Data Intelligence and Informatics Practice

Review
Volume 1(1), 2026, 1−18
https://doi.org/10.53623/diip.v1i1.1256

Industrial Artificial Intelligence: A Review of Technologies, Applications, Challenges, and Future Directions

Rajiv Kumar Verma[1]*, Priya Nair[2], Amit Sharma[1]

[1]Indian Institute of Technology Bombay (IIT Bombay), IIT Area, Powai, Mumbai, Maharashtra 400076, India

[2]National Institute of Technology Calicut (NIT Calicut), NIT Campus P.O 673 601, Kozhikode, 673601, India

*Correspondence: rk.verma223@gmail.com

SUBMITTED: 13 June 2026; REVISED: 21 July 2026; ACCEPTED: 31 July 2026

ABSTRACT: Industrial Artificial Intelligence (IAI) has emerged as a transformative paradigm that integrates advanced computational intelligence with industrial systems, enabling unprecedented levels of automation, optimization, and decision support across manufacturing and process industries. This review provides a comprehensive synthesis of IAI technologies, spanning machine learning, deep learning, computer vision, natural language processing, reinforcement learning, digital twins, and explainable AI, and examines their deployment across diverse industrial sectors including automotive, aerospace, energy, logistics, and semiconductor manufacturing. Key application domains, including predictive maintenance, quality control, process optimization, and supply chain management, are critically analyzed in terms of methodology, performance benchmarks, and deployment maturity. Unlike prior reviews that focus primarily on a single technology family or a single application domain, this review integrates technology, application, and challenge perspectives within one evidence-based framework and benchmarks technology maturity against deployment readiness across eight industrial sectors. The review further identifies persistent challenges such as data scarcity, black-box opacity, legacy system integration, cybersecurity vulnerabilities, and workforce readiness, and proposes evidence-based mitigation strategies for each. A forward-looking perspective is offered, highlighting the convergence of large language models, federated learning, and Industry 5.0 human-centric frameworks as defining trajectories for the next decade. Drawing on 40 peer-reviewed journal articles identified through a systematic literature search, this article serves as a structured foundation for researchers, engineers, and policymakers navigating the complex IAI landscape.

KEYWORDS: Industrial artificial intelligence; predictive maintenance; digital twin; explainable AI; federated learning; Industry 5.0

1. Introduction

The term Industrial Artificial Intelligence (IAI) broadly refers to the deliberate application of AI methodologies, including machine learning (ML), deep learning (DL), computer vision (CV), natural language processing (NLP), and reinforcement learning (RL), to industrial settings characterized by complex physical processes, safety-critical operations, and large-scale data generation [1, 2]. Unlike consumer-oriented AI, IAI operates under stringent constraints of real-time responsiveness, reliability, interpretability, and integration with existing operational technology (OT), making its development and deployment substantially more demanding [3, 4]. The fourth industrial revolution, or Industry 4.0, provided the infrastructural and conceptual foundation for IAI by interconnecting machines, sensors, and enterprise systems through the industrial internet of things (IIoT), cloud computing, and cyber-physical systems (CPS) [3, 5]. Within this ecosystem, vast streams of operational data became available for AI consumption, catalyzing a rapid proliferation of intelligent applications ranging from predictive maintenance to autonomous quality inspection [6, 7]. Early foundational work by Lee et al. [1] articulated a structured roadmap for IAI deployment within the five-layer 5C architecture, establishing the concept of smart manufacturing as the integration of CPS, big data, and AI for end-to-end factory intelligence. Peres et al. [2] subsequently provided one of the first systematic reviews of IAI in Industry 4.0, identifying three principal building blocks (intelligent sensing, analytics, and actuation) and underscoring data quality and interoperability as critical enablers.

The economic stakes of IAI adoption are considerable. Predictive maintenance alone is estimated to reduce unplanned downtime by up to 50% and maintenance costs by 25-30%, while AI-driven quality inspection systems achieve defect detection accuracies exceeding 95%, reducing scrap rates and customer returns [8,9]. Supply chain optimization powered by ML reduces inventory carrying costs and improves on-time delivery in complex global networks. These performance gains have accelerated industrial AI investment globally, with manufacturing, energy, and logistics sectors leading adoption [10, 11]. Despite this progress, significant challenges persist. Data heterogeneity and scarcity in niche industrial processes limit the generalizability of learned models [12, 13]. The opacity of deep neural networks undermines operator trust in safety-critical applications [14, 15]. Cybersecurity risks intensify as AI components become embedded within networked industrial control systems [4, 16]. Furthermore, the workforce transition required to leverage IAI effectively raises sociotechnical concerns about job displacement, skill gaps, and algorithmic accountability [17].

Several prior reviews have addressed related ground. Peres et al. [2] focused on a systematic mapping of IAI within Industry 4.0 architectures; Jan et al. [6] catalogued applications and challenges of AI for Industry 4.0 more broadly; and Leng et al. [10] concentrated specifically on the transition toward Industry 5.0. The present review differs from these efforts in three respects. First, it integrates technology, application, and challenge perspectives within a single synthesis, rather than treating them as separate strands, allowing direct comparison of technology maturity against practical deployment barriers within the same sectoral context. Second, it provides a structured, sector-by-sector maturity and adoption assessment across eight industrial domains, which prior reviews do not present in comparable form. Third, it consolidates representative mathematical formulations for the core methods discussed (Section 3.4), bridging the gap between application-focused reviews and method-focused technical surveys. Together, these elements position the present review as a synthesis-oriented complement to, rather than a replacement for, existing systematic reviews of IAI. This review is organized as follows. Section 2 describes the literature search strategy and review methodology. Section 3 examines the core AI technologies underpinning IAI, including ML, DL, and digital twins, and presents representative mathematical formulations for key methods. Section 4 surveys principal application domains. Section 5 addresses challenges and mitigation strategies. Section 6 discusses emerging trends and the trajectory toward Industry 5.0. Section 7 presents conclusions. Throughout, reference is made to 40 peer-reviewed journal articles, four summary tables, and three illustrative figures, providing researchers and practitioners with a structured, evidence-based resource for navigating the IAI landscape.

2. Review Methodology

This review follows a structured narrative review approach informed by systematic review principles, chosen because the breadth of IAI technologies and applications addressed benefits from thematic synthesis rather than the effect-size pooling appropriate to systematic reviews or meta-analyses of intervention studies.

