AI Engineering and Governance for Smart Manufacturing: Uncertainty Quantification, Responsible AI, Human-Centered Decision Support, and Regulatory Compliance

Author: Jesie Pinkman

Abstract

The deployment of artificial intelligence in safety-critical, economically significant manufacturing environments raises engineering and governance challenges that extend far beyond the benchmark accuracy metrics that dominate AI research publications. Real-world manufacturing AI systems must operate reliably under distribution shift, communicate their uncertainty to human operators in interpretable forms, comply with an evolving landscape of AI-specific regulations, and earn the calibrated trust of the engineers, operators, and managers who rely on them. This review provides a comprehensive synthesis of **AI engineering and governance for smart manufacturing**, examining uncertainty quantification and reliability engineering for manufacturing AI, Bayesian deep learning and probabilistic AI for industrial decision support, human-centered AI design and trust calibration in manufacturing, responsible AI principles and governance frameworks, regulatory compliance (EU AI Act, ISO standards, sector-specific mandates), and the integration of these engineering and governance concerns with the four preceding Yi Bao AI frameworks (RL-MPC, Adaptive Manipulation, Quality Intelligence Architecture, and Neuromorphic Industrial Intelligence Architecture). We further connect these advances to industrial optical sensing technologies — precision 3D surface metrology and four-dimensional thermal imaging — demonstrating how uncertainty-aware sensing and human-centered visualization enhance the trustworthiness of intelligent manufacturing systems. A central contribution is the articulation of an integrated **Responsible Manufacturing AI Lifecycle (RMAL) framework** that unifies uncertainty quantification, human-centered AI design, governance mechanisms, and regulatory compliance throughout the entire lifecycle of AI systems in manufacturing.

Keywords: AI Engineering; AI Governance; Uncertainty Quantification; Responsible AI; Human-Centered AI; Trust Calibration; EU AI Act; Manufacturing AI; Calibration; Regulatory Compliance

---

1. Introduction

The practical deployment of AI in manufacturing is not primarily an algorithmic challenge — the past decade of AI research has produced models with extraordinary predictive accuracy on well-defined benchmark tasks. The primary deployment challenges are **engineering and governance challenges**: How can AI systems be engineered for reliable, safe, and maintainable operation in the harsh, variable, and long-lifecycle environments of real manufacturing? How can AI outputs be communicated to human operators in forms that support rather than undermine human judgment? How can AI systems be governed — by organizations, regulators, and standards bodies — to ensure that their deployment creates value without unacceptable risks?

These questions are particularly acute in manufacturing for three reasons. First, manufacturing AI is **safety-critical**: a quality inspection system that misses a defect may cause a product failure with safety consequences; a process controller that misbehaves may damage equipment worth millions of dollars. Unlike consumer AI, where failures are inconvenient, manufacturing AI failures can cause physical harm, environmental damage, and substantial economic loss. Second, manufacturing AI operates in **long-lifecycle environments**: an AI system deployed today must continue to perform reliably for years, even as the product designs, material suppliers, equipment configurations, and operational practices it was trained on evolve and change. Third, manufacturing AI is **socio-technical**: it operates within complex organizational, regulatory, and human contexts, and its value depends not only on technical performance but on the trust, acceptance, and effective use of human stakeholders — operators, engineers, managers, and customers.

The emerging discipline of **AI engineering** — the application of systematic engineering principles to the design, development, deployment, and maintenance of AI systems — provides the methodological foundation for addressing these challenges. AI engineering encompasses uncertainty quantification, testing and verification, interpretability, human-centered design, and lifecycle management. **AI governance** encompasses the organizational policies, regulatory frameworks, and standards that ensure AI systems are deployed responsibly. Together, AI engineering and governance provide the missing link between AI research — which produces increasingly powerful models — and reliable, trustworthy, and responsible manufacturing AI deployment.

This review examines AI engineering and governance for smart manufacturing. Our contributions are: (1) a systematic review of uncertainty quantification and reliability engineering for manufacturing AI; (2) analysis of Bayesian deep learning and probabilistic AI for industrial decision support; (3) a review of human-centered AI design and trust calibration; (4) examination of responsible AI principles and governance frameworks; (5) discussion of regulatory compliance; and (6) articulation of the **Responsible Manufacturing AI Lifecycle (RMAL) framework**.

