Horizons

S&P Global
Energy

The dawn of cognitive manufacturing: The AI factory

The Take

We are entering the next phase of AI adoption. The massive build-out of AI infrastructure continues apace, foundational models grow more capable, and mass market inference services are now ubiquitous. Yet behind all this visible activity, the industry is shifting. Alongside general-purpose, horizontal models and cloud services, highly specialized, domain-specific applications are emerging that are deeply embedded into industry-specific workflows and automate complex tasks.

The manufacturing sector is a prime example: Industry 4.0 manufacturing was well underway before AI became a primary driver, with advances in always-on connectivity, data gathering through smart sensors and digitization. Now, AI can supercharge those efforts with its ability to analyze the often messy, unstructured data of the manufacturing world in real time, enabling new levels of data-driven decision-making, automation and optimization.

Industrial AI requires different modalities

Modality		LLMs		Industrial AI
Text	✓	Primary	✓	Logs, manuals
Images	✓	(in multimodal models)	✓	Visual inspection, IR/thermal
Audio	✓	Speech, sounds	✓	Machine diagnostics
Video	✓	(emerging capability)	✓	Surveillance, visual tracking
Time-series	—	(limited)	✓	Core for sensors & operations
Sensor data	—	Limited direct input	✓	Essential for industrial processes
Code	✓	Programming tasks	—	Rare, except for automation logic
CAD/3D models	—	Not native	✓	Used in design and simulation
Geospatial	—	Emerging	✓	Fleet, geographic information systems, smart grids

Source: 451 Research from S&P Global Energy Horizons.

Industrial work involves interpreting information across multiple modalities, which requires AI models that can integrate and reason across visual, numeric and textual inputs (see figure above). The drivers and incentives to make the transition include cost reduction (from next-generation automation and waste reduction), operational resilience (predictive maintenance), improved quality control (vision systems) and supply chain optimization. Yet the sector also faces challenges not commonly seen in other domains where generative AI has proliferated more quickly. Industrial systems must handle data types largely absent from the LLM training corpus, such as multivariate time-series, equipment tags and control logic, making direct application of general-purpose foundation models often impractical. Compared with sectors like software, financial services and retail, this has slowed the adoption of generative AI in manufacturing, as well as in telecommunications, utilities and energy.

Large, cloud-centric architectures common in generative AI are not always suitable for industrial environments where latency and uptime are critical, and adapting them into compact, interpretable models for edge deployment remains difficult. The lack of comprehensive domain-specific datasets, particularly those capturing rare events and anomalies, also limits the fine-tuning of models for industrial use cases.

Deep industry expertise is essential to address these issues, but many technology providers and AI startups are poorly positioned to meet such requirements. Another challenge is reconciling the rapidly evolving refresh cycles of bleeding-edge AI technology (typically measured in months) with the long lifespan of manufacturing equipment, which often remains in place for a decade or more.

Business impact

Next-generation automation. Traditional automation excels at high-volume, repetitive tasks but requires manual reprogramming for every process change. AI can power smarter machines and robots capable of performing complex, adaptive tasks, turning deterministic systems into probabilistic and adaptive systems that can learn and self-optimize.

Predictive maintenance. IoT sensors can be augmented with machine learning algorithms, enabling analysis of both real-time and historical data to forecast equipment failures before they occur and identify the optimal window for just-in-time maintenance.

Quality control. Organizations can use machine learning and computer vision for real-time inspection of all items in production along with continuous monitoring of production processes. Edge computing — processing data directly on the manufacturing equipment — enables the low latency needed for near-instantaneous quality decisions.

Design and simulation. Generative AI tools can accelerate product innovation, including by helping to design products that are easier and more efficient to manufacture. Teams can use digital twins — virtual replicas of physical machines — to simulate operation in production environments and to test alternatives before building any physical equipment.

Supply chain optimization. AI can shift supply chain management from reactive, manual adjustments to proactive, automated orchestration, incorporating demand forecasting and market trends, with shipments autonomously rerouted to avoid bottlenecks.

Looking ahead

The manufacturing sector is cautious and measured by nature, but the potential benefits of AI, along with a fear of being left behind, are motivating a more rapid transition than might have been expected. Advanced data governance is essential. Data must be standardized and available at the edge for real-time analytics. Data quality frameworks should be in place to ensure accuracy for AI model training and inference. And organizations must also manage the convergence of IT and operational technology, connecting legacy equipment to modern, intelligent software. The human workforce must be upskilled to work with and manage agentic AI tools, shifting from manual operations to supervising intelligent systems.

AI in manufacturing requires dense networks of sensors and actuators at the core of industrial IoT, high-bandwidth networking for the large amounts of data produced by AI vision systems, and convertors that understand legacy OT protocols. These systems rely on edge computing and hybrid cloud infrastructure, powered by high-performance CPUs, GPUs and AI accelerators behind the scenes.

A fundamental transition is underway on the factory floor. Agentic AI, while still an emerging technology, plays a growing role in actively managing inventory, avoiding supply chain bottlenecks and coordinating end-to-end production with minimal human intervention. The next phase will include more widespread use of digital twins to optimize product design, as well as deeper integration of "physical AI" — the combination of machine learning, sensory controls and robotics — to achieve higher efficiency while operating safely alongside human workers. The goal is to enable manufacturers to achieve both higher productivity and greater agility in an increasingly volatile world.