SIEMENS

TECH TRENDS 2030: A SIEMENS FORESIGHT SERIES

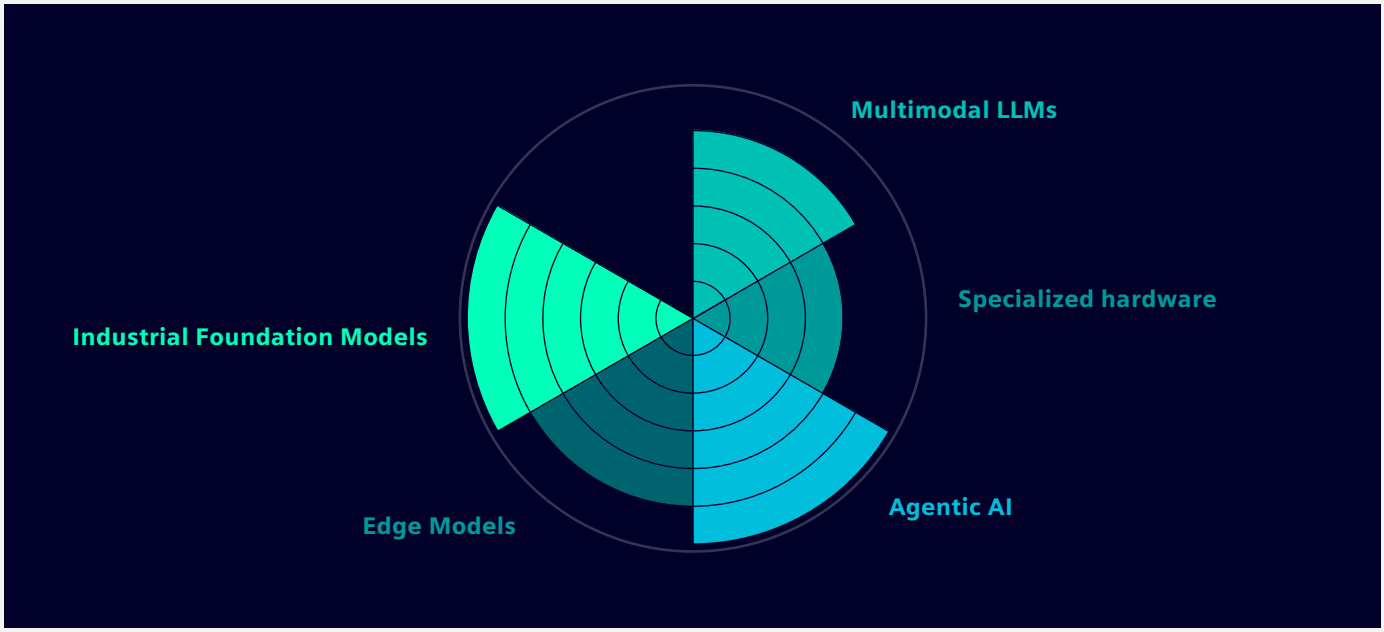
The next era of **generative AI**

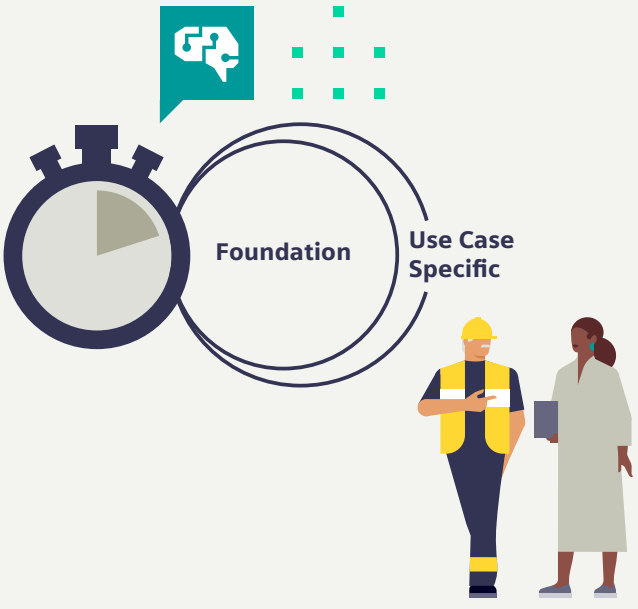
Siemens, a global technology leader, scans for emerging tech trends and translates insights into actionable intelligence. In this briefing, we will explore developments in generative industrial AI and their implications for industry.