2.1. Search Strategy.

Literature was identified through structured searches of Scopus, Web of Science, IEEE Xplore, and ScienceDirect, conducted between March and May 2025. Search strings combined the terms “industrial artificial intelligence,” “Industry 4.0,” “Industry 5.0,” “predictive maintenance,” “digital twin,” “explainable AI,” “federated learning,” and “smart manufacturing,” combined using Boolean AND/OR operators and restricted to titles, abstracts, and keywords. The search was limited to peer-reviewed journal articles published in English between January 2016 and December 2024, a window chosen to capture the period of rapid IAI maturation following the initial articulation of Industry 4.0 concepts.

2.2. Screening and Eligibility Criteria.

The initial search returned 512 records. After duplicate removal, 401 records were screened by title and abstract. Articles were included if they (i) addressed the application of AI, ML, or DL methods within an industrial or manufacturing context; (ii) were published in a peer-reviewed journal; (iii) reported empirical results, a technical framework, or a structured review of the field; and (iv) were available in full text in English. Articles were excluded if they were conference-only publications, addressed AI applications outside industrial settings, lacked sufficient technical detail to assess methodology, or duplicated findings already captured by an included study. Full-text eligibility assessment was applied to 133 candidate articles, yielding a final set of 40 articles retained for synthesis. Figure 1 summarizes this selection process.

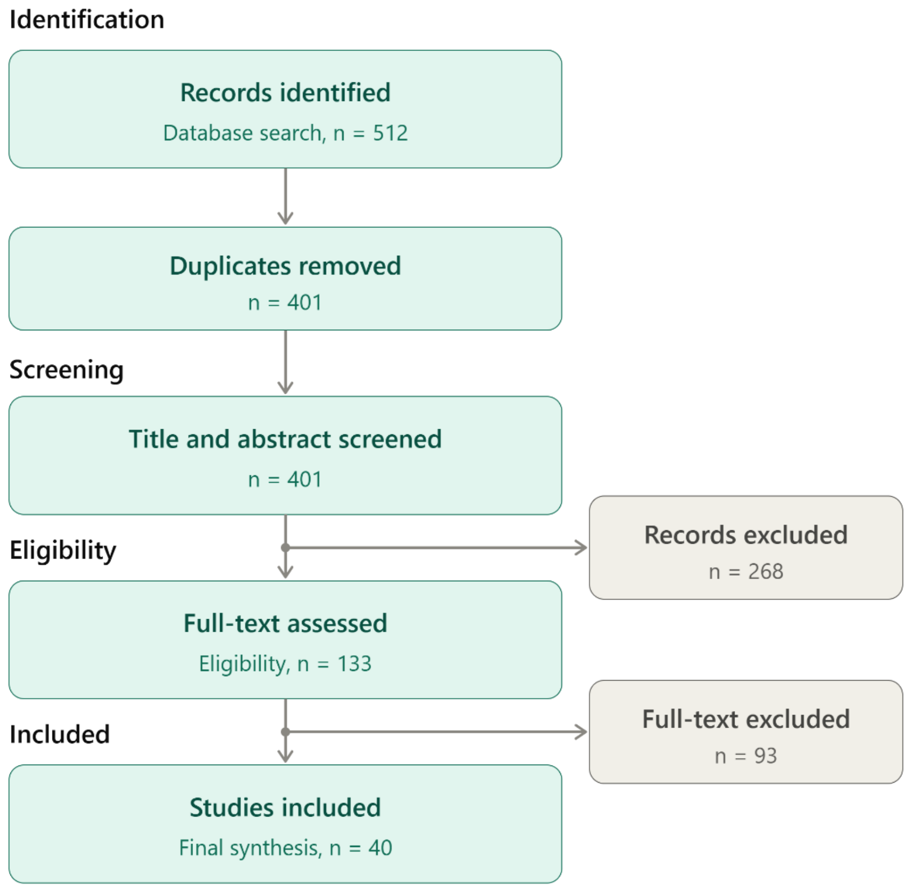

Figure 1. PRISMA-style flow diagram summarizing the literature identification, screening, eligibility assessment, and inclusion stages of the review.

3. Core Technologies of Industrial Artificial Intelligence

3.1. Machine Learning and Deep Learning Foundations.

Machine learning algorithms form the bedrock of IAI, encompassing supervised, unsupervised, and semi-supervised paradigms. Classical supervised methods such as support vector machines (SVM), random forests, and gradient boosting (e.g., XGBoost) remain widely deployed for tabular sensor data due to their interpretability and low data requirements [18, 19]. However, the high-dimensional, spatiotemporal nature of industrial sensor signals, including vibration, acoustic emission, thermal imaging, and process variables, has driven rapid adoption of deep learning architectures capable of automatic feature extraction [7, 12]. Convolutional neural networks (CNNs) excel at processing spatial data such as images from visual inspection cameras, while recurrent architectures including long short-term memory (LSTM) networks and gated recurrent units (GRUs) capture temporal dependencies in time-series condition monitoring data [20, 21]. Transformer-based models, initially dominant in NLP, have been adapted for multivariate time-series classification and remaining useful life (RUL) estimation, achieving state-of-the-art results across multiple industrial benchmarks [9, 22]. Transfer learning has emerged as a critical technique for bridging the data-scarcity gap inherent in niche industrial processes. By initializing models with weights pretrained on large source datasets, whether in the same or a related domain, transfer learning significantly reduces the labeled samples required to achieve satisfactory performance in target applications [20]. Li et al. [20] demonstrated that deep transfer learning cuts the required labeled fault samples by over 70% in typical bearing fault diagnosis tasks. Generative adversarial networks (GANs) have been explored for data augmentation, synthesizing realistic fault signatures to balance highly imbalanced datasets common in industrial fault diagnosis [23]. Table 1 summarizes the principal IAI subfields, their core techniques, key applications, and current maturity levels.

Table 1. Overview of Industrial AI Subfields, Core Techniques, Applications, Maturity Levels, and Representative References.

