

Introduction: Industrial AI for Industry 5.0

Pál Varga, Kārlis Freivalds, Gian Antonio Susto, and Péter Olaszi

I. AIMS5.0 – PROVING GROUND FOR ENGINEERED AI

Industry 4.0 connected machines, sensors and enterprise systems. Industry 5.0 asks what that connectivity is for, and adds three requirements to efficiency: production must be human-centric, sustainable and resilient. Artificial intelligence is the most promising tool for meeting these requirements, and also the easiest to misapply.

The gap between a promising model and a dependable production system is wide. A detector that scores well on a public benchmark may fail when the camera tilts by a few degrees. A scheduling policy trained in simulation may be opaque to the engineers who must trust it. A prognostic model may be accurate yet cost more to run than the downtime it prevents. Closing this gap is an engineering problem as much as a modelling one: it concerns data, integration, tolerances, cost and human acceptance. This book is written for the engineers who must close it. It combines methodological reviews with results from real industrial deployments, and it reports what did not work alongside what did.

AIMS5.0 (Artificial Intelligence in Manufacturing leading to Sustainability and Industry 5.0) is a European research and innovation project supported by the Chips Joint Undertaking and national funding authorities under grant agreement no. 101112089. The target here is to accelerate the transition from Industry 4.0 to Industry 5.0 by integrating AI systematically into industrial processes. Rather than treating algorithm development as an isolated activity, AIMS5.0 addresses the complete adoption lifecycle, from the definition of industrial use cases through data management, algorithm development, deployment, validation and standardisation.

AIMS5.0 (<https://aims50.eu/>) is built around twenty industrial use cases spanning semiconductor manufacturing, machine tools, robotics, logistics and production systems. These use cases provide the empirical foundation of this book and make it possible to examine AI at very different levels of technological maturity. Some cases remain exploratory, with the problem formulation and relevant performance indicators still being refined, whereas others have progressed to production pilots operating on live industrial lines. Taken together, they reveal not only what industrial AI can achieve but also the conditions that determine whether an approach can move beyond experimentation.

Semiconductor manufacturing appears repeatedly throughout the book because it provides an especially demanding environment for industrial AI. Few sectors combine such high degrees of automation with equally strict quality requirements

Pál Varga is with the Department of Telecommunications and Artificial Intelligence, Budapest University of Technology and Economics.

Kārlis Freivalds is with the Institute of Electronics and Computer Science.

Gian Antonio Susto is with the Department of Information Engineering.

Péter Olaszi is with AITIA International Inc.

and such costly consequences of downtime or process deviation. At the same time, the book extends well beyond this domain. Applications in cosmetics production, lighting, battery manufacturing and additive manufacturing demonstrate that many of the underlying engineering challenges are shared across sectors. The resulting lessons are therefore not tied to a single production technology, but reflect broader patterns in the adoption of AI for Industry 5.0.

II. ORGANISATION OF THE BOOK, AND SOME LESSONS LEARNED

The core chapters follow a common structure, moving from the background and state of the art in each field to AIMS5.0 applications and the research gaps that remain. Their sequence also reflects the functional loop of an intelligent industrial system: sensing the physical process, diagnosing its condition, deciding how to respond and acting on the plant. This progression is framed by a methodological foundation at the beginning of the book and by emerging language-model-based capabilities at the end.

Chapter 2, *Machine Learning in Industry 5.0*, provides this foundation. It shows how Industry 5.0 broadens the productivity-oriented perspective of Industry 4.0 by placing greater emphasis on human-centricity, resilience and sustainability. The chapter also explains why established non-neural methods remain important in industrial settings, particularly when interpretability, data efficiency or limited computing resources matter. The AIMS5.0 use cases show how these methods support tasks ranging from condition monitoring and predictive maintenance to production planning. Across them, the same principle recurs: industrial AI is most effective when domain expertise continues to shape both the model and the decisions made from its output.

Chapter 3, *Image Processing: a Review on Industrial Applications*, turns to the sensing stage. It examines industrial vision for quality control, automation and real-time monitoring, where model accuracy must be balanced against processing speed and the consequences of incorrect decisions. Modern vision methods are discussed through AIMS5.0 applications in robotic grasping, automated insertion, semiconductor inspection and assembly support. The chapter also shows how synthetic data and continual learning can reduce annotation effort and help vision systems adapt as products and operating conditions change. These cases present perception not as an isolated recognition problem, but as the first step in a wider decision and control process.

