snowflake®

AI + DATA PREDICTIONS 2025

Our AI-accelerated future is here and it means operationalizing AI, rewriting the leader's skill set and harnessing the next killer app

TABLE OF CONTENTS

THE BIG STORY: ARTIFICIAL INTELLIGENCE (STILL)

The most important thing happening in enterprise technology in 2025 will be the most important thing that's been happening since late 2022: Advanced artificial intelligence, particularly LLMs and generative AI. But where 2023 was the year of panic and wonder, and 2024 was the year of experimentation, 2025 is when enterprises start to get real about what AI can specifically do for their organizations.

Enterprises have begun graduating their AI proofs of concept from the sandbox to full production just as some of AI's biggest cheerleaders have turned a bit dour. "Where's the ROI?" cried the barons of Wall Street (Goldman Sachs, under the headline "Gen AI: too much spend, too little benefit?") and "This get-rich-quick scheme won't make us all rich quick!" gasped the lords of Silicon Valley (Sequoia Capital, adding that plenty of people will benefit, in time).

We've seen CIO-focused news sites tout rising AI budgets for 2025, while CFO-oriented sites have said CFOs will become more conservative about AI spending. (And those articles were written by the same reporter!) The conflicting vibes reflect how all-over-the-place every pundit, consultant and professional expert has been as the hype beast demands new opinions on the hour.

Snowflake CEO Sridhar Ramaswamy agrees with the perspective of Wired, which offered the arguably deflating headline, "The Unsexy Future of Generative AI Is Enterprise Apps."

"The hype cycle will continue, but for me it's a tale of two cities — one is Hype and the other is Let's Go Get Some Work Done," he says. "And we are firmly in the latter."

So while the investor perspective may be that some of us are the bull market and some of us are the china shop, the enterprise challenge is separating obstacle from opportunity. So let's give that a shot. Nearly two dozen Snowflake leaders and technology experts have contributed to this report to try to describe how — and how much — the transformative promise of AI will be realized in the year ahead.

AI MOVES INTO PRODUCTION: VALUE AND OBSERVABILITY

In 2023 and ’24, board directors seemed to have a recurring nightmare: They’re running down the dock just as a big steamer called AI is pulling out of port. Afraid of missing the boat in real life, they woke up demanding that their C-suites figure out a strategy around generative AI and LLMs immediately, and get started. At something.

“In 2023-24,” says Snowflake SVP of Engineering Vivek Raghunathan, “the investor would call the board and say, ‘What are you doing about AI?’ And the board asked the CEO, ‘What are you doing about AI?’ and the CEO asked the CDO or CIO, and money got spent on a lot of prototypes and demos.”

“Organizations were really sandwiched,” says Jennifer Belissent, Snowflake’s Principal Data Strategist. “You had pressure to ‘figure out AI’ coming from above, and you had rogue use of open AI tools like ChatGPT from below.”

That experimentation has been important, the former Forrester analyst says, for exploring what’s possible and determining what’s practical. As a result, “Organizations have solidified their data strategy and have become more systematic in their pursuit of AI value.”

Which is good, Raghunathan says, because 2025 is the year when CFOs start asking about ROI.

“There will be a focus on which AI projects allow a company to do well in terms of price, governance and actual results,” Raghunathan says. “The drivers of what gets put into production are what enterprises always demand: high quality, low cost, safe and secure.”

> **The next focus of AI strategy will be operationalizing LLMs, requiring evolved approaches to observability, governance and security.**

Bringing AI out of the sandbox and into production raises new questions — or, rather, familiar concerns around security, governance and observability.

“You need new layers of supportive infrastructure for security and governance,” says Mona Attariyan, Director of Machine Learning at Snowflake. “You need a new top layer of observability, which was not really needed with the static data set during the pilot phase. But it’s essential to reliability, transparency and explainability, especially over time.”

Platforms will roll out AI observability solutions as a core element of bringing AI to production.

In enterprise systems generally, observability refers to the ability to see and understand the state of the system. It's a performance-management discipline. The emerging field of AI observability examines not only the performance of the system, but also the quality of the outputs of an LLM, including accuracy, ethical/bias issues, and security problems such as data leakage. It helps ensure that the guardrails placed around the model continue to work.

"As soon as you have an app that is in production for real use cases and taking in new data, you need to be able to assess performance and to detect, diagnose and solve any problems," Attariyan says. "Without observability, you're flying blind."

While there are many AI observability startups and open source initiatives gaining traction, Snowflake Head of AI Baris Gultekin says that observability will end up in the hands of data platforms and the large cloud providers.

"It's hard to do observability as a standalone startup," he says. "And companies that adopt LLMs are going to need AI observability solutions, so big cloud providers will be adding the capability."

While LLMs and generative AI tools are new technologies that raise new concerns, CIOs can be trusted to manage the rollouts. At least, that's what our CIO tells us.

"IT knows how to roll out applications," says Sunny Bedi, CIO and CDO of Snowflake. AI, he says, has been a progression, not a sudden arrival. "A few years ago, we were learning to use ML to power automation and solve business challenges. In the last year or so, that has evolved to introducing AI-powered applications."

The overall challenges around governance, security and compliance are understood, he says, though AI brings new technical and regulatory challenges. The degree of complexity is directly related, he adds, to the quality of your data and AI strategies.

Operationalizing AI will be harder at first, then it'll be much easier.

"Operationalizing AI does require a data platform and AI-focused tools," he says. "If you have a unified data strategy, you solve half the problem by bringing all your data together. Then with AI tools, you can address the other half: the security, compliance and governance challenges. You need both to protect and empower your AI deployments and alleviate the burden of stitching everything together in an ad hoc way."

In addition to outfitting your IT teams with the right platform and tools, the AI-forward enterprise must pitch a large tent. A major trajectory for failure in operationalizing and integrating AI, says Raghunathan, would be failing to incorporate all the right stakeholders.

HOW CDOS SUCCEED WITH AI

In 2024, Jennifer Belissent talked to a dozen CDOs across industry and geography to ask them how they were approaching the advent of advanced AI. She details their collective insights in "The Data Executive's Guide to Effective AI." Their firsthand experience covers the entire growth cycle, from initial experimentation through operationalization to transformation. Through it all, and amid waves of AI excitement and anxiety, they maintain a sharp focus on utility and value.

"Companies are going back to the basics and really thinking about what they want to achieve."

"Adopting new platforms and technologies is as much a social construct as a tech construct," he says. "Organizations will fail if they don't bring along all the stakeholders who should have a voice, especially around governance, security and ethics."

And if bringing AI into production seems difficult today, Raghunathan has reassuring words about the future: "In the long run, the successful AI vendors will make AI powerful, simple, safe and secure. Many of the issues organizations might struggle with today will vanish into the background as an implementation detail."

> **The next evolution in data platforms focuses not only on formatting, storing and accessing data but on contextualizing it as well.**

Getting there depends not only on those AI vendors delivering miraculous advances. Enterprises will have to continue to make their data ready for these advanced systems. Gultekin says that data platforms will continue to help organizations optimize their data.

"In the very near future," he says, "data will be more AI-ready, so AI can easily and safely plug it into existing agents and applications."

The evolution of the data platform is essential to the evolution of AI, says Benoit Dageville, Snowflake Co-Founder and President of Product. He predicts breakthroughs in terms of making LLMs better understand the data they're working with.

"What's missing from today's data platforms is the semantic layer, the understanding of what the data means," he says. "When you have financial data in a table, it's the developer or the analyst who understands where the data came from, how it was calculated, what it means."

Having to build that understanding into every application you develop on top of your data would be simply too burdensome, he says.

"The semantic layer is a critical aspect that needs to be pushed down close to the data so the AI can understand the nature of the data and do a much better job," he says. "You don't want to reinvent the semantic concepts for each app. You want to push down to the data layer — that's the next evolution."

KILLER APP: AGENTS OF CHANGE

What is the best use of AI? Finding the killer app among the countless promises of game-changing gen AI and LLM capabilities can be difficult when there's so much potential.

"We encourage our customers to focus on what creates value," says CEO Sridhar Ramaswamy. "Is it that your customers get the information that they need faster? Is your decision-making faster? Are you able to get more done? Is it the efficiency of being able to get your hands on information without having a ton of intermediaries and wasted time?"

"Organizations are going through the cases where an LLM could replace complex processes or outdated tooling," says Director of ML Mona Attariyan. "Evaluating the trade-offs takes time and should involve a good amount of rigor. LLMs are amazing, but they're not a clear-cut win for every single thing that they might be able to do."

