International Information and
Engineering Technology Association
Advancing the World of Information and Engineering

Artificial Intelligence Techniques for Industrial Predictive Maintenance: A Systematic Review of Recent Advances

Khalid Lefrouni[*] , Saoudi Taibi

Mohammadia School of Engineers, Mohammed V University in Rabat, Rabat 10000, Morocco

Corresponding Author Email: lefrouni@emi.ac.ma

ABSTRACT

Modern industrial operations are under growing pressure to maximize asset performance, reduce expensive downtime, and improve safety. Exploiting progress in sensor technology and data analytics, Predictive Maintenance (PdM) presents a forward-looking maintenance strategy, moving past conventional reactive or scheduled approaches. This review explores the innovative use of Artificial Intelligence (AI), encompassing Machine Learning (ML) and Deep Learning (DL), to significantly boost PdM efficiency. Drawing upon a methodical analysis of 29 peer-reviewed articles from the last decade, this review consolidates the current landscape, major trends, obstacles, and future outlook regarding the deployment of AI methods for industrial Predictive Maintenance. The analysis indicates a significant tendency towards utilizing DL techniques for sophisticated tasks such as Remaining Useful Life (RUL) estimation and anomaly identification. Developing fields include employing Deep Reinforcement Learning (DRL) for optimal maintenance scheduling, methods for explainability (XAI) for fostering trust, and the convergence of PdM with data-driven production planning and emerging Digital Twins. Despite substantial advancements, significant hurdles remain concerning data quality and accessibility, model interpretability, scalability, system integration, and cybersecurity. This review offers a thorough, holistic overview for researchers and industry professionals, underscoring the game-changing possibilities of AI within PdM and pinpointing key domains that warrant deeper exploration.

1. INTRODUCTION

Industrial assets, ranging from manufacturing machinery and power generation equipment to transportation systems like aircraft and vehicles, represent significant capital investments and are critical to operational continuity and economic productivity [1-3]. Maintaining these assets effectively is paramount, yet traditional maintenance strategies often fall short. Reactive maintenance, performed only after a failure occurs, leads to unplanned downtime, potential secondary damages, safety hazards, and significant economic losses [3-6]. Preventive maintenance (PvM), on the other hand, operates on predetermined schedules, such as fixed time intervals or usage counts, regardless of the actual condition of the equipment. While it aims to mitigate some failures by intervening before they are expected to occur based on these schedules, PvM often results in unnecessary interventions on healthy equipment or, conversely, fails to prevent unexpected breakdowns occurring between scheduled services [2, 3, 5, 7, 8]. The core limitation of PvM is its reliance on general statistical lifespans rather than real-time operational health.

These shortcomings of both reactive and traditional preventive strategies have spurred the transition to Predictive Maintenance (PdM). Distinct from PvM's schedule-based approach, PdM is a proactive strategy focused on anticipating potential failures by continuously monitoring the actual equipment condition and scrutinizing operational data [2, 4-6, 9, 10]. The emergence of Industry 4.0, characterized by the widespread adoption of sensors (Internet of Things - IoT), sophisticated data analysis techniques, and Artificial Intelligence (AI), has greatly enhanced the capabilities of PdM [4, 7-9, 11-15]. AI, notably Machine Learning (ML) and Deep Learning (DL), provides robust tools to handle immense volumes of intricate, high-dimensional sensor data, detect subtle indicators signaling degradation, determine fault categories, forecast the Remaining Useful Life (RUL) of parts, and refine maintenance timelines [6, 10, 13, 16]. By harnessing AI, organizations can transition to smarter, data-informed maintenance strategies, potentially realizing significant gains in asset uptime, operational effectiveness, safety, and cost savings [1-3, 11, 15, 16].

Despite the promise, the practical implementation of AI-driven PdM faces numerous challenges related to data acquisition and quality, algorithm selection and validation, model interpretability, system integration, cybersecurity, and organizational adoption [8, 9, 15, 17]. Understanding the current landscape of AI applications in PdM, identifying successful approaches, recognizing persistent challenges, and discerning future trends is crucial for both researchers developing new methods and practitioners seeking to implement these technologies effectively.

This systematic review endeavors to synthesize the peer-

reviewed literature published within the past 10 years concerning the deployment of AI techniques for industrial PdM. Adhering to a defined search and screening process, 29 relevant publications were selected to explore the following research inquiries:

- Which principal AI methods (ML/DL algorithms and strategies) are presently utilized and under development for industrial PdM?
- What are the key industrial application areas and specific use scenarios (e.g., fault diagnosis, RUL prognosis, scheduling optimization) where AI-driven PdM is being implemented?
- What principal difficulties, constraints, and obstacles are faced during the creation and deployment of AI-powered PdM solutions?
- What developing patterns, models, and potential avenues for future study are highlighted in the existing research?

Through answering these inquiries, this review aims to deliver a thorough, empirically grounded summary of the current status of AI applications for enhancing PdM, yielding useful perspectives for both academic researchers and industry practitioners. The rest of this paper is organized as follows: Section 2 gives background information on PdM and relevant AI concepts. Section 3 describes the conceptual underpinnings for AI-driven PdM. Section 4 outlines the systematic review methodology applied. Section 5 presents the synthesized results from the literature analysis. Section 6 examines the implications, challenges, and future paths. Lastly, Section 7 brings the review to a close.

2. BACKGROUND: PREDICTIVE MAINTENANCE AND ARTIFICIAL INTELLIGENCE

This section defines the core concepts of Predictive Maintenance (PdM) and Artificial Intelligence (AI) as they relate to the scope of this review, based on the analyzed literature.

2.1 Predictive Maintenance (PdM)

PdM represents a shift from reactive or scheduled maintenance towards condition-based, proactive interventions [2-5]. It involves monitoring the operational state and degradation of equipment using various sensors and data sources to forecast the probable timing of a failure [1, 5, 6, 10, 18]. The primary goal is to perform maintenance only when necessary, just before failure, thereby minimizing unscheduled downtime, reducing maintenance costs associated with unnecessary tasks, extending equipment lifespan, and improving safety [2, 5-7, 9, 19]. Key components of a PdM system typically include:

- Data Acquisition: Collecting relevant data from sensors (e.g., vibration [13], temperature [20], acoustic emission [8], pressure, current [21]) and operational systems (e.g., process parameters [17], event logs [22], quality data [23]).
- Data Processing: Cleaning, normalizing, synchronizing, and transforming raw data into a usable format; extracting or selecting relevant features [4, 12, 13, 15].
- Condition Monitoring (CM): Tracking the evolution of key parameters or calculated health indicators over time to assess the current condition of the asset [6, 18,

24].
- Diagnostics: Detecting deviations from normal operation (anomaly detection [12, 25-27]) and identifying the specific type and location of faults (fault classification [4, 8, 28]).
- Prognostics: Predicting the future degradation trend and estimating the RUL of the component or system [5, 6, 16, 18, 19, 26, 29].
- Decision Support: Recommending optimal maintenance actions and schedules based on diagnostic/prognostic information, costs, risks, and operational context [1, 3, 11, 16, 19].

While traditional PdM often relied on expert knowledge, predefined rules, or simpler statistical models [2], the integration of AI has enabled more sophisticated analysis and prediction capabilities [7, 8].

2.2 Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL)

AI represents an extensive domain concerned with developing systems able to execute functions that usually demand human cognitive abilities [7, 8]. ML, a branch of AI, involves systems that discern patterns and forecast outcomes using data, without needing direct instructions for the specific task [4, 6, 8, 10, 13, 15, 29]. DL, a specialized area within ML, employs multi-layered artificial neural networks (deep architectures) to acquire intricate representations directly from unprocessed data, frequently demonstrating superior performance on tasks involving extensive, high-dimensional datasets [5, 6, 16, 26, 29]. In the PdM context, AI/ML/DL methodologies are utilized throughout different phases:

- Anomaly Detection: Algorithms like Autoencoders [26], Self-Organizing Maps (SOM) [27], Isolation Forest [13], or clustering techniques [24] are used to identify unusual operating conditions.
- Fault Identification/Categorization: Techniques employing supervised learning – including Support Vector Machines (SVM) [8, 11, 12], Random Forests (RF) [4, 8, 11, 13], Decision Trees (DT) [11, 12], k-Nearest Neighbors (k-NN) [11], Gradient Boosting [11], AdaBoost [4], and DL approaches like Convolutional Neural Networks (CNN) [4, 6, 8, 16, 19] – are developed to categorize distinct failure modes based on sensor data patterns.
- RUL Estimation/Forecasting: Models designed for time-series prediction, especially Recurrent Neural Networks (RNNs) such as Long Short-Term Memory (LSTM) [5, 6, 13, 16, 19] and Gated Recurrent Units (GRUs) [16], alongside CNNs [19] and combined approaches [16, 20], are commonly employed to estimate the residual service duration according to deterioration patterns. The potential of generative models is also under investigation [29].
- Maintenance Scheduling Optimization: Techniques like Deep Reinforcement Learning (DRL) [3, 19] learn optimal maintenance policies through interaction with a simulated or real environment, considering complex trade-offs. Multi-Criteria Decision Making (MCDM) frameworks can also incorporate ML model outputs [11].

The ability of AI, especially DL, to handle complex, non-linear relationships in large datasets makes it particularly suitable for addressing the challenges of modern industrial

PdM [5, 6, 26].

