IBM Institute for Business Value | Research Insights

Winning the silicon race

Three strategies to secure AI advantage

IBM
semi

How IBM can help

AI strategy is at the core of every industry today. Securing the computing power enabled by competitive semiconductors is essential for AI resiliency. IBM helps increase supply chain transparency, transfer knowledge, and build critical skills by applying AI and hybrid cloud infrastructure. For semiconductor operations, IBM uses AI to improve product development efficiency and optimize manufacturing. By combining advanced data analytics and predictive maintenance, IBM helps boost yields and ensure reliable operations.

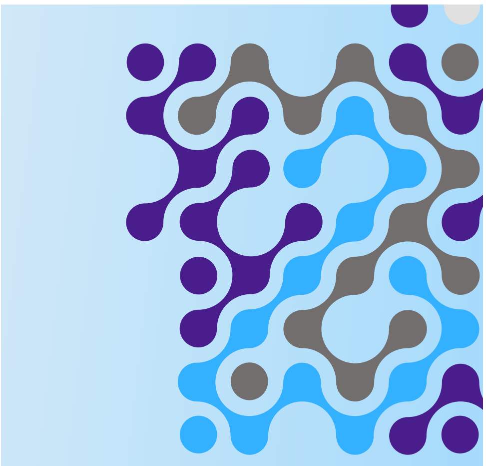
For more information, please visit:
https://www.ibm.com/consulting

How SEMI can help

As AI demand surges and supply chains strain, SEMI provides the platform for industry resilience. Our member-led initiatives address critical priorities: Smart Data & AI, Supply Chain Resilience, Workforce Development, and Sustainability. Over 30 technical communities—from Advanced Packaging to Fab Owners Association—enable the expertise sharing essential for next-gen technologies.

SEMI's global events and specialized conferences forge the supplier-buyer partnerships this report identifies as crucial. We deliver market intelligence, standards development, and proof-of-concept projects that accelerate innovation. With 84% of executives depending on government incentives, SEMI shapes supportive policy frameworks.

Through "Connect, Collaborate, Innovate," SEMI helps members diversify sources, build regional ecosystems, and maintain the innovation velocity AI advancement demands.



Key takeaways

Chip suppliers are struggling to keep up with intensifying AI demand.

Demand for AI accelerator chips is expected to grow 50% to 70% by 2028. 83% of buyers say they've already experienced supply issues—and new industrial uses will drive up demand even more.

Chip buyers are prioritizing regional sources.

About 80% of executives say it's essential to have access to local AI chips, AI talent pools, and AI platforms. But the transition is not happening as quickly as buyers and suppliers would like.

Energy efficiency will be a top priority as AI matures.

82% of chip buyers are pursuing specialized chips for specific tasks to optimize energy usage. Nearly 90% of suppliers anticipate increased demand for custom systems on a chip (SoCs) and chiplets that balance performance, cost, and energy efficiency.

Next-gen tech will help optimize chip performance.

88% of chip suppliers expect alternative computing tech such as photonic, neuromorphic, and quantum to emerge within three years to support AI demands.

Contents

Introduction

Success in a scarcity market

Compute power is becoming the world's most valuable commodity—and AI has sent demand skyrocketing.

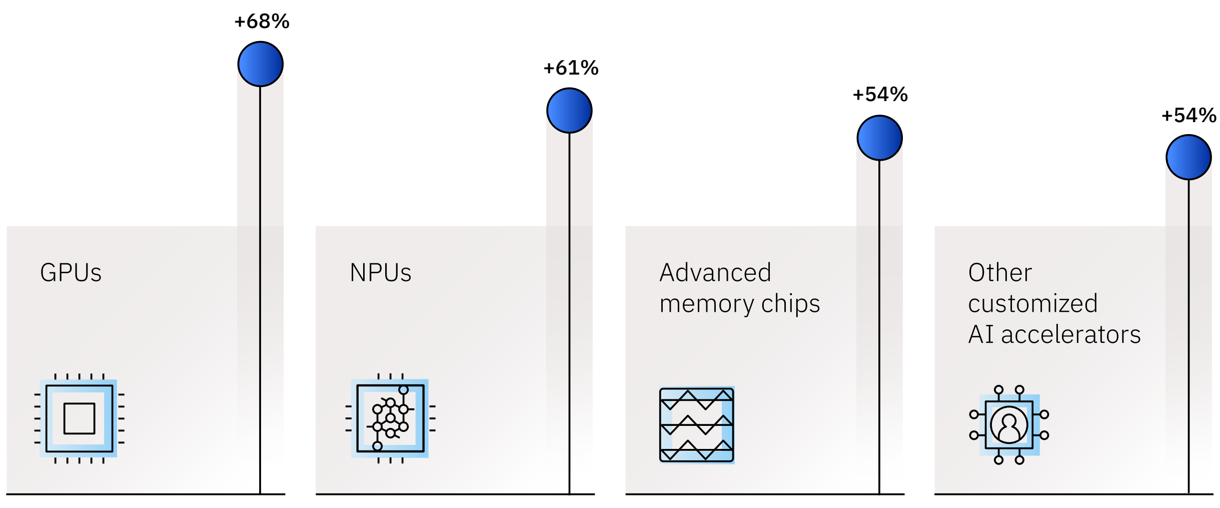
As companies race to capture AI-powered business opportunities, their ambitions rely on nanoscale semiconductors. Many organizations have integrated advanced AI chips into their products and services. Yet global production and resource bottlenecks mean supply cannot meet demand.

How is this uncertain semiconductor market impacting chip users and chip sellers? In exclusive, proprietary research, the IBM Institute for Business Value (IBM IBV), in partnership with SEMI, surveyed 800 executives from semiconductor buyers and suppliers in May 2025 (see "Research methodology" on page 25). Our findings indicate that the already constrained AI chip market is about to get even tighter.

Demand for the most advanced semiconductors—AI accelerator chips such as GPUs, NPUs, TPUs, and ASICs—will grow 50% to 70% by 2028, according to the chip buyers we surveyed (see Figure 1). By 2028, 62% say high-performance AI computing infrastructure will be central to competitive advantage.

FIGURE 1

Demand for advanced AI chips is expected to surge over the next three years



Q. By how much do you expect the demand for semiconductors to evolve (increase/decrease) for your organization in the next three years?

Percentages include only responses from chip buyers.

See Appendix on page 23 for full glossary of chip types.