"Not every application is going to be AI-assisted," Ramaswamy agrees. "But many will combine a large language model, a knowledge repository and human input in various ways, while getting smarter and smarter over time. And it's pretty exciting, because you're going to see a number of practical applications in surprising contexts."

One argument would be that the breadth of capabilities puts LLMs and gen AI in a category beyond that singular, dominant use case. In other words, advanced AI doesn't need a "killer app."

"What an LLM does is itself the killer app," says SVP of Engineering Vivek Raghunathan. "Look, if you ask me what was the original killer application of spreadsheets, maybe it was finance. But that's maybe 10% of use cases throughout the enterprise. The killer app for the spreadsheet was the spreadsheet. It was a new way for people to be productive, and I guarantee you that there are industries now where you couldn't live without spreadsheets. Similarly, the killer app of a LLM is the LLM itself. It is a new way for people to be productive, and people are discovering new use cases of it every day."

He points to a solid handful of powerful applications of LLMs that are already here, and improving steadily, such as assistants for coding, writing and information retrieval. But it doesn't stop with "copilots."

"The next step is that these copilots will become pilots," he says. "Autonomous agents. And the impact of that will be considerable. There's always a long tail of people using a technology for things you don't expect. LLMs are a bicycle for the mind that will unlock creativity for things we haven't even conceived of yet."

So, in sum, there's no need to think in terms of a killer app for artificial intelligence, because it's — whoa, wait, back up there a moment. What was that about copilots becoming pilots?

> **AI's "killer app" will be autonomous agents.**

Chatbots and assistants work best when they're built for a relatively focused task or corpus of information. The approach tends to be more cost- and resource-efficient, and the error rate goes down significantly. But that means that these tools are rather limited in their capabilities.

Until you get them talking to each other.

"We are at the beginning of that — building the layers and capabilities that compose an agent," says Snowflake Co-Founder Benoit Dageville. The result, he says, will be an interface that can turn a broad prompt for a complex action ("Build a marketing campaign around our new product," "Plan and book everything for my weekend getaway to Vancouver") and break it down to a discrete set of subtasks that are farmed out to a chain of focused assistants. Such a powerful application is not all that far away.

"Many of the capabilities needed to implement these agents in your enterprise exist now, to some degree," Dageville says. "You need a lot of data in a unified platform, an application framework, and key toolkits such as RAG, or retrieval-augmented generation. Snowflake has those things now, via our platform, Streamlit and Cortex. So while creating a truly powerful agent will require dramatic improvement, it's not a vision for something that doesn't exist yet."

"We'll start surfing the copilot-to-pilot trend in 2025," Raghunathan says. "But it will take awhile longer to get to a truly independent agent, because to be independent it must be nearly perfect."

Today, Raghunathan says, a copilot is a tool that's wrong easily 20% of the time. "Your single-task coding assistant or enterprise search assistant does one thing and gets it, at best, 80% correct, and then humans fix the outputs. Such models perform about as well as an intern."

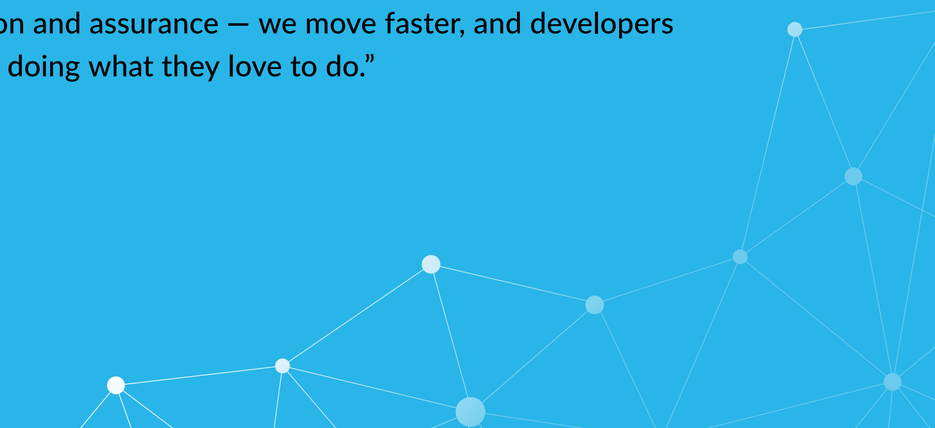
In other words, they can handle something repetitive that needs to be simplified and automated, but they're not up for reasoning, planning, and a high degree of sophistication.

"That's going to be the measure of next-generation performance," Raghunathan says. "Not how well a model does on the SAT, but how well it does multistep planning and problem solving. Every startup and big player is working on this with current models, with the expectation that next-generation foundational models will do it with much higher accuracy."

COPILOT CODING TOOLS WILL HELP DEVELOPERS DO WHAT THEY LOVE.

The idea of autonomous agents able to execute complex sets of tasks is exciting, but don't sell the copilot short. Snowflake CIO/CDO Sunny Bedi notes that coding assistants, for instance, don't just let humans work faster. They also foster greater engagement and creativity, making sure the drudgery gets done quickly and well, without crushing a developer's will to code.

"If you talk to developers about the software development lifecycle, across the design, development and testing phases, you'll learn that pretty much no one likes QA," says Bedi, who is working to help his developers offload more of that quality assurance work to AI assistants. "Good QA is very cumbersome and time consuming. If we can offload 40% or more of the testing process to an AI-powered assistant — with human supervision and assurance — we move faster, and developers spend more time doing what they love to do."



THE GOOD, THE BAD AND THE HALLUCINATORY

As noted, the hype cycle will rollercoaster its way through the media and the board room for a good while. An important way to detach from those ups and downs, says Jennifer Belissent, is to think less about the bright, shiny objects themselves, and more about what you actually want to do.

"We've now entered the era of purpose-built tools delivered to serve specific functions," she says. "As a result, we're less enthralled by gen AI itself and trying to apply it to everything. We're increasingly focused on what actually needs to be done, and identifying the tools to do it."

Generative AI and LLMs will be infrastructure, an underlying enabler to many things.

> **A lot of AI "backlash" or negativity will be mitigated one successful use case at a time.**

Although the hype cycle may not be one of them, there *are* things to be concerned about with AI. LLMs "hallucinate" and, really, "The machine will sometimes just make stuff up" is not a small caveat. Especially if you're going to be held accountable for the outputs.

"Hallucinations are the biggest blocker to getting generative AI tools in front of our customers' customers," says Baris Gultekin, Snowflake's Head of AI. "Right now, a lot of gen AI is being deployed for internal uses, because it's still scary to make it available to the public when you feel like you can't control what the AI is going to say."

But there will be improvements, especially in terms of keeping AI outputs within acceptable boundaries.

"There's increasing ability to run guardrails on the output of these models," Gultekin says. "That's one of the most important product requirements. Guardrails will constrain what generative AI can or can't say, what tone is or isn't allowed. Models increasingly understand these guardrails and can be tuned to protect against things like bias, for instance."

In addition to establishing guardrails, access to more data, to diverse data and to more relevant sources will improve accuracy. "Data diversity is one means of mitigating the risk of AI models capturing internal bias and 'hallucinating' or, plainly speaking, making mistakes," says Jennifer Belissent. "Getting a second opinion is a common practice for humans — it must be true for automation, too. We need data diversity."

She adds that processing of unstructured data, for example, will continue to increase significantly as companies tap into previously inaccessible sources. So will data sharing both within organizations ("in-sourcing") and across broader business ecosystems. Yet it's a new technology development, retrieval-augmented generation, that will drive significant reduction of hallucinations.

> **RAG will effectively mitigate the hallucination problem.**

"RAG in particular can dramatically reduce hallucinations," says Snowflake Co-Founder and President of Product Benoit Dageville. RAG is an AI framework that works to improve the accuracy and timeliness of AI outputs by combining information retrieval systems with the language generation capability of a large language model. By further grounding a response in factual information and linking to public sources, RAG represents a tremendous advance in the reliability and explainability of AI outputs.

"If you ask an open-ended question to ChatGPT, it doesn't use data, it uses its training and can make up a false answer," Dageville says. "In the enterprise, you want to ground generative AI with real facts."

RAG, he says, draws information from relevant sources and can cite them for verification. It just makes for a better answer-generator. "It can analyze structured and unstructured data and give you a summary that is more human, more accurate and more transparent."