3. CONCEPTUAL FRAMEWORK OF AI-DRIVEN PREDICTIVE MAINTENANCE

The integration of AI fundamentally reshapes the conceptual basis of PdM, shifting from reliance on predefined schedules or simple threshold rules towards a continuous, data-driven learning and adaptation cycle. The underlying principle is that historical and real-time operational data contain implicit information about the health state and degradation processes of industrial assets [5, 6, 8]. AI algorithms provide the means to extract this information, model complex system dynamics, and generate actionable insights for maintenance optimization [13, 16, 19, 26]. This AI-driven approach can be conceptualized as an iterative loop (See Figure 1), distinct from purely human-centric observation or basic condition monitoring:

- Data Acquisition & Processing: Continuous streams of data from diverse sources are collected, cleaned, integrated [17, 29], and prepared for analysis. Contextual information is crucial [17].
- AI Model Training & Validation: ML/DL models are trained on historical data to learn normal behavior, failure patterns, and degradation trajectories. Validation ensures model accuracy and generalization [5, 11, 12].
- AI Analysis: Trained models are deployed to monitor live data, detect anomalies [25-27], diagnose faults [4, 12, 29], and predict RUL [5, 16, 19, 29]. Handling uncertainty [5, 19] and providing explanations [8, 26] are increasingly important aspects.
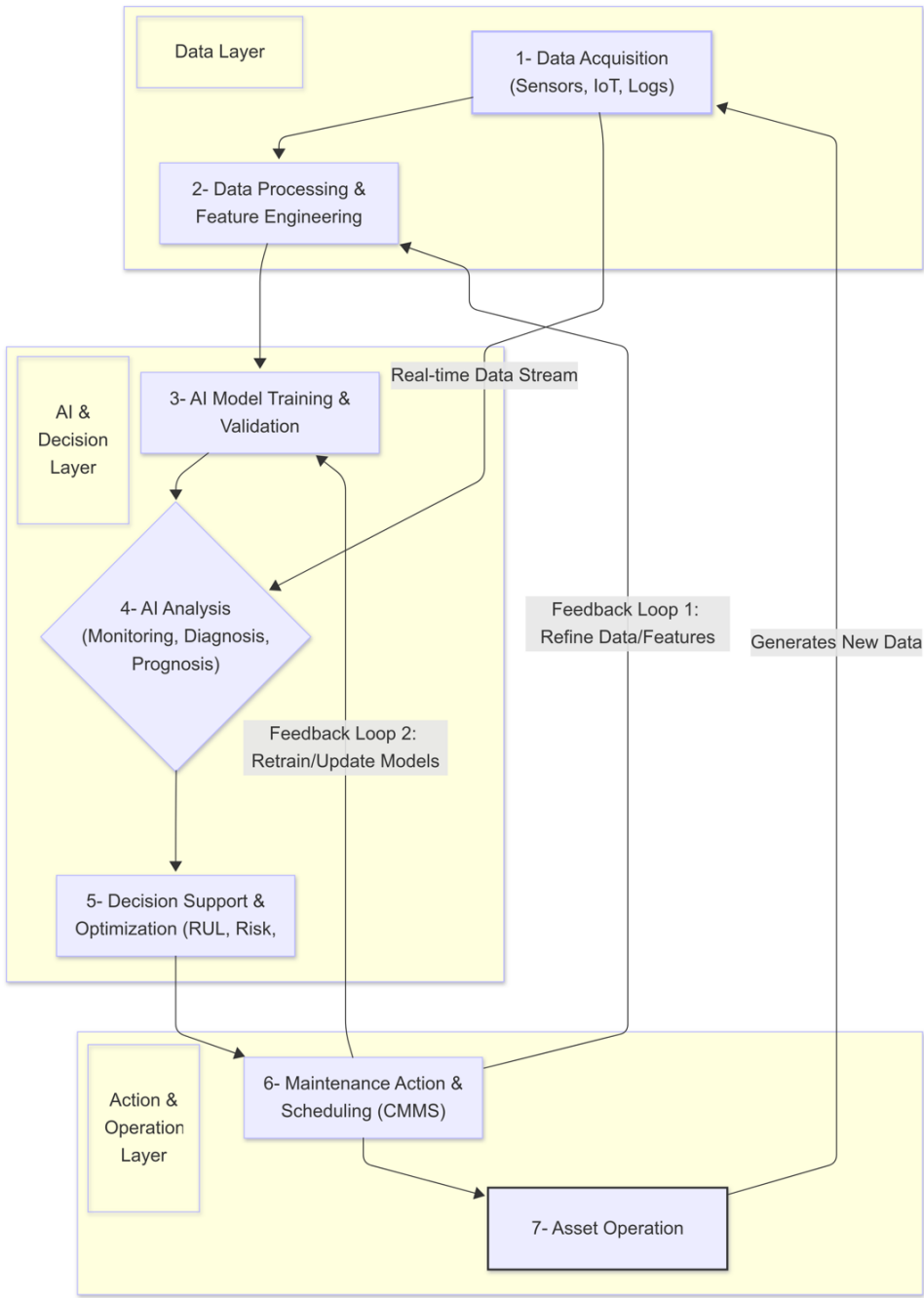

Figure 1. Conceptual loop of AI-driven Predictive Maintenance

- Decision Support & Optimization: Insights from AI analysis are translated into optimal maintenance recommendations, considering economic factors [1, 16], operational constraints [16, 29], and risk assessments. DRL can automate policy generation [19, 3].
- Action & Feedback: Recommended maintenance tasks are scheduled and executed. Crucially, the outcomes of these actions provide feedback to refine data understanding and retrain AI models, enabling continuous learning and adaptation [21].

This conceptual model underscores the pivotal function of AI algorithms in converting unprocessed data into predictive understanding and optimized decisions, surpassing the constraints of conventional methods. The prevalent emphasis within the reviewed literatures [1-29] is largely placed upon enhancing and refining the elements inside this technological cycle, especially the phases involving AI modeling and analysis.

4. METHODOLOGY

This research utilizes a Systematic Literature Review (SLR) approach for locating and integrating studies concerning the use of Artificial Intelligence to enhance Predictive Maintenance within industrial settings. The complete procedure followed to carry out this review is illustrated in Figure 2.

4.1 Search strategy and study selection

This research utilizes a Systematic Literature Review (SLR) approach for locating and integrating studies concerning the use of Artificial Intelligence to enhance Predictive Maintenance within industrial settings. The search and selection procedure involved distinct phases, depicted in Figure 3.

4.1.1 Search strategy

The primary database used for identifying relevant literature was Scopus. The search was restricted to articles published within the last 10 years from the date of the search to capture recent advances.

The following search query was applied specifically to the article titles: TITLE ("Predictive Maintenance" OR "proactive maintenance"). This initial search yielded 413 results.
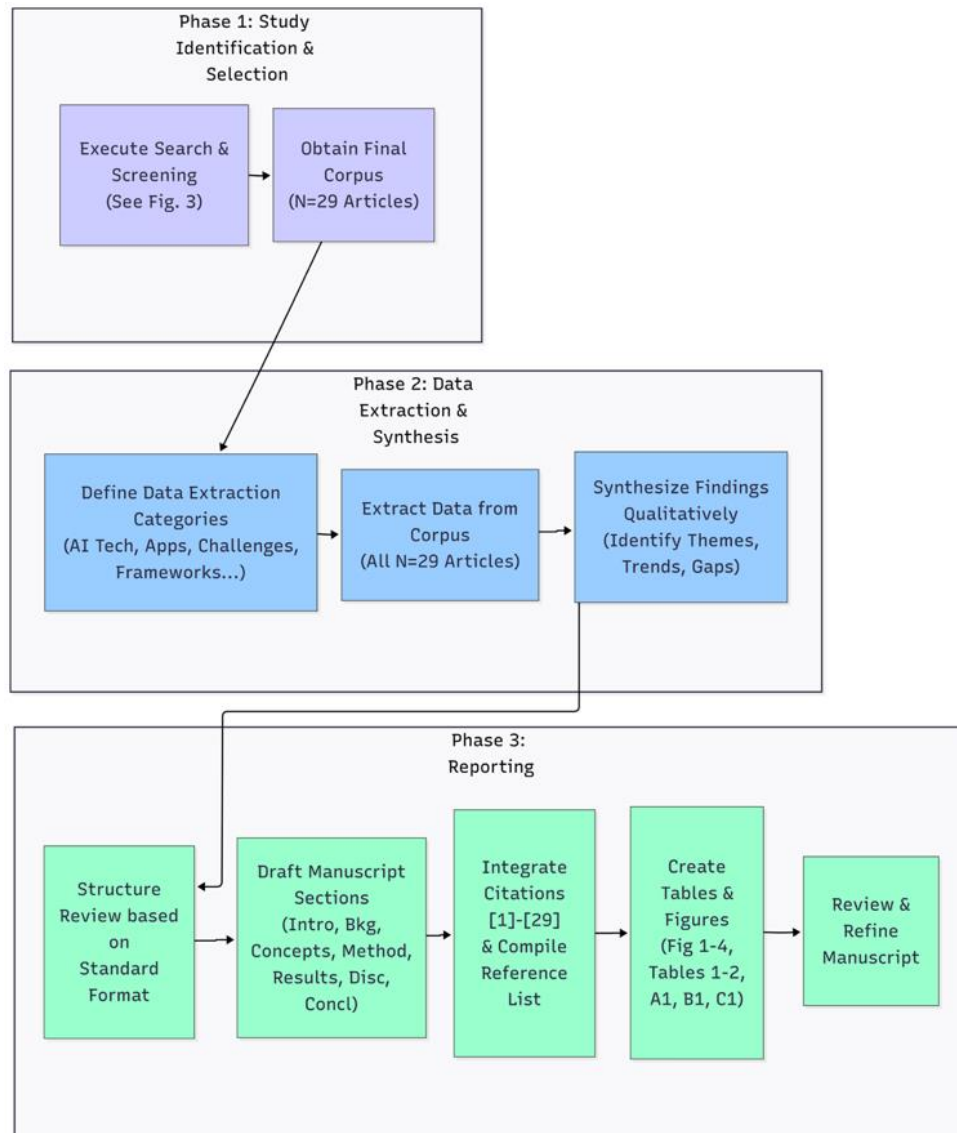