So far, most of the demand has come from applications in the virtual world, such as large language model (LLM) training and mobile applications. As AI matures and moves deeper into physical domains—including autonomous cars, intelligent manufacturing equipment, robotics, and medical devices—the scale of demand will change entirely. We'll see greater need for domain-specific AI training and embedded AI inference at the edge.

Already, chip supply chain players—including fabless and foundry, integrated device manufacturers, machinery and materials suppliers, and electronic design automation and intellectual property vendors—are struggling to keep up. Some 83% of buyers say they've experienced supply issues with AI accelerator chips. An increased business appetite for agentic AI and other AI applications will exacerbate the situation.

Chip buyers are in a bind, with 75% saying dependence on a narrow pool of semiconductor vendors is a major strategic challenge. Today, only a few companies can design and fabricate the latest high-performance AI chips, including AI training GPUs and highly integrated smartphone chips. Essential machinery, materials, and tools are also produced by a small number of firms. New players are entering the market, often with government backing—but skills and technical limitations still block true competition.

The solution, as our research reveals, will require creativity and fresh approaches. In each of the sections that follow, we identify strategies that can enable organizations to win in a highly volatile AI chip market. In part one, we explore how remaking your relationship with chip suppliers right now can move you up on the AI chip delivery list. Part two delves into building AI resilience by boosting local capability. In part three, we explain how the advanced AI of tomorrow could spur the development of more specialized chips that can strike a better balance between performance, cost, and energy consumption.

These strategies mature at different rates. To build resilience, organizations must begin all three now.

Part
1
Become an irreplaceable partner

Part
2
Boost AI resilience with regional ecosystems

Part
3
Max out the next wave of AI chips

> "AI, powered by accelerated computing for both training and inference, will transform every industry like the internet and mobile phones did—reaching from cloud to edge, enabling rapid model development, and impacting every market globally."

Director
Nvidia

Part one: Become an irreplaceable partner

In a seller's market, buyers have to compete for a favored position. That starts with understanding the challenges their suppliers face—and how they might be able to help.

The suppliers we surveyed face no shortage of obstacles. More than 80% point to four key struggles: geopolitical tensions, investment constraints, tech skills shortages, and limitations in supply chain technologies (see Figure 2). For chip buyers, these are areas of opportunity to deepen and improve their relationships with chip supply chain players.

Geopolitically, chip companies are building new factories on different continents in response to changing trade environments and potential military conflicts—but the lead time to start production and transfer knowledge exceeds two years. Financially, 81% of chip suppliers say investment needed for new chip technologies is out of reach; 74% cite limited access to capital.

Chip companies report bottlenecks for key tech, such as high bandwidth memory and chip-on-wafer substrates. Meanwhile, 61% say they have a shortage of supply chain management skills, and only half say they have the requisite AI integration skills (see Figure 3).

> "Pursuing cutting-edge technologies requires continuous and substantial investment."
>
> **Toshihiro Awa**
> President
> SUMCO

FIGURE 2

Chip suppliers face hurdles on many fronts

Investment requirements	Geopolitical tensions	Supply side technologies	Skills
81%	85%	83%	84%

Q: How significant are the following challenges for the semiconductor industry in meeting market demands and expectations? Percentages include responses of "moderate challenge" and "significant challenge."

Chip buyers can step up and help ease these challenges. Already, 80% of chip buyers say they're strengthening strategic alliances with chip companies—and 84% are looking to form new alliances.

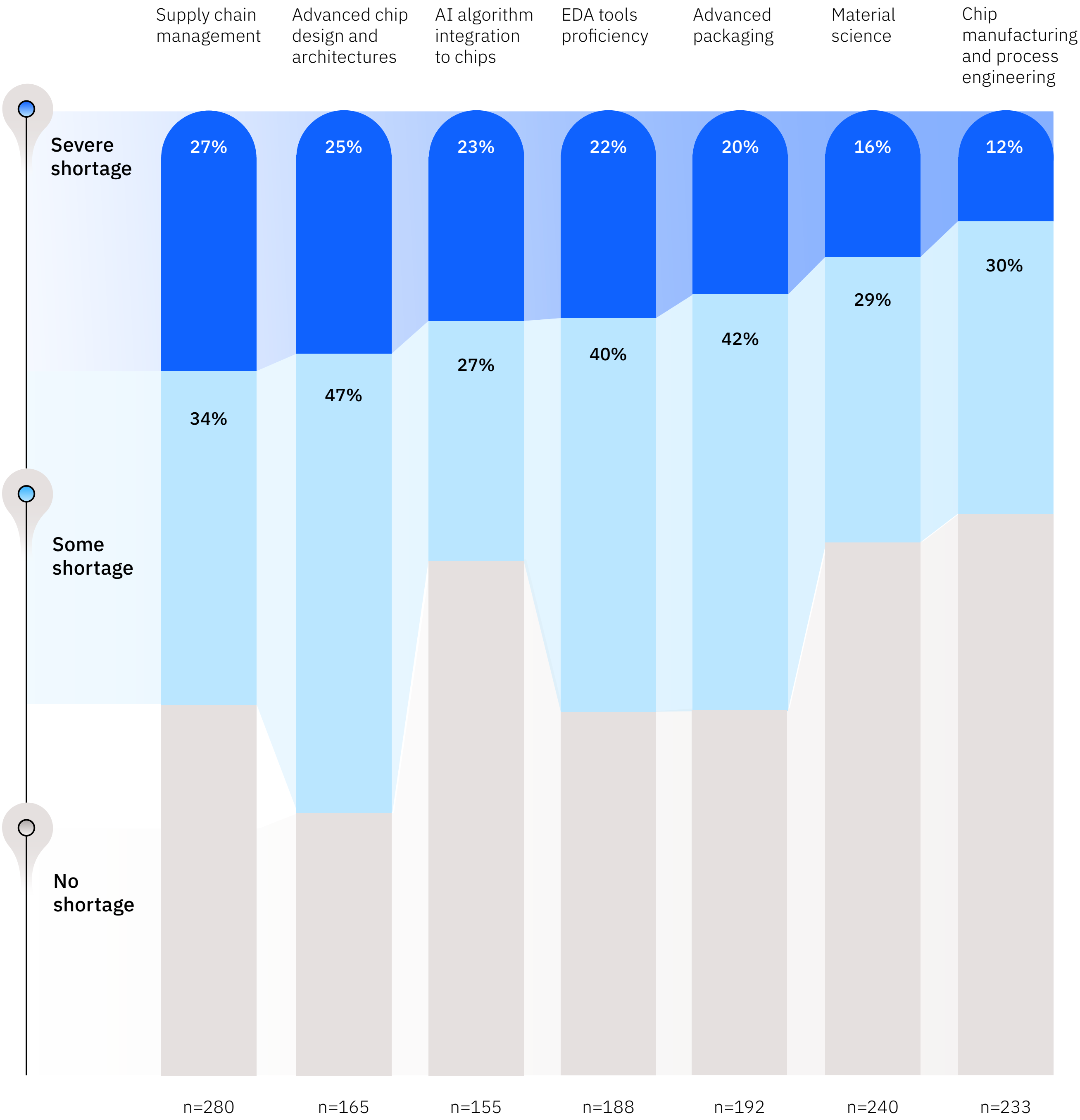
Because chip buyers own the use cases that drive chip design, they can offer insights into the technical requirements to optimize semiconductor specifications. As more advanced AI demands more specialized chips, buyers can play a more active role in chip development.