AI has delivered tremendous value in industries over the past decades. Innovations in machine learning and neural networks enabled solutions like predictive maintenance or generative design. However, with the recent breakthrough in generative AI new opportunities emerged, which – beyond all the hype and excitement – are delivering real value to industries. With Industrial Copilots domain know-how in industries plagued by a lack of skilled labor is being democratized and AI-powered human-machine interfaces are making technology more accessible. Additionally, large language models (LLMs) are playing a transformative role in industrial applications by acting as “translators” between APIs. This capability simplifies communication and integration across diverse tools and systems, breaking down silos and enabling more efficient workflows. At the same time, there are many obstacles that need to be overcome. For example, the data hunger of LLMs cannot be satisfied as long as data sovereignty and trust between partners remain unresolved. Developments are fragile and difficult to predict. That said, here's what our experts think, in a nutshell.

We will, for example, examine Industrial Foundation Models for robust AI solutions, how Agentic AI gets more and more autonomy from human guidance and discuss Multimodal LLMs for processing diverse data types simultaneously. Finally, we will look at various future scenarios and discuss how a holistic strategic approach can help to tackle the challenges lying ahead.

I. The evolution of Industrial generative AI: **Key trends** on our radar





1. Industrial Foundation Models:
A trend that our experts consider to be one of the most important today. Industrial Foundation Models are pre-trained on industry-specific data, enabling faster and more accurate deployment of AI solutions. They provide a standardized starting point, saving time, resources and energy by scaling effects. The models capture industry complexities, leading to informed decision-making. They also facilitate knowledge transfer and collaboration across sectors.

What does this mean? An Industrial Foundation Model is trained on the "language of engineering" and goes beyond typical LLMs. It supports not only text and images but also 3D models, 2D drawings, and other complex structures like time-series data specific to the industry (see trend 3: multimodal LLMs).

Tailoring AI solutions to industry needs improves efficiency and performance. However, challenges include data availability and privacy concerns. Industrial partners need to team up and share use case and their specific training data – trust is essential here. Industrial Foundation Models are reshaping industries by accelerating AI adoption and addressing specific challenges.

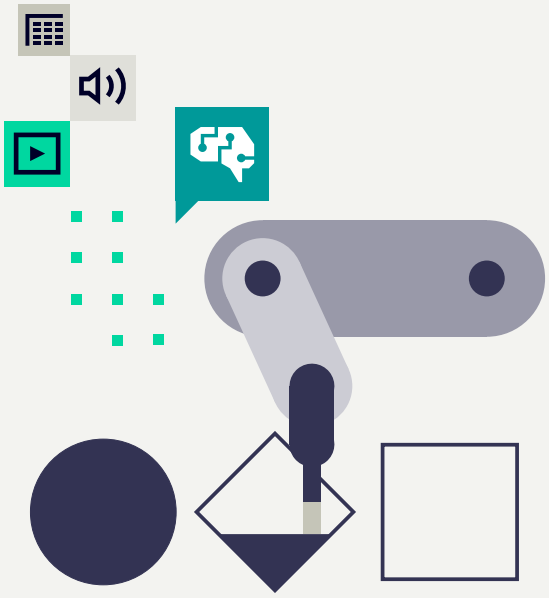
2. Agentic AI:

Agentic AI, especially in combination with generative AI, will be a game changer. Agentic AI refers to the use of AI systems that possess a certain level of autonomy and decision-making capabilities in the industrial context. These AI systems are designed to act as agents, performing tasks and making decisions on behalf of human operators or organizations.

Agentic AI in the industrial context involves deploying AI systems that can independently monitor, analyze, and control various aspects of industrial operations, such as predictive maintenance, quality control, inventory management, or optimization of production processes.

By leveraging agentic AI, industries can benefit from increased efficiency, reduced costs, improved safety, and enhanced decision-making capabilities.

Generative AI combined with agents will change the way we engineer and design. AI will take over more tasks that were previously done by humans, such as designing a specific element or product. With agents, we will first simulate and optimize the process, then derive the parameters for the factory or plant, and then design the system, which will then be built.