This leads to simple tweaks that make generative AI outputs more dependable, such as making the tools cite their sources, which goes a long way toward making outputs verifiably accurate.

Snowflake Distinguished AI Software Engineer Yuxiong He, who previously co-founded and led DeepSpeed, the open source deep learning optimization library developed by Microsoft Research, says there will also be qualitative choices on the market.

"You'll be able to balance hallucination reduction against cost and other performance metrics," she says. "Different industries, use cases or workloads have different tolerances for error. Models and AI systems will be developed with different levels of accuracy, and their costs will reflect that. Businesses will select models based on their specific needs and budget."

> **The GPU market will self-correct (in most places), allowing companies to better manage their AI-related costs and goals.**

One of the most talked-about AI problems in 2024 was the shortage of AI-powering GPU chips. All our experts say the limited chip supply will be short-lived.

"Demand exploded in a way that no one had predicted," Dageville says. "It's a supply chain issue. I'm not worried about it persisting."

"Nvidia is going to make more chips both in the high end and the low end. And there are inventive ways to mitigate the current GPU shortage. At Snowflake, for example, we offer options for customers to route traffic across deployments to places where there is GPU capacity," CEO Sridhar Ramaswamy says. "The cloud service providers and new players like Groq are busy creating new chips to meet demand as well, not to mention new models that are both highly capable and cheaper to run."

The problem, says Gultekin, is that the super-chip future will not be evenly distributed at first. Europe is more worried about the GPU shortage than companies in the United States, where there is greater capacity.

"Regional availability will be a longer-term problem," he says. And even the kind of solution Ramaswamy mentions may give organizations pause at first. The need to access GPU capacity in another country may incite legal complications or organizational anxiety.

"There may be regional data laws to consider, and even without that, it's still a mental shift for security architects," Gultekin says.

"Sending your data to a different region, even through a secure connection and within the same platform, is something the security architects will need to vet and get comfortable with."

Technology aside, there are also legal and ethical hurdles with the potential to obstruct AI development, notes Snowflake EVP of Product Christian Kleinerman. "Intellectual property and the rights associated with AI inputs and outputs are key considerations for AI development in some organizations," he says. "There are solutions for some aspects of it, such as licensing agreements, but there's no easy button."

> **Ethical AI will become a top priority for enterprises, driving improvements in transparency, fairness and accountability.**

Even if progress is made on IP, other ethical questions around advanced AI will continue to vex the industry. Adequately addressing regulatory and ethical concerns is difficult, in part because there are many stakeholders, many philosophical approaches and many potential outcomes. And a lot of competitive pressure, because there's a lot of money at stake.

"The challenge for the enterprise is the lack of transparency in the models," says Snowflake Head of Product Security Anoosh Saboori. "There is no food label listing the ingredients, in terms of data sources. AI providers don't want to disclose all their secret sauce, but until they solve this, they're not going to completely unblock the enterprise space."

"Across the industry, we're seeing more openness," says Yuxiong He. "You can look at the release of Llama 3.1 last summer to see that there's a very comprehensive report, with a good amount of information about data sources, training recipe and so on."

She notes that Snowflake has also moved toward greater clarity. "Openness and transparency were critical to our release of Snowflake Arctic in 2024," she says. "Unlike many models, we're sharing not only the model weights and fine-tuning codes, but also data sources and training recipes. That's our approach, and overall the whole industry is moving more and more toward openness."

That's not to say that industry will provide all the necessary solutions. Ramaswamy says the hands-off approach that industry leaders often demand is not going to work.

"I think it's very hard to take a blanket 'no laws, no regulation' approach when we have technologies that are so broadly applicable," he says. "Absolutely, there's a real risk in regulating a field too early. But I do think that smart regulation and smart legislation are possible. And industry and government should continue to work together."

Even in the most cooperative environment, regulation may be slow in coming. Governments are notorious for moving at glacial speed against blizzards of innovation, but sorting out standards for AI will be especially hard, says Kleinerman.

"The good news is that existing regulation already covers many of the aspects people may be concerned about regarding uses of AI," he says. "But AI will be pervasive throughout technology we use every day, and that will complicate things."

AI will affect every industry short of a kindergartener's summer lemonade stand (then again, knowing how kids embrace tech …), while areas of potential concern range from privacy, bias and accuracy to IP, cross-border data transit, security and accountability. Fortunately, a lot of misbehavior is already covered. Much like there are no laws about hammers, but there's a law that generally covers hitting someone with a hammer.

"In that sense," Jennifer Belissent says, "regulation would be best applied to use cases and outcomes, as opposed to trying to prevent development of the technology itself."

> **Legal restrictions against AI misuse will largely come from existing laws rather than a broad raft of AI-specific regulation.**

Of course, that doesn't address every potential case. There was a lot of news coverage starting in late 2023 about the fact that in San Francisco, a continuous testing ground for autonomous vehicles, police could not issue traffic tickets when a driverless car broke a traffic law. Largely because the law speaks of issuing the ticket to the driver — not, for instance, the owner or manufacturer of the car. By September 2024, a bill facilitating ticketing autonomous scofflaws had cleared the California Assembly and was headed for the governor's desk.

"There needs to be a combination of efforts that recognize these legitimate problems, including blind spots in existing laws and regulation," Ramaswamy says. "And certainly regulation will be enacted if industry doesn't act."

WORKING — AND LEADING — IN AN AI WORLD

There has been much talk about what AI will do to the workforce across industries. There is consensus that jobs will be lost and jobs will be created, but most will be transformed by AI-infused tools.

Obviously, workers will face a learning curve. There will be less demand for repetitive, tactical execution (increasingly the AI's sweet spot) and more for data-driven strategic thinking (a human's job, at least for now). That's the future of work. What about the future of leadership? How will bosses have to change or improve their game to keep up with an AI revolution?

> **Leaders, as much as workers, will be challenged to adapt to an AI-accelerated world.**

"The big leadership challenge in the AI era is how managers uplevel their teams' productivity with AI tools," says Product EVP Christian Kleinerman. "There's a big opportunity to democratize AI by making it easier for less technical people to leverage, securely and responsibly, but it takes vision and leadership to get there."

Leaders may have trouble keeping pace with their accelerated workforce. If a dozen direct reports are now doing the work of a 30-person team, that's more to manage and review.

"Leaders will need to be very smart about actively adopting AI tools themselves, as well as to empower their workers," says CEO Sridhar Ramaswamy. "What's good for the goose is good for the gander."

One thing both Kleinerman and Ramaswamy say is that the concept of "human in the loop" means that even when AI performs the work or suggests a course of action, humans will have to take responsibility.

"At the end of the day, you have to verify your work," Ramaswamy says. "I have little sympathy for someone who says they used an AI model and got their work wrong. It behooves us to verify the work that we do."

So cross "blame it on the AI" off your list of workplace excuses.

MANAGE AI AS A WORKER, NOT A TOOL.

Rather than perceiving AI as a mere tool, the advice of experts is to envision AI as a remote team member — an additional employee capable of improving business operations or customer experience. Doing that changes the dynamics of the relationship. We know how to manage an employee and we have tools and processes in place to do so.

- **Define the role with an explicit "job description."**
- **Identify the right AI model "candidate," either to be bought or built.**
- **Invest in "onboarding" by training models with appropriate data.**
- **Establish output and outcome performance metrics.**
- **Monitor performance against benchmarks to attribute value generation.**

These methods ensure models are kept up to date and continue to deliver value.

— From The Data Executive's Guide to Effective AI.

Asked for his best tip or secret to evolving as a leader in the high-tech space, Ramaswamy, who also founded the search platform Neeva and led Google AdWords to industry-defining success, offered a philosophy rather than a technique.

"Endless curiosity," he says. "In both my work and personal lives, I'm using all kinds of AI models extensively. We all need to get better at using tools, and the more comfortable we feel with different kinds of AI models, the more we understand their strengths and weaknesses, the better we are going to be in dealing with the world that's ahead of us."

One outcome of giving workers and leaders an AI boost is that routine tasks become automated, leaving more time for the kind of visionary efforts that really move the bottom line.

"We'll have more cycles to devote to being more imaginative," says Snowflake CIO/CDO Sunny Bedi, "and then applying AI to create those outcomes."

Snowflake Head of AI Baris Gultekin provides an example of putting AI-driven productivity gains to good use. "If, for instance, you generate SQL faster with Snowflake Cortex, you can have more SQL analysts doing deep investigative work instead of answering simpler questions from business users."