"The lack of affordable alternatives to high-cost manufacturing equipment poses a challenge for all players. Our goal is to present a viable, cost-effective solution."

Hiroaki Takeishi
Senior Managing Director
Head of Industrial Group
Canon

FIGURE 3

Chip suppliers are struggling to address critical skills gaps



Q. Which critical semiconductor skills are lacking at your organization? Percentages indicate proportion of executives that identified a skill as lacking in their organization.

Through deeper engagement and longer-term purchase commitments, buyers can become preferred partners. For example, during a crisis, suppliers may prioritize customers with multiyear commitments over those with shorter-term agreements.

Partnerships may involve building local support for new chip operations, joint research and development, or funding for new technologies. For example, buyers could leverage their supply chain management expertise and networks to aid chip-making partners. They can also share technical expertise: 65% of chip buyers are actively expanding in-house AI chip design capabilities.

By creatively bringing resources to the table, chip buyers can become indispensable to their semiconductor suppliers—and put themselves in a better position to access chips now and make their long-term AI strategy more resilient.

What to do

Make chip partnerships a priority. Invest to acquire the semiconductor expertise needed to collaborate more effectively with chip suppliers. Learn about industry bottlenecks and pursue alternative partnership approaches to offer expertise chip suppliers might lack. Identify creative ways to acquire diverse, but valuable talent in a scarce semiconductor skills market, such as developing AI agents to fill talent gaps.

Own product requirements and get involved in chip design. Create clearly defined and powerful use cases that help suppliers optimize chip architecture for cost, performance, and energy consumption. Be a predictable business partner to help chip partners better navigate the silicon cycle.

Invest in supply chain transparency to manage business disruption risks. Diversify suppliers and improve inventory management. Conduct AI-assisted scenario simulations. Develop multiple backup strategies to navigate uncertainties. Ensure you have traceability down to material level to maintain regulatory compliance and social responsibility.

"Supply chain transparency is critical in equipment operations. Both solid-state drives (SSD) and semiconductors are composed of numerous parts and materials, and the production lines will stop if even a single component is missing."

Koichiro Shibayama
CEO
KIOXIA Iwate

Part two: Boost AI resilience with regional ecosystems

Today, just a few regions are able to supply the most in-demand AI chips and supporting key technologies. This creates supply chain choke points that are difficult to work around when disruption strikes. In this environment, 91% of chip buyers say they're working to diversify suppliers.

Geographic diversity is the top priority. Roughly 80% of executives surveyed say access to locally produced AI chips and accelerators is essential for their region. Similarly, 80% cite essential access to regional AI talent pools, and 77% to AI platforms. This type of regional sourcing can help increase ecosystem-wide innovation and reduce reliance on global supply chains, which can be more vulnerable to disruptions and tariffs.

So far, progress has been slow. By 2028, executives expect regional AI resources to address about half of regional demand in most areas of concern (see Figure 4). The local sourcing projections for AI chips and data center hardware are among the lowest, despite the fact that companies consider high-performance AI computing infrastructure as a top competitive differentiator. This highlights a critical supply chain vulnerability. Without regional access, geopolitical turmoil puts access to advanced AI capabilities at greater risk.

Our analysis shows that effective financial strategy is a key enabler that can accelerate local sourcing for AI computing hardware (see "Research methodology" on page 25). Because data centers and AI chips require substantial capital investment, early financial commitments will be critical.

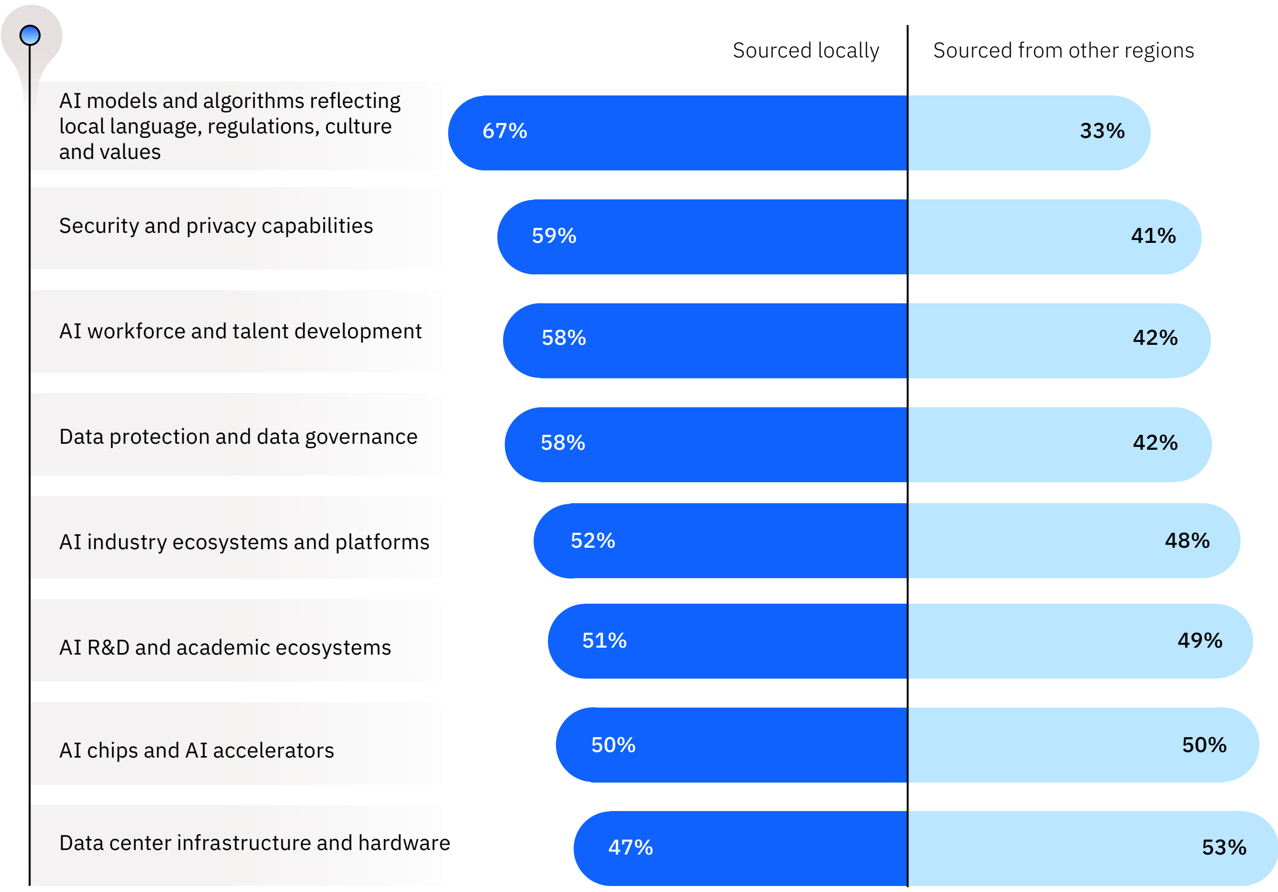
To put the financial demands of semiconductor operations in context, the average payback period for the development of a chip fabrication plant is 75 months, but for a new-generation AI accelerator chip, it's 41 months. Similarly, it's 24 months for a new data center and 22 months for a new smartphone.