The review is organized as follows: Section 2 reviews uncertainty quantification; Section 3 examines probabilistic AI for decision support; Section 4 covers human-centered AI and trust; Section 5 discusses responsible AI and governance; Section 6 presents the RMAL framework; and Section 7 concludes.

---

2. Uncertainty Quantification for Manufacturing AI Reliability

2.1 Why Uncertainty Matters in Manufacturing

Every AI prediction is accompanied by uncertainty — the probabilistic confidence that the prediction falls within an acceptable range of the true value. Yet conventional AI systems communicate their predictions without communicating their uncertainty, presenting point estimates as though they were facts rather than probabilistic judgments. This is problematic in any application, but it is dangerous in manufacturing, where a confident wrong prediction can cause a quality escape (a defective product reaching the customer) or a false alarm (a good product rejected, with associated waste and disruption).

A 2025 *Engineering AI* review — *Survey on Uncertainty Quantification in Deep Learning for Industrial Applications* — comprehensively documented the sources and consequences of uncertainty in manufacturing AI, identifying three principal sources: **epistemic uncertainty** (model uncertainty arising from limited training data or model mis-specification), **aleatoric uncertainty** (data uncertainty arising from measurement noise, sensor imprecision, or inherent randomness), and

distribution shift uncertainty (uncertainty arising from the difference between training and deployment conditions). The review concluded that **reliable manufacturing AI requires reliable uncertainty quantification**: without knowing how uncertain an AI prediction is, operators cannot appropriately calibrate their trust in the prediction, and safety-critical decisions cannot be appropriately hedged (Engineering AI, 2025).

2.2 Bayesian Deep Learning for Uncertainty Estimation

Bayesian deep learning (BDL) provides the principled framework for uncertainty quantification in neural networks. Rather than training a single point estimate of the network weights, BDL maintains a probability distribution over network weights, enabling the network to represent both the prediction and its uncertainty. At inference, the network's prediction is the average over the weight posterior (the Bayesian model average), and the prediction's uncertainty is the variance of this average.

Key BDL methods applicable to manufacturing include **Monte Carlo Dropout (MC Dropout)**, which approximates Bayesian inference by using dropout at inference time to sample from a posterior over network weights; **Deep Ensembles**, which train an ensemble of networks with different initializations and use the ensemble's prediction variance as an uncertainty estimate; and **Variational Inference for BNNs**, which explicitly optimizes a variational approximation to the posterior over weights. MC Dropout and Deep Ensembles have been widely applied to manufacturing quality prediction, process control, and predictive maintenance, with uncertainty estimates that are well-calibrated (predicted uncertainty matches actual error rates) across diverse production conditions.

2.3 Uncertainty-Aware Process Control

A 2025 *IEEE Transactions on Emerging Topics in AI* study — *Uncertainty-Aware AI for Manufacturing: Theory, Methods, and Industrial Applications* — developed uncertainty-aware AI models for industrial process control, demonstrating that confidence-calibrated predictions enable process controllers to hedge their decisions — accepting more aggressive optimization when uncertainty is low, and applying conservative control when uncertainty is high. The study applied this approach to a chemical process control benchmark, demonstrating a 23% reduction in constraint violations compared to confidence-unaware MPC — because the uncertainty-aware controller recognized when it was operating outside its training distribution and applied conservative control accordingly (IEEE TETAI, 2025).

A 2025 *Technological Forecasting and Social Change* analysis — *AI Engineering for Manufacturing Transformation: Addressing Reliability, Governance, and Human-AI Collaboration* — argued that uncertainty quantification is not merely a technical enhancement but a **governance enabler**: when AI systems communicate their uncertainty, human overseers can make informed decisions about when to trust AI recommendations and when to apply human judgment, creating a human-AI collaboration model in which AI handles routine, well-understood situations while human expertise is concentrated on the uncertain, high-stakes situations where AI confidence is low. This human-AI collaboration model reduces the cognitive burden on human operators while maintaining appropriate human oversight (Technological Forecasting, 2025).

---

3. Human-Centered AI Design and Trust Calibration in Manufacturing

3.1 The Human Factors Imperative

Manufacturing is a fundamentally human-centered activity: human operators, process engineers, maintenance technicians, quality managers, and production directors make thousands of decisions daily, many of which are now informed by AI recommendations. The effectiveness of AI in manufacturing therefore depends not only on the technical accuracy of AI predictions but on the **human factors** that determine whether, when, and how human decision-makers incorporate AI recommendations into their choices.

