv	$143.13
Zoom Video Communications, Inc. (Outperform)	$78.54

IMPORTANT DISCLOSURES

William Blair or an affiliate beneficially own or control (either directly or through its managed accounts) 1% or more of the equity securities of Pure Storage, Inc., AAON, Inc., Clearwater Analytics Holdings, Inc., Grid Dynamics Holdings, Inc. and nVent Electric plc as of the end of the month ending not more than 40 days from the date herein.

William Blair or an affiliate is a market maker in the security of Arista Networks, Inc., Arm Holdings plc, Broadcom Inc., Cisco Systems, Inc., Microsoft Corporation, NVIDIA Corporation, Oracle Corporation, Pure Storage, Inc., Snowflake Inc., AAON, Inc., Amazon.com, Inc., AppFolio, Inc., Axon Enterprise, Inc., CCC Intelligent Solutions Holdings Inc., Salesforce, Inc., Cognizant Technology Solutions Corporation, Clearwater Analytics Holdings, Inc., Datadog, Inc., Elastic N.V., Grid Dynamics Holdings, Inc., GE Vernova Inc., Globant S.A., Alphabet, Inc., Infosys Technologies Limited, Meta Platforms, Inc., Modine Manufacturing Company, Marvell Technology Group Ltd., Motorola Solutions, Inc., Cloudflare, Inc., Netflix, Inc., ServiceNow, Inc., nVent Electric plc, Palo Alto Networks, Inc., UiPath, Inc., Palantir Technologies Inc., Shopify Inc., Sterling Infrastructure, Inc., Atlassian Corporation Plc, Toast, Inc., Tesla, Inc., Twilio Inc. and Zoom Video Communications, Inc.

William Blair or an affiliate expects to receive or intends to seek compensation for investment banking services from Arista Networks, Inc., Arm Holdings plc, Broadcom Inc., Cisco Systems, Inc., Microsoft Corporation, NVIDIA Corporation, Oracle Corporation, Pure Storage, Inc., Snowflake Inc., AAON, Inc., Amazon.com, Inc., AppFolio, Inc., Axon Enterprise, Inc., CCC Intelligent Solutions Holdings Inc., Salesforce, Inc., Cognizant Technology Solutions Corporation, Clearwater Analytics Holdings, Inc., Datadog, Inc., Elastic N.V., Grid Dynamics Holdings, Inc., GE Vernova Inc., Globant S.A., Alphabet, Inc., Infosys Technologies Limited, Meta Platforms, Inc., Modine Manufacturing Company, Marvell Technology Group Ltd., Motorola Solutions, Inc., Cloudflare, Inc., Netflix, Inc., ServiceNow, Inc., nVent Electric plc, Palo Alto Networks, Inc., UiPath, Inc., Palantir Technologies Inc., Shopify Inc., Sterling Infrastructure, Inc., Atlassian Corporation Plc, Toast, Inc., Tesla, Inc., Twilio Inc. and Zoom Video Communications, Inc. or an affiliate within the next three months.

Officers and employees of William Blair or its affiliates (other than research analysts) may have a financial interest in the securities of Arista Networks, Inc., Arm Holdings plc, Broadcom Inc., Cisco Systems, Inc., Microsoft Corporation, NVIDIA Corporation, Oracle Corporation, Pure Storage, Inc., Snowflake Inc., AAON, Inc., Amazon.com, Inc., AppFolio, Inc., Axon Enterprise, Inc., CCC Intelligent Solutions Holdings Inc., Salesforce, Inc., Cognizant Technology Solutions Corporation, Clearwater Analytics Holdings, Inc., Datadog, Inc., Elastic N.V., Grid Dynamics Holdings, Inc., GE Vernova Inc., Globant S.A., Alphabet, Inc., Infosys Technologies Limited, Meta Platforms, Inc., Modine Manufacturing Company, Marvell Technology Group Ltd., Motorola Solutions, Inc., Cloudflare, Inc., Netflix, Inc., ServiceNow, Inc., nVent Electric plc, Palo Alto Networks, Inc., UiPath, Inc., Palantir Technologies Inc., Shopify Inc., Sterling Infrastructure, Inc., Atlassian Corporation Plc, Toast, Inc., Tesla, Inc., Twilio Inc. and Zoom Video Communications, Inc.

This report is available in electronic form to registered users via R*Docs™ at https://williamblairlibrary.bluematrix.com or www.williamblair.com.

Please contact us at +1 800 621 0687 or consult https://www.williamblair.com/equity-research/coverage for all disclosures.

Jason Ader attests that 1) all of the views expressed in this research report accurately reflect his/her personal views about any and all of the securities and companies covered by this report, and 2) no part of his/her compensation was, is, or will be related, directly or indirectly, to the specific recommendations or views expressed by him/her in this report. We seek to update our research as appropriate. Other than certain periodical industry reports, the majority of reports are published at irregular intervals as deemed appropriate by the research analyst.