> *"Geopolitical factors can abruptly disrupt business operations, posing a significant risk to companies."*

Masato Goto
President and CEO
SCREEN Holdings

FIGURE 4

Regional AI resources are critical for AI resilience by 2028



Q. What percentage of your region's AI needs do you expect to be met through local sourcing/capabilities in three years?

The discrepancy in investment time horizons makes managing silicon cycles difficult. Buyers want to adjust orders based on shifting market conditions, but chip suppliers require long lead times and planning cycles to hit profit-earning production targets. This creates a timing mismatch. Synching and aligning financial expectations are essential to make partnerships successful.

An example: Tesla and Samsung have recently partnered to share the financial burden of bringing a new advanced AI chip fabrication facility to the US. The $16.5 billion, eight-year deal stipulates that Samsung will manufacture Tesla's proprietary A16 chip at the new Samsung plant currently under construction in Texas.[1] Other organizations are also likely to benefit from the US-based capacity provided by the new facility. Apple has announced that it will source chips from the Texas fab as part of its push to increase US production—and avoid potential tariffs.[2]

Not every company has Tesla's resources. Organizations may need to build coalitions of partners willing to contribute capital, including industry associations interested in regional chip supply chains. Bringing along government agencies is another key lever: 84% of semiconductor executives say government incentives play a crucial role in their organization's growth and competitiveness.

> "Before, very complex phases were executed only in California, but now the process has become more global."
>
> **Venkat Mattela**
> Founder and CEO
> Ceremorphic

Organizations must consider how AI-driven innovation can unlock new revenue streams—which help sustain the growth of regional semiconductor hubs. As revenues grow, organizations can invest more capital in diversifying chip supply chains. Homegrown capabilities take time to mature, but they have the potential to disrupt the status quo and boost resilience across regional ecosystems.

What to do

Seek out regional partners with shared AI ambitions. Identify diverse funding sources to support the development of regional chip supply chains, including government grants, industry investments, and private sector partnerships. Lean into AI-enabled revenue streams that add value to your ecosystem.

Realign sourcing, market, and partnership strategies to new global realities. Improve or acquire your own global intelligence capabilities. Prepare for unexpected business disruptions by strategically reallocating investments. Use AI to make scenario planning more accurate and comprehensive.

Capture local AI demand with competitive AI computing capabilities. Develop compelling, strong AI business cases supported by vertically integrated AI capabilities and your proprietary data. Tap new AI revenue pools to nurture local AI ecosystems and collectively cross the minimum ROI threshold of the capital-intensive chip industry.

> "If semiconductor production moves to higher cost countries, we need to reduce labor intensive processes. Robotics and automation will play critical roles."

Tsunetoshi Oba
Executive Officer
Industrial Automation Company
OMRON

Perspective

A virtuous cycle: How AI can optimize semiconductor operations

The complexity and high costs associated with designing and fabricating cutting-edge AI chips stem from the need for specialized knowledge that leading companies have cultivated over decades. This industry is renowned for its continuous efforts to improve operations through proprietary methods, which are hard to replicate beyond one's corporate and ecosystem context. These methodologies are the accumulation of countless minor enhancements, resulting in a competitive advantage or "moat."

Historically, the acquisition of expertise has been undocumented, passed between people within the organization. This analog process is in sharp contrast to the high-tech nature of the rest of the industry. However, semiconductor businesses are now using AI to refine their internal operations and improve knowledge transfer across semiconductor design, manufacturing, yield and quality control, supply chain, and machine maintenance—a top priority for organizations looking to stand up operations in new regions. AI helps gather past records, analyze them, and support or train workers so that valuable corporate know-how is passed on to the next generation.

Semiconductor suppliers anticipate that AI will reduce operating costs by 17% over the next three years and contribute to a 32% increase in revenue. Key focus areas include AI-powered advanced packaging and assembly optimization, where 73% of suppliers see potential for a 21% productivity improvement.

> "The use of AI-driven simulation is extremely important for material development under tight time constraints."
>
> **Toru Kimura**
> Senior Officer, Electronic Materials Business
> JSR

Semiconductor R&D, particularly for advanced node chips, is another area AI can add value. With 72% of suppliers reporting challenges in finding skilled talent for advanced chip design, it's unsurprising that 72% are turning to AI to optimize electronic design automation (EDA), with an expected productivity gain of 26%.

> "We have set an ambitious goal of doubling the productivity of our facilities through advanced solutions such as digital transformation."

Kazunori Tsukada
CEO
KOKUSAI ELECTRIC

For example, Taiwan Semiconductor, a leading fab, cites "Fab Runs On Code" as a secret to its competitive advantage. This approach leverages DevOps, cloud, and AI to improve yield and performance.[3]

Companies are exploring every avenue to enhance semiconductor operations using AI, from yield analysis to visual inspection and preventive maintenance to optimizing supply chain and logistics. Those that successfully leverage AI to streamline operations will be better positioned to survive and meet the surge in AI chip demand—which is only just beginning.