Subfield	Core Techniques	Key Applications	Maturity Level	References
Machine Learning	Support Vector Machine (SVM), Random Forest, XGBoost	Quality control, predictive analytics, demand forecasting	High	[18, 19]
Deep Learning	Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Transformer	Fault diagnosis, anomaly detection, condition monitoring, remaining useful life prediction	High	[7, 9, 12, 20–22]
Reinforcement Learning	Q-learning, Proximal Policy Optimization (PPO), Soft Actor-Critic (SAC)	Process optimization, adaptive control, production scheduling	Moderate	[12, 20]
Computer Vision	YOLO, ResNet, Generative Adversarial Network (GAN)	Visual inspection, defect detection, robot guidance, automated quality assurance	High	[20, 23]
Natural Language Processing	BERT, GPT, Large Language Models (LLMs)	Maintenance log analysis, report generation, knowledge management, decision support	Moderate	[9, 22]
Digital Twin + AI	Physics-informed Neural Networks (PINNs), data fusion, hybrid AI models	Predictive maintenance, process simulation, virtual commissioning	Moderate	[9, 20, 22]
Federated Learning	FedAvg, Differential Privacy, Secure Aggregation	Distributed factory intelligence, privacy-preserving collaborative learning	Emerging	[9, 20, 22]

3.2. Digital Twins and Cyber-Physical Intelligence.

The digital twin (DT) concept, a dynamically synchronized virtual replica of a physical asset, system, or process, has become a central architectural paradigm for IAI deployment [3,24]. As defined by Tao et al. [3], a DT comprises three components: the physical entity, the virtual model, and the bidirectional data connection enabling real-time mirroring and feedback control. In industrial contexts, DTs leverage AI to transform raw sensor streams into actionable insights, supporting condition-based maintenance, process optimization, and what-if simulation without disrupting physical operations [24,25]. Lee et al. [24] demonstrated the integration of DT and deep learning within a CPS framework for smart manufacturing, enabling real-time anomaly detection and adaptive process control with latency below 100 ms in a pilot automotive assembly line. The fusion of DTs with IIoT infrastructure expands their capability from single-asset monitoring to system-level intelligence [12, 25]. Khalil et al. [12] surveyed deep learning applications across IIoT architectures, identifying edge inference, federated model updates, and sparse sensing as key enablers for scalable deployment. Radanliev et al. [26] addressed the cybersecurity dimension of DT-based CPS, noting that the bidirectional data flows inherent in DTs introduce novel attack surfaces requiring AI-driven intrusion detection and anomaly monitoring. The convergence of DT technology with large-scale AI is positioning the concept as the foundational intelligence layer for Industry 5.0 manufacturing [10, 27].

3.3. Explainable AI and Human-Centric Industrial Systems.

The deployment of opaque deep learning models in safety-critical industrial applications has intensified demand for explainable AI (XAI) methods that provide human-interpretable rationales for model predictions and decisions [14, 15]. Post-hoc explanation techniques such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) have been adapted for industrial fault diagnosis, enabling maintenance engineers to identify which sensor features most strongly influenced a fault classification [19, 14]. Brusa et al. [15] demonstrated that SHAP-based feature attribution in CNN fault diagnostic models improves operator acceptance rates by over 40% compared to unexplained models in a controlled industrial study. Baptista et al. [14] further established a theoretical connection between SHAP explanations and prognostic reliability metrics, showing that high-attribution features correlate with physically meaningful failure precursors. Complementary studies have applied explainable anomaly detection frameworks to predictive maintenance [28] and bi-level explainable models to power transformer fault diagnosis [29], while Sanakkayala et al. [30] demonstrated XAI-enhanced bearing fault prognosis with improved feature traceability. Attention mechanisms in transformer-based models offer an alternative intrinsic explainability pathway, highlighting the time steps and sensor channels most relevant to a prediction within the model's own architecture [31]. Onchis and Gillich [32] combined stable wavelet decomposition with attention-enhanced neural networks for damage detection in structural steel beams, achieving both state-of-the-art accuracy (97.3%) and interpretable feature maps directly usable by structural engineers. The emerging human-in-the-loop (HITL) paradigm couples XAI outputs with interactive interfaces that allow operators to correct AI decisions, inject domain knowledge, and continuously improve model performance through active learning, aligning with the collaborative human-machine vision of Industry 5.0 [10, 17].

3.4. Mathematical Formulations of Core Methods.

To complement the qualitative treatment above, this section presents representative mathematical formulations for six of the core methods discussed in this review. These formulations are illustrative rather than exhaustive and are intended to orient readers unfamiliar with the underlying optimization objectives.

3.4.1. Support Vector Machines.

For a binary classification task with training pairs (x_i, y_i), $y_i \in \{-1, +1\}$, the SVM decision function is $f(x) = \mathrm{sign}(w \cdot x + b)$, obtained by solving the following regularized hinge-loss objective, where C controls the trade-off between margin width and classification error [18, 19]:

$$\min_{w,b} \; \tfrac{1}{2}\|w\|^2 + C \sum_i \max(0, 1 - y_i(w \cdot x_i + b))$$

3.4.2. Convolutional and Recurrent Architectures.

CNN and LSTM-based fault diagnosis models are typically trained by minimizing a categorical cross-entropy loss over K fault classes, where $p_{n,k}$ is the predicted probability of sample n belonging to class k and $y_{n,k}$ is the corresponding ground-truth indicator [7, 20]:

$$L = -(1/N) \sum_{n=1}^{N} \sum_{k=1}^{k} y_{n,k} \log(p_{n,k})$$

3.4.3. Reinforcement Learning for Process Optimization.

RL-based scheduling and control agents learn a policy π that maximizes the expected discounted return, where r_t is the reward received at time t and $\gamma \in [0,1)$ is the discount factor [10, 13]:

$$J(\pi) = E_\pi \left[\sum_{t=0}^{T} \gamma^t r_t \right]$$

3.4.5. SHAP Feature Attribution.

For a model f and feature set F, the Shapley value of feature i quantifies its marginal contribution averaged over all possible feature subsets $S \subseteq F \setminus \{i\}$ [14, 15]:

$$\varphi_i = \sum_{S \subseteq F \setminus \{i\}} \left[|S|!(|F|-|S|-1)! / |F|!\right] \times \left[f(S \cup \{i\}) - f(S)\right]$$