That again points to the need to uplevel the workforce, from top to bottom. And at the cutting edge of AI development, the pressure is considerable.

"The rate of innovation has been incredible in the last nine to 12 months," says Vivek Raghunathan, Snowflake's SVP of Engineering (and Ramaswamy's co-founder at Neeva). "My peers and I feel a constant combination of exhilaration and whiplash — in a good way — of going into the week and not knowing what comes next. And if anything, that pace is still accelerating."

> **Exhaustion is going to plague many in the AI space.**

"The notion of keeping up with AI is interesting, because of how fast everything is changing," Gultekin says. "There's a hot new thing, or new development, every month, and leaders can't chase them all."

There's a danger, he says, of a project always being current but never being finished. "If you've just deployed Llama 3, and now 3.1 is available, do you do the upgrade? Can you afford not to?"

Gultekin says the competition and the sense that there's so much opportunity and value to be captured with AI drives everyone to the point of exhaustion. "Everyone I talk to who works in AI says they're working much harder than they ever have, because there's so much innovation and change that what you did 10 days ago can now be done better."

The pace of innovation isn't likely to slow down, he says, so leaders in the AI age have to resist getting carried away by the excitement.

"Leaders need to focus on goals and ROI," he says, "rather than chase either the shiny object or every upgrade."

AI TRANSFORMATION IS A DECADE'S WORK.

Yes, AI is going to change everything. But not all at once. Amanda Kelly, Snowflake's Director of Product Experiences and Co-Founder of Streamlit, says that for most companies, broad workplace transformation driven by AI is still years out.

"A lot of companies don't really know how they operate or what makes them successful today," she notes. "Patterns evolve organically, and decision-making structures are often convoluted, and therefore slow. It's going to take time for most companies to incorporate AI and understand how it can truly elevate them.

"With AI, the companies that I see having the most success have one of two things," she adds. "Either they have very strong top-down leadership with a vision for the transformation, or they don't force a vision, but empower individual teams to figure out how to use LLMs to transform their team."

OPEN SOURCE ACCELERATES AI INNOVATION

The open source movement is important to the software industry in general, and it will be crucial to the development of advanced AI. Open source projects play important roles in the broad ecosystem of software vendors, data platform providers and AI makers. For AI, open source democratizes access to ideas, fosters diversity of thinking, accelerates innovation, creates communities to enhance understanding of AI systems, can reduce costs, and more.

A discussion about the future of data, AI and enterprise strategy would be remiss without considering where open source is going in the years ahead. Generally, says Distinguished AI Software Engineer Yuxiong He, open source will continue to play a critical role, particularly in the tools used for training AI systems.

"Other aspects of model development — such as highly optimized inference systems and integrating models into production environments — are more appropriate for a closed source approach, and will likely remain that way," she says. "But when you talk about AI training tools, to be able to make continuous innovation, to be able to attract adoption, open source will remain the best way to go."

> **AI-driven low-code and no-code breakthroughs will make developers more important than ever.**

Another place where open source advances will be most strongly felt is in coding assistants. AI-driven low-code and no-code development tools will elevate the role a skilled developer plays in an organization. With time-consuming but basic coding challenges handed off to assistive tools, devs will focus on more complex, high-value tasks, system architecture, and integration of AI capabilities.

Amanda Kelly, Co-Founder of Streamlit, says that rather than devaluing developers by equipping any half-trained coder with an AI sidekick, highly skilled devs will be freed to turn their deep expertise to solving bigger, business-critical problems. They'll become more strategic. But Kelly warns that it won't be an overnight transformation.

"This is happening already, but the seismic shift will take time," she says. "It really takes innovative thinkers to fundamentally change the approach to development. Right now most companies are mostly using AI coding tools for small efficiency gains. They're not really thinking about a complete paradigm shift, in part because the tools are not there — yet."

"But when they get here, they'll create a bigger tent," says Jeff Hollan, Head of Applications and Developer Platform at Snowflake. It's not just about speed or freeing a dev from a tedious task. "Developers with different skills and domains will be able to work together, harnessing more creativity and brainpower."

Open source projects also make significant contributions to the data analytics space, and one that Snowflake has been involved with seems poised to go mainstream. Apache Iceberg is an open source table format meant to bring SQL-like simplicity and reliability to large-scale data analytics.

> **Iceberg will go mainstream and finally combine operational and analytical data.**

Organizations increasingly seek efficient ways to manage and analyze their ever-larger datasets, and the robust community driving Iceberg's features and flexibility have positioned it well for mainstream adoption. We could list the benefits — a unified data lake with a consistent table format for managing diverse data sources, real-time data ingestion, schema evolution — and clearly will, at the drop of a hat.

As data volumes continue to explode and real-time analytics become increasingly important, Iceberg's ability to handle petabyte-scale datasets efficiently makes it an attractive choice. Its vendor-agnostic nature aligns well with modern data architectures, promoting interoperability and reducing vendor lock-in. These factors, combined with Iceberg's robust feature set, position it as a strong contender for widespread adoption in the evolving big data landscape.

"It is going mainstream," Kelly says. "Many, many folks are sold on the vision, and we're seeing a shift."

As always, true widespread adoption will take time, but Hollan agrees that Iceberg is the consensus choice. "It's all but unstoppable," he says.

> **The demand for interoperability will accelerate adoption of open standards, data democratization and ecosystem collaboration. It'll also increase the focus on data governance and security.**

The snowballing success of Iceberg (see what we did there) will also help drive other key trends in the open source space. Part of Iceberg's appeal is that it meets the demand for interoperability between data processing engines and storage systems. This capability will give further momentum to the adoption of open standards that allow enterprises the benefits of vendor-agnostic solutions.

Iceberg is also aces at addressing the ever-increasing needs around governance and security. Those benefits plus the democratization angle — ease of use is a key feature — will help drive data literacy initiatives that will assist with the workforce transformation needed in this dawning AI era. And because Iceberg is cloud-agnostic, it'll allow enterprises to pursue whichever cloud-native, multi-cloud and cross-cloud strategies they need.

Hollan adds that, for him, the real value is not the vendor-agnostic nature of Iceberg, because that would be only one of the considerations in making a major vendor move.

"The interoperability of Iceberg and other open source projects is truly great," he agrees, "but the value that will be most evident to organizations is not 'I can leave this cloud anytime,' but the robust and fast-moving developer community."

GEN AI AND LLMS: SECURING THE NEW ATTACK SURFACE

Because of AI advancements, 2024 was the year that cybersecurity finally became a high-stakes, rapidly evolving digital battleground that pitted defenders against criminal and nation-state actors. Kidding; it has always been like that and probably always will be. Even so, the development and deployment of LLMs and generative AI do give stressed-out security teams more to worry about — and also things to be hopeful about.

Leading with the bad news, gen AI makes it easier for criminals to build convincing social engineering attacks. Hard-to-spot deepfakes and phishing schemes will sound more like a workplace colleague or a helpful support person from your favorite big-box retailer, and less like a soulless scammer in a basement or business park half a planet away. And those copilot tools that help your developers write code faster and safer (if they check the work) also help the bad guys write better malware.

The good news is that new AI-based tools continue the advantages brought by automation and machine learning in recent years. The idea has always been to help short-staffed teams of human defenders sort the malicious signal from the hazy noise, surfacing the incidents that require attention and intervention. New AI-based tools promise to be better at that than ever. And 2024 saw vendors emerging with tools specifically designed to secure advanced models.

"Overall," says Anoosh Saboori, Snowflake's Head of Product Security, "the driving change in security is the amount of data that needs to be protected, and the size of the attack surface. This is shaping how AI is adopted by security teams."

That's an aspect of AI to focus on, Saboori says: AI as an attack surface.

> **The model itself is the next focus of AI-centered attacks.**

"Last year, we talked about attacks at the container layer — the less-secured developer playgrounds," he says. "Now the attackers are moving up a layer to the ML infrastructure. We'll start seeking patterns like attackers injecting themselves into part of the pipeline to make the model provide the wrong answer, or worse, reveal the information it was trained with."

Forewarned = Forearmed

Risks of a successful attack on an AI model include training data leakage, exposure of private data, data poisoning, IP theft (including theft of the model itself), bias issues, and underlying vulnerabilities in the software packages that support an ML project.

Anoosh Saboori points the concerned reader to the Snowflake white paper “An AI Security Framework” for a deeper dive into the dangers and mitigation strategies.