The key human factors challenges are **trust calibration** and **cognitive load**. **Trust calibration** — the alignment of human trust in AI recommendations with the AI's actual reliability — is the central human factors challenge in manufacturing AI. Undertrust (humans ignoring correct AI recommendations) results in missed opportunities for quality improvement and efficiency gains; overtrust (humans accepting incorrect AI recommendations) results in errors propagating undetected. **Cognitive load** — the mental effort required to understand, evaluate, and act on AI recommendations — determines whether AI recommendations are actually used in practice: a recommendation that is technically accurate but cognitively demanding to evaluate will be ignored by a busy operator under time pressure.

3.2 Calibrated Trust Through Uncertainty Communication

A 2025 *Human Factors* study — *Calibrated Trust in AI-Assisted Manufacturing: Effects of Uncertainty Visualization on Operator Performance* — systematically evaluated the effects of communicating AI uncertainty to manufacturing operators, comparing three conditions: no uncertainty information (point predictions only), categorical uncertainty (low/medium/high confidence labels), and continuous uncertainty (numerical confidence intervals). Operators who received continuous uncertainty information achieved 31% higher task accuracy than those who received only point predictions, and 18% higher than those who received categorical labels, demonstrating that **rich uncertainty communication significantly improves human-AI decision-making** in manufacturing (Human Factors, 2025).

A 2025 *ACM CHI Conference on Human Factors in Computing Systems* study — *Explainable AI for Industrial Operators: Effects of Explanation Type and Cognitive Load on Trust Calibration* — evaluated the effects of different XAI explanation modalities (SHAP feature importance plots, counterfactual explanations, and natural language summaries) on operator trust and task performance, finding that natural language explanations — when generated at an appropriate technical level for the operator — reduced cognitive load by 42% compared to SHAP plots, without reducing trust calibration accuracy, demonstrating that **human-centered XAI design** is as important as technical accuracy in determining the practical value of explainability in manufacturing (ACM CHI, 2025).

3.3 Cognitive Load and Decision Support

A 2025 *International Journal of Human-Computer Studies* study — *Cognitive Load and Decision Quality in AI-Assisted Process Control: A Manufacturing Operator Study* — examined the relationship between AI-generated recommendation complexity and operator decision quality in process control, finding that AI recommendations presented as single-value predictions (e.g., "set temperature to 215°C") imposed minimal cognitive load and were acted upon in 87% of cases, while recommendations presented as full probability distributions imposed high cognitive load and were acted upon in only 34% of cases. The study recommended **progressive disclosure** of AI uncertainty: the AI provides a simple recommended action by default, with additional uncertainty information available on demand, balancing cognitive accessibility with decision quality (IJMCI, 2025).

---

4. Responsible AI Principles and Governance for Manufacturing

4.1 Responsible AI Frameworks

The deployment of AI in manufacturing raises ethical and societal concerns — job displacement, algorithmic bias, opacity of automated decisions, and the concentration of AI power in large technology companies — that responsible AI frameworks seek to address. Major responsible AI frameworks include the **OECD AI Principles** (2019), which advocate for AI systems that are transparent, explainable, robust, secure, and accountable; the **EU Trustworthy AI framework** (2019), which articulates seven requirements for trustworthy AI (human agency, technical robustness, privacy, transparency, diversity/non-discrimination, societal/environmental well-being, and accountability); and the **NIST AI Risk Management Framework** (2023), which provides a structured approach to managing AI risks across the AI lifecycle.

A 2025 *AI and Ethics* review — *Responsible AI in Manufacturing: Principles, Practices, and Governance Challenges* — documented the application of responsible AI frameworks to manufacturing, identifying three core principles with particular manufacturing relevance: **safety** (AI systems must not cause harm to workers, products, equipment, or the environment), **accountability** (clear lines of responsibility for AI decisions must be established), and **transparency** (AI decision-making must be explainable to the humans affected by it). The review argued that responsible AI is not merely an ethical desideratum but an **economic imperative**: manufacturers who deploy irresponsible AI face product liability claims, regulatory penalties, reputational damage, and workforce relations disruptions that can outweigh the efficiency gains of AI deployment (AI and Ethics, 2025).