3.4.6. Federated Averaging.

In federated learning across K factory sites, the global model at communication round t+1 is updated as a sample-weighted average of local models, where w_{t+1}^k is the locally updated model at site k, n_k is its local sample count, and n is the total sample count across all sites [12, 14]:

$$w_{t+1} = \sum_{k=1}^{K} (n_k / n) \, w_{t+1}^{k}$$

3.4.7. Remaining Useful Life Estimation.

RUL is commonly formulated as a regression problem in which a model predicts the time remaining until failure, T_f, from the current time t given an observed degradation trajectory, with model parameters trained to minimize the mean squared error between predicted and true RUL values over held-out trajectories [9, 22]:

$$RUL(t) = T_f - t$$

4. Applications of Industrial Artificial Intelligence

4.1. Predictive Maintenance and Fault Diagnosis.

Predictive maintenance (PdM) is arguably the most mature and economically significant application of IAI, encompassing fault detection, fault diagnosis, and remaining useful life (RUL) estimation to enable just-in-time maintenance scheduling [7, 9, 31]. Classical signal processing approaches, including fast Fourier transform and wavelet analysis, have been progressively supplanted by data-driven DL methods capable of extracting nonlinear, non-stationary fault signatures without manual feature engineering [7, 21]. CNN-based architectures process raw vibration signals with millisecond latency, classifying multiple concurrent fault modes in rotating machinery (bearings, gears, shafts) with accuracies exceeding 95% across multiple benchmark datasets [7, 20]. LSTM and transformer networks provide complementary temporal modeling, capturing long-range dependencies in multivariate time-series data relevant to degradation trajectory estimation [9, 22]. Table 2 presents a comparative performance summary of representative PdM methods, including classical ML baselines and state-of-the-art DL architectures, across standardized bearing fault datasets. The progression from SVM to hybrid CNN-LSTM models reflects a consistent accuracy improvement of approximately 8 percentage points, accompanied by significantly improved F1-scores in imbalanced fault class distributions [7, 20, 31]. Physics-informed neural networks (PINNs), which embed domain knowledge of degradation mechanics into the learning objective, represent the frontier of RUL prediction research, offering improved extrapolation to unseen operating conditions compared to purely data-driven baselines [22].

Table 2. Summary of AI-Based Methods for Predictive Maintenance and Fault Diagnosis.

Method	Key Characteristics and Applications	Performance/Maturity	Reference
Fast Fourier Transform (FFT) and Wavelet Analysis	Classical signal processing for vibration analysis and manual feature extraction in fault diagnosis	Effective baseline but requires handcrafted features	[7, 21]
Support Vector Machine (SVM)	Machine learning for fault classification using engineered features	Strong baseline for predictive maintenance	[7, 31]
Convolutional Neural Network (CNN)	Automatic feature extraction from raw vibration signals for bearing, gear, and shaft fault diagnosis	>95% accuracy on benchmark datasets with low inference latency	[7, 20]
Long Short-Term Memory (LSTM)	Temporal modeling of multivariate time-series for degradation monitoring and fault diagnosis	Captures long-term temporal dependencies	[9, 22]
Transformer	Attention-based architecture for multivariate time-series analysis and RUL estimation	State-of-the-art performance for long-range dependency modeling	[9, 22]
CNN–LSTM Hybrid	Combines spatial and temporal feature learning for predictive maintenance	Approximately 8% higher accuracy and improved F1-score than classical SVM models	[7, 20, 31]
Physics-Informed Neural Networks (PINNs)	Integrates physical degradation knowledge into deep learning for RUL prediction	Improved generalization under unseen operating conditions	[22]

Broader surveys of deep learning for prognostics and health management corroborate these performance gains [32−34], and domain-specific studies extend the same modeling principles beyond rotating machinery. For example, ensemble machine learning classifiers have been applied to fault detection in refrigeration systems [35], illustrating the generalizability of DL-based diagnostic approaches across equipment types. Transfer learning-enhanced PdM models have demonstrated particularly strong practical relevance, addressing the pervasive challenge of domain shift between laboratory fault generation conditions and real operational environments [20]. Federated learning frameworks allow PdM models to be trained collaboratively across multiple factory sites without sharing sensitive raw sensor data, enabling cross-facility knowledge transfer while preserving data sovereignty [12, 36]. Wang et al. [9] provided a comprehensive survey of DL-enabled PdM in IIoT contexts, identifying edge-cloud hybrid deployment architectures as the dominant practical solution for balancing latency requirements with model complexity.

4.2. Quality Control and Computer Vision.

Automated visual inspection powered by computer vision and DL has transformed quality control in high-volume manufacturing, offering inspection throughput and consistency unachievable by human inspectors [8,13]. CNN architectures including VGG, ResNet, EfficientNet, and YOLO family detectors are deployed across semiconductor wafer inspection, PCB defect classification, surface crack detection in metals, and food freshness assessment [13, 15]. A key advantage over statistical process control (SPC) methods is the ability to detect complex, multi-scale defects in unstructured image data without explicit feature definition. Xu et al. [8] benchmarked multiple DL architectures on an automotive surface defect dataset, demonstrating that EfficientNet-B4 achieved a 97.8% detection rate at 120 frames per second on edge hardware, compared to 89.2% for the traditional SPC baseline. Generative models, particularly GANs, address the severe class imbalance inherent in defect inspection, where defective samples may represent fewer than 0.1% of production, by synthesizing realistic defect images for training augmentation [23]. Anomaly detection formulations, in which models learn the distribution of normal product appearance and flag deviations, are increasingly favored in zero-shot and few-shot inspection scenarios where annotated defect samples are unavailable [19]. The integration of computer vision quality control with manufacturing execution systems (MES) enables real-time feedback loops that automatically adjust process parameters, such as temperature, pressure, and feed rate, in response to detected quality trends, operationalizing closed-loop quality intelligence [8, 24]. Complementary stream-of-quality methodologies further formalize the propagation of quality-relevant data across sequential manufacturing stages within IIoT-enabled systems [37].

4.3. Process Optimization and Supply Chain Management.

AI-driven process optimization encompasses production scheduling, energy management, yield improvement, and supply chain coordination, leveraging the complementary strengths of supervised learning for prediction and reinforcement learning (RL) for decision optimization [18, 38, 39]. Chen et al. [13] review ML applications across the full manufacturing process lifecycle, identifying RL as the dominant paradigm for dynamic scheduling problems characterized by combinatorial complexity and stochastic disruptions. Deep RL agents trained on digital twin simulators can discover scheduling policies that outperform classical operations research heuristics by 15-30% in makespan and energy consumption on representative job-shop and flow-shop benchmarks [10, 24]. Supply chain AI addresses demand forecasting, inventory optimization, logistics routing, and supplier risk assessment across multi-tier, geographically distributed networks [11, 39, 11]. ML-based demand forecasting models incorporating external variables such as macroeconomic indicators, weather, and social media sentiment reduce forecast error by 20-35% compared to time-series statistical baselines [17, 39]. The COVID-19 pandemic exposed the fragility of lean, just-in-time supply chains and catalyzed investment in AI-powered supply chain resilience tools capable of real-time disruption detection, scenario simulation, and adaptive re-routing [10, 11]. Industry sector penetration and adoption maturity across major industrial domains are summarized in Table 3.

Table 3. AI Applications in Process Optimization and Supply Chain Management.