There is concern that threat actors could poison LLMs with vulnerabilities that can be later exploited. Think about the way scammers create fake pages with fraudulent customer service numbers so that when someone searches for a number to, say, rebook a flight, the scam page surfaces and connects the caller to identity thieves posing as customer service reps. Then remember that the same pages that fool human customers can be scraped to train AI.

Similarly, developers are using AI coding tools to accelerate productivity, but AI-delivered code may contain vulnerabilities. Last summer, university researchers demonstrated that LLMs could be poisoned with flawed code that would produce vulnerable outputs. And they weren’t the first. To mitigate this risk, code scans need to be integrated into the developer experience to flag vulnerabilities and recommend remediation earlier in the development lifecycle.

Snowflake CISO Brad Jones says that all the questions around AI security mean that organizations will have to establish a rigorous, formal approach to how advanced AI is operationalized.

> **AI security will become more formalized as we bring LLMs and generative AI into production.**

“There need to be whole levels of controls around how a model was trained and what data was used,” Jones says. “There needs to be an auditable trail, certifications, governance structure and controls around the model. And there will be a bigger push toward more formality around that.”

The ad hoc approach to date will crystalize around industry standards. Jones points to the new ISO 42001 standard, which he says will soon be table stakes in the AI space. Saboori mentions data security posture management — a comprehensive approach to continuously monitoring and improving an organization’s approach to protecting its data. DSPM, he says, will help reassure organizations feeling competitive pressure to pour their data into complex AI models they don’t quite understand.

“The tooling around AI gateways is going to be key to fostering a level of comfort with bringing up new models,” Jones adds. “‘Having standard controls and protections built in at the gateway will allow for rapid testing and adoption of new models’ — that is going to be key.”

As security leaders develop their best practices around cybersecurity in the AI era, a significant element will be the security data lake. A security data lake stores large volumes of security-related data from diverse sources, consolidating logs, events and other security information for advanced analytics, threat detection, incident response and long-term retention. Flexibility and scalability are key benefits.

> **The security data lake, as part of a modular approach to cybersecurity tooling, will be essential to security in an AI era.**

"It's a question of cost and scale," Saboori says. "The interest is driven by the greater trend in the market, which is to be able to prevent, detect and respond to attacks when the amount of data involved is now much bigger."

Security data lakes are supporting the rise of more modular and data centric security strategies. Snowflake customers have said they've been unhappy with monolithic SIEMs because controlling costs forces compromises around data collection and retention. Security data lakes let customers achieve full visibility into all of their security data at a small fraction of the cost, while working with a rich ecosystem of best-of-breed cyber apps.

Best-of-breed thinking will also let security teams test and incorporate AI-driven security assistants as they come to market. Such tools have been garnering attention at least since the RSA Conference last spring. As an Axios report noted, "Executives and government officials have been hopeful that generative AI tools will help alleviate cyber workers who are stretched thin and burned out."

> **AI security copilots will extend human teams.**

New copilot tools that will apply AI to the inhuman volume of potential incidents a security operations center faces promise to make junior analysts more effective, but Jones notes that the tools also bring concerns.

"Copilot tools could bring a new risk of containing within the model code or copy that may have some level of copyright," Jones says. "Some discipline needs to be there."

In the long run, as AI tools become more versatile and more accurate, security assistants will become a significant part of the SOC, easing the perennial manpower shortage.

"The benefit of AI will be to summarize incidents at a higher level," Saboori says. Rather than an alert that requires analysts to go through all the logs to connect the dots, he says, they'll get a high-level summary that makes sense to a human, and is actionable. "These systems are getting more accurate, and as the false-positive rate trends much closer to zero, you'll have fewer false-alarm 'emergencies.' Not only will your security team be more effective, but job satisfaction will go up because when analysts do have to respond, it's actually an attack."

Jones emphasizes that though AI may bring new attack vectors and defensive techniques, the cybersecurity field will rise to the occasion, as it always does. The tech may be new, but the basic concerns — data loss, reputational risk and legal liability — are well understood. The specific new risks, such as gen AI hallucinations or inaccuracies, will be addressed.

"Generative AI will never be 100% accurate, and organizations will have to know where that risk is tolerable," he says. "There may be more risk tolerance for using a particular solution to ask customer support questions versus making financial transactions. There's going to have to be human oversight and validation for a long period of time."

FOCUSED FORWARD MOTION

In 2024, AI hit the lowest point of the enthusiasm cycle (it's not all hype) since the explosive advent of ChatGPT. A skeptical eye is good and necessary. It forces deeper thinking and broader discussion. It can prevent budget-killing impulsivity and reputation-damaging mistakes. It can encourage us to think about guardrails and responsibilities as much as profit and potential.

But don't let the need for clear-eyed vision raise doubts about AI's overall potential. Remember that people will tend to overestimate a technology shift in the short term, and underestimate it in the long term. Take mobile, says Jennifer Belissent.

"It wasn't until we had significant infrastructure built out, and sufficient bandwidth to do interesting and productive things, that mobile really took off," she notes. "It also helped that the technology became much easier and more affordable for the average person to use."

When gen AI hit the public consciousness, the general public was genuinely stunned, says Mona Attariyan. "People were in awe of what generative AI and LLMs can do and were overlooking some of the downsides and difficulties," she says.

Those downsides and difficulties may be easier to see now, but they'll be diminished with new advances, and as a greater infrastructure around AI is built out.

> **Hold onto your hats, folks. We're just getting started.**

"It's going to take a few more years, but there are a lot of startups out there with a ton of new ideas," Attariyan says. "It will be a life-changing wave of innovation."

Through continuing evolution, the technology will have to become smarter, and perhaps so will its creators, as we develop new ethical guidelines and regulations to make the technology successful, equitable, transparent, and good for society as a whole. We'll definitely have to keep getting better at recognizing AI's implications and impacts in the world.

"AI will make us smarter, driving improvements in health, education, urban planning and much more," says Sridhar Ramaswamy. "And all of us — technologists, policy makers and everyday citizens — must learn to use every new technology effectively and responsibly."

A LOOK AT KEY INDUSTRIES

Along with our broad discussions about how AI, in particular, will shape the year(s) ahead, we talked to our experts in seven specific sectors to understand what their industries can expect in 2025 and beyond.

ADVERTISING, MEDIA AND ENTERTAINMENT

EXPANDING CREATIVITY, MAXIMIZING BUDGETS

Advertising, media and entertainment companies are already data-driven businesses, which has made it easier to explore and embrace technologies like generative AI.

"They're an experimental industry to begin with," says Bill Stratton, Global Head of Media, Entertainment and Advertising at Snowflake. "Advertising, media and entertainment tends to be a leader in adoption."

But, while they may not be under the same level of regulatory scrutiny as financial services or healthcare, the creative nature of their work can run them into IP issues more often.

> **Big IP holders (and their big legal teams) will take the lead in forging a licensing regime.**

Response to ongoing legal challenges made by independent artists against text-to-image-based AI models will drive model makers to license content from publishers. The growing concerns around fair use and compensation for IP will also lead media titans to develop their own proprietary models.

"Companies that own IP — whether media companies or creative agencies or image companies like Getty — will develop their own models so they can marry their IP with the deployment of generative AI," Stratton explains.

AI will let advertisers' budgets go further, because there will be fewer middle men.

Another benefit of AI tools will be efficiency across the entire process of getting messages to consumers. As it stands, marketers lose nearly half of their dollars somewhere between launching an advertisement and getting in front of the customer. A clear benefit that AI will bring to advertisers is more bang for their budget.

"For every dollar that's spent by a marketer, let's say 60 cents goes to the actual distribution and optimization of marketing content across different channels. The remaining 40 cents is spent in the ecosystem between the seller and the buyer," Stratton explains. As the adtech/martech convergence continues and truncates the value chain by eliminating redundancies, less of an advertising dollar will disappear before it reaches the consumer.

While any corporation will be concerned with how far budgets can stretch, the ultimate measure of success for anyone in advertising, media and entertainment is conversion to consumer. The market wants to more clearly draw a straight line from marketing efforts to revenue.

Data gravity — and the ability to leverage data where it resides — will give marketers new and better opportunities to drive conversions.

"What prevented that in the past was this data movement or data copying," Stratton says. "Now that you don't have to do that, with technology solving that data movement problem, we'll see more marketers able to gauge their campaigns' success through conversions."