> "Improving yield through AI is a critical objective. Even a 0.1% increase in yield can have a significant impact on the business."

Toshihiro Awa
President
SUMCO

Part three:
Max out the next wave of AI chips

Organizations prefer to control their own fate, particularly when it comes to essential resources. For in-demand computer chips, that isn't often possible. But prospects will improve in the years ahead—and, as our research reveals, the seeds of change are already sprouting.

The key data point: 83% of chip buyers say technology advancement is included in their product strategy. These organizations need innovation to continue at a dizzying pace to remain competitive—and they're looking at all options on the table.

Specialization is one way chip suppliers are keeping up, as 82% of chip buyers pursue specialized chips for specific tasks to optimize energy usage. By focusing on specific tasks, rather than general purpose computing, they can better balance performance, cost, and energy efficiency as use cases become more sophisticated.

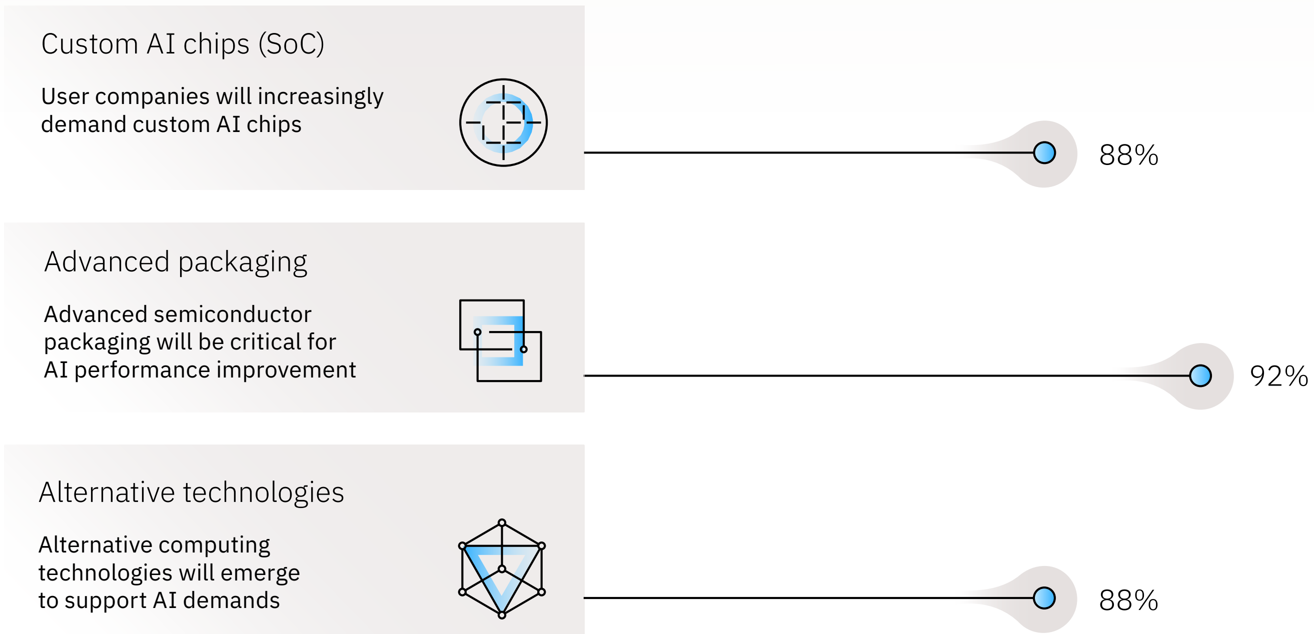
Nearly 90% of chip suppliers say their buyers will increasingly demand custom AI chips—primarily custom systems on chips (SoCs), which provide a complete computing system on a single chip, instead of separate chips or circuit boards. And 92% say advanced packaging technologies, which are key enablers of chiplets advancement, will be critical (see Figure 5). That is a dramatic acceleration compared with today, when 24% of AI accelerators and 21% of SoCs are customized.

> "I believe that demand for application-specific, multi-variety AI chips will grow rapidly over the next few years and will surge starting around 2027 or 2028."

Atsuyoshi Koike
CEO
Rapidus

FIGURE 5

Executives say three trends promise to transform the global chip market by 2028



Q. To what degree do you agree with with the following statements in three years (by 2028)?

Percentages indicate proportion of executives that responded "agree" or "strongly agree."

> “Customizing a base chip for custom SoC may take less than two years, while building from scratch takes longer. Using chiplets and standard interfaces can significantly shorten development time.”

Takao Fukuda
Senior Director, Product Planning, Strategy Group
Astemo

SoC will be the preferred choice for the near term, but organizations are increasingly looking at chiplets as packaging technologies advance. Chiplets offer a number of benefits, including improved performance, reduced power consumption, and increased design flexibility. And while SoC’s monolithic architecture requires the most advanced node for the entire system, that level of processing power may be overkill in some cases. Chiplets allow for more customized approaches and enable better optimization.

In this environment, suppliers must strike the right balance between performance, cost, and energy consumption for every application, for every customer (see “Solving AI’s energy problem” on page 22). Suppliers are employing a range of strategies, including new chip designs, new manufacturing techniques—and breakthrough technologies.

Within the next three years, 88% of chip suppliers expect alternative computing technologies to emerge to support AI demands. This includes advanced packaging technologies to support chiplet advancements, such as optical and photonic computing and 3D stacked architectures.

Then there are the breakthrough capabilities (see Figure 6). Quantum AI, which leverages the principles of quantum mechanics to enhance machine learning algorithms, could enable organizations to address complex problems that have been out of reach with traditional computing technologies. There's high hope for artificial general intelligence (AGI)—AI that can understand, learn, and apply knowledge across a wide range of tasks. If realized, AGI would create new possibilities for applications such as robotics, natural language processing, and computer vision.

> "There certainly is a demand for high-performance SoCs but we need to balance performance with affordability. We often compromise on specs when costs exceed what customers are willing to pay."