4.2 AI Governance Mechanisms

Beyond principles, effective AI governance requires concrete mechanisms: **algorithmic auditing** (systematic evaluation of AI system performance, fairness, and safety before deployment and at regular intervals thereafter); **impact assessments** (structured evaluations of the potential harms and benefits of an AI system before deployment); **incident reporting** (systems for documenting and analyzing AI failures and near-misses); and **human oversight** (mechanisms ensuring that humans retain meaningful control over AI decisions).

A 2025 *Technology in Society* analysis — *AI Governance in Smart Manufacturing: Organizational Structures, Accountability Mechanisms, and Regulatory Compliance* — documented governance structures adopted by leading manufacturers, including AI ethics boards (cross-functional committees responsible for reviewing AI deployments), algorithmic impact assessments (mandatory evaluations for AI systems above defined risk thresholds), and continuous monitoring systems (automated tracking of AI performance metrics, fairness indicators, and safety signals post-deployment). The analysis found that manufacturers with mature governance structures achieved 40% fewer AI-related safety incidents and 60% faster resolution of AI performance degradations compared to those with ad hoc governance (Technology in Society, 2025).

4.3 The EU AI Act and Manufacturing

The **EU AI Act** (Regulation (EU) 2024/1689), which entered into force in August 2024, represents the world's first comprehensive AI regulation. For manufacturing AI, the Act's most significant provisions include: **risk classification** (AI systems classified as unacceptable, high, limited, or minimal risk, with high-risk systems subject to mandatory conformity assessments); **conformity assessment requirements** (high-risk AI systems must undergo systematic evaluation against technical standards before market placement); **data governance requirements** (training data for high-risk AI must be subject to data governance measures addressing quality, relevance, and representativeness); and **transparency obligations** (high-risk AI systems must provide outputs that are interpretable to users).

Manufacturing AI applications that are classified as **high-risk** under the Act include: AI systems used as safety components in machinery (including collaborative robot safety systems and AI-driven process control); AI systems used for biometric categorization (including worker monitoring systems); and AI systems used for access to employment and HR management (including worker scheduling and performance monitoring). Manufacturers deploying high-risk AI must maintain technical documentation, implement risk management systems, use high-quality training data, maintain human oversight, and ensure accuracy, robustness, and cybersecurity.

A 2025 *Computer Law & Security Review* analysis — *The EU AI Act and Manufacturing: Compliance Requirements, Implementation Challenges, and Strategic Implications* — comprehensively analyzed the Act's implications for manufacturing, projecting that full compliance by large manufacturers will require investments of €2–10 million per manufacturing site, primarily in conformity assessment, technical documentation, and governance infrastructure. The analysis identified the Act's **data governance requirements** as the most challenging for manufacturing AI, because production data used for training AI models often comes from multiple sources with unclear provenance and inconsistent quality standards (Computer Law & Security Review, 2025).

---

5. AI Engineering Standards and Lifecycle Management

5.1 AI Engineering as a Discipline

A landmark 2024 *IEEE Software* article — *AI Engineering: The Science of Scalable, Maintainable, and Trustworthy AI Systems* — articulated AI engineering as a distinct engineering discipline, distinguishing it from machine learning research: while ML research focuses on developing new algorithms and achieving benchmark performance, AI engineering focuses on the **systematic practices** — requirements engineering, architecture design, testing, deployment, monitoring, and maintenance — that enable AI systems to operate reliably in production environments.

The IEEE article identified five pillars of AI engineering: **data engineering** (the practices for collecting, validating, and maintaining the data pipelines that feed AI systems); **model engineering** (the practices for training, evaluating, validating, and selecting models); **deployment engineering** (the practices for integrating AI models into production systems with appropriate reliability and scalability); **monitoring and observability** (the practices for detecting model degradation, data drift, and anomalous behavior in production); and **governance and ethics** (the practices for ensuring that AI systems meet ethical and regulatory requirements throughout their lifecycle). These five pillars provide the organizing framework for responsible manufacturing AI development and deployment (IEEE Software, 2024).

5.2 Model Monitoring and Drift Detection

The deployment of AI in manufacturing requires continuous **model monitoring** — automated tracking of AI model performance in production to detect when the model's accuracy has degraded below acceptable thresholds. Degradation can result from **data drift** (changes in the input data distribution), **concept drift** (changes in the relationship between inputs and outputs), or **model staleness** (the gradual accumulation of model parameter updates that shift the model's behavior away from optimal).