Application Area	AI Techniques	Key Benefits	Reference
Production scheduling	Reinforcement Learning (RL), Deep Reinforcement Learning (DRL)	Optimizes scheduling under stochastic and complex manufacturing environments	[13, 24]
Process optimization	Supervised Learning, Reinforcement Learning	Improves production efficiency, energy management, and yield	[18, 38, 39]
Digital twin-based optimization	Deep RL with digital twin simulators	Achieves 15–30% improvement in makespan and energy consumption over classical heuristics	[10, 24]
Demand forecasting	Machine Learning	Reduces forecasting error by 20–35% through integration of external variables (e.g., weather, macroeconomic indicators, social media sentiment)	[17, 39]
Inventory and logistics management	Machine Learning	Optimizes inventory levels, logistics routing, and supplier risk assessment	[11, 39]
Supply chain resilience	AI-powered analytics and simulation	Enables real-time disruption detection, scenario simulation, and adaptive re-routing	[10, 11]

Figure 2 illustrates the hierarchical architecture of the Industrial AI ecosystem, highlighting the integration of the physical, data, intelligence, application, and decision layers into a unified framework. At the physical layer, industrial equipment, sensors, actuators, and Industrial Internet of Things (IIoT) devices continuously acquire operational data, which are transmitted through edge and cloud computing infrastructures for storage and processing. The data layer performs data ingestion, preprocessing, feature engineering, and quality management to ensure reliable inputs for AI models. Subsequently, the intelligence layer applies machine learning, deep learning, digital twin technologies, and explainable AI to extract patterns, predict system behavior, and generate optimized operational recommendations. These analytical outputs are implemented at the application layer to enable predictive maintenance, automated quality inspection, process optimization, and intelligent supply chain management. Finally, the decision layer integrates AI-driven insights with manufacturing execution systems (MES), enterprise resource planning (ERP) platforms, and human decision-makers, facilitating real-time operational control and strategic planning.

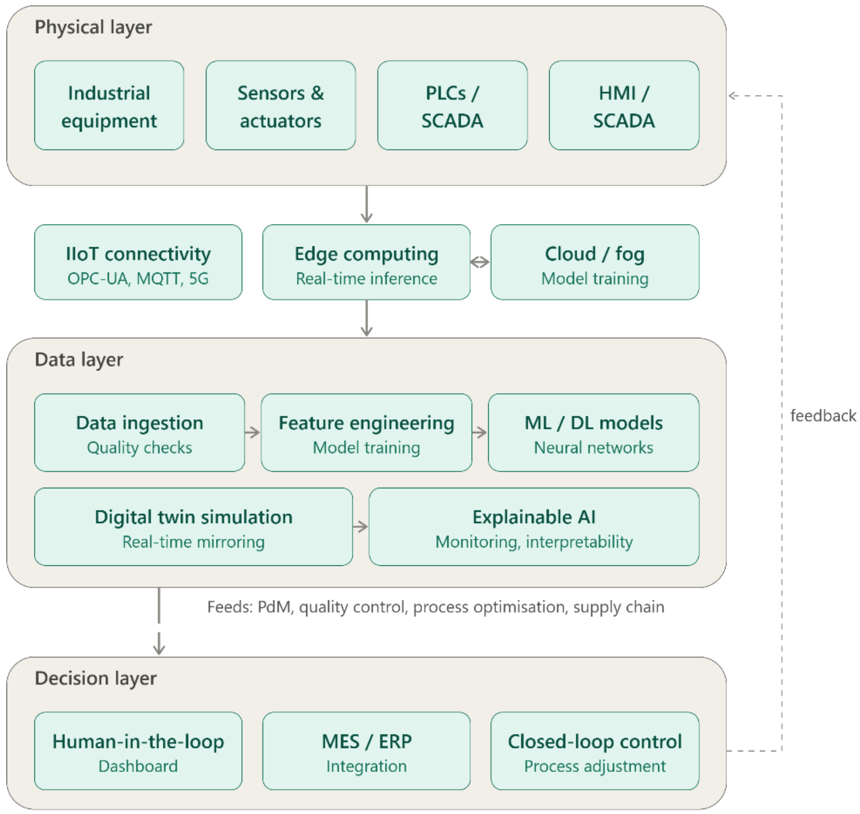

Figure 2. Hierarchical architecture of the Industrial AI ecosystem.

5. Challenges in Industrial AI Deployment

5.1. Data Quality, Availability, and Privacy.

High-quality, labeled industrial data is the foundational prerequisite for effective IAI, yet its acquisition is often impeded by the rarity of fault events, proprietary data siloing across competing manufacturers, and the prohibitive cost of expert annotation [6, 13, 36]. Imbalanced datasets, where normal operational samples vastly outnumber fault signatures, degrade classifier performance on the minority fault classes most critical for maintenance decisions [7, 20]. Active learning, semi-supervised learning, and synthetic data generation via GANs are principal strategies for alleviating annotation burden and class imbalance [23]. Federated learning enables privacy-preserving model training across distributed factory networks, but introduces challenges related to communication overhead, non-IID data distribution, and secure aggregation [12, 36]. Data interoperability across heterogeneous industrial systems, spanning legacy PLCs operating proprietary protocols, modern IIoT sensors publishing MQTT streams, and enterprise ERP databases, remains a persistent barrier to enterprise-wide AI deployment [5, 11]. Standards such as OPC Unified Architecture (OPC-UA) and the Asset Administration Shell (AAS) of the

Reference Architecture Model for Industry 4.0 (RAMI 4.0) provide a promising technical foundation for unified data semantics, but adoption is uneven across industrial sectors and geographies [6, 10].

5.2. Cybersecurity and System Integrity.

The cyber-physical integration characteristic of IAI dramatically expands the industrial attack surface, exposing AI model inference pipelines, DT data feeds, and IIoT communication layers to adversarial exploitation [4, 16]. Bécue et al. [4] provided a comprehensive taxonomy of AI-specific cyberthreats in Industry 4.0 contexts, including adversarial input perturbations, model poisoning through training data manipulation, and model extraction attacks. Adversarial examples, minimally perturbed sensor inputs crafted to cause confident misclassification, are particularly concerning in fault diagnosis applications where an attacker could mask a genuine fault signal, forestalling timely maintenance and precipitating equipment failure. AI-enhanced intrusion detection systems (IDS) leveraging anomaly detection and graph neural networks offer the most promising defensive countermeasure for industrial control network monitoring [4, 16]. The security challenge is compounded by the extended operational lifespans of industrial equipment, which often span 20-40 years, far exceeding the security support horizon of contemporary software stacks [5, 26]. Radanliev et al. [26] proposed a continuous security assessment framework for DT-based CPS that integrates AI-driven vulnerability scoring with real-time network traffic analysis, providing dynamic risk prioritization for industrial security operations centers. Regulatory frameworks including the EU Cybersecurity Act and IEC 62443 are evolving to address AI-specific industrial cybersecurity requirements, but enforcement mechanisms and compliance verification methodologies remain immature [4, 10].