A common marketplace (aka Agentic Economy) will form where agents can be ordered/bought. Agents will also need to be able to search for other agents in such stores, so a standard for agents and their meta-description may become important. Agents will also have the ability to pay for 3rd party services. At some point, you will have a network of agents, orchestrated and controlled by a „master agent".

End-to-end quality management across multiple agents is one of the biggest challenges our experts see. This is especially true when an agent economy is established that combines a wide variety of agents from different companies.

3. Multimodal LLMs:

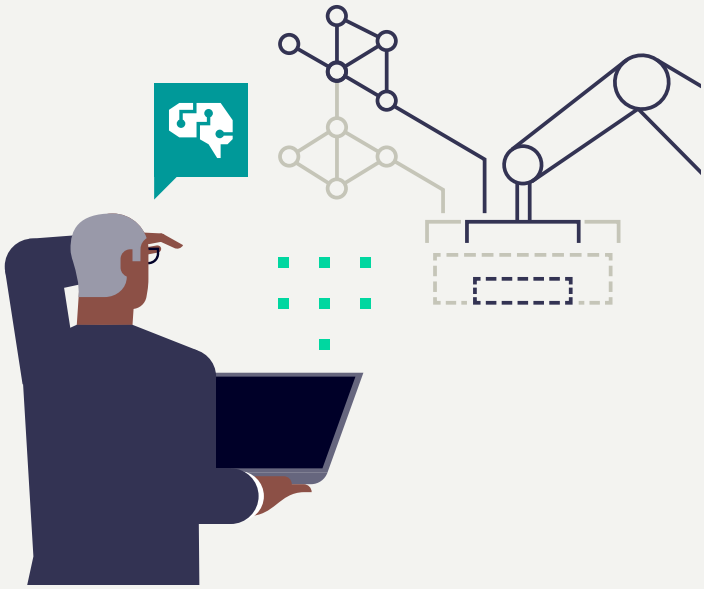
Multimodal LLMs represent an exciting trend in AI, enhancing language processing, improving computer vision, and enabling machines to better interact with the real world. They combine language understanding with visual perception. These models process data from text, images, and videos to gain a deeper understanding. They can generate accurate descriptions, analyze visual data with textual information, and enhance natural language understanding.

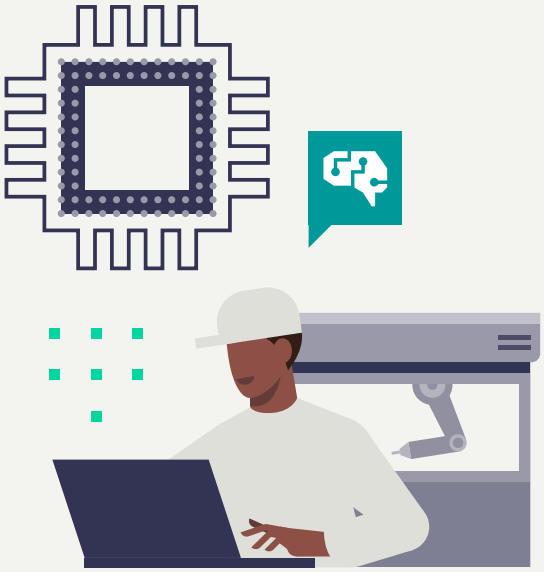
Multimodal LLMs have applications in computer vision, autonomous vehicles, and robotics. They improve object recognition, scene understanding, and enable machines to follow complex instructions. Multimodal LLMs have the potential to revolutionize the processing and generation of industry-specific data such as time series, 2D and 3D models, or machine vision in the same way that conventional LLMs have revolutionized text and speech processing.

Challenges include acquiring and labeling multimodal datasets and addressing biases. A good approach would be to pre-train multimodal LLMs to the extent that the local effort required by the end user remains low. This increases acceptance and safety and enables faster scaling.