Figure 2. Methodology flowchart for this systematic literature review

4.1.2 Screening and selection

The initial results underwent a screening process:

1. Duplicate Removal: Automated and manual checks were performed to remove duplicate entries, resulting in 285 unique articles.
2. Title/Abstract/Keyword Screening: Each unique manuscript was initially assessed based on its title, abstract, and keywords to determine potential relevance to the review's scope (AI techniques for industrial PdM).
3. Full-Text Screening and Eligibility Criteria: Articles passing the initial screening were subjected to a full-text review to determine final inclusion based on the following criteria:
   o Publication Type: Inclusion was limited solely to articles from peer-reviewed journals. Exclusions comprised conference papers, book chapters, theses, and informally published works (grey literature).
   o Language: Only articles published in English were considered. This language restriction, while necessary for the feasibility of this review, represents a potential limitation as relevant studies published in other languages may have been excluded.
   o Focus: Articles must primarily focus on the study of Predictive Maintenance using AI techniques within an industrial or manufacturing domain.
   o Exclusion: State-of-the-art reviews, systematic literature reviews, and meta-analyses were excluded to focus on primary research contributions.

Following this rigorous screening process, 29 publications were deemed eligible and retained for this systematic review. These 29 articles form the basis of the analysis presented herein.

4.2 Data extraction and synthesis

A structured data extraction approach was used (Phase 3 in Figure 2). Each of the final 29 included articles was analyzed to identify information pertaining to the review's research questions. Information was systematically extracted and organized according to the primary analytical categories derived from the research questions and the scope of the reviewed literature, as listed in Table 1. These categories provided a framework for structuring the data before qualitative synthesis. Subsequently, the specific concepts and findings extracted within these broad categories were mapped to the core synthesis themes used to organize the results of this review (Section 5), as illustrated conceptually in Table 2. This mapping facilitated the identification of key trends, challenges, and areas of focus within the literature.
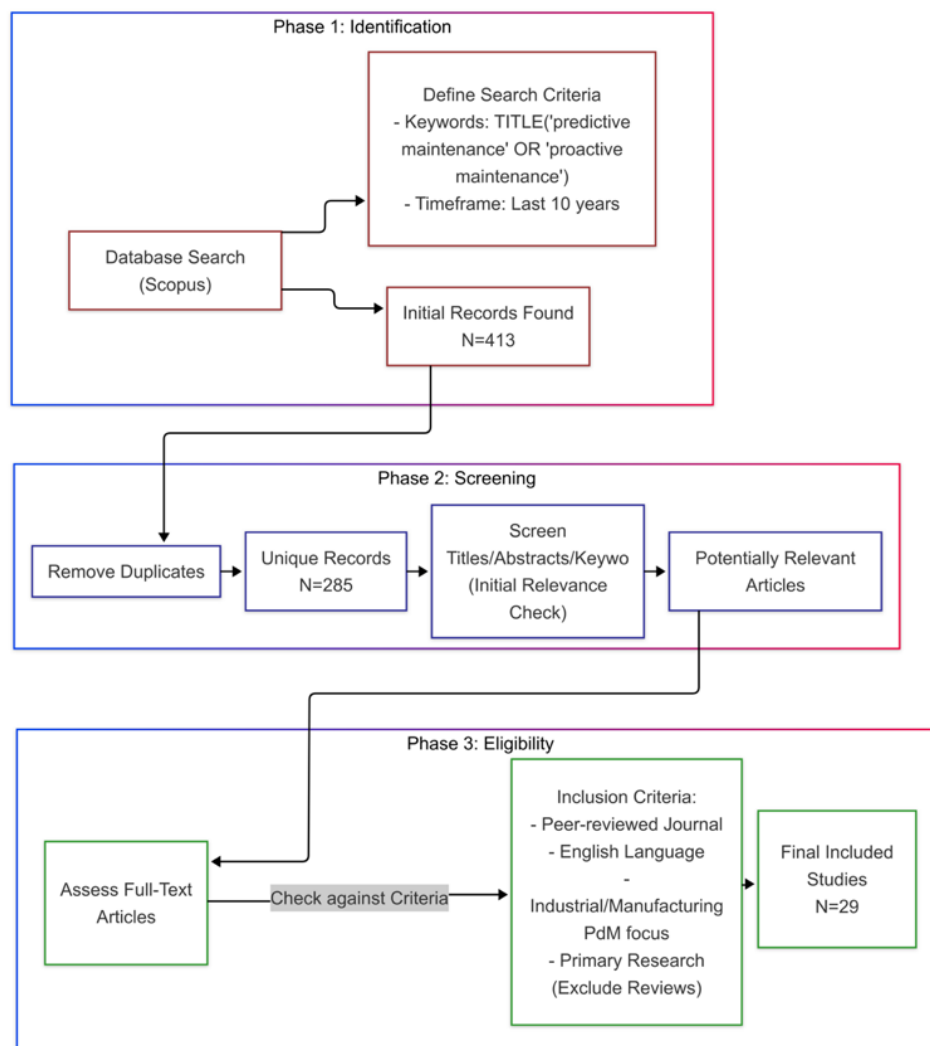

Figure 3. Flowchart of the literature search and screening process

Table 1. Primary data extraction and synthesis categories

Category	Definition / Scope
AI Techniques Applied	Specific ML, DL, DRL, Ensemble, or Hybrid algorithms and architectures used or proposed.
PdM Task Focus	The primary maintenance task addressed (e.g., Diagnosis, Fault Classification, Prognosis/RUL, Scheduling, Monitoring).
Application Domain/Context	The specific industry (e.g., Manufacturing, Automotive, Energy) and/or equipment type (e.g., Motors, Bearings, Turbines).
Data Aspects	Types of data sources used (sensors, logs, etc.), preprocessing steps, and data-related challenges encountered.
Integration Approaches	Methods for integrating PdM with other systems (Production, Quality, DT, IoT, Cloud).
Implementation Frameworks/Methods	Proposed methodologies, roadmaps, or systematic approaches for developing or deploying AI-driven PdM.
Performance/Validation Aspects	Metrics used for evaluation, reported results (accuracy, cost savings), validation methods (simulations, case studies).
Identified Challenges/Limitations	Explicitly stated technical, organizational, or data-related barriers and limitations.
Future Research Directions	Gaps identified and suggestions made by the authors for subsequent research.

Table 2. Conceptual mapping of extracted concepts to synthesis themes

Theme 1: AI Techniques Applied	- SVM, RF, DT, k-NN, NB [4, 8, 11, 12, 13] - CNN, LSTM, GRU, AE [4-6, 8, 13, 16, 19, 26] - Deep Reinforcement Learning (DRL) [3, 19] - Ensemble Methods (Boosting, Stacking) [4, 9, 11] - Hybrid Models (DL+Physics, CNN+LSTM...) [10, 16, 18] - Unsupervised (Clustering, SOM, AE) [13, 21, 24, 26, 27] - Generative Models (GAN, VAE) [29] - Uncertainty Quantification [5, 19] - Explainable AI (XAI) Methods [8, 16, 26]
Theme 2: Application Domains & PdM Tasks	- Manufacturing (General) [1, 3, 15, 16, 29] - Machine Tools [21, 23, 24, 28, 29] - Motors/Engines [5, 13, 19, 20] - Bearings [4, 6, 10] - Automotive [8] - Aerospace [19] - Electrical Systems [12] - Robotics [18] - Fault Diagnosis/Classification [4, 9, 12, 21, 25, 26, 27, 28] - RUL Prognosis [5, 6, 8, 13, 16, 18, 19, 29] - Scheduling/Optimization [1, 3, 16, 17, 19, 29] - Condition Monitoring/Health Index [18, 24, 26, 29]
Theme 3: Data Considerations	- Sensor Data (Vibration, Temp, Current...) [4, 12, 13, 18, 20, 21] - Process Parameters [17, 21, 24] - Event Logs [29] - Quality Data [23] - Multi-Source Fusion [12, 28] - Data Preprocessing (Normalization, Outliers...) [4, 13, 15, 23] - Feature Engineering/Selection [4, 8, 12] - Data Imbalance [8, 25] - Data Security/Privacy [9, 17]
Theme 4: Integration & Frameworks	- Digital Twin Integration [7, 14, 18] - Production Planning/Scheduling Integration [16, 29] - Quality Control Integration [23] - Semantic Frameworks [17] - Cloud Platforms [17] - Implementation Methods (SMEs) [15] - Human-AI Collaboration (DIAs) [14] - Standardization Roadmaps [7]
Theme 5: Challenges & Limitations	- Data Availability/Scarcity [7, 18, 19, 23] - Data Quality (Noise, Missing, Labels) [17, 23, 25] - Data Heterogeneity/Integration Complexity [7, 17, 21, 28] - Model Interpretability/Explainability (XAI) [7, 19, 26] - Model Generalization/Scalability [7, 19, 21] - System Integration Hurdles [7, 17] - Cybersecurity Risks (Adversarial Attacks) [9] - Implementation Cost & ROI Justification [7, 12, 23] - Skills Gap & Training Needs [7, 10, 23] - False Positive Management [25]

The extracted information, organized under these categories and mapped to themes, was then synthesized qualitatively (Phase 3 in Figure 2). Key themes, trends, common challenges, and divergent approaches were identified through iterative reading and comparison across the articles. The synthesis focused on building a coherent narrative addressing the research questions, supported by evidence cited from the corpus [1-29].