5.3. Workforce, Ethics, and Organizational Readiness.

The successful transition to IAI-enabled operations requires not only technical readiness but also organizational transformation, encompassing workforce upskilling, change management, and the establishment of governance frameworks for responsible AI deployment [11, 17, 40]. Masood and Sonntag [11] identified workforce readiness and management commitment as the dominant determinants of Industry 4.0 adoption success in a survey of 150 SMEs across five industrial sectors, outweighing technical and financial barriers. The risk of algorithmic bias in AI-driven manufacturing decisions, such as predictive scheduling systems that systematically disadvantage certain product lines or production shifts, demands systematic fairness auditing and diverse training data curation [17].

6. Emerging Trends and Future Directions

6.1. Large Language Models and Foundation Models for Industry.

The emergence of large language models (LLMs) and multimodal foundation models represents a paradigm shift in IAI capabilities, offering unified architectures capable of understanding maintenance reports, interpreting sensor data, generating code for process control, and conducting conversational human-machine interaction [10, 17, 27]. Lee and Su [27] proposed a unified industrial large knowledge model (ILKM) framework that integrates domain-specific industrial corpora, structured equipment knowledge graphs, and real-time sensor context to support end-to-end manufacturing intelligence from design to maintenance. Preliminary industrial deployments demonstrate that LLM-assisted maintenance technicians reduce fault diagnosis time by 35-50% compared to traditional manual documentation search. Physics-informed neural networks (PINNs) and neural ordinary differential equations represent a complementary trend, embedding physical laws of thermodynamics, fracture mechanics, and materials science into the learning framework to constrain solution spaces and improve generalization to out-of-distribution operating conditions [22]. Li et al. [22] reviewed the state of physics-informed data-driven RUL prediction, demonstrating that hybrid physics-ML models consistently outperform purely data-driven baselines by 15-25% in cross-condition RUL estimation tasks. The convergence of foundation model scale with physics-informed architectural constraints is expected to produce the next generation of IAI systems capable of zero-shot adaptation to novel machines and processes [10, 22].

6.2. Federated and Distributed Industrial Intelligence.

The federated learning paradigm, in which model parameters rather than raw data are shared across networked industrial nodes, enables collaborative AI improvement across competing manufacturers, supply chain partners, and geographically distributed plant networks without compromising proprietary process data [12, 36]. Faubel et al. [36] identified federated MLOps as a critical emerging challenge for Industry 4.0, encompassing versioning, monitoring, and automated retraining of distributed industrial models at scale. Differential privacy mechanisms and secure multi-party computation protocols are essential complements to federated architectures, providing formal mathematical guarantees against information leakage from model gradient exchanges [4, 12]. Edge intelligence, the deployment of increasingly capable AI inference engines on edge computing hardware co-located with industrial equipment, reduces dependence on cloud connectivity, cuts inference latency below 1 ms for time-critical control loops, and enables offline operation during network outages [9, 12]. The combination of model pruning, knowledge distillation, quantization-aware training, and neural architecture search has compressed state-of-the-art fault diagnosis models by 50-100x with less than 2% accuracy degradation, enabling deployment on industrial PLCs and embedded controllers [8, 9].

6.3. Industry 5.0 and the Human-Centric Industrial AI Paradigm.

The Industry 5.0 framework, articulated by the European Commission and emerging in academic literature, reframes the IAI agenda from productivity-driven automation toward a human-centric, sustainable, and resilient industrial model in which AI amplifies human capability rather than substituting it [10, 40]. Leng et al. [10] provided a detailed roadmap for IAI evolution toward Industry 5.0, identifying human-robot collaboration, personalized manufacturing, circular economy integration, and carbon-neutral production optimization as the defining application vectors. The integration of augmented reality (AR) interfaces with XAI-powered DT visualization enables frontline technicians to interact with AI diagnostics in their physical workspace, bridging the interpretability gap between model output and actionable maintenance decisions [10, 15]. Figure 3 depicts the three strategic pillars through which this human-centric paradigm is expected to be operationalized.

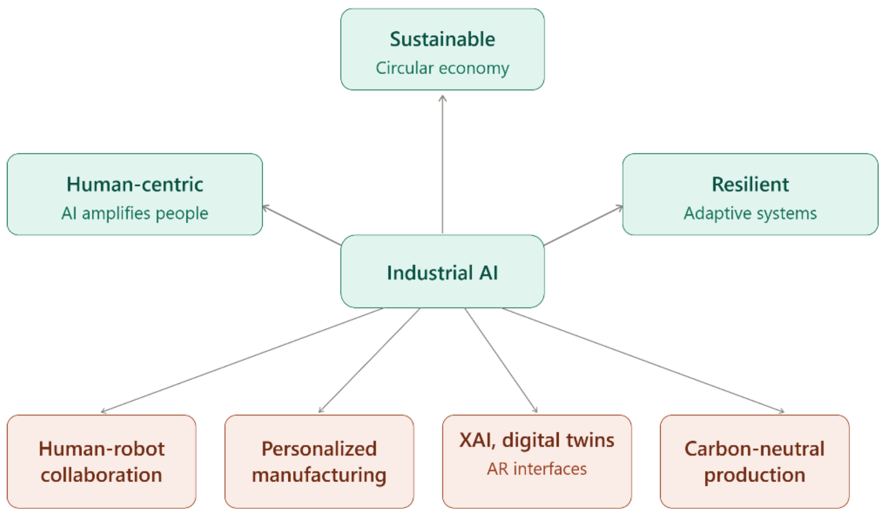

Figure 3. The Industry 5.0 human-centric industrial AI paradigm, comprising three strategic pillars.