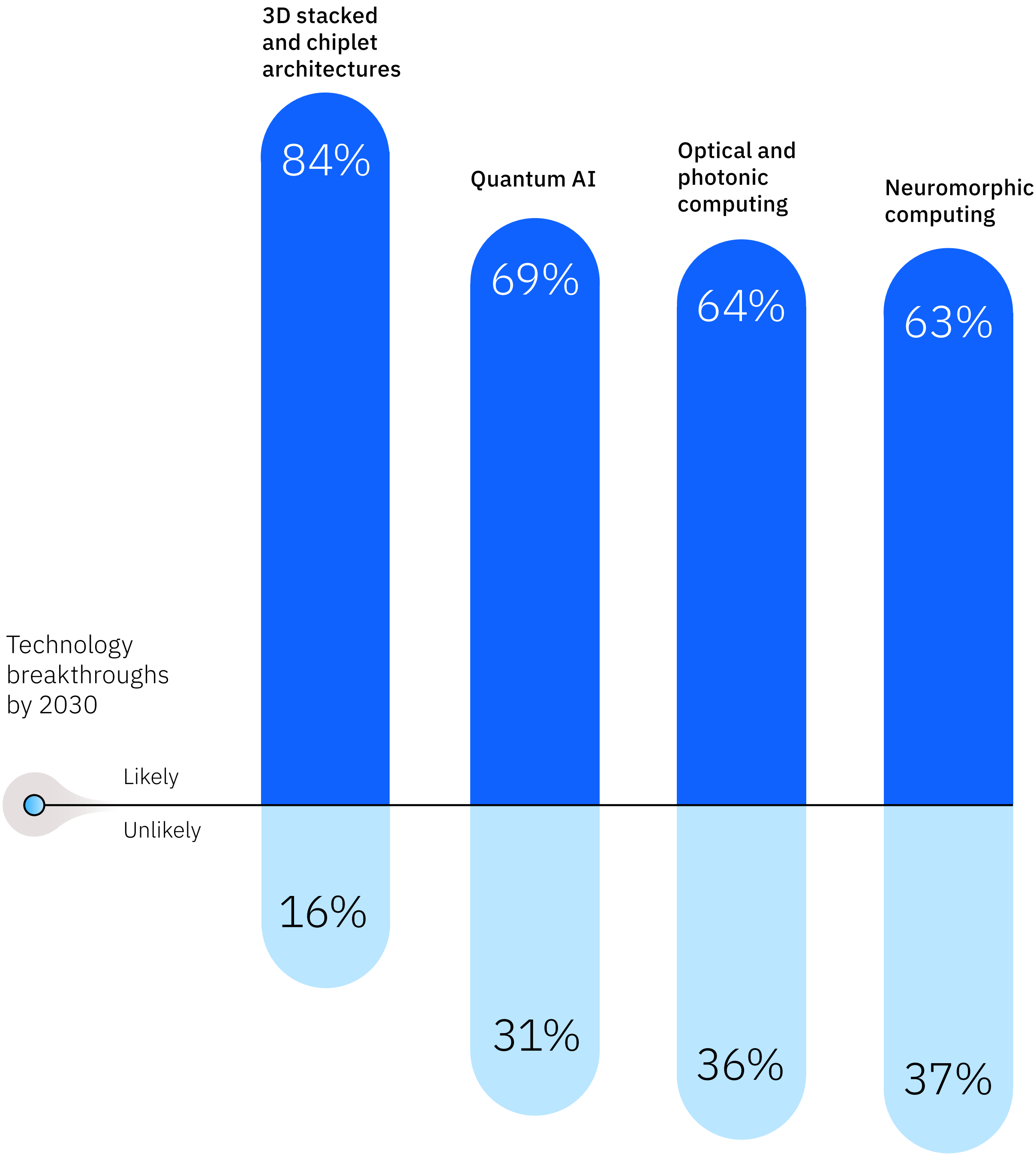
Honggul Jun
VP, Automotive Head Unit
LG Electronics

However, such breakthroughs depend on developments in semiconductor technologies themselves. 83% of supply side semiconductor executives say the feasibility of AGI is highly or critically dependent on semiconductor advancements.

These technologies, still in the early stages of development, may be closer to reality than it seems. Organizations that can evolve at the pace of the market will be in position to capitalize on these technologies when they come of age—and gain a lucrative advantage.

FIGURE 6

Next-gen tech could help address AI demands



What to do

Find where customization can deliver a competitive advantage. Assess the AI capabilities that benefit most from custom components. Weigh those capabilities in the context of whether they will deliver differentiating outcomes. Pursue alternative technology approaches with strategic partnerships, such as advanced packaging chiplets, to gain AI competitiveness.

Test the water early. Identify high-potential use cases for emerging technologies and volunteer to test or provide feedback on solutions as they evolve. Identify innovative ways to collaborate, such as sharing simulation environments or co-investing in pooled testing facilities, to share costs, reduce risk, and accelerate learning.

Imagine AI use cases that aren't possible today. Anticipate the emergence of breakthrough technologies and prepare for faster deployments. Equip your workforce with the most advanced AI and delegate repetitive tasks to AI agents to unleash your employees' creativity.

> "Drug discovery demands advanced supercomputing and AI. We're developing custom silicon chips—using the same hardware but optimized with specialized AI software and analog circuits—to accelerate this process."

Venkat Mattela
Founder and CEO
Ceremorphic

Perspective

Solving AI’s energy problem

The world is racing to train AI models using massive amounts of data, but this comes with a significant energy cost. AI training consumes an enormous amount of electricity, and it is increasingly common for new AI training data centers to be built alongside new power generation plants. According to the International Energy Agency (IEA), global electricity consumption by data centers is expected to more than double by 2030, reaching levels comparable to Japan’s current annual electricity use.[4]

> “To reduce energy consumption, we are working on specialized AI chips and 2nm processes, cutting power by roughly 75%. “
>
> **Atsuyoshi Koike**
> CEO
> Rapidus

Our survey shows that while 57% of data center providers are satisfied with chipmakers’ efforts to improve energy efficiency, only 40% of suppliers are satisfied with their own progress. This may be because the immediate demand for AI is heavily focused on training large models. Investments in AI training are expected to peak around 2028, after which the focus will shift toward AI inference by 2030.[5]

While the current competitive race to train AI models has pushed energy efficiency concerns to the sidelines, the problem remains. And as AI demands shift toward inference and deployment on physical edge devices, energy efficiency will become a critical issue.

Both chip buyers and suppliers are taking steps to tackle this challenge from different angles. For instance, 83% of data center and high-performance computing manufacturers are developing specialized AI processors optimized for specific tasks, while 78% of industrial users are adopting smaller AI models. Additionally, 70% of semiconductor machinery and materials suppliers are working on energy-efficient fabrication processes, with 65% expecting that AI will accelerate these efforts.