4.3 Quality assessment

All 29 articles were sourced as peer-reviewed publications (journal articles), implying a baseline level of scientific quality. Concerns regarding methodological limitations of the primary studies, as reported or inferred, are considered within the CERQual assessment of the review findings (see Section 4.4).

4.4 Confidence assessment of review findings (CERQual)

For evaluating the level of certainty regarding the principal conclusions drawn from the qualitative analysis of the included studies, the GRADE-CERQual (Confidence in the Evidence from Reviews of Qualitative research) method was utilized.

In applying the CERQual approach, we also acknowledge certain limitations inherent to the scope of this systematic review. Notably, the exclusion of non-English literature (as stated in Section 4.1.2) constitutes a potential source of selection bias. This could theoretically impact the overall 'Data Sufficiency' if significant bodies of work from non-Anglophone regions were missed, or the 'Applicability' of our synthesized findings if research trends or challenges differ markedly in those contexts. While the current review draws upon a substantial corpus of English-language peer-reviewed articles, this language restriction is considered when assessing the overall confidence in each review finding, particularly if a finding appears to be based on a narrow range of geographical or research group origins evident within the English literature. The CERQual assessment for each principal conclusion is detailed in Table A1 (Appendix A).

CERQual evaluates confidence using four elements:
- Methodological Limitations: Evaluating the extent to which the included studies contributing to a finding have limitations in their design or conduct that might affect confidence.
- Coherence: Assessing the extent to which the finding is consistent across the included primary studies contributing to it.
- Data Sufficiency: Assessing the depth and volume of evidence provided by the original research underpinning the finding.
- Applicability: Determining the degree to which findings from the source research pertain to the specific context defined for this review (AI for industrial PdM).

Considering these factors, a comprehensive confidence assessment for every consolidated review finding is assigned (High, Moderate, Low, or Very Low).

5. RESULTS: SYNTHESIS OF FINDINGS

This part consolidates the principal discoveries derived from reviewing the 29 selected articles, tackling the research inquiries concerning AI techniques, applications, challenges,

and frameworks within industrial PdM. The key results are detailed subsequently, accompanied by a certainty evaluation for each finding using the CERQual method.

5.1 AI techniques applied in Predictive Maintenance

It is important to note that a direct, universal comparison of performance (e.g., accuracy, efficiency) across all listed AI techniques is challenging to distill from the reviewed literature. This is due to the heterogeneity in datasets, specific PdM tasks (diagnosis, RUL estimation, scheduling), evaluation metrics, and implementation details across the primary studies. However, this section synthesizes the commonly employed AI techniques, highlighting their typical applications, and where possible, general characteristics related to their performance and computational considerations in the context of PdM as reported in the reviewed corpus [1-29].

The analyzed studies demonstrate a broad spectrum of AI techniques utilized for PdM, indicating a definite shift towards more advanced ML and particularly DL models. (CERQual Confidence: High) (Confidence is assessed as high due to the substantial volume, consistency, and depth of evidence across the body of work detailing specific AI methods employed [See Appendix A, Finding 1]).

5.1.1 Machine Learning algorithms
Traditional ML algorithms remain relevant, particularly for classification and baseline comparisons. Commonly cited techniques include:
- Support Vector Machines (SVM): Used for fault classification [11, 12] and anomaly detection [8]. SVMs are often chosen for their effectiveness in high-dimensional spaces and their ability to model non-linear decision boundaries with appropriate kernels. However, their training time can scale significantly with the size of the dataset, and performance is sensitive to kernel selection and parameter tuning.
- Random Forests (RF) & Decision Trees (DT): Applied for classification [4, 11, 12], feature importance analysis [4], and sometimes regression tasks [13]. RF is often favored for its robustness [21]. Decision Trees offer high interpretability, while Random Forests, by ensembling multiple trees, generally achieve higher accuracy and robustness against overfitting, often at the cost of some interpretability and increased computational load for training large ensembles. RFs are also valuable for their inherent ability to rank feature importance.
- k-Nearest Neighbors (k-NN): Employed for categorization [11] and outlier detection [27]. k-NN is a simple, non-parametric method that can be effective when local data structure is important. Its computational cost during inference can be high for large datasets as it requires computing distances to all training points, and its performance is sensitive to the choice of 'k' and the distance metric.
- Naive Bayes (NB): Employed for classification, though sometimes showing lower performance compared to others [12].
- Ensemble Methods: Techniques like AdaBoost [4], Gradient Boosting [11], and stacking ensembles [9] are used to combine multiple models, often improving accuracy and robustness, including resilience against adversarial attacks [9]. While often leading to superior

predictive performance and robustness, ensemble methods inherently increase computational complexity both in training and inference compared to single models.

- Clustering Algorithms: Unsupervised methods like k-Means [21], DBSCAN [21], and Self-Organizing Maps (SOM) [27] are used for identifying operational states, detecting anomalies, or localizing faults [24, 27]. These unsupervised methods are crucial for discovering inherent structures or anomalies in unlabeled data. Their computational efficiency varies greatly depending on the algorithm (e.g., k-Means is relatively efficient, while density-based methods can be more demanding).
- Time Series Models: Traditional models like ARIMA are sometimes used as baselines or components in hybrid approaches [13].

To further summarize the characteristics of these traditional Machine Learning algorithms as applied within the reviewed PdM literature, Table 3 provides a comparative overview. This table highlights their typical tasks, common strengths and weaknesses, and general computational efficiency considerations.

5.1.2 Deep Learning architectures

DL models have become increasingly popular owing to their capacity for managing intricate, high-dimensional sensor data and automatically discerning relevant features [5, 6, 26]. Key architectures include:

- Convolutional Neural Networks (CNNs): Proficient at extracting spatial patterns from sensor data (like vibration signals viewed as images) or time-series sections; frequently employed for fault categorization [4, 6, 19] and identifying key features within more extensive models [16]. CNNs have shown strong performance in tasks involving spatial hierarchies in data, like processing vibration signals as images or spectrograms. They can be computationally intensive to train, especially with deep architectures, but can be very efficient at inference with optimized implementations.
- Recurrent Neural Networks (RNNs) - LSTM & GRU: Are particularly adept at modeling sequential information and temporal dependencies, rendering them highly suitable for RUL estimation [5, 6, 13, 16, 19] and health index forecasting [29]. LSTMs and GRUs are powerful for modeling long-term dependencies in sequential data, crucial for RUL estimation. Their training can be computationally demanding and slower than feedforward networks due to their recurrent nature, but they excel where temporal context is key.
- Autoencoders (AE): Primarily used in unsupervised anomaly detection by learning a compressed representation of normal data and identifying deviations [6, 7, 26]. Convolutional AEs are common [26]. Explainable variants are emerging [26]. AEs are effective for unsupervised anomaly detection and feature learning by learning compressed data representations. Their complexity and training time depend on the depth and width of the encoder/decoder networks.
- Hybrid Models: Combinations like CNN-LSTM [16], CNN-LSTM-Attention [16], LSTMs integrated with physical models [18], or generative models combined with health indicators [29] aim to leverage the strengths of different architectures for improved performance, particularly in complex RUL prediction or integrated planning tasks [16, 20, 29]. Hybrid models aim to combine the strengths of different architectures, potentially leading to better performance but also increased model complexity and training requirements.
- Generative Models: Used for data augmentation or directly for health prognostics [29].

Deep Reinforcement Learning (DRL): Applied to optimize maintenance scheduling and decision-making policies in dynamic and uncertain environments, learning through interaction [3, 19]. Multi-agent DRL addresses coordination in complex systems [3]. DRL shows promise for optimizing complex, sequential decision-making tasks like maintenance scheduling. However, DRL typically requires significant computational resources for training, often involving numerous simulation iterations, and careful environment design.

Table 3. Summary of traditional Machine Learning techniques in PdM from reviewed literature

AI Technique	Typical Predictive Maintenance Task(s)	Common Strengths Cited/Observed	Common Weaknesses/Considerations	Computational Efficiency (General)	Key References
Support Vector Machine (SVM)	Fault Classification, Anomaly Detection	Good with high-dimensional data, non-linear problems (using kernels)	Training time can be significant, sensitive to kernel and parameter selection	Moderate to High (Training)	[7, 11, 12]
Random Forests (RF) / Decision Trees (DT)	Fault Classification, Feature Importance Assessment	Interpretable (Decision Tree), Robust, Good accuracy (Random Forest)	Random Forest less interpretable, single Decision Tree can overfit	Moderate (Random Forest training can be high)	[4, 11, 12, 13, 21]
k-Nearest Neighbors (k-NN)	Categorization, Outlier Detection	Simple, non-parametric method	High inference cost for large datasets, sensitive to 'k' and distance metric	Low (Training), High (Inference)	[11, 27]
Ensemble Methods	Classification, Regression	Improved accuracy and robustness	Increased model complexity and computational requirements	High	[4, 9, 11]
Clustering (e.g., k-Means, SOM)	Anomaly Detection, System State Identification	Unsupervised, finds inherent data structures	Sensitive to algorithm parameters, efficiency varies by algorithm	Varies (k-Means is relatively efficient)	[21, 24, 27]

Note: This table provides a general synthesis. The actual performance and efficiency of these algorithms are highly dependent on the specific dataset, implementation details, and the nature of the PdM task. Computational efficiency refers to general trends; for instance, SVM training can be demanding for very large datasets, while k-NN inference can be slow without appropriate indexing structures.