4. Edge Models:

Industrial Edge architectures represent a growing trend in the field of AI, which involves the deployment of AI algorithms and processing power at the edge of industrial networks, in closer proximity to the data source. This approach facilitates real-time data analysis, decision-making, and automation, thereby reducing reliance on cloud-based processing. It addresses latency and bandwidth limitations, enabling industrial systems to respond promptly to critical events and make timely decisions. This approach enhances trustworthiness and cybersecurity by keeping sensitive data and AI algorithms within the local network, thereby reducing exposure to potential cyber threats associated with transmitting data to the cloud. Additionally, it offers cost savings by reducing the need for continuous cloud connectivity and associated data transfer costs. However, the decision to adopt Industrial Edge AI depends on factors such as algorithm complexity, data volume, and the need for centralized monitoring and control. Achieving a balance between edge and cloud computing is crucial for ensuring optimal performance and scalability. Additionally, federated learning can be employed in Industrial Edge AI to train AI models on decentralized data sources while preserving data privacy.





5. Specialized hardware:

Specialized hardware, like Graphics Processing Unit (GPU) or Language Processing Units (LPUs) -enabled edge devices, is an emerging trend in industrial AI. These devices provide high-performance computing power at the edge, enabling real-time processing of AI algorithms. Its integration allows for parallel processing and accelerated performance, resulting in faster execution of complex AI tasks. This local processing reduces latency and reliance on cloud resources, making it crucial for time-sensitive applications. Specialized hardware also supports advanced AI models, leading to enhanced insights and improved performance. Moreover, it reduces costs by minimizing the need for extensive cloud infrastructure and data transfer. One example are LPUs that are specialized for processing (not training) LLMs, developed and produced by Groq, a semiconductor startup valued at $2.8 bn now. Their platform, launched in Feb 2024, is used already by about 800,000 developers. However, careful consideration is needed for factors like power consumption, scalability, and compatibility with existing systems to ensure seamless integration and optimal performance.

II. The impact: Future scenarios

Scenario I

From assistance/recommendation systems to autonomous systems

A future industrial AI has experienced a significant transformation, shifting from assistance and recommendation systems to an increasing number of highly / more autonomous systems in the realm of industrial automation. This trend is expected to extend to the automation domain in the next five years, as experts foresee the need for more autonomy in factories due to a lack of skilled labor. Advanced task planning systems, such as Robotic Task Planning, are likely to play a crucial role in this transformation.

In manufacturing, autonomous robots will have taken over repetitive and labor-intensive tasks on the assembly line. These robots work collaboratively with humans, seamlessly integrating into production processes. Equipped with advanced AI algorithms, they can analyze data in real-time, make decisions, and adapt to changing production demands. This level of automation has drastically improved efficiency, productivity, and product quality.

The logistics sector has also witnessed a revolution through the implementation of autonomous systems. Smart robots navigate autonomously through warehouses, retrieving and delivering goods with precision. They optimize inventory placement, track stock levels, and streamline order fulfillment processes. As a result, order processing has become faster, errors have been reduced, and customer satisfaction has improved.

In industries such as energy and utilities, autonomous systems play a crucial role in monitoring and controlling complex infrastructure. AI-powered systems continuously analyze data from sensors, predict maintenance needs, and automatically schedule repairs or adjustments. This proactive approach to maintenance ensures the smooth operation of critical systems and minimizes downtime.

Besides providing support at individual steps in existing workflows, next-generation industrial generative AI solutions are expected to enable shortcuts in the workflow process, spanning across multiple individual tasks, from intent to outcome. This advancement will further enhance productivity and efficiency in industrial processes.

The rise of autonomous systems in industrial automation has led to increased productivity, cost savings, and improved safety. These advancements bring us closer to the progress levels projected by economists for a world filled with autonomous systems, compensating for the limitations of human labor.

However, it is crucial to also address the challenges of potential job displacement and ethical considerations that arise from this transition. It must be clear that AI is not intended to replace humans, but to support them. Less challenging, less complex tasks can be handled by AI, leaving humans to focus on tasks that require deeper understanding and creativity. Careful planning and management are necessary to ensure a smooth integration of autonomous systems while addressing these concerns.

Scenario II

Autonomous industrial AI for highly flexible, customized and high-speed processes

Autonomous industrial AI is transforming production processes, outperforming human intervention in terms of speed and flexibility. These intelligent systems analyze vast amounts of data in real time, enabling split-second decisions to optimize operations. Achieving this level of automation requires trust in AI's performance. Manufacturers must have confidence in the accuracy and reliability of AI algorithms to delegate control. By placing trust in AI, businesses can unlock the full potential of speed, flexibility, and efficiency.