> “It’s not just the chip that changes, but the software in the chip becomes more optimized and more efficient. And we’re continually improving that.”
>
> **Director**
> Nvidia

Appendix

Glossary of chip types and use cases

AI depends on different types of semiconductors for different purposes. During the training phase, AI relies on powerful logic and memory chips to absorb massive amounts of data and adjust its internal parameters millions—or billions—of times to improve the accuracy of its outputs. During inference, when the trained model performs desired tasks, AI needs more specialized chips to optimize its performance. Use this glossary to help determine which chip type is best suited for different AI needs and industry-specific use cases.

	Role in AI development and usage	Primary use cases
Logic chips		
GPU (Graphics processing unit)	Dominant in training but also used in inference	Data centers
NPU (Neural processing unit)	Primarily inference	Smartphones and edge devices
ASIC (Application specific integrated circuit)	Training and inference, specific to workload	AI services provided by hyperscalers
TPU (Tensor processing unit)	Training and inference	Google Cloud AI
AI SoC (System on a chip)	Primarily inference	Full chip integration for smartphones and edge devices
Memory chips		
DRAM (Dynamic random-access memory)	Plays a support role in training and inference	Feeds AI accelerators
HBM (High bandwidth memory)	Critical for training, support role for inference	Training frontier AI models

About the authors

Rami Ahola, PhD

Global Industry Leader, Industrial Manufacturing
IBM Consulting
linkedin.com/in/ramiahola/

Rami has over 30 years of experience with semiconductor, telecommunications, and consumer electronics industries. He combines deep industry knowledge with first-hand experience in R&D, manufacturing, supply chain, and marketing processes to help clients with business strategies, processes, and execution.

Pushkar Apte, PhD

Strategic Technology Advisor
SEMI
linkedin.com/in/pushkar-apte-5363/

Pushkar is focused on the intersection of semiconductors and AI, specifically on two critical areas: building energy-efficient AI systems and applying AI-powered digital twins to mission-critical applications.

Stephen Pierce

Partner, Discrete Manufacturing 4.0, Industrial Manufacturing
IBM Consulting
linkedin.com/in/spierceibm/

Stephen has 35 years of experience in driving strategic growth agendas with leading global high-tech manufacturers. He leads IBM Consulting's manufacturing transformation solutions and services for autonomous manufacturing, including AI-enabled real-time production control, predictive asset management, intelligent product quality, circular supply chain, and IT-OT integration.

Noriko Suzuki

Electronics and Automotive Industry Leader
IBM Institute for Business Value
linkedin.com/in/norikosuzuki/

Noriko is responsible for developing thought leadership for the automotive, electronics, and energy industries. She has more than 20 years of experience working with global manufacturing customers on technology strategies and implementation. Her expertise includes Industry 4.0, digital transformation of operations, mobility solutions, and sustainable transportation.

Contributors

Mukesh Khare
General Manager
IBM Semiconductors

Noboru Itoh
VP, Japan Semiconductors
IBM Technology

Masahito Terakado
Japan Semiconductor Leader
IBM Consulting

Norishige Morimoto
VP, Japan
IBM Research

Gaurav Kapoor
US Lead Client Partner, Semiconductors
IBM Consulting

Michael Kuo
Taiwan TSMC Lead
Client Partner
IBM Consulting

Eric Yang
China Electronics Leader
IBM Consulting

Yoshihiro Shimada
Japan Electronics Leader
IBM Consulting

Yoshifumi Sakamoto
DE, Edge Computing
IBM Research

Yuji Tomita
SiView
IBM Research

IBM Institute for Business Value

Sara Aboulhosn, Associate Creative Director

Angela Finley, Design Lead

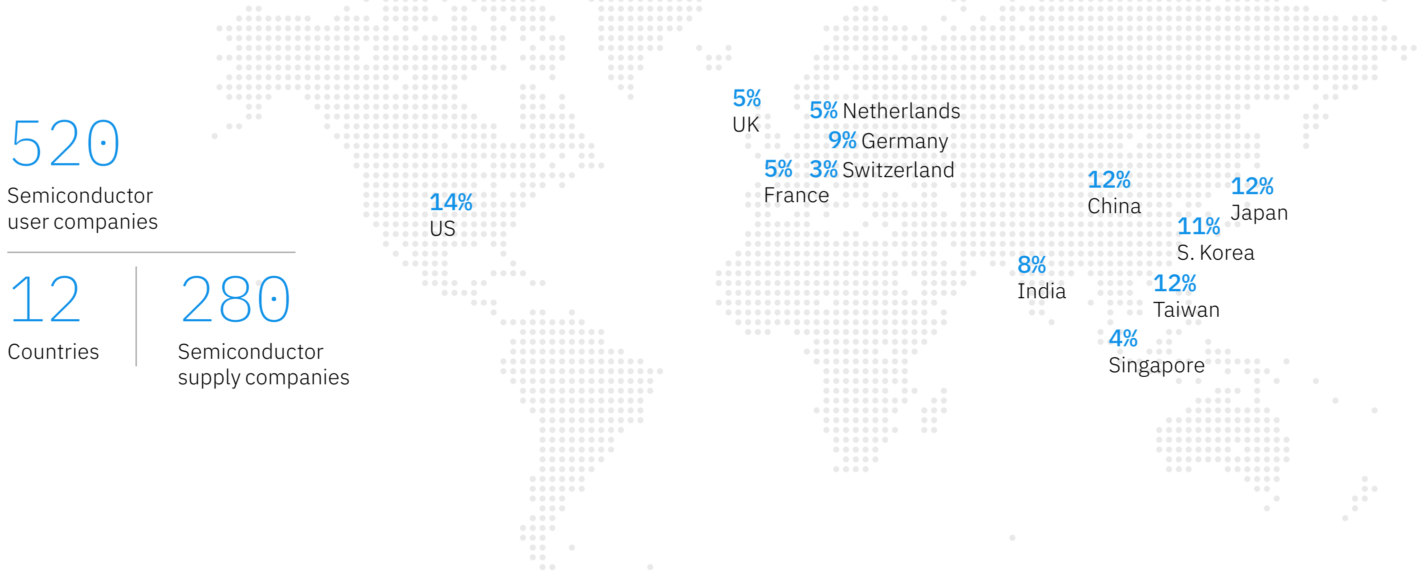
Jiu Hanashima, Consultant

Tegan Jones, Editorial Lead

Christopher Nowak, Research Leader

Tihomir Trifonov, Visual Designer

Michael Tucker, Editorial Lead



Research methodology

The IBM Institute for Business Value (IBM IBV), in cooperation with Phronesis Partners, surveyed C-level executives from 520 semiconductor user companies ("chip buyers") and 280 supply companies ("chip suppliers") in 12 countries in the second quarter of 2025.