A summary of these Deep Learning architectures, detailing their application in PdM tasks, along with their common strengths, weaknesses, and computational aspects as observed in the reviewed corpus, is presented in Table 4.

5.2 Key application domains and use cases

AI-driven PdM is being applied across a diverse range of industrial sectors and equipment types, demonstrating its versatility. (CERQual Confidence: High) (Evidence for this finding is abundant across the corpus, covering numerous industries and equipment types consistently, indicating high coherence, adequacy, and relevance [See Appendix A, Finding 1]).

- Manufacturing: This domain is a prominent area for AI-driven PdM, with numerous studies leveraging empirical data from real enterprises. One such study presents a data-driven drift detection and diagnosis framework applied to a multiple tapping process in a real industrial setting [21]. It utilized sensor data (current, voltage, process parameters) and employed clustering algorithms (k-Means, DBSCAN) alongside Random Forests to identify and classify operational drifts, demonstrating the practical application of AI for improving the stability of heterogeneous production processes. Another study, conducted in the context of an automotive plant, linked product quality control from a machine tool to its predictive maintenance requirements [23]. It illustrated how machine degradation, monitored through sensor data and process parameters, could be correlated with part quality deviations, offering a framework for joint product–process–machine optimization using real-world manufacturing data. Further illustrating the application of AI in critical industrial infrastructure, a recent study developed an IoT and Machine Learning-based system for anomaly detection in electrical panels aimed at Predictive Maintenance and fire prevention

[12]. The work involved an experimental setup using Arduino and Raspberry Pi platforms to collect real-time data (3,478 data points) including gas, temperature, humidity, current, and voltage from electrical panels under various simulated fault conditions leading to potential fire hazards. Several ML classifiers were evaluated (Decision Tree, Gaussian Naive Bayes, Support Vector Machine, and Gaussian Process Classifier). The study demonstrated that the Gaussian Process Classifier (GPC) achieved the highest performance, with an accuracy of 99.56% and an AUC of 0.99, in effectively distinguishing between normal and anomalous (pre-fire) conditions. This case study underscores the practical value of integrating IoT sensor data with robust ML algorithms for Predictive Maintenance of critical electrical infrastructure, enabling early anomaly detection and proactive safety interventions based on empirical evidence. Other applications within manufacturing include general systems [1, 14, 16, 17, 22, 23], specific equipment such as injection molding machines [1], industrial robots [18], power press motors [13], and textile knitting machines [4].
- Aerospace: Aircraft turbofan engines (widely used benchmark datasets) [5, 16, 19].
- Automotive: Vehicle components and systems [8].
- Energy & Utilities: Electrical distribution panels [12], electric motors [13, 20].
- Transportation: Railcar bearings [10].
- The primary use cases reported align with the core PdM tasks:

Fault Detection and Diagnosis: Identifying anomalous behavior and classifying specific fault types [4, 7, 9, 12, 13, 21, 24, 25, 27, 28].

- Prognosis and RUL Prediction: Estimating the remaining time until failure [4-6, 13, 16, 18, 19, 29]. This is often considered the most challenging but valuable task.

Table 4. Summary of Deep Learning architectures in PdM from reviewed literature

AI Technique	Typical Predictive Maintenance Task(s)	Common Strengths Cited/Observed	Common Weaknesses/Considerations	Computational Efficiency (General)	Key References
Convolutional Neural Networks (CNNs)	Fault Diagnosis (e.g., from vibration signals, images)	Automatic extraction of spatial features	Data-hungry, computationally intensive to train	High (Training), Varies (Inference)	[4, 6, 7, 16, 19]
Recurrent Neural Networks (RNNs) (e.g., LSTM, GRU)	Remaining Useful Life Estimation, Health Index Forecasting	Models temporal dependencies and long-term context	Data-hungry, slower training, risk of vanishing/exploding gradients (mitigated in LSTM/GRU)	High (Training)	[5, 6, 13, 16, 19, 29]
Autoencoders (AE)	Anomaly Detection, Feature Learning	Unsupervised learning, effective for dimensionality reduction	Reconstruction error may not capture all anomalies	Moderate to High (Training)	[6, 7, 26]
Deep Reinforcement Learning (DRL)	Maintenance Scheduling Optimization	Learns optimal sequential decision-making policies	Very data-hungry, computationally intensive, needs complex environment modeling	Very High (Training)	[17, 19]
Hybrid Models	Complex Remaining Useful Life Estimation or Diagnosis	Can combine strengths of different architectures, potentially better performance	Increased model complexity and longer training times	High (Training)	[16, 18, 20, 29]

Note: This table offers a general overview. Deep Learning models are generally data-hungry and computationally intensive to train, often requiring specialized hardware (e.g., GPUs). Their inference efficiency can vary. Performance is contingent upon architecture design, hyperparameter tuning, dataset size and quality.

- Maintenance Scheduling and Optimization: Determining the optimal time and type of maintenance intervention [1, 14, 17, 19, 23, 29]. This increasingly involves integration with production schedules [16, 29].
- Condition Monitoring and Health Assessment: Tracking the overall health status using calculated indices or learned representations [18, 22, 26, 27].
- Linking Machine Health to Process Outcomes: Connecting machine degradation signals to impacts on product quality [23] or process efficiency [21].

5.3 Data aspects: Sources, processing, and challenges

Data is the cornerstone of AI-driven PdM. The reviewed articles highlight the use of diverse data sources and the critical importance of data processing, while consistently identifying data-related issues as major challenges. (CERQual Confidence: High) (Confidence is high, mirroring Finding 3 in Appendix A, due to strong coherence and adequacy of evidence across numerous studies detailing data sources, processing steps, and challenges).

- Data Sources: Vibration sensors [4, 13, 18, 19], temperature sensors [12, 13, 20], acoustic sensors [8], current/voltage sensors [12, 21], pressure sensors, process parameters (speeds, loads, flow rates) [21, 22, 24], event logs [22], quality measurements [23], thermal imaging [12, 20], and operational context data are commonly used. Multi-sensor fusion approaches are frequent [12, 28].
- Data Processing: Standard steps include cleaning, handling missing values, normalization/scaling [4], noise reduction [6], outlier detection/removal [13, 26], and time-series segmentation. Feature engineering (e.g., statistical features from time/frequency domains [8]) is still common, although DL aims to reduce this need [5]. Feature selection techniques (e.g., RFE [12], Permutation Importance [4]) are used to identify relevant inputs.
- Data Challenges: These remain a significant bottleneck:
  - Data Availability/Scarcity: A recurrent challenge is the lack of sufficient historical data, particularly comprehensive run-to-failure datasets, which critically hinders robust model training and validation [7, 8, 15, 18]. For instance, the scarcity of failure data in real-world automotive fleets often necessitates reliance on simulated data or limits the complexity of deployable ML models, potentially affecting their generalization to unseen failure modes, as highlighted in recent research [8]. Similarly, in the context of industrial robots, it has been noted that obtaining extensive degradation data can be time-consuming and costly, often leading to models being trained on limited operational histories which may not capture the full spectrum of wear patterns [18]. To mitigate such scarcity, the use of generative models for data augmentation is an emerging research avenue. One such effort aims to create synthetic but realistic operational data to supplement sparse real-world datasets [29].
  - Data Quality: The performance of AI models is also significantly impacted by issues such as noise, missing values, sensor drift, and inconsistent labeling in the available data [9, 15, 25]. For

example, recent work on anomaly detection has emphasized that noisy sensor data or ambiguous labeling can lead to a high rate of false positives, thereby diminishing the trustworthiness and practical utility of PdM systems; a specific mitigation methodology to address this impact has been proposed [25]. In the context of resilient PdM model development, it has been acknowledged that adversarial noise—even slight perturbations—can drastically degrade model performance, illustrating the sensitivity of AI techniques to data integrity [9]. Furthermore, obtaining high-quality, consistently labeled data remains a major hurdle for Small and Medium-sized Enterprises looking to implement ML-based PdM solutions [15].
  - Data Heterogeneity & Integration: Combining data from diverse sources with different formats, semantics, and sampling rates is complex [7, 11, 21, 28]. Semantic frameworks are proposed [17].
  - Imbalanced Data: Failure events are often rare compared to normal operation, requiring specific techniques during training [8, 25].
  - Data Confidentiality & Security: Protecting sensitive operational data is crucial [9, 17].

5.4 Integration approaches and frameworks

Recognizing that PdM does not operate in isolation, several articles focus on integration and systematic implementation frameworks, representing a key trend towards more holistic asset management. (CERQual Confidence: Moderate) (Confidence is moderate, mirroring Finding 2 in Appendix A, reflecting a clear trend but with varied approaches and fewer large-scale validated implementations reported in this corpus compared to core AI techniques).

