A Knowledge Graph-Driven Approach for Predictive Maintenance and Fault Diagnosis in Offshore Wind Turbines

[1]Yuxuan Chen , [2]Sofia Lindberg , [3]Jianping Wang
Department of Wind and Energy Systems, Technical University of Denmark (DTU), Kongens Lyngby, Denmark

ABSTRACT

Offshore wind turbines (OWTs) represent critical assets in the global transition to renewable energy, yet their operation and maintenance (O&M) costs can constitute up to 30% of the levelized cost of energy due to harsh marine environments, limited accessibility, and complex failure modes in gearboxes, blades, generators, and drivetrains. This review examines a knowledge graph (KG)-driven paradigm for predictive maintenance (PdM) and fault diagnosis (FD) as an exemplar of artificial intelligence (AI) empowerment within intelligent engineering systems (IES). Grounded in cyber-physical systems (CPS) theory, systems engineering principles, and information-theoretic perspectives on uncertainty reduction, the approach integrates heterogeneous data sources—including SCADA streams, vibration signatures, maintenance logs, weather data, and OEM documentation—into semantically structured graphs. These graphs enable causal reasoning, multi-label fault classification, and maintenance action prediction through graph neural networks (GNNs) such as time-domain relational graph convolutional networks (TRGCNs). Concrete implementations, including simulation-augmented KGs achieving over 94% fault-mode accuracy and 90–93% maintenance-action prediction precision for gearboxes, demonstrate substantial reductions in unplanned downtime and vessel deployment costs. The framework exemplifies broader IES transformations: from reactive to proactive strategies, isolated monitoring to integrated CPS ecosystems, and manual interventions to semi-autonomous control augmented by digital twins (DTs) and edge intelligence. Challenges include data heterogeneity, ontology incompleteness in rare-event offshore scenarios, cybersecurity vulnerabilities, and the tension between model expressiveness and real-time latency. Prospects encompass neuro-symbolic hybrids with foundation models, physics-informed embeddings, and multi-agent cognitive CPS architectures. By bridging data-driven pattern recognition with explicit domain knowledge, KG-driven methods enhance explainability, resilience, and lifecycle value in distributed energy systems while highlighting the necessity of standardized data pipelines and human-AI collaboration protocols.

Keywords: Knowledge Graph; Predictive Maintenance; Fault Diagnosis; Offshore Wind Turbine; Digital Twin; Cyber-Physical Systems; Graph Neural Network; Intelligent Engineering Systems

1. INTRODUCTION

Offshore wind power has emerged as a cornerstone of decarbonization strategies, with global installed capacity projected to exceed 500 GW by 2035. However, the remote and corrosive marine setting imposes severe constraints on O&M, where corrective interventions often require specialized vessels or helicopters, incurring costs orders of magnitude higher than onshore equivalents. Gearbox failures, blade erosion, generator bearing degradation, and hydraulic system faults frequently propagate from subtle sensor anomalies to catastrophic downtime, eroding asset availability and revenue. Conventional time-based or reactive maintenance proves inadequate for minimizing these losses, prompting a paradigm shift toward AI-augmented PdM and FD.

A particularly promising direction centers on knowledge graph-driven approaches. By explicitly modeling entities (components, fault modes, sensor signals, maintenance actions) and typed relations (causes, affects, requires, co-occurs), KGs furnish a semantic layer that fuses multi-modal data while supporting traceable inference. When coupled with GNN variants such as TRGCNs and physics-informed digital twins, these graphs enable simultaneous fault classification, severity assessment, and optimal intervention planning. Empirical demonstrations on gearbox subsystems have reported average accuracies of 94.06% for fault modes and 97.10% for degradation states, alongside maintenance-action prediction exceeding 90% across inspection, minor repair, major repair, and replacement categories [1]. Complementary reviews document the rapid evolution from physics-based and classical machine-learning methods toward hybrid DT–AI frameworks that address data scarcity, class imbalance, and the need for edge-deployable, privacy-preserving solutions [2,3].

This trajectory aligns with the wider transformation of intelligent engineering systems. OWT farms constitute complex CPS in which physical dynamics (aerodynamics, structural loading, electrical conversion) are tightly coupled with computational layers for sensing, modeling, prediction, and control. The present review situates the KG-driven PdM/FD approach within the theoretical foundations of IES, traces the socio-technical transformation paths enabled by AI, surveys concrete applications with emphasis on fault diagnosis and predictive maintenance, and critically examines challenges alongside future prospects. Two recurring concrete exemplars anchor the discussion: (i) simulation-augmented maintenance KGs integrated with TRGCNs for drivetrain subsystems, and (ii) hybrid physics–data digital twins synchronized via IoT/edge–cloud architectures for remaining-useful-life (RUL) estimation and fleet-level optimization. The analysis maintains a deliberate balance between demonstrated benefits—reduced unplanned downtime, enhanced explainability, and lifecycle cost savings—and persistent limitations arising from environmental harshness, data quality, and integration complexity.