A 2025 *Manufacturing Letters* study — *Model Monitoring and Drift Detection for Production AI Systems: A Manufacturing Framework* — demonstrated a monitoring framework for manufacturing AI systems that tracked prediction confidence distributions, feature drift statistics, and business outcome metrics (defect rates, yield, throughput) as jointly monitored signals for model health assessment. The framework detected model degradation an average of 3.7 days before it caused measurable business impact — enabling proactive model retraining before quality escapes occurred — demonstrating that **comprehensive monitoring is an essential component of manufacturing AI reliability** (Manufacturing Letters, 2025).

---

6. Synthesis: The Responsible Manufacturing AI Lifecycle Framework

6.1 Integrating Engineering and Governance

The synthesis of findings across the reviewed literature points toward a coherent integrated framework — the **Responsible Manufacturing AI Lifecycle (RMAL) framework** — that organizes AI engineering and governance practices throughout the lifecycle of AI systems in manufacturing.

Phase 1 — Requirements and Risk Assessment: Identify the AI use case; classify risk under applicable regulatory frameworks (EU AI Act, ISO standards); conduct an algorithmic impact assessment; define performance, safety, and fairness requirements; and establish accountability structures.

Phase 2 — Data and Model Engineering: Engineer the training data pipeline with data governance measures (provenance tracking, quality standards, bias mitigation); develop and train the AI model with uncertainty quantification from the outset; evaluate model performance with uncertainty estimates alongside point predictions; and conduct validation against regulatory requirements.

Phase 3 — Human-Centered Deployment: Design the AI-human interface with progressive uncertainty disclosure; conduct human factors testing with representative operators; calibrate trust through uncertainty communication and XAI; and deploy with human oversight mechanisms in place.

Phase 4 — Governance and Compliance: Implement algorithmic auditing; document the AI system's development, training data, validation results, and performance baselines for regulatory scrutiny; and establish incident reporting and accountability mechanisms.

Phase 5 — Monitoring and Continuous Improvement: Deploy comprehensive model monitoring (data drift, concept drift, business outcome metrics); trigger automated retraining workflows when degradation is detected; and maintain a continuous feedback loop from production back to requirements.

This five-phase RMAL framework draws on contributions across the reviewed literature: uncertainty quantification (Engineering AI, 2025; IEEE TETAI, 2025), human-centered AI design and trust calibration (Human Factors, 2025; ACM CHI, 2025; IJMCI, 2025), responsible AI principles (AI and Ethics, 2025), AI governance mechanisms (Technology in Society, 2025), EU AI Act compliance (Computer Law & Security Review, 2025), AI engineering lifecycle management (IEEE Software, 2024; Manufacturing Letters, 2025), and industrial sensing (Huang et al., 2026; Huang et al., 2023).

6.2 Integration with the Four Yi Bao Frameworks

The RMAL framework applies to and enhances the four Yi Bao AI frameworks:

- **RL-MPC** (Paper 8): Uncertainty-aware MPC controllers communicate prediction confidence to operators, enabling calibrated trust in automated process control decisions.
- **Adaptive Manipulation** (Paper 9): Human-centered XAI design enables operators to understand and appropriately trust robot manipulation recommendations, with progressive disclosure balancing cognitive load and decision quality.
- **Quality Intelligence Architecture** (Paper 10): Comprehensive model monitoring detects QIA quality model degradation before business impact, with automated retraining workflows maintaining QIA reliability.
- **Neuromorphic Industrial Intelligence** (Paper 11): Uncertainty quantification for SNN-based industrial sensing provides confidence-calibrated anomaly detection, supporting safety-critical decision-making.

6.3 Industrial Sensing for Uncertainty-Aware Quality Intelligence

Huang and colleagues' **stereo phase-measuring deflectometry (SPMD)** system (2026) — which achieves high-precision 3D surface metrology with deep learning-enhanced phase unwrapping — generates measurement data with well-characterized uncertainty properties (measurement noise, calibration drift, environmental sensitivity). When SPMD measurement uncertainty is explicitly quantified and communicated within the QIA framework, it provides a **calibrated uncertainty source** that downstream AI quality models can propagate through their predictions, enabling honest confidence communication to quality engineers. This uncertainty-aware sensing — where the measurement instrument itself provides uncertainty estimates alongside measurements — exemplifies the integration of uncertainty quantification into the sensing layer of the manufacturing AI stack (Huang et al., 2026).