- Integration with Production Planning: Optimizing maintenance schedules jointly with production lot-sizing to minimize overall costs and disruption [16, 29]. This requires predicting health based on operational parameters [29].
- Integration with Quality Control: Linking machine health degradation directly to potential impacts on manufactured part quality [23].
- Integration via Digital Twins (DTs): The use of DTs as comprehensive platforms for integrating real-time data, physics-based models, AI algorithms, and visualization capabilities is increasingly advocated to support holistic PdM [7, 14, 18]. DTs can facilitate advanced simulation and 'what-if' analysis for maintenance decisions [18], and offer a dynamic, up-to-date representation of asset health. Standardization roadmaps for DT-based PdM are being proposed to guide their development and interoperability [7]. However, the practical implementation of DTs, especially for Small and Medium-sized Enterprises (SMEs), faces significant cost and technical barriers. The development and maintenance of high-fidelity DTs can be resource-intensive, requiring substantial investment in sensorization, data infrastructure, modeling expertise, and computational power [7, 8]. Technical barriers include challenges in data integration from heterogeneous sources, ensuring real-time synchronization between the physical asset and its

digital counterpart, and the complexity of validating and maintaining the twin over the asset's lifecycle [7, 17]. In the context of advanced ML solutions, some studies have emphasized that these challenges are often amplified for SMEs due to limited financial resources, a skills gap in areas like data science and complex system modeling, and the lack of standardized, cost-effective DT platforms tailored to their needs [15]. Therefore, while DTs offer transformative potential for PdM, a careful feasibility analysis considering these cost, technical, and organizational aspects is crucial, particularly for SMEs, before embarking on full-scale DT integration.
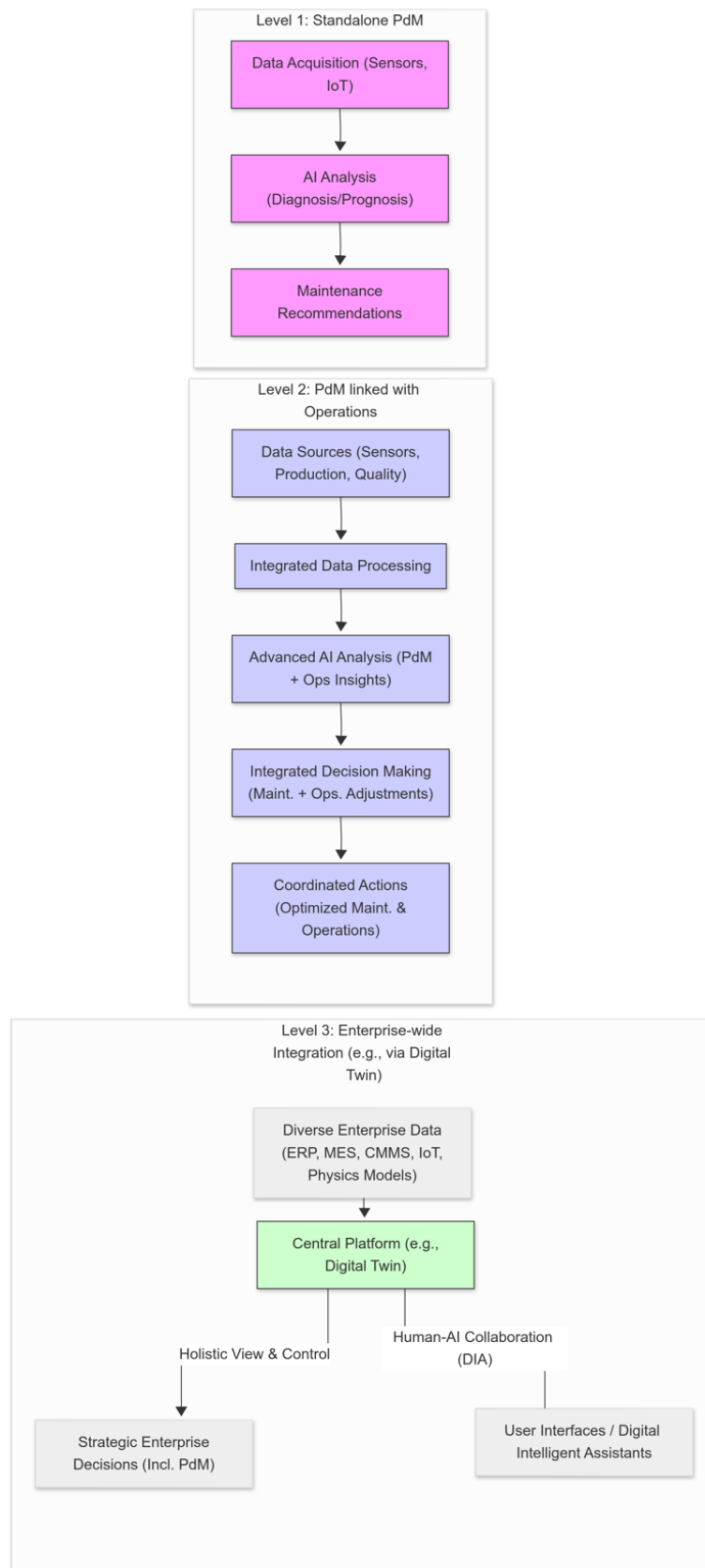
Figure 4. Levels of PdM integration approaches

- Semantic Frameworks: Utilizing ontologies and semantic technologies to integrate heterogeneous data sources and provide context for analysis in cloud environments [17].
- Implementation Methodologies: Structured methods are proposed to guide the implementation of ML-based PdM, especially for SMEs, covering data acquisition, model selection, validation, and deployment [15].
- Hybrid Intelligence Systems: Frameworks combining AI capabilities with human expertise through interfaces like Digital Intelligent Assistants (DIAs) for collaborative decision-making [14].

Figure 4 illustrates different levels of integration discussed in the literature, moving from standalone PdM to deeply integrated enterprise solutions.

5.5 Emerging trends and techniques

Beyond the established applications, the literature highlights several emerging research frontiers aimed at overcoming current limitations and enhancing PdM capabilities. (CERQual Confidence: Moderate) (Confidence is moderate, mirroring Finding 4 in Appendix A, as these are identifiable trends but represent newer research areas with a comparatively smaller or less mature evidence base within this specific corpus). These include a strong focus on:
- Uncertainty Quantification: Developing models that not only predict RUL but also provide associated confidence levels or probability distributions [5, 19].
- Deep Reinforcement Learning (DRL): Applying DRL for complex, adaptive scheduling policies that learn optimal actions through interaction [3, 19].
- Explainable AI (XAI): Developing approaches that render the predictions from AI models transparent and comprehensible to human experts [8, 26].
- Cybersecurity Resilience: Designing AI-based PdM systems robust against adversarial data manipulation or model attacks [9].
- Systematic Implementation: Devising practical methodologies and frameworks to guide successful adoption, particularly addressing the needs and constraints of SMEs [15].
- Hybrid Intelligence: Integrating human knowledge and reasoning capabilities with AI systems, potentially via DIAs [14].

6. DISCUSSION

The synthesized findings reveal a dynamic and rapidly evolving field where AI is increasingly integral to realizing the full potential of PdM. This section discusses the key trends, persistent challenges, implications, and future research directions emerging from the reviewed literature [1-29].

6.1 Key trends and developments

The review highlights a significant shift towards data-driven, AI-powered PdM. The dominance of Deep Learning (Finding 1, High Confidence) is undeniable, particularly for handling complex sensor data in RUL prediction and diagnostics [5, 6, 16, 23, 26]. This trend is coupled with a growing awareness of the need for uncertainty quantification and explainability (Finding 4, Moderate Confidence) [5, 7, 16,

23]. As models become more complex, ensuring trust and enabling informed decision-making requires understanding prediction confidence and model reasoning.

Another major trend is the push towards integration (Finding 2, Moderate Confidence). Standalone PdM models are evolving into components of larger, interconnected systems that link maintenance insights with production scheduling [16, 29], quality control [23], and enterprise-level planning, often orchestrated through Digital Twin platforms [7, 14, 18] or semantic frameworks [17]. This reflects a move towards optimizing overall system performance rather than just individual asset maintenance.

Furthermore, there is increasing interest in automation and optimization using techniques like DRL [3, 19] and the development of hybrid approaches combining AI with physics [18] or human intelligence [14] to overcome limitations of purely data-driven methods, especially in data-scarce situations. Finally, research is actively addressing practical implementation challenges, including security [9] and methods tailored for SMEs [15].

6.2 Major challenges and limitations

Despite significant progress, the review confirms that several major challenges persist, hindering the seamless adoption and scaling of AI-driven PdM. Data-related issues remain the most critical bottleneck (Finding 3, High Confidence). Obtaining sufficient volumes of high-quality, well-labeled data, especially run-to-failure data, is a universal challenge [7, 8, 15, 18]. Integrating and managing heterogeneous data from multiple sources also presents significant technical difficulties [7, 15, 22, 28].

Model clarity and trustworthiness (also related to Finding 3, High Confidence) constitute a further significant obstacle [7, 8, 23]. The opaque characteristics of intricate DL models impede the ability of engineers and upkeep teams to comprehend and rely on their results, thereby obstructing adoption and effective application. Although XAI techniques are emerging as a means to mitigate this opacity and build trust, these methods themselves come with limitations that need critical consideration [26]. For instance, many post-hoc XAI methods (e.g., LIME, SHAP) provide local explanations for specific predictions, which might not always offer a complete global understanding of the model's behavior, especially for highly non-linear and complex Deep Learning models often used in PdM. Furthermore, the "explanations" generated by XAI can sometimes be approximations or simplifications of the true model logic, and their faithfulness to the underlying model can be difficult to verify [30]. The applicability and scalability of certain XAI techniques to very large datasets or extremely deep architectures also remain a challenge, potentially adding computational overhead. This need is emphasized in recent work focusing on explainable approaches tailored to specific complex architectures [26]. Therefore, while XAI is a crucial step forward, the development of truly workable, dependable, transparent, and computationally efficient XAI solutions that are readily deployable and genuinely insightful for end-users in industrial PdM remains a significant research endeavor.