DOW JONES: 44424.20
S&P 500: 6101.24
NASDAQ: 19954.30

Additional information is available upon request.

Current Rating Distribution (as of January 27, 2025):

Coverage Universe	Percent	Inv. Banking Relationships *	Percent
Outperform (Buy)	71	Outperform (Buy)	9
Market Perform (Hold)	28	Market Perform (Hold)	1
Underperform (Sell)	1	Underperform (Sell)	0

*Percentage of companies in each rating category that are investment banking clients, defined as companies for which William Blair has received compensation for investment banking services within the past 12 months.

The compensation of the research analyst is based on a variety of factors, including performance of his or her stock recommendations; contributions to all of the firm's departments, including asset management, corporate finance, institutional sales, and retail brokerage; firm profitability; and competitive factors.

OTHER IMPORTANT DISCLOSURES

Stock ratings and valuation methodologies: William Blair & Company, L.L.C. uses a three-point system to rate stocks. Individual ratings reflect the expected performance of the stock relative to the broader market (generally the S&P 500, unless otherwise indicated) over the next 12 months. The assessment of expected performance is a function of near-, intermediate-, and long-term company fundamentals, industry outlook, confidence in earnings estimates, valuation (and our valuation methodology), and other factors. Outperform (O) - stock expected to outperform the broader market over the next 12 months; Market Perform (M) - stock expected to perform approximately in line with the broader market over the next 12 months; Underperform (U) - stock expected to underperform the broader market over the next 12 months; not rated (NR) - the stock is not currently rated. The valuation methodologies include (but are not limited to) price-to-earnings multiple (P/E), relative P/E (compared with the relevant market), P/E-to-growth-rate (PEG) ratio, market capitalization/revenue multiple, enterprise value/EBITDA ratio, discounted cash flow, and others. Stock ratings and valuation methodologies should not be used or relied upon as investment advice. Past performance is not necessarily a guide to future performance.

The ratings and valuation methodologies reflect the opinion of the individual analyst and are subject to change at any time.

Our salespeople, traders, and other professionals may provide oral or written market commentary, short-term trade ideas, or trading strategies-to our clients, prospective clients, and our trading desks-that are contrary to opinions expressed in this research report. Certain outstanding research reports may contain discussions or investment opinions relating to securities, financial instruments and/or issuers that are no longer current. Always refer to the most recent report on a company or issuer. Our asset management and trading desks may make investment decisions that are inconsistent with recommendations or views expressed in this report. We will from time to time have long or short positions in, act as principal in, and buy or sell the securities referred to in this report. Our research is disseminated primarily electronically, and in some instances in printed form. Research is simultaneously available to all clients. This research report is for our clients only. No part of this material may be copied or duplicated in any form by any means or redistributed without the prior written consent of William Blair & Company, L.L.C.

This is not in any sense an offer or solicitation for the purchase or sale of a security or financial instrument. The factual statements herein have been taken from sources we believe to be reliable, but such statements are made without any representation as to accuracy or completeness or otherwise, except with respect to any disclosures relative to William Blair or its research analysts. Opinions expressed are our own unless otherwise stated and are subject to change without notice. Prices shown are approximate. This report or any portion hereof may not be copied, reprinted, sold, or redistributed or disclosed by the recipient to any third party, by content scraping or extraction, automated processing, or any other form or means, without the prior written consent of William Blair. Any unauthorized use is prohibited.

If the recipient received this research report pursuant to terms of service for, or a contract with William Blair for, the provision of research services for a separate fee, and in connection with the delivery of such research services we may be deemed to be acting as an investment adviser, then such investment adviser status relates, if at all, only to the recipient with whom we have contracted directly and does not extend beyond the delivery of this report (unless otherwise agreed specifically in writing). If such recipient uses these research services in connection with the sale or purchase of a security referred to herein, William Blair may act as principal for our own account or as riskless principal or agent for another party. William Blair is and continues to act solely as a broker-dealer in connection with the execution of any transactions, including transactions in any securities referred to herein.

For important disclosures, please visit our website at williamblair.com.

This material is distributed in the United Kingdom and the European Economic Area (EEA) by William Blair International, Ltd., authorised and regulated by the Financial Conduct Authority (FCA). William Blair International, Limited is a limited liability company registered in England and Wales with company number 03619027. This material is only directed and issued to persons regarded as Professional investors or equivalent in their home jurisdiction, or persons falling within articles 19 (5), 38, 47, and 49 of the Financial Services and Markets Act of 2000 (Financial Promotion) Order 2005 (all such persons being referred to as "relevant persons"). This document must not be acted on or relied on by persons who are not "relevant persons."

“William Blair” and “R*Docs” are registered trademarks of William Blair & Company, L.L.C. Copyright 2025, William Blair & Company, L.L.C. All rights reserved.

All statements in this report attributable to Gartner represent William Blair’s interpretation of data, research opinion or viewpoints published as part of a syndicated subscription service by Gartner, Inc., and have not been reviewed by Gartner. Each Gartner publication speaks as of its original publication date (and not as of the date of this report). The opinions expressed in Gartner publications are not representations of fact, and are subject to change without notice.