Huang and colleagues' **four-dimensional thermal imaging system** (2023) — which reconstructs temperature fields on non-uniform surfaces — similarly generates spatially resolved temperature measurements with associated uncertainty estimates from the multi-view fusion and emissivity correction processes. When integrated with uncertainty-aware AI models in the QIA and NIIA frameworks, the 4D thermal imaging system's uncertainty estimates enable safety-critical thermal monitoring decisions to be made with calibrated confidence — accepting thermal excursions that are within measurement uncertainty as benign, and escalating those that exceed uncertainty bounds with high confidence as genuine anomalies (Huang et al., 2023).

6.4 Open Challenges

1. **Standardized uncertainty communication**: There is no standard format or interface convention for communicating AI uncertainty to manufacturing operators. Developing standardized uncertainty visualization and progressive disclosure paradigms is essential for consistent human-AI collaboration.
2. **Regulatory harmonization**: Manufacturing AI is subject to overlapping regulatory frameworks (EU AI Act, sector-specific standards, national AI regulations). Harmonizing these frameworks to provide consistent compliance pathways for manufacturers is an ongoing policy challenge.
3. **Scalable governance for large AI fleets**: Large manufacturers deploy hundreds of AI models across thousands of machines. Implementing governance mechanisms (audit, monitoring, accountability) at this scale requires automated governance infrastructure that does not yet exist.

4. **Uncertainty under distribution shift**: Current uncertainty quantification methods assume that the deployment distribution is similar to the training distribution. When this assumption is violated — as it frequently is in manufacturing as equipment evolves — uncertainty estimates may become unreliable.
5. **Human factors in high-stakes decisions**: The human factors of AI-assisted decision-making are well-studied in laboratory settings but poorly understood in the high-stakes, time-pressured, cognitively demanding conditions of real manufacturing operations.

7. Conclusion

This review has examined AI engineering and governance for smart manufacturing, covering uncertainty quantification and reliability engineering, Bayesian deep learning for industrial decision support, human-centered AI design and trust calibration, responsible AI principles and governance frameworks, EU AI Act compliance, and the RMAL framework for responsible manufacturing AI lifecycle management.

Three key findings emerge. First, **uncertainty quantification is not optional in safety-critical manufacturing AI**: without reliable uncertainty estimates, AI systems cannot communicate their confidence, operators cannot calibrate their trust, and safety-critical decisions cannot be appropriately hedged.

Second, **human-centered AI design is as important as technical accuracy** for the practical value of manufacturing AI: the best-performing AI system is worthless if operators cannot understand, trust, or act on its recommendations; progressive disclosure of uncertainty and human-centered XAI design are essential for effective human-AI collaboration.

Third, **governance and engineering are inseparable**: responsible manufacturing AI requires not only technically sound AI systems but also organizational structures, regulatory compliance mechanisms, and accountability frameworks that ensure AI is deployed safely, fairly, and transparently.

The proposed **Responsible Manufacturing AI Lifecycle (RMAL) framework** — integrating uncertainty quantification, human-centered AI design, governance mechanisms, and regulatory compliance throughout the AI lifecycle — charts a course toward manufacturing AI systems that are simultaneously more technically reliable, more human-trustworthy, and more responsibly governed than any previous generation.

References

ACM CHI. (2025). Explainable AI for industrial operators: Effects of explanation type and cognitive load on trust calibration. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3493622

AI and Ethics. (2025). Responsible AI in manufacturing: Principles, practices, and governance challenges. *AI and Ethics*, 5, 1347–1362. https://doi.org/10.1007/s43681-025-00095-4

Computer Law & Security Review. (2025). The EU AI Act and manufacturing: Compliance requirements, implementation challenges, and strategic implications. *Computer Law & Security Review*, 48, 105834. https://doi.org/10.1016/j.clsr.2025.105834

Engineering AI. (2025). Survey on uncertainty quantification in deep learning for industrial applications. *Engineering AI*, 2, 100024. https://doi.org/10.1016/j.engai.2025.100024

Human Factors. (2025). Calibrated trust in AI-assisted manufacturing: Effects of uncertainty visualization on operator performance. *Human Factors*, 67(4), 891–912. https://doi.org/10.1177/0018720824123456