Other significant challenges identified include:
- Ensuring model generalization across different operating conditions or similar assets [7, 8, 21].
- Achieving seamless system integration with legacy IT/OT systems [7, 17].

- Addressing cybersecurity vulnerabilities inherent in interconnected, data-driven systems [9].
- Overcoming the high implementation costs and the need for specialized skills [7, 10, 15].
- Managing false positives in anomaly detection systems [25].

6.3 Implications

The successful implementation of AI-driven PdM carries significant implications across multiple dimensions:
- Operational: Substantial improvements in asset uptime, reliability, and overall equipment effectiveness (OEE); optimized maintenance resource allocation; potential for condition-based adjustments to operating parameters [1, 6, 14, 15, 17, 23].
- Economic: Reduced maintenance costs (labor, parts), minimized production losses due to unplanned downtime, potentially extended asset life reducing CapEx, and enabling new service-based business models [1, 2, 7, 11, 22].
- Safety & Environmental: Enhanced safety through proactive identification and mitigation of potential failures; prevention of catastrophic failures with potential environmental impact [6, 7, 28].
- Workforce: A necessary evolution of maintenance roles towards data analysis, system interpretation, and collaboration with AI tools, demanding new skill sets and training approaches [7, 10, 14, 15].

6.4 Future research directions

The challenges and trends identified point towards several key areas for future research:
- Robust AI for Imperfect Data: Developing AI/ML techniques that are inherently more robust to noise, missing data, limited labeled data (few-shot learning, transfer learning), and data imbalance [8, 12, 25]. This includes exploring self-supervised learning approaches to leverage large unlabeled datasets prevalent in industry, and developing generative models (e.g., GANs, VAEs) specifically tailored for augmenting scarce run-to-failure data in PdM contexts [29].
- Trustworthy AI (XAI & Uncertainty): Advancing practical XAI methods tailored for PdM time-series data and improving the reliability and communication of uncertainty estimations [5, 7, 16, 23]. Future work should focus on developing XAI techniques that are not only post-hoc but also inherently interpretable (e.g., attention mechanisms in DL models), and on quantifying different sources of uncertainty (aleatoric vs. epistemic) to provide actionable confidence scores for maintenance decisions [5].
- Hybrid & Physics-Informed AI: Methodically combining domain knowledge and physics-based models with data-driven AI approaches to improve accuracy, interpretability, and data efficiency [8, 18]. This could involve, for example, developing hybrid models that integrate first-principle degradation models with Deep Learning techniques [18] to better capture both known physical phenomena and unmodeled data-driven patterns, particularly in systems with well-understood failure physics but noisy sensor data.

- Scalable & Standardized Platforms: Developing standardized architectures, data models (e.g., semantic frameworks [17]), and MLOps practices to facilitate scalable deployment and integration, potentially leveraging Digital Twins [7]. A key industry need is the development of open-source or interoperable platforms for PdM that reduce vendor lock-in and facilitate the integration of AI tools from different providers. Research into MLOps for PdM is crucial for managing the lifecycle of deployed models, including continuous monitoring, retraining, and versioning in dynamic industrial environments.
- Cybersecurity & Resilience: Designing inherently secure AI algorithms and PdM systems resistant to data poisoning or adversarial attacks [9]. This includes research into robust training strategies against adversarial examples, developing differential privacy techniques for PdM data, and designing secure architectures for distributed PdM systems, for instance, leveraging blockchain for data integrity.
- DRL & Autonomous Systems: Exploring DRL further for fully autonomous maintenance decision-making and multi-agent coordination in complex industrial environments [3, 19]. Specific needs include developing DRL agents that can handle partial observability, learn from sparse rewards common in maintenance tasks, and ensure safe exploration during learning in safety-critical systems. Multi-agent DRL approaches for coordinating maintenance across multiple machines or production lines [3] also warrant further investigation.
- Human-AI Collaboration: Researching effective human-computer interfaces (like DIAs [14]) and workflows for synergistic collaboration between maintenance personnel and AI systems. This involves designing intuitive interfaces that can present complex AI insights (e.g., XAI outputs, uncertainty measures) to non-expert users and facilitate human-in-the-loop learning where operator feedback refines AI models. This approach has been suggested in recent research on Digital Intelligent Assistants [14].
- Real-World Validation & Benchmarking: Increased focus on validating proposed methods on diverse, real-world industrial datasets beyond standard benchmarks, and establishing clear benchmarking protocols [7, 11, 15]. There is a pressing need for more publicly available, high-quality industrial datasets that include diverse fault modes and operational conditions. Furthermore, developing standardized benchmarking methodologies, including common metrics and evaluation procedures, will be crucial for objectively comparing different PdM approaches.
- Federated Learning for Collaborative and Privacy-Preserving PdM: Exploring Federated Learning (FL) frameworks to enable multiple organizations (or multiple plants within an organization) to collaboratively train robust PdM models without sharing their raw sensitive data. This addresses both data scarcity (by leveraging larger, more diverse datasets) and data privacy concerns, which are significant barriers in many industrial sectors [9, 17]. Research should focus on FL algorithms robust to data heterogeneity (non-IID data) across participants and communication efficiency.

7. CONCLUSION

This systematic review has synthesized the findings from 29 peer-reviewed articles published in the last 10 years on the application of Artificial Intelligence to improve Predictive Maintenance, identified through a targeted Scopus search. The analysis validates that AI, notably Deep Learning, is quickly reshaping the PdM field, facilitating higher accuracy in fault diagnosis, more precise Remaining Useful Life estimation, and better-planned maintenance scheduling throughout diverse industrial sectors.

There is a clear trend towards leveraging complex models like CNNs, LSTMs, and Autoencoders, often integrated within broader systems involving IoT data streams and Digital Twins. Key advancements focus not only on predictive accuracy but also on addressing critical aspects like model uncertainty, explainability (XAI), integration with production planning, and automated decision-making through techniques like Deep Reinforcement Learning.

However, significant challenges persist, primarily centered around data availability and quality, model interpretability and trust, system integration complexity, cybersecurity, and the practical hurdles of implementation, especially for SMEs. Despite these obstacles, the potential benefits of AI-driven PdM – improved operational efficiency, reduced costs, enhanced safety, and extended asset life – are substantial, driving continued research and development.

Future efforts should focus on creating more robust, trustworthy, integrated, and scalable AI solutions, alongside developing the necessary skills and methodologies for successful real-world deployment. The convergence of AI with other Industry 4.0 technologies holds the promise of enabling truly intelligent, adaptive, and ultimately autonomous maintenance systems, marking a fundamental shift in industrial asset management. This review provides a structured overview of this progress, highlighting the achievements while emphasizing the critical areas requiring continued research and innovation.

REFERENCES

[1] Dinh, D.H., Do, P., Iung, B., Bang, T.Q. (2025). Predictive maintenance optimization for manufacturing systems considering perfect and imperfect inspections: Application to injection molding machine. CIRP Annals. https://doi.org/10.1016/j.cirp.2025.03.027

[2] Le Xuan, Q., Adhisantoso, Y.G., Munderloh, M., Ostermann, J. (2023). Uncertainty-aware remaining useful life prediction for predictive maintenance using deep learning. Procedia CIRP, 118: 116-121. https://doi.org/10.1016/j.procir.2023.06.021

[3] Daniyan, I., Mpofu, K., Oyesola, M., Ramatsetse, B., Adeodu, A. (2020). Artificial intelligence for predictive maintenance in the railcar learning factories. Procedia Manufacturing, 45: 13-18. https://doi.org/10.1016/j.promfg.2020.04.032

[4] Li, Z., Jana, C., Pamucar, D., Pedrycz, W. (2025). A comprehensive assessment of machine learning models for predictive maintenance using a decision-making framework in the industrial sector. Alexandria Engineering Journal, 120: 561-583. https://doi.org/10.1016/j.aej.2025.02.010

[5] Elkateb, S., Métwalli, A., Shendy, A., Abu-Elanien, A.E.B. (2024). Machine learning and IoT-based predictive maintenance approach for industrial applications. Alexandria Engineering Journal, 88: 298-309. https://doi.org/10.1016/j.aej.2023.12.065

[6] Wellsandt, S., Klein, K., Hribernik, K., Lewandowski, M., Bousdekis, A., Mentzas, G., Thoben, K.D. (2022). Hybrid-augmented intelligence in predictive maintenance with digital intelligent assistants. Annual Reviews in Control, 53: 382-390. https://doi.org/10.1016/j.arcontrol.2022.04.001

[7] Silveira, N.N.A., Meghoe, A.A., Tinga, T. (2024). Quantifying the suitability and feasibility of predictive maintenance approaches. Computers & Industrial Engineering, 194: 110342. https://doi.org/10.1016/j.cie.2024.110342

[8] Lee, J., Mitici, M. (2023). Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics. Reliability Engineering and System Safety, 230: 108908. https://doi.org/10.1016/j.ress.2022.108908

[9] López, D., Aguilera-Martos, I., García-Barzana, M., Herrera, F., García-Gil, D., Luengo, J. (2023). Fusing anomaly detection with false positive mitigation methodology for predictive maintenance under multivariate time series. Information Fusion, 100: 101957. https://doi.org/10.1016/j.inffus.2023.101957