William Blair & Company, L.L.C. licenses and applies the SASB Materiality Map® and SICSTM in our work.

William Blair

Equity Research Directory

John Kreger, Partner Director of Research +1 312 364 8612
Kyle Harris, CFA, Partner Operations Manager +1 312 364 8230

Scott Hansen Associate Director of Research +1 212 245 6526

CONSUMER

Sharon Zackfia, CFA, Partner +1 312 364 5386
Group Head–Consumer
Lifestyle and Leisure Brands, Restaurants, Automotive/E-commerce

Jon Andersen, CFA, Partner +1 312 364 8697
Consumer Products

Phillip Blee, CPA +1 312 801 7874
Home and Outdoor, Automotive Parts and Services, Discount and Convenience

Dylan Carden +1 312 801 7857
E-commerce, Specialty Retail

ECONOMICS

Richard de Chazal, CFA +44 20 7868 4489

ENERGY AND SUSTAINABILITY

Jed Dorsheimer +1 617 235 7555
Group Head–Energy and Sustainability
Generation, Efficiency, Storage

Tim Mulrooney, Partner +1 312 364 8123
Sustainability Services

FINANCIAL SERVICES AND TECHNOLOGY

Adam Klauber, CFA, Partner +1 312 364 8232
Group Head–Financial Services and Technology
Financial Analytic Service Providers, Insurance Brokers, Property & Casualty Insurance

Andrew W. Jeffrey, CFA +1 415 796 6896
Fintech

Cristopher Kennedy, CFA +1 312 364 8596
Fintech, Specialty Finance

Jeff Schmitt +1 312 364 8106
Wealthtech, Wealth Management, Capital Markets Technology

GLOBAL SERVICES

Tim Mulrooney, Partner +1 312 364 8123
Group Head–Global Services
Commercial and Residential Services

Andrew Nicholas, CPA +1 312 364 8689
Consulting, HR Technology, Information Services

Trevor Romeo, CFA +1 312 801 7854
Staffing, Waste and Recycling

HEALTHCARE

Biotechnology

Matt Phipps, Ph.D., Partner +1 312 364 8602
Group Head–Biotechnology

Sami Corwin, Ph.D. +1 312 801 7783

Lachlan Hanbury-Brown +1 312 364 8125

Andy T. Hsieh, Ph.D., Partner +1 312 364 5051

Myles R. Minter, Ph.D. +1 617 235 7534

Sarah Schram, Ph.D. +1 312 364 5464

Healthcare Technology and Services

Ryan S. Daniels, CFA, Partner +1 312 364 8418
Group Head–Healthcare Technology and Services
Healthcare Technology, Healthcare Services

Margaret Kaczor Andrew, CFA, Partner +1 312 364 8608
Medical Technology

Brandon Vazquez, CFA +1 212 237 2776
Dental, Animal Health, Medical Technology

Life Sciences

Matt Larew, Partner +1 312 801 7795
Life Science Tools, Bioprocessing, Healthcare Delivery

Andrew F. Brackmann, CFA +1 312 364 8776
Diagnostics

Max Smock, CFA +1 312 364 8336
Pharmaceutical Outsourcing and Services

INDUSTRIALS

Brian Drab, CFA, Partner +1 312 364 8280
Co-Group Head–Industrials
Advanced Manufacturing, Industrial Technology

Ryan Merkel, CFA , Partner +1 312 364 8603
Co-Group Head–Industrials
Building Products, Specialty Distribution

Louie DiPalma, CFA +1 312 364 5437
Aerospace and Defense, Smart Cities

Ross Sparenblek +1 312 364 8361
Diversified Industrials, Robotics, and Automation

TECHNOLOGY, MEDIA, AND COMMUNICATIONS

Jason Ader, CFA, Partner +1 617 235 7519
Co-Group Head–Technology, Media, and Communications
Infrastructure Software

Arjun Bhatia, Partner +1 312 364 5696
Co-Group Head–Technology, Media, and Communications
Software

Dylan Becker, CFA +1 312 364 8938
Software

Louie DiPalma, CFA +1 312 364 5437
Government Technology

Jonathan Ho, Partner +1 312 364 8276
Cybersecurity, Security Technology

Maggie Nolan, CPA, Partner +1 312 364 5090
IT Services

Jake Roberge +1 312 364 8056
Software

Ralph Schackart III, CFA, Partner +1 312 364 8753
Internet and Digital Media

Stephen Sheldon, CFA, CPA, Partner +1 312 364 5167
Vertical Technology – Real Estate, Education, Restaurant/Hospitality

EDITORIAL AND SUPERVISORY ANALYSTS

Steve Goldsmith, Head Editor and SA +1 312 364 8540

Audrey Majors, Editor and SA +1 312 364 8992

Beth Pekol Porto, Editor and SA +1 312 364 8924

Lisa Zurcher, Editor and SA +44 20 7868 4549