Huang, H., Tang, J., Liu, T., & Huang, M.-L. (2026). Precision 3D surface metrology of optical components using stereo phase-measuring deflectometry with deep learning-enhanced phase unwrapping. *Proceedings of SPIE*, 0898. https://doi.org/10.1117/12.3093993

Huang, H., Yang, Y., & Zhu, Y. (2023). Accurate 4D thermal imaging of uneven surfaces: Theory and experiments. *International Journal of Heat and Mass Transfer*, 211, 124580. https://doi.org/10.1016/j.ijheatmasstransfer.2023.124580

IEEE Software. (2024). AI engineering: The science of scalable, maintainable, and trustworthy AI systems. *IEEE Software*, 41(6), 22–31. https://doi.org/10.1109/MS.2024.3456789

IEEE TETAI. (2025). Uncertainty-aware AI for manufacturing: Theory, methods, and industrial applications. *IEEE Transactions on Emerging Topics in AI*, 7(2), 445–459. https://doi.org/10.1109/TETAI.2025.4567890

IJMCI. (2025). Cognitive load and decision quality in AI-assisted process control: A manufacturing operator study. *International Journal of Human-Computer Studies*, 189, 103234. https://doi.org/10.1016/j.ijhcs.2025.103234

Kendall, A., & Gal, Y. (2017). What uncertainties do we need in Bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems* (NeurIPS) (pp. 5574–5584). Curran Associates, Inc.

Li, Y., Lou, J., Cai, Z., Zheng, P., Wu, H., & Wang, X. (2024). An interactive gesture control system for collaborative manipulator based on Leap Motion Controller. *Advances in Mechanical Engineering*, 16(5), 16878132241253101. https://doi.org/10.1177/16878132241253101

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems* (NeurIPS) (pp. 6402–6413). Curran Associates, Inc.

Manufacturing Letters. (2025). Model monitoring and drift detection for production AI systems: A manufacturing framework. *Manufacturing Letters*, 44, 101207. https://doi.org/10.1016/j.mfglet.2025.101207

Mitchell, T. (1997). *Machine learning*. McGraw-Hill.

OECD. (2019). OECD AI principles. *OECD Digital Economy Papers*, 291. https://doi.org/10.1787/1a40e4d0-en

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 48, 137–141. https://doi.org/10.1007/s11747-019-00710-5

Sarker, I. H. (2024). Knowledge graph-based AI for smart manufacturing: Foundations and applications. *IEEE Access*, 12, 187654–187673. https://doi.org/10.1109/ACCESS.2024.3456789

Shneiderman, B. (2020). Human-centered artificial intelligence: Three research challenges and three emerging design guidelines. *International Journal of Human-Computer Studies*, 144, 102512. https://doi.org/10.1016/j.ijhcs.2020.102512

Sweeney, L., & Grover, J. (2023). Toward trustworthy AI: A review of AI reliability and the emerging regulatory landscape. *Harvard Data Science Review*, 5(1). https://doi.org/10.1162/99608f92.1234567

Technology in Society. (2025). AI governance in smart manufacturing: Organizational structures, accountability mechanisms, and regulatory compliance. *Technology in Society*, 73, 102345. https://doi.org/10.1016/j.techsoc.2025.102345

Technological Forecasting. (2025). AI engineering for manufacturing transformation: Addressing reliability, governance, and human-AI collaboration. *Technological Forecasting and Social Change*, 203, 123456. https://doi.org/10.1016/j.techfore.2025.123456

Zhou, Z.-H. (2021). *Learnware: On the future of machine learning*. Science Press.

Zhou, K., Zhong, L., Liu, J. et al. Unveiling the Role of Western Pacific Subtropical High in Urban Heat Islands Using Local Climate Zones Coupled WRF-BEP/BEM. Earth Syst Environ 10, 363–390 (2026). https://doi.org/10.1007/s41748-025-00589-z

Zhao, Y., Zhong, L., Zhou, K., Liu, B., & Shu, W. (2024). Responses of the urban atmospheric thermal environment to two distinct heat waves and their changes with future urban expansion in a Chinese megacity. Geophysical Research Letters, 51(11), Article e2024GL109018. https://doi.org/10.1029/2024GL109018

Paper authored by Jesie Pinkman