User companies include nine segments that use advanced semiconductors for their products and services—automotive, data center, consumer electronics, industrial machinery, telecom carrier, medical equipment, smartphone, PC/tablet/gaming device, and HPC/mainframe. Semiconductor supply companies include nine segments—fabless, foundry, materials, machinery, integrated device manufacturers, advanced packaging, distributor, and electronic design automation/intellectual property vendors.

Participants were asked a range of questions in various formats (multiple choice, numerical, and Likert scale). They were asked about their organization's expectations, results, and outlooks related to AI advancements, as well as product and semiconductor strategies that incorporate AI technologies.

The analysis of this comprehensive dataset necessitated a sophisticated, multidimensional analytical framework. The research methodology incorporated several complementary analytical techniques, each selected to reveal distinct insights within the data structure.

Group-level comparative analysis was conducted to systematically evaluate supply-side versus demand-side factors, establishing their relative importance in driving AI advancement trajectories. To capture complex, nonlinear relationships within the data, multilayer perceptron neural networks and ordinal logistic regression were deployed to identify intricate patterns governing local sourcing enablement mechanisms.

Subsequently, the K-means clustering algorithm was applied to systematically categorize participating organizations into homogeneous segments based on their respective AI maturity levels and organizational resilience characteristics. Additionally, segmentation analysis was performed on R&D budget allocation to identify distinct investment patterns across organizations.

This integrated analytical approach ensured comprehensive examination of the dataset while maintaining methodological rigor throughout the research process.

In addition, authors conducted 14 in-depth interviews with industry executives to understand more nuanced strategic insights. Those qualitative insights are reflected in the analysis, as well.

The right partner for a changing world

At IBM, we collaborate with our clients, bringing together business insight, advanced research, and technology to give them a distinct advantage in today's rapidly changing environment.

IBM Institute for Business Value

For two decades, the IBM Institute for Business Value has served as the thought leadership think tank for IBM. What inspires us is producing research-backed, technology-informed strategic insights that help leaders make smarter business decisions.

From our unique position at the intersection of business, technology, and society, we survey, interview, and engage with thousands of executives, consumers, and experts each year, synthesizing their perspectives into credible, inspiring, and actionable insights.

To stay connected and informed, sign up to receive IBV's email newsletter at ibm.com/ibv. You can also follow us on LinkedIn at https://ibm.co/ibv-linkedin.

About SEMI

SEMI® is the global industry association connecting over 3,000 member companies and 1.5 million professionals worldwide across the semiconductor and electronics design and manufacturing supply chain. We accelerate member collaboration on solutions to top industry challenges through Advocacy, Workforce Development, Sustainability, Supply Chain Management and other programs. Our SEMICON® expositions and events, technology communities, standards and market intelligence help advance our members' business growth and innovations in design, devices, equipment, materials, services and software, enabling smarter, faster, more secure electronics. Visit www.semi.org, contact a regional office, and connect with SEMI on LinkedIn and X to learn more.

Notes and sources

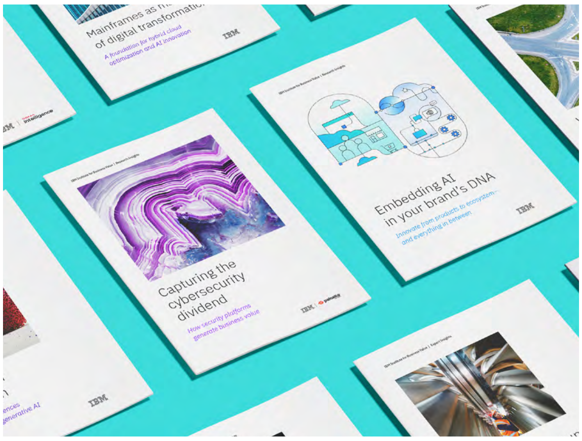
1 Jones, Cristela. "Samsung's massive Texas plant scores $16B AI-chip deal from Elon Musk." MSN. July 28, 2025. https://www.msn.com/en-us/news/technology/samsungs-massive-texas-plant-scores-16b-ai-chip-deal-from-elon-musk/ar-AA1JrJJw?apiversion=v2&noservercache=1&domshim=1&renderwebcomponents=1&wcseo=1&batchservertelemetry=1&noservertelemetry=1-telcs

2 "Apple says Samsung will supply chips from Texas factory." Reuters. August 7, 2025. https://www.reuters.com/business/apple-says-samsung-will-supply-chips-texas-factory-2025-08-06/

3 "Fab Runs On Code." TSMC IT Careers. Accessed September 4, 2025. https://www.tsmc.com/static/english/careers/it_career/index.html

4 "AI is set to drive surging electricity demand from data centres while offering the potential to transform how the energy sector works." International Energy Agency. April 10, 2025. https://www.iea.org/news/ai-is-set-to-drive-surging-electricity-demand-from-data-centres-while-offering-the-potential-to-transform-how-the-energy-sector-works?utm_source=chatgpt.com

5 Walters, Asya, Andy Walker, and Moe Kelley. "Rethinking AI Demand Part 1: AI Data Centers Are Experiencing a Surge of Training Demand - What Happens When the Surge is Over?" A&M. February 25, 2025. https://www.alvarezandmarsal.com/insights/rethinking-ai-demand-part-1-ai-data-centers-are-experiencing-surge-training-demand-what?utm_source=chatgpt.com



Subscribe to our IdeaWatch newsletter

Just the insights. At your fingertips. Delivered monthly.

Brought to you by the IBM Institute for Business Value, ranked #1 in thought leadership quality by Source Global Research for the second consecutive year.

Research-based thought leadership insights, data, and analysis to help you make smarter business decisions and more informed technology investments.

Subscribe now: ibm.co/ideawatch

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America | October 2025

This document is printed on chlorine-free 100% post-consumer paper meeting Forest Stewardship Council (FSC) responsible forestry certification. The energy used to manufacture this paper was generated through renewable green energy. Please recycle.

This document is printed on chlorine-free 100% post-consumer paper meeting Forest Stewardship Council (FSC) responsible forestry certification. The energy used to manufacture this paper was generated through renewable green energy. Please recycle.

ibm.biz/semiconductors-ai-race