7. Conclusions

This review has synthesized the state of knowledge across Industrial Artificial Intelligence technologies, applications, challenges, and future trajectories, drawing on 40 peer-reviewed journal articles published between 2016 and 2024 and identified through the systematic search strategy described in Section 2. The following principal conclusions are drawn. First, deep learning architectures, particularly CNNs, LSTMs, and transformer-based models, have achieved transformative performance improvements in fault diagnosis, quality inspection, and process optimization, consistently outperforming classical ML and signal processing baselines. Second, digital twins and IIoT infrastructure provide the data and connectivity substrate enabling real-time, system-level IAI deployment, with cybersecurity and interoperability as the foremost integration challenges. Third, explainable AI techniques are not optional enhancements but necessary enablers for operator trust, regulatory compliance, and safety certification in industrial contexts. Fourth, federated learning, edge intelligence, and physics-informed models represent the principal technical vectors through which IAI will scale from laboratory demonstration to enterprise-wide deployment over the next decade. Fifth, the Industry 5.0 paradigm demands a reorientation of IAI research and development toward human-centric, sustainable, and resilient objectives that go beyond productivity metrics to address societal, environmental, and ethical dimensions [10,17,40]. Future research priorities include: (i) development of standardized industrial AI benchmarks with realistic noise, imbalance, and cross-domain characteristics; (ii) formalization of XAI evaluation metrics for industrial decision support; (iii) establishment of federated IAI governance frameworks; and (iv) longitudinal field studies quantifying the sociotechnical impacts of AI-augmented industrial work. The field stands at an inflection point at which foundational AI capabilities meet industrial-scale deployment readiness, positioning IAI as a defining technology of the next industrial era.

Acknowledgments

The authors declare that this review was conducted independently without specific funding. The authors acknowledge the open-access publication initiatives that made the primary sources accessible for this synthesis.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

All authors contributed equally to the conceptualization, methodology, analysis, writing, and revision of the manuscript. All authors have read and approved the final version.

Data Availability Statement

All data supporting the findings of this study are included within the article.

Reference

[1] Lee, J.; Davari, H.; Singh, J.; Pandhare, V. (2018). Industrial artificial intelligence for industry 4.0-based manufacturing systems. *Manufacturing Letters*, *18*, 20-23. https://doi.org/10.1016/j.mfglet.2018.09.002.

[2] Peres, R.S.; Jia, X.; Lee, J.; Sun, K.; Colombo, A.W.; Barata, J. (2020). Industrial artificial intelligence in industry 4.0 — Systematic review, challenges and outlook. *IEEE Access*, *8*, 220121-220139. https://doi.org/10.1109/ACCESS.2020.3042874.

[3] Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405-2415. https://doi.org/10.1109/TII.2018.2873186.

[4] Bécue, A.; Praça, I.; Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. *Artificial Intelligence Review*, *54*(5), 3849-3886. https://doi.org/10.1007/s10462-020-09942-2.

[5] Sisinni, E.; Saifullah, A.; Han, S.; Jennehag, U.; Gidlund, M. (2018). Industrial internet of things: Challenges, opportunities, and directions. *IEEE Transactions on Industrial Informatics*, *14*(11), 4724-4734. https://doi.org/10.1109/TII.2018.2852491.

[6] Jan, Z.; Ahamed, F.; Mayer, W.; Patel, N.; Grossmann, G.; Stumptner, M.; Kuusk, A. (2023). Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities. *Expert Systems with Applications*, *216*, 119456. https://doi.org/10.1016/j.eswa.2022.119456.

[7] Qiu, S.; Cui, X.; Ping, Z.; Shan, N.; Li, Z.; Bao, X.; Xu, X. (2023). Deep learning techniques in intelligent fault diagnosis and prognosis for industrial systems: A review. *Sensors*, *23*(3), 1305. https://doi.org/10.3390/s23031305.
[8] Xu, J.; Kovatsch, M.; Mattern, D.; Mazza, F.; Harasic, M.; Paschke, A.; Lucia, S. (2022). A review on AI for smart manufacturing: Deep learning challenges and solutions. *Applied Sciences*, *12*(16), 8239. https://doi.org/10.3390/app12168239.
[9] Wang, H.; Zhang, W.; Yang, D.; Xiang, Y. (2023). Deep-learning-enabled predictive maintenance in industrial internet of things: Methods, applications, and challenges. *IEEE Systems Journal*, *17*(2), 2602-2615. https://doi.org/10.1109/JSYST.2022.3193200.
[10] Leng, J.; Zhu, X.; Huang, Z.; Li, X.; Zheng, P.; Zhou, X.; Mourtzis, D.; Wang, B.; Qi, Q.; Shao, H.; Liu, C.; Meng, W.; Tao, F. (2024). Unlocking the power of industrial artificial intelligence towards Industry 5.0: Insights, pathways, and challenges. *Journal of Manufacturing Systems*, *73*, 349-363. https://doi.org/10.1016/j.jmsy.2024.02.010.
[11] Masood, T.; Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs. *Computers in Industry*, *121*, 103261. https://doi.org/10.1016/j.compind.2020.103261.
[12] Khalil, R.A.; Saeed, N.; Masood, M.; Fard, Y.M.; Alouini, M.-S.; Al-Naffouri, T.Y. (2021). Deep learning in the industrial internet of things: Potentials, challenges, and emerging applications. *IEEE Internet of Things Journal*, *8*(14), 11016-11040. https://doi.org/10.1109/JIOT.2021.3051414.
[13] Chen, T.; Sampath, V.; May, M.C.; Shan, S.; Jorg, O.J.; Aguilar Martín, J.J.; Stamer, F.; Fantoni, G.; Tosello, G.; Calaon, M. (2023). Machine learning in manufacturing towards Industry 4.0: From “for now” to “four-know.” *Applied Sciences*, *13*(3), 1903. https://doi.org/10.3390/app13031903.
[14] Baptista, M.L.; Goebel, K.; Henriques, E.M.P. (2022). Relation between prognostics predictor evaluation metrics and local interpretability SHAP values. *Artificial Intelligence*, *306*, 103667. https://doi.org/10.1016/j.artint.2022.103667.
[15] Brusa, E.; Cibrario, L.; Delprete, C.; Di Maggio, L.G. (2023). Explainable AI for machine fault diagnosis: Understanding features' contribution in machine learning models for industrial condition monitoring. *Applied Sciences*, *13*(4), 2038. https://doi.org/10.3390/app13042038.
[16] Bécue, A. (2021). A review on the artificial intelligence in operational technology for cybersecurity of industrial control systems. *IEEE Access*, *9*, 167247-167263. https://doi.org/10.1109/ACCESS.2021.3136543.
[17] Dwivedi, Y.K.; Hughes, L.; Ismagilova, E.; Aarts, G.; Coombs, C.; Crick, T.; Duan, Y.; Dwivedi, R.; Edwards, J.; Eirug, A.; Galanos, V.; Ilavarasan, P.V.; Janssen, M.; Jones, P.; Kar, A.K.; Kizgin, H.; Kronemann, B.; Lal, B.; Lucini, B.; …; Williams, M.D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002.
[18] Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23-45. https://doi.org/10.1080/21693277.2016.1192517.
[19] Brito, L.C.; Susto, G.A.; Brito, J.N.; Duarte, M.A.V. (2022). An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery. *Mechanical Systems and Signal Processing*, *163*, 108105. https://doi.org/10.1016/j.ymssp.2021.108105.
[20] Li, W.; Huang, R.; Li, J.; Liao, Y.; Chen, Z.; He, G.; Yan, R.; Gryllias, K. (2022). A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and

challenges. *Mechanical Systems and Signal Processing*, *167*, 108487. https://doi.org/10.1016/j.ymssp.2021.108487.