The imminent integration of generative and agentic AI with advanced hardware, such as AI accelerators designed for alternative neural network architectures, is poised to transform the landscape of AI technology. This advancement will lead to a significant reduction in AI costs across all levels. The seamless integration of generative AI with existing services will empower organizations to respond promptly and effectively to unforeseen engineering and operational scenarios, enhancing customization and flexibility. Furthermore, automation will accelerate the design process, enabling the exploration of a broader range of alternatives to identify efficient solutions.

As AI capabilities evolve beyond our current understanding and experience, the importance of human involvement will diminish, with the extent of autonomy granted to AI systems determining the possibilities. The key question is whether we will limit what is achievable, or whether we will embrace the realm beyond our current understanding.

Scenario III

Industrial AI democratizes industrial expertise in less developed countries

Industrial AI is democratizing access to industrial expertise in less developed countries and propelling them past traditional industrial revolution steps. By adopting AI technologies, these countries can bypass costly and time-consuming infrastructure development, moving directly to advanced, intelligent systems. This enables them to quickly catch up with more developed nations in terms of industrial capabilities and competitiveness. Industrial AI optimizes processes, improves quality, and facilitates data-driven decisions, accelerating industrial growth and contributing to global economic advancement.

However, it is important to note that while industrial AI makes many technologies more accessible, it also creates new specializations that may only be available to a select few players. This is not only true for less industrialized countries but also for countries like Germany, where a shortage of skilled labor and the need for fast decision-making will drive the adoption of AI. Easier access to information, domain knowledge, and automation supports the accessibility of sophisticated industrial solutions, particularly in scenarios with a less skilled workforce. Additionally, intelligent devices with generative AI capabilities at the edge reduce the need for complex IT/OT infrastructures.

On the other hand, there is a possibility of access to AI becoming more separated due to geopolitical conflicts. Less developed countries may face situations where their access to AI is restricted and traded for other compromises. This highlights the importance of ensuring equitable access to AI talent, compute power, and industrial domain knowledge for global progress.

III. Mastering the new era of generative AI: A holistic strategy

To ensure readiness for the advancements and challenges of industrial AI in 2030, it is essential that stakeholders adopt a comprehensive strategic approach, incorporating the following key elements:

1. General considerations:

- Fostering a culture of innovation within the organization that embraces AI technology is crucial. This can be achieved by promoting collaboration, providing training and resources for employees to upskill in AI-related areas, and fostering an environment that encourages experimentation and learning from failures.
- Explore which support processes (like marketing or sales) and primary processes (like software development or production) can benefit most from the power of AI. Increasing efficiency, reducing material and energy consumption, and shortening lead times are also key to achieving sustainability goals.
- Industrial AI differs significantly from end-user AI. It must meet the rigorous requirements and standards of industrial environments - it must be industrial-grade. Cybersecurity, harm reduction, legal compliance and the mitigation of bias in training data should be ensured through binding standards and rules.

2. Enable an Industrial AI ecosystem-centric approach:

- What is true for digitalization is all the more true for AI: no one can do it alone!
- True creativity and value creation can only happen in a space where partners, customers, vendors, developers and experts come together to explore and develop best-in-class solutions.
- Industrial AI in particular, which can only be as good as its training data, benefits from collaboration and co-creation.
- The old formula for success of protecting your data to realize your interests (such as intellectual property) will change in the long run. In the emerging age of AI, those who share their data with partners, customers and experts in the best way will succeed. A clear strategy for data sovereignty, data pricing and mutual trust must now take center stage. Quite a few of our experts think this is the hardest nut to crack.
- Siemens Xcelerator is a growing ecosystem with a strong community of partners that has set itself precisely this goal: to combine knowledge and data for co-creation. Only in such an environment can a future agentic economy, for example, be realized.

For more information, please contact our experts:

Michael May, Head of Technology Field,
may.michael@siemens.com

Boris Scharinger, AI Strategist at Siemens Digital Industries,
boris.scharinger@siemens.com

Laura Engelhardt, Head of Strategy and Transformation,
laura.engelhardt@siemens.com

Published by
Siemens AG

Siemens AG
Werner-von-Siemens-Str. 1
80333 Munich
Germany

www.siemens.com

© Siemens 2025