FINANCIAL SERVICES

BALANCING INNOVATION WITH FISCAL PRUDENCE

Financial services organizations tend to be the leaders in adopting new technology, and because they're so important to economies, they influence wider adoption. Yet the industry is heading for a brief pause in its adoption of advanced AI, says Rinesh Patel, Global Head of Financial Services at Snowflake.

> The era of unchecked AI spending will come to a close in financial services.

"The momentum around AI may slow as the financial services industry level sets and evaluates the demonstrable ROI," he says. "Everyone's got a line item in their balance sheet saying GPUs, but no one's able to actually explain what it's delivering."

The industry's approach to AI will continue to strike a balance between innovation and fiscal prudence. "Boards don't care about experiments; they care about profit sustainability," he says. "There's a bright line between firms that can explain what ROI is being delivered by AI and firms that are merely experimenting."

He sees AI adding value in two ways: through augmenting workers and improving their efficiency, and through autonomous operation, without human involvement. The latter, he says, will remain in the experimentation stage for some time. A technological development with strong bottom-line potential can't come with an outsized risk.

"The board is going to care about ROI, of course, but not without safeguarding the organization," he says. And, naturally, regulators are also vigilant about protecting firms, consumers and the financial system.

"Regulators have already started paying closer attention to the evolution of this technology," he says. Looking ahead, they'll aim to protect investors, ensure there's enough liquidity in the market, etc.

> **There will be a disastrously negative headline or event, caused by AI, that is significant enough to jumpstart regulation.**

Regulatory frameworks and compliance requirements will adapt as they always have, with reactivity. "Regulators are still trying to grasp this technology and keep up with the momentum," Patel says. "Regulation will either proactively stop the worst facets of technology — like biased advisory based on race, gender, religion or political inclination — or regulation will be a consequence of some black swan event."

A black swan, he says, could be similar to the 2010 flash crash. Fifteen years ago, the financial services industry was in the throes of a debate around the merits of human-driven, discretionary hedge fund strategies versus computer-driven, systematic hedge fund strategies. Then, in 2010, the Dow Jones fell more than 1,000 points in just 10 minutes — over $1 trillion dollars lost. It was a flash crash later attributed to just one London-based futures trader quickly selling off thousands of small contracts to manipulate the market. From a regulatory perspective, first came the crash, then came regulations to prevent it. Expect something similar with an LLM going rogue.

"When that hallucination happens and someone gets burned — me, you or some institution — we will see regulatory acceleration," Patel says.

Given the financial industry's long-term outlook, how does such a crisis happen? Cutting corners could trigger a crisis if one organization puts short-term gain ahead of long-term responsibility.

"You can have a use case with AI, but if you have not put the right guardrails around that and understood governance and responsible AI, then obviously you leave yourself exposed as an organization," Patel says. "It's really all about governance and transparency."

> **There will be winners and losers in the race to operationalize AI in financial services but the gap separating them will be massive, not marginal.**

Data and culture will play a huge role in the adoption of AI. Governance and transparency are moot points in the absence of a modern data strategy and well-established data culture.

"When AI is deployed alongside a solid data strategy and surrounded by a culture of openness and creativity, it has the potential to be the greatest accelerant we've seen since the rise of cloud computing," Patel says. "When those conditions don't exist, AI can seem a quixotic investment. The firms that get those conditions right, which is difficult, will run far ahead of the competition. There won't be a photo finish in this race."

HEALTHCARE AND LIFE SCIENCES

EMBRACING AI WITH CAUTION

The healthcare and life sciences industry approached the gen AI hype cycle with its usual pragmatism. While the industry is embracing AI — with pharma and research institutions grabbing headlines like "AI Gains Ground with Breast Cancer Diagnosis and Prevention" — the patient care side will continue to proceed with caution.

Expect healthcare and life sciences organizations to maintain a measured pace that will prioritize evidence-based proof points and tangible results. With high standards of care, stringent regulatory requirements and physicians' concerns about potential lawsuits, they will need to see AI-driven results in other highly regulated spaces before they dive in.

"People never want to be a first mover in this space, especially on the patient side," says Jesse Cugliotta, Global Head of Healthcare and Life Sciences at Snowflake. "We saw some providers break away from that mentality during COVID-19 because everyone was forced to be innovative, but that risk appetite is coming back to normal levels."

On the product development side, big pharma makes its margins on a handful of products or drugs. So, it will continue to be more willing and able to take the risks associated with gen AI in pursuit of the next big revenue stream.

> **The healthcare industry will rely on existing regulation to guide smart, responsible AI rollouts.**

Of all the questions surrounding AI, how to use AI ethically may be the biggest one. Healthcare and life sciences professionals operate against a series of core ethical principles guiding their decisions, so, they have a better starting point than most.

"There's precedent for what this would look like if we use AI to help drive clinical decisions," Cugliotta says. "There's an entire legal and ethical infrastructure established to protect against medical malpractice. The buck has to stop with someone, and that's generally a licensed physician who is held liable for care decisions."

So while other industries fear that people will be replaced entirely by AI, in healthcare and life sciences, people will always be involved. They just now have the added responsibility of ensuring AI results are correct.

The existing ethical and regulatory frameworks within healthcare and life sciences will make transparency, reliability and explainability all the more crucial. As AI technology matures, hallucinations will become less frequent overall, and businesses will choose which models to implement, based in part on their tolerance of risk. In healthcare and life sciences, obviously, that tolerance will always be extremely low.

"[There are a] lot of the black-box models that are probably never going to be used for patient care," Cugliotta says.

> **But they'll also move faster than others to bring industry leaders and regulators together to sort out ethical challenges.**

And while existing regulations will cover most scenarios, there will inevitably be new ethical challenges brought to a head by AI.

"There's a certain level of responsibility that the end user has when asking a search engine a question," Cugliotta says. But when generative AI delivers a single answer, that same end user may feel less motivated to weigh opinions and feel more emboldened to take a wrongheaded — and potentially harmful — action. "We haven't come up with the liability infrastructure for that yet," he says.

To address those gaps, consortiums will come together to develop the rules of engagement.

"There will be a combination of voices — from governments, academia, clinicians and the tech community — trying to come up with standards for something like this," Cugliotta says. "These efforts are often used to influence regulations around the world."

MANUFACTURING

LESS HYPE, ALL ROI

Rather than get caught up in the swings of the AI hype cycle, the manufacturing industry will refocus on value-based enterprise applications that can prove ROI.

Tim Long, Global Head of Manufacturing at Snowflake, says that the fastest-moving organizations are tightly focused on results.

"They're not starting with IT gains, but with the business value of an AI solution, then taking iterative steps to achieve it," Long says. "They achieve value, measure the impact, and use the ROI to invest in the next project. That's the flywheel."

> **Large vision models will be a major focus for manufacturers.**

One of the first AI technologies to have a big effect in manufacturing will be large vision models — an AI model designed to process and interpret visual inputs which will revolutionize quality control in manufacturing, significantly improving product design, reducing inspection time, and accelerating production throughput.

"Gen AI can be used to simulate product defects which are then used to train computer-vision defect detection models. Or, models can simulate product design improvements, exploring different tradeoffs," Long says. "I see AI tools lifting the workforce to tackle higher-level tasks overall."

But large vision models and generative AI, on the whole, won't eliminate human involvement. "I think of it as a step forward in terms of workforce productivity and how they use these tools to perform at a higher level," Long explains. AI won't replace you, he says, but if you don't keep up, you could be replaced by someone who knows how to use AI.