[10] Wu, H., Huang, A., Sutherland, J.W. (2020). Avoiding environmental consequences of equipment failure via an LSTM-based model for predictive maintenance. Procedia Manufacturing, 43: 666-673. https://doi.org/10.1016/j.promfg.2020.02.174

[11] Ma, S., Flanigan, K.A., Bergés, M. (2024). State-of-the-art review and synthesis: A requirement-based roadmap for standardized predictive maintenance automation using digital twin technologies. Advanced Engineering Informatics, 62: 102800. https://doi.org/10.1016/j.aei.2024.102800

[12] Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., Elger, G. (2021). Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. Reliability Engineering and System Safety, 215: 107864. https://doi.org/10.1016/j.ress.2021.107864

[13] von Birgelen, A., Buratti, D., Mager, J., Niggemann, O. (2018). Self-organizing maps for anomaly localization and predictive maintenance in cyber-physical production systems. Procedia CIRP, 72: 480-485. https://doi.org/10.1016/j.procir.2018.03.150

[14] Chapelin, J., Voisin, A., Rose, B., Iung, B., Steck, L., Chaves, L., Lauer, M., Jotz, O. (2025). Data-driven drift detection and diagnosis framework for predictive maintenance of heterogeneous production processes: Application to a multiple tapping process. Engineering Applications of Artificial Intelligence, 139: 109552. https://doi.org/10.1016/j.engappai.2024.109552

[15] Voisin, A., Laloi, T., Iung, B., Romagne, E. (2018). Predictive maintenance and part quality control from joint product-process-machine requirements: Application to a machine tool. Procedia Manufacturing, 16: 147-154. https://doi.org/10.1016/j.promfg.2018.10.166

[16] Ruiz Rodríguez, M.L., Kubler, S., de Giorgio, A., Cordy, M., Robert, J., Le Traon, Y. (2022). Multi-agent deep reinforcement learning based predictive maintenance on

parallel machines. Robotics and Computer-Integrated Manufacturing, 78: 102406. https://doi.org/10.1016/j.rcim.2022.102406

[17] Hanifi, S., Alkali, B., Lindsay, G., Waters, M., McGlinchey, D. (2024). Advancements in predictive maintenance modelling for industrial electrical motors: Integrating machine learning and sensor technologies. Measurement: Sensors, 101473. https://doi.org/10.1016/j.measen.2024.101473

[18] Aivaliotis, P., Arkouli, Z., Georgoulias, K., and Makris, S. (2021). Degradation curves integration in physics-based models: Towards the predictive maintenance of industrial robots. Robotics and Computer-Integrated Manufacturing, 71: 102177. https://doi.org/10.1016/j.rcim.2021.102177

[19] Pekşen, M.F., Yurtsever, U., Uyaroğlu, Y. (2024). Enhancing electrical panel anomaly detection for predictive maintenance with machine learning and IoT. Alexandria Engineering Journal, 96: 112-123. https://doi.org/10.1016/j.aej.2024.03.106

[20] Dehghan Shoorkand, H., Nourelfath, M., Hajji, A. (2024). A hybrid deep learning approach to integrate predictive maintenance and production planning for multi-state systems. Journal of Manufacturing Systems, 74: 397-410. https://doi.org/10.1016/j.jmsy.2024.04.005

[21] Sharifi, T., Eikani, A., Mirsalim, M. (2024). Heat transfer study on a stator-permanent magnet electric motor: A hybrid estimation model for real-time temperature monitoring and predictive maintenance. Case Studies in Thermal Engineering, 63: 105286. https://doi.org/10.1016/j.csite.2024.105286

[22] Schmidt, B., Wang, L. (2018). Predictive maintenance of machine tool linear axes: A case from manufacturing industry. Procedia Manufacturing, 17: 118-125. https://doi.org/10.1016/j.promfg.2018.10.022

[23] Hajgató, G., Wéber, R., Szilágyi, B., Tóthpál, B., Gyires-Tóth, B., Hős, C. (2022). PredMaX: Predictive maintenance with explainable deep convolutional autoencoders. Advanced Engineering Informatics, 54: 101778. https://doi.org/10.1016/j.aei.2022.101778

[24] Gungor, O., Rosing, T., Aksanli, B. (2022). STEWART: Stacking ensemble for white-box adversarial attacks towards more resilient data-driven predictive maintenance. Computers in Industry, 140: 103660. https://doi.org/10.1016/j.compind.2022.103660

[25] Uhlmann, E., Pontes, R.P., Geisert, C., Hohwieler, E. (2018). Cluster identification of sensor data for predictive maintenance in a selective laser melting machine tool. Procedia Manufacturing, 24: 60-65. https://doi.org/10.1016/j.promfg.2018.06.013

[26] Welte, R., Estler, M., Lucke, D. (2020). A method for implementation of machine learning solutions for predictive maintenance in small and medium sized enterprises. Procedia CIRP, 93: 909-914. https://doi.org/10.1016/j.procir.2020.04.052

[27] Schmidt, B., Wang, L., Galar, D. (2017). Semantic framework for predictive maintenance in a cloud environment. Procedia CIRP, 62: 583-588. https://doi.org/10.1016/j.procir.2016.06.047

[28] Zhai, S., Gehring, B., Reinhart, G. (2021). Enabling predictive maintenance integrated production scheduling by operation-specific health prognostics with generative deep learning. Journal of Manufacturing Systems, 61: 830-855. https://doi.org/10.1016/j.jmsy.2021.02.006

[29] Ruschel, E., Santos, E.A.P., Loures, E.D.F.R. (2017). Mining shop-floor data for preventive maintenance management: Integrating probabilistic and predictive models. Procedia Manufacturing, 11: 1127-1134. https://doi.org/10.1016/j.promfg.2017.07.234

[30] Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58: 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

APPENDIX

Table A1. CERQual qualitative evidence profile for key review findings

Finding 1: AI, particularly DL (CNNs, LSTMs, AEs), dominates recent PdM research for tasks like RUL prediction and fault diagnosis, showing potential for high accuracy.	
Studies Contributing to Finding	[4, 5, 6, 7, 13, 14, 16, 18, 19, 29]
Assessment of CERQual Components	Methodological Limitations: Moderate concerns (validation often on benchmark datasets [e.g., 5, 6, 19] or specific cases [e.g., 4, 13]; varying reporting quality). Coherence: High (consistent trend reported across numerous studies). Adequacy: High (rich data from many studies detailing methods/results). Relevance: High (directly addresses RQ1 & RQ2).
CERQual Assessment of Confidence	High
Explanation for Confidence Assessment	Confidence upgraded to High. Despite moderate methodological concerns in some primary studies regarding validation scope, the sheer volume, consistency (coherence), and richness (adequacy) of evidence across the corpus strongly supports this finding's representativeness.
Finding 2: Integration of AI-PdM with broader systems (Production Planning, Quality Control, Digital Twins) is a key trend, aiming for holistic optimization.	
Studies Contributing to Finding	[7, 14, 16, 18, 22, 24, 23]
Assessment of CERQual Components	Methodological Limitations: Moderate concerns (many studies propose frameworks [7, 14, 18, 24] or simulations [16, 29] with limited large-scale industrial validation reported within these papers). Coherence: Moderate (general trend towards integration is clear, but specific approaches vary significantly). Adequacy: Moderate (sufficient studies discuss integration, but detailed implementation data is less rich). Relevance: High (directly addresses RQ2 & RQ4).
CERQual Assessment of Confidence	Moderate
Explanation for Confidence	Confidence remains Moderate. While the trend towards integration is coherent and

Assessment	relevant, the evidence base contains many conceptual frameworks or specific simulations, with fewer detailed accounts of fully validated, large-scale industrial implementations within this corpus. Coherence on how to integrate varies.
Finding 3: Data-related issues (availability, quality, heterogeneity, labeling) and model interpretability (XAI) remain major, persistent challenges in AI-driven PdM.	
Studies Contributing to Finding (Examples)	[4, 7, 9, 13, 14, 15, 16, 18, 19, 21, 22, 23, 28]
Assessment of CERQual Components	Methodological Limitations: Minor concerns (these challenges are widely reported across studies with diverse methods, strengthening the finding). Coherence: High (very consistent reporting of these issues across almost all relevant studies). Adequacy: High (rich descriptions and examples provided in numerous studies). Relevance: High (directly addresses RQ3).
CERQual Assessment of Confidence	High
Explanation for Confidence Assessment	Confidence remains High due to the strong coherence and consistency with which data and interpretability challenges are reported across a large number of diverse studies within the corpus. The evidence is rich and

	directly relevant to identifying key barriers.
Finding 4: Emerging research focuses on uncertainty quantification, DRL for scheduling, cybersecurity resilience, and systematic implementation methods.	
Studies Contributing to Finding (Examples)	[1, 2, 5, 6, 9, 10, 15, 17, 23]
Assessment of CERQual Components	Methodological Limitations: Moderate concerns (some areas like DRL [3, 19] and security [9] are relatively newer with fewer validation studies compared to core ML/DL). Coherence: Moderate (different emerging areas addressed by distinct subsets of studies, coherence within each area is good). Adequacy: Moderate (sufficient studies to identify trends, but data for some emerging areas is less extensive than core themes). Relevance: High (directly addresses RQ4).
CERQual Assessment of Confidence	Moderate
Explanation for Confidence Assessment	Confidence remains Moderate. These trends are clearly identifiable and relevant, supported by adequate data. However, some areas are more nascent, with fewer contributing studies or less extensive validation reported within this specific corpus compared to the more established findings, slightly reducing confidence.

668