Chapter 4, *Anomaly and Fault Detection and Predictive Maintenance*, moves from sensing to diagnosis. It considers how AI and connected sensing can be used to identify abnormal behaviour, detect faults and anticipate maintenance needs. Particular attention is given to explainability and uncertainty,

since industrial users must understand not only what a model predicts but also how much confidence can be placed in that prediction. The use cases span production processes, machining and edge-based monitoring, including the transition from experimental anomaly detection to production-ready deployment. This transition is important because a diagnostic model becomes useful only when its warning can lead to timely action without overwhelming operators with unreliable or poorly contextualised alerts.

The next stage is decision making. Chapter 5, *Reinforcement Learning, Optimization and Scheduling*, examines planning and scheduling problems in which conventional approaches may struggle with changing conditions or real-time requirements. Reinforcement learning and related optimisation methods are illustrated through semiconductor scheduling, supply-chain planning and robot motion planning. The central message is that a useful industrial decision must do more than optimise an abstract objective: it must fit the production environment, remain understandable and adapt when operating conditions change.

Chapter 6, *Robotics, Automation, and Process Management*, completes the sense–diagnose–decide–act loop by focusing on execution in the physical system. It connects robotics and automation with the broader management of production processes, using AIMS5.0 examples from wafer fabrication, intralogistics, material handling and automated production. These cases show that deployment is not a final step added after model development. The way an AI system will be integrated, operated and maintained influences its (re-)design.

Chapter 7, *Large Language Models in Industry 5.0*, extends the discussion toward knowledge access and human–AI interaction. In engineering environments, relevant knowledge is often distributed across documents, software tools and organisational boundaries. Large language models can help engineers work with this information, but their value depends on reliable grounding, appropriate safeguards and continued human oversight. The AIMS5.0 examples range from engineering copilots and knowledge retrieval to compliance support and natural language interfaces for robot programming, illustrating both the potential and the limits of these emerging systems.

Although the chapters focus on different technological functions, their use cases in AIMS5.0 overlap. These connections reinforce the main argument of the book: industrial AI creates value not through isolated algorithms, but through the careful integration of sensing, reasoning, action and human expertise.

Across the AIMS5.0 use cases, the strongest recurring lesson is that industrial AI depends as much on the structure of the engineering problem as on the learning algorithm itself. Data are often scarce, imbalanced or expensive to label, while maintenance and production knowledge may be fragmented across systems and documents. Improving the data pipeline can therefore matter more than increasing model complexity. At the same time, purely data-driven solutions are rarely sufficient on their own. The most robust approaches tend to combine learning with established engineering knowledge, using physical, geometric or process constraints to guide the model where possible. This also changes how performance should be judged. A higher model score has practical value only when it

improves an engineering outcome, such as detecting a relevant defect, respecting a process constraint or supporting a better maintenance decision. In some cases, the limiting factor lies not in the model but in the sensor arrangement, mechanical setup or surrounding production process. Industrial AI must therefore be evaluated as part of the complete system rather than as an isolated prediction task.

A second set of lessons concerns what happens when a promising model leaves the laboratory. Performance can deteriorate when the product, machine, operating regime or production site changes, which makes generalisation an explicit engineering requirement rather than something that is just assumed. Methods such as transfer learning and domain adaptation can help, but in some applications a deliberately specialised model with a clearly defined scope is the more dependable choice. Human factors are equally important. Operators and engineers need enough transparency to judge when an AI output can be trusted. Finally, deployment itself requires sustained engineering. Models must fit existing manufacturing systems, operate within practical computing constraints and remain maintainable as data and processes evolve. The AIMS5.0 cases show that these considerations cannot be postponed until the end of development. They shape the choice of data, models, interfaces and evaluation criteria from the beginning, and they ultimately determine whether an experimental result can become a reliable in industrial operation.

III. HOW TO READ THIS BOOK

The chapters are written so that each can be read independently, but the book is most informative when the connections between them are kept in view. Read as a whole, the research gaps identified in the individual chapters form a broader agenda for the next stage of industrial AI. Questions of transferability across machines and sites, interpretability for operators, scalability under real-time and edge constraints, data privacy and effective human–AI collaboration appear in different forms throughout the book. Their recurrence suggests that the central challenge of Industry 5.0 is no longer simply to demonstrate that AI can solve isolated industrial tasks. The more demanding objective is to create AI-enabled systems that continue to perform reliably when they encounter real production variability, organisational constraints and human decision makers.

The chapters that follow document what has already been demonstrated under industrial conditions while remaining explicit about the limitations that still prevent broader adoption. In this sense, the book is intended neither as a catalogue of algorithms nor as a collection of success stories. It is an engineering account of how industrial AI is being developed, tested and integrated into the transition toward Industry 5.0, and of the work that still remains before such systems can be considered dependable at scale.

ACKNOWLEDGMENT

The research leading to these results was conducted in the AIMS5.0 project, which received funding from the Chips Joint Undertaking under grant agreement No. 101112089 and from the partners' national funding programmes.