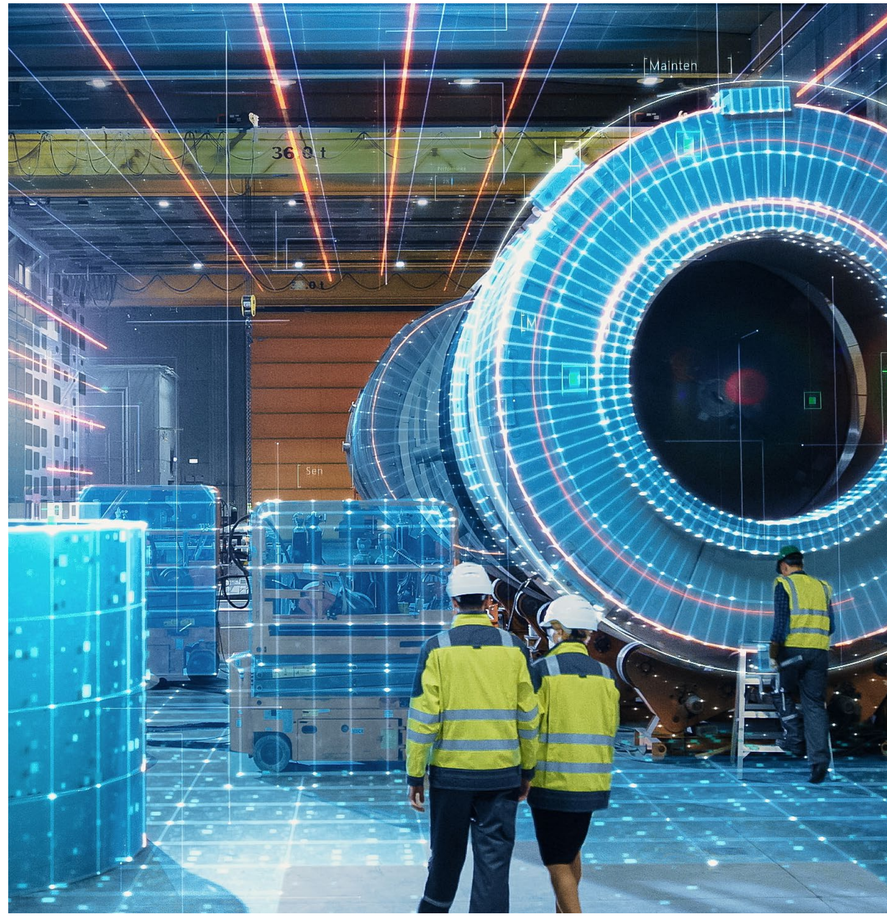
AI will also help manufacturers reach environmental goals. That may seem counterintuitive, given the ongoing concerns about AI's intense energy demands, but our industry experts are optimistic that, in the long term, AI will optimize resource usage, enhance product design, predict supply chain downtime, and help identify opportunities for manufacturers to recycle reclaimed materials from their processes.

AI will make the fourth industrial revolution real in the next decade.

AI's effect on manufacturing will be about more than simple efficiencies. It will be transformative. "The fourth industrial revolution kind of became a marketing term that people didn't really believe could meet the hype," Long says. "But with all these tools and cloud capacity, dreams of this higher performance in a factory will be a reality for the many manufacturers who invested in unsiloing their data and implementing AI/ML solutions."

The expected shift toward unified data, more fully optimized for AI consumption, will accelerate this. "An integrated, standardized data strategy will allow manufacturers to deploy one AI solution across an entire factory network," Long says. "That's where we go from incremental efficiencies to true digital transformation."

Beyond AI, Long says that environmental concerns are driving change in manufacturing. "There's both social and regulatory pressure on organizations to really reduce carbon emissions not just within their own facilities, but within their entire supply chains," Long said. "European markets are driving that change in manufacturing, and data is the key enabler."



TELECOMMUNICATIONS

LEANING INTO AI, GEOSPATIAL DATA AND MONETIZATION

Gen AI and 5G will be major priorities in 2025. The complexity of 5G networks — the physical infrastructure, the strain on virtualized networking functions, and the cloud requirements — represents a major leap forward in terms of network capacity and availability for consumers accompanied by a hefty maintenance and development price tag for providers.

In their Mobile Economy Report, GSMA projects that the number of connections on 5G networks will grow from 18% in 2023 to 56% by 2030. To accommodate the high-speed, low latency and massive connectivity promises of 5G, telecommunications organizations are turning to generative AI to expand and optimize their networks.

AI and geospatial data will revolutionize network planning.

Data analysis for network optimization is a longstanding industry challenge. The integration of native language models and geospatial data is set to transform network planning.

“Things previously were just too complex,” says Phil Kippen, Global Head of Telecommunications at Snowflake. “You were trying to run data sets and come to conclusions on very specific topics or questions, but you may not even be evaluating the right data sets. With gen AI, you’re able to make that interaction with data much easier.”

Now, he says, network operators can ask pointed questions, like where can I put capacity to maximize revenue or uptime or resilience?

Geospatial data adds another layer of insight so that telecommunications providers can fully understand network performance in different environments. With geospatial, vendors can see that they're experiencing network volatility in a specific area and analyze terrain, weather conditions, building density, and other factors that impact service quality. Providers will be better informed and better able to allocate resources to fix whatever problem is at hand.

With these two technologies at their disposal, telecommunications providers can create digital twins — virtual representations of network infrastructure — offering unprecedented insights into network performance and optimization opportunities.

Service providers will turn to open integration platforms.

As the industry navigates through ever-increasing complexity and speeds up on data-driven transformation requirements, it will look toward open integration as a means to reduce deployment effort and time to market. "There's always a push toward openness," Kippen says. "It's important because as we look forward, open platforms and open infrastructures dramatically reduce complexity, lowering the threshold so more players can enter the game. This will be better for end users and will have a significant impact on operational costs."

Open integrations create a common environment where various vendors can connect their solutions, breaking down silos between network components and vendors to make more flexible, efficient and cost-effective network solutions.

More telecom providers will sell subscriber insights to advertisers as a new revenue source.

In a bid to diversify revenue streams and replenish the huge investments required for 5G, more telecom providers will sell their troves of anonymized subscriber insights to advertisers.

"Providers are anonymizing and enriching subscriber insight data and selling it to customers who then advertise on their networks and to their customers," Kippen says. "They're monetizing data in multiple ways."

This would be a massive shift for telecommunications providers, reframing them in the data economy as both service providers and invaluable sources of consumer insights.

PUBLIC SECTOR

HIGH STANDARDS FOR DATA PRIVACY WHILE EMBRACING AI

It's well known that the public sector — federal civilian, national security and education — tends to lag when it comes to embracing digital transformation initiatives. But that reality may actually position the industry to benefit the most from generative AI. Given the ongoing shortage of IT and data science professionals in the public sector, AI's promise to help nontechnical staff work with data may be the light at the end of the tunnel.

The public sector faces particular challenges in adopting AI. Budget limitations make it hard to afford; lack of key technical staff makes it hard to implement; and the public-service mission makes assuring the right levels of privacy protection and transparency all the more critical.

> **It'll be all about data privacy and data residency for federal governments.**

In federal civilian government, conversations around AI focus largely on privacy and data residency. Winston Chang, Global Head of Public Sector at Snowflake, explains that, while it's three to five years in the making, "data ownership really is at the individual constituent level. We are starting to see models come out where I, as an individual, can decide to grant access to my data or not." While that will inevitably have massive implications for regulation around privacy, it also brings up the question of data residency.

The major leak of Social Security numbers in April was not a breach of a public agency, but of a private entity, National Public Data. Chang says that if the data had been understood as belonging to the individuals it identified rather than to the government, the breach would have been less likely.

"If we had a concept that that data is not theirs, that I allowed them to access that data," he says, "it's a wholly different security, policy, privacy and data residency framework."

This concept will bleed into public-based healthcare systems. Government officials in Canada, the U.K., the EU, and other single-payer systems increasingly recognize their roles as stewards of the data, not the owners of the data.

AI's biggest impact on national security will be drones and simulations.

In national security, one of the ongoing military tactics that will be hugely impacted by AI is the use of drones.

"That seems odd, because it's hardware, but there's so much software that plays into that," Chang says. Next year we'll see AI enable drones to operate more autonomously, improve data analysis for the vast amounts of information gathered by drones, and more. As gen AI continues to evolve, we can expect increasingly sophisticated applications of AI-powered drones.

When it comes to national security, one of the biggest challenges is preparing for the unknown. Simulation isn't new. But the support of the computing requirements and power training programs with AI certainly is. Before the widespread use of simulations, governments had to train workers with traditional methods.

"There was no way to actually run a training program on war, economic engagements, or cyber attacks on critical infrastructure," Chang says. "But you can run those sorts of things in simulation now."

And that synthetic data — representing economies, traffic patterns, weather systems and more — will be used to inform future national security decisions.

Edtech will need to tackle data silos.

Education will have challenges leveraging AI because it's "plagued with tons and tons of applications that are single-use applications, and they tend to buy all of them," Chang says. "So their data silo issue is tremendous."

Over the next year, we'll see edtech continue to chip away at that problem, but more importantly, start to coalesce all that data to help inform the industry's approach to larger problems, like total student lifecycle tracking.

Across the public sector, he expects gen AI's arrival to be delayed. "Government has a higher order of standard," he says, "because it is of the people and for the people."

But he notes that gen AI will act as a catalyst for data literacy and while the transformation may be slower, it should be hugely impactful.