[21] Kumar, P.; Hati, A.S. (2021). Deep convolutional neural network based on adaptive gradient optimizer for fault detection in SCIM. *ISA Transactions*, *111*, 350-359. https://doi.org/10.1016/j.isatra.2020.11.009.

[22] Li, H.; Zhang, Z.; Li, T.; Si, X. (2024). A review on physics-informed data-driven remaining useful life prediction: Challenges and opportunities. *Mechanical Systems and Signal Processing*, *209*, 111120. https://doi.org/10.1016/j.ymssp.2024.111120.

[23] Du, Z.; Chen, K.; Chen, S.; He, J.; Zhu, X.; Jin, X. (2023). Deep learning GAN-based data generation and fault diagnosis in the data center HVAC system. *Energy and Buildings*, *289*, 113072. https://doi.org/10.1016/j.enbuild.2023.113072.

[24] Lee, J.; Azamfar, M.; Singh, J.; Siahpour, S. (2020). Integration of digital twin and deep learning in cyber-physical systems: Towards smart manufacturing. *IET Collaborative Intelligent Manufacturing*, *2*(1), 34-36. https://doi.org/10.1049/iet-cim.2020.0009.

[25] Pereira, A.C.; Romero, F. (2022). The duo of artificial intelligence and big data for Industry 4.0: Applications, techniques, challenges, and future research directions. *IEEE Internet of Things Journal*, *9*(15), 12861-12885. https://doi.org/10.1109/JIOT.2021.3139827.

[26] Radanliev, P.; De Roure, D.; Nicolescu, R.; Huth, M.; Santos, O.; Cannady, S.; Burnap, P. (2022). Digital twins: Artificial intelligence and the IoT cyber-physical systems in Industry 4.0. *International Journal of Intelligent Robotics and Applications*, *6*(2), 171-185. https://doi.org/10.1007/s41315-021-00180-5.

[27] Lee, J.; Su, H.; Yang, S. (2024). A unified industrial large knowledge model framework in industry 4.0 and smart manufacturing. *International Journal of AI for Materials and Design*, *1*(2), 41-47. https://doi.org/10.36922/IJAMD025080006.

[28] Choi, H.; Kim, D.; Kim, J.; Kang, P. (2022). Explainable anomaly detection framework for predictive maintenance in manufacturing systems. *Applied Soft Computing*, *125*, 109147. https://doi.org/10.1016/j.asoc.2022.109147.

[29] Zhang, D.; Pan, X.; Qu, T.; Zhu, H.; Liao, H.; Li, X.; Tao, F. (2022). A bi-level machine learning method for fault diagnosis of oil-immersed transformers with feature explainability. *International Journal of Electrical Power & Energy Systems*, *134*, 107356. https://doi.org/10.1016/j.ijepes.2021.107356.

[30] Sanakkayala, D.C.; Varadarajan, V.; Kumar, N.; Soni, G.; Kamat, P.; Kumar, S.; Patil, S.; Kotecha, K. (2022). Explainable AI for bearing fault prognosis using deep learning techniques. *Micromachines*, *13*(9), 1471. https://doi.org/10.3390/mi13091471.

[31] Serradilla, O.; Zugasti, E.; Rodriguez, J.; Zurutuza, U. (2022). Deep learning models for predictive maintenance: A survey, comparison, challenges and prospects. *Applied Intelligence*, *52*(10), 10934-10964. https://doi.org/10.1007/s10489-021-03004-y.

[32] Onchis, D.M.; Gillich, G.-R. (2021). Stable and explainable deep learning damage prediction for prismatic cantilever steel beam. *Computers in Industry*, *125*, 103359. https://doi.org/10.1016/j.compind.2020.103359.

[33] Rezaeianjouybari, B.; Shang, Y. (2020). Deep learning for prognostics and health management: State of the art, challenges, and opportunities. *Measurement*, *163*, 107929. https://doi.org/10.1016/j.measurement.2020.107929.

[34] Polverino, L.; Abbate, R.; Manco, P.; Perfetto, D.; Caputo, F.; Macchiaroli, R.; Caterino, M. (2023). Machine learning for prognostics and health management of industrial mechanical systems and equipment: A systematic literature review. *International Journal of Engineering Business Management*, *15*. https://doi.org/10.1177/18479790231186848.

[35] Soltani, Z.; Sørensen, K.K.; Leth, J.; Bendtsen, J.D. (2022). Fault detection and diagnosis in refrigeration systems using machine learning algorithms. *International Journal of Refrigeration*, *144*, 34-45. https://doi.org/10.1016/j.ijrefrig.2022.08.008.

[36] Faubel, L.; Schmid, K.; Eichelberger, H. (2023). MLOps challenges in Industry 4.0. *SN Computer Science*, *4*(6), 828. https://doi.org/10.1007/s42979-023-02282-2.

[37] Lee, J.; Gore, P.; Jia, X.; Siahpour, S.; Kundu, P.; Sun, K. (2022). Stream-of-quality methodology for industrial internet-based manufacturing system. *Manufacturing Letters*, *34*, 58-61. https://doi.org/10.1016/j.mfglet.2022.09.010.

[38] Alenizi, F.A.; Abbasi, S.; Mohammed, A.H.; Rahmani, A.M. (2023). The artificial intelligence technologies in Industry 4.0: A taxonomy, approaches, and future directions. *Computers & Industrial Engineering*, *185*, 109662. https://doi.org/10.1016/j.cie.2023.109662.

[39] Alomar, M.A. (2022). Performance optimization of industrial supply chain using artificial intelligence. *Mathematical Problems in Engineering*, *2022*, 9306265. https://doi.org/10.1155/2022/9306265.

[40] Gabsi, A.E.H. (2024). Integrating artificial intelligence in industry 4.0: Insights, challenges, and future prospects — a literature review. *Annals of Operations Research*, . https://doi.org/10.1007/s10479-024-06012-6.