RETAIL AND CONSUMER GOODS

BUILDING QUICKLY ON INCREMENTAL SUCCESS WITH GEN AI

Retailers' data strategy will have three primary elements in 2025: Securely collaborating internally and externally, operationalizing generative AI, and more deeply integrating traditional ML and deep learning, says Prabhath Nanisetty, Global Head of Industry, Retail Data and Technology at Snowflake.

"Gen AI will continue to be a focus area," Nanisetty says. "But we're also going to see more adoption of ML and data science across more use cases, in addition to an explosion of privacy-preserving technologies, as marketing becomes more central to a retailer's growth strategy."

Retailers continue to face significant supply chain challenges, disrupting inventory maintenance, on-time deliveries, operational costs and customer sentiment.

"From ships crossing the ocean to the trucks delivering to your distribution centers and to stores, there could be dozens of different companies involved, and if retailers don't establish ways to securely collaborate, there's an upper limit on the efficiency they can possibly achieve," Nanisetty says. "Further, driving data accessibility internally is critical to ensure that all systems have visibility into key data and metrics."

> **Frontline productivity will be a main focus as retailers get into gen AI.**

As retail leaders try to influence more positive customer sentiment, many are focused on in-store experiences. For these leaders, having the right items available in the right place can be the most important part of the experience, and AI and advanced technologies will significantly transform in-store inventory management.

Where it makes sense, frontline retail workers can focus on high-value tasks requiring human involvement, like customer engagement and assistance, promotional displays and store design.

Leading retailers will build quickly on incremental AI successes.

As retailers navigate which AI use cases are worth the investment, they should take into account three things — where good data exists, where decisions can be made quickly, and where impact can be quantified rapidly. Last year was a year of AI experimentation. Many retailers found themselves spending money on AI and not necessarily seeing the return. Nanisetty explains that retailers and brands weren't always as intentional as they could have been — many threw AI at the wall to see what would stick.

"Leaders will start to place bets where there is already good data that they have access to," Nanisetty says. "Where they've got completeness of that data in a workable format, and a use case that is connected to the business's priorities. If those three things are true, it's good to test and publicize the success there."

A lot of disappointment around the ROI of generative AI is a data strategy problem, not an AI problem, he says. As retailers improve their data foundations and align AI strategy with key growth levers, frustration should ease.

"Over the next three years, cooler heads will prevail," he says. "What we want is sustainable growth using the right technologies. Generative AI is one tool in a bag that also includes machine learning, data science, data collaboration and better data."

He adds that a solid strategy should take into account the people and processes involved including upskilling employees and removing unwarranted barriers to data accessibility.

CONTRIBUTORS

MONA ATTARIYAN

Mona is Snowflake's Director of Machine Learning, and holds a Ph.D. in computer science from the University of Michigan. Previously, she served as Head of Analytics and Machine Learning at Redfin and held various technical roles at Google.

SUNNY BEDI

Sunny is the Chief Information Officer and Chief Data Officer at Snowflake. Before that, he spent over a decade at NVIDIA, where he was VP of Corporate IT and Head of Global IT, among other roles. Sunny also serves on the board of DevSecOps company GitLab.

JENNIFER BELISSENT

Jennifer is Principal Data Strategist at Snowflake. She spent over a decade at Forrester Research as an internationally recognized expert in data sharing, the data economy and data leadership, including best practices for building world-class data organizations. She has an M.A. and Ph.D. in political science and organizational theory from Stanford.

WINSTON CHANG

Winston serves as Snowflake's Global Head of Public Sector. He supports global government and education ecosystems for modernizing data practices. He is an expert in organizational transformation derived from data, AI/ML and innovation. His personal mission is to help government and educational institutions leverage data for maximum societal impact.

JESSE CUGLIOTTA

Jesse is the Global Head of Healthcare and Life Sciences at Snowflake. He has 20 years of experience working with clients in the healthcare and life sciences industries with a specific focus on driving new capabilities with data. He holds a graduate degree in engineering from the University of Pennsylvania and has spent his entire career in data and analytics.

BENOIT DAGEVILLE

Benoit is the Co-Founder of Snowflake, and today serves as the company's President of Product. Before co-founding Snowflake in 2012, Benoit spent more than 16 years as a data architect and developer at Oracle.

BARIS GULTEKIN

As Head of AI, Baris leads AI and ML initiatives and drives Snowflake's AI product roadmap and strategy. He joined Snowflake following its acquisition of nxyz in August 2023, where he was co-founder and CEO. Before that, he spent over a decade at Google across various product leadership roles.

YUXIONG HE

Yuxiong is a Distinguished AI Engineer at Snowflake, spearheading LLM development. At Microsoft, Yuxiong co-founded and led the groundbreaking DeepSpeed project. She has published more than 100 papers in major computer science conferences and journals, and her research is widely applied in diverse systems and products.

JEFF HOLLAN

Jeff is Head of Applications and Developer Platform, leading the product vision and strategy for Snowpark, Ecosystem and Developer platforms. Jeff was Head of Product for Microsoft Azure's PaaS and Serverless portfolio, and co-created and maintains KEDA, a CNCF graduated project for serverless container scaling in Kubernetes.

BRAD JONES

Brad is the Chief Information Security Officer at Snowflake and VP of Information Security. Prior to Snowflake, he spent more than six years as CISO and VP of Information Security at Seagate. He has actively participated in a number of customer advisory boards and is currently part of the CISO Advisor Council at NightDragon.

AMANDA KELLY

Amanda is Director of Product Experiences and Co-Founder of Streamlit, which was acquired by Snowflake in 2022. Amanda has also led product and operations in autonomous vehicles at Zoox and for several stealthy Google X projects around Google's natural language processing capabilities. She holds an MBA from Stanford.

PHIL KIPPEN

Phil Kipen is Global Head of Telecommunications at Snowflake. Phil has over 20 years of experience working with more than 200 wireless and wireline telecommunications services providers globally, advising on transformation and technology evolution strategies, driving customer and industry thought leadership, transforming network and cloud services delivery architectures, and designing new, innovative enterprise and subscriber services.

CHRISTIAN KLEINERMAN

As Snowflake's EVP of Product, Christian brings more than two decades of experience working for the world's largest enterprise tech companies. He worked at Google as YouTube's Senior Director of Product Management and held a number of management roles at Microsoft.

TIM LONG

Tim is Global Head of Manufacturing at Snowflake and a technical leader with education and experience practicing advanced analytical methods including ML, data mining, optimization, simulation, natural language processing, small and big data warehousing, and the creation of compelling data visualizations.

PRABHATH NANISETTY

Prabhath has more than 20 years of experience in product management and analytics, working with technology startups and large consumer brands. He is currently the Global Head of Industry, Retail Data and Technology, where he advises retail and consumer goods companies in leading product strategies and data-driven decision-making.

RINESH PATEL

Rinesh is Global Head of Financial Services at Snowflake. With over 20 years of experience in financial services, he has held several leadership roles across enterprise data and analytics, consultancy and investment banking. He's based in London where he lives with his wife and two daughters.

VIVEK RAGHUNATHAN

Vivek is SVP of Engineering at Snowflake. He co-founded Neeva (acquired by Snowflake in 2023) where he pioneered the development of a next-generation AI search engine for consumers. He also spent over a decade at Google as a VP of engineering in various technical leadership roles.

SRIDHAR RAMASWAMY

Sridhar was Co-Founder of Neeva, acquired in 2023 by Snowflake. Not quite a year later, he was named Snowflake CEO. He spent more than 15 years at Google, where he started as a software engineer and rose to SVP of Ads and Commerce. Sridhar earned a Ph.D. in computer science from Brown University.

ANOOSH SABOORI

Anoosh is Snowflake's Head of Product Security. He holds a Ph.D. in computer and information systems security/information assurance from the University of Illinois. Prior to Snowflake, he helped lead Google Cloud's zero trust portfolio and led the Windows identity and access management team at Microsoft.

BILL STRATTON

Bill is the Global Head of Media, Entertainment and Advertising at Snowflake. He works with major media companies, ad tech, measurement and identity companies, and sports properties as well as major advertisers to leverage Snowflake to execute all mission-critical data initiatives.

Snowflake makes enterprise AI easy, efficient and trusted. Thousands of companies around the globe, including hundreds of the world's largest, use Snowflake's AI Data Cloud to share data, build applications, and power their business with AI. The era of enterprise AI is here.

Learn more at **snowflake.com** (NYSE: SNOW)