2. The Theoretical Foundation of Artificial Intelligence and Intelligent Engineering Systems

2.1 The Connotation and Characteristics of Intelligent Engineering Systems

Intelligent engineering systems are socio-technical constructs that integrate physical assets, computational intelligence, human actors, and organizational processes into coherent wholes exhibiting emergent properties such as adaptability, resilience, and autonomous decision-making. Drawing on systems engineering theory, IES are characterized by hierarchical decomposition with cross-layer feedback, explicit handling of uncertainty, and continuous co-evolution of structure and function. In the offshore wind domain, an IES encompasses individual turbine CPS (rotor–nacelle–tower–foundation), farm-level supervisory control, supply-chain logistics, grid interfaces, and regulatory oversight.

Key characteristics include semantic interoperability (enabled by ontologies and KGs), real-time situational awareness through sensor–edge–cloud pipelines, and value-driven optimization balancing energy yield, structural integrity, safety, and economic return. CPS theory emphasizes the tight coupling between continuous physical dynamics and discrete cyber computations; information theory complements this view by framing intelligence as the systematic reduction of entropy through structured representation and inference. KGs exemplify this principle: raw high-dimensional SCADA and vibration streams possess high Shannon entropy, whereas instantiated graph entities and relations compress domain knowledge into queryable, causal structures that support both diagnostic and prescriptive reasoning.

2.2 The Theoretical Logic of Artificial Intelligence Empowering Engineering Systems

AI empowers IES by supplying four interlocking capabilities: perception (feature extraction from heterogeneous signals), cognition (causal reasoning and knowledge integration via KGs or neuro-symbolic hybrids), decision (optimization and planning, e.g., via reinforcement learning), and action (closed-loop control or human–machine task allocation). Within CPS architectures, AI algorithms operate across edge (low-latency anomaly detection), fog (regional coordination), and cloud (fleet-level KG reasoning and DT synchronization) layers.

The logical progression proceeds from data to information to knowledge to actionable intelligence. Information-theoretic gains arise when KGs explicitly encode conditional dependencies (e.g., "bearing temperature rise precedes gearbox fault under high wind shear"), thereby lowering the uncertainty associated with downstream maintenance decisions. Systems engineering further demands verification and validation of emergent behaviors; KG-augmented DTs facilitate this by maintaining synchronized virtual replicas whose predictions can be traced back to specific sensor evidence and ontological axioms. An original insight here is that KGs function as "semantic middleware," mediating between low-level signal processing and high-level enterprise optimization, thereby enabling the transition from component-centric to ecosystem-centric intelligence in distributed energy assets.

3. The Transformation Path of Intelligent Engineering Systems Mode Driven by Artificial Intelligence

3.1 Transition from Standardized Production to Personalized and Flexible Manufacturing

Although OWTs are manufactured to standardized designs, their in-service behavior exhibits pronounced individuality arising from site-specific wind regimes, installation variances, and progressive degradation. AI-driven IES enable a shift toward personalized asset management. Component-level digital profiles—continuously updated via KG instantiation of sensor observations, inspection reports, and maintenance history—support tailored PdM schedules rather than fleet-wide fixed intervals. For example, a gearbox whose degradation trajectory deviates from the population mean can trigger earlier oil analysis or vibration thresholding, while a turbine operating in a milder micro-climate may safely extend inspection windows. This personalization reduces both over-maintenance and catastrophic failures.

3.2 Transition from Reactive Response to Proactive Prediction and Optimization

The most consequential transformation lies in replacing reactive repair with proactive prediction. Traditional SCADA threshold alarms frequently generate excessive false positives, prompting costly and unnecessary vessel mobilizations. KG-driven methods address this by embedding domain knowledge (failure mode and effects analysis outcomes, physics-based degradation models) alongside time-series patterns. The TRGCN architecture processes temporal relational graphs to jointly predict fault modes, degradation severity, and recommended actions, achieving multi-label classification accuracies above 94% and maintenance-action prediction above 90% in gearbox case studies [1]. Complementary DT frameworks integrate environmental, operational, and market data to estimate failure probabilities and optimize O&M task scheduling under weather-window constraints [3,4]. Reinforcement-learning policies further optimize dynamic decisions such as crew routing or opportunistic maintenance during low-wind periods, converting weather and health forecasts into cost-minimizing actions.

3.3 Transition from Manual Operation to Autonomous Intelligent Control

Edge-deployed lightweight models perform instantaneous anomaly detection and protective actions (e.g., derating or safe shutdown) with sub-second latency, while the central KG maintains fleet-wide context and supplies explainable recommendations to remote operations centers. Human–machine collaboration is elevated: operators receive not merely alerts but causally annotated pathways (“high vibration at bearing frequency → lubricant degradation probable → recommend oil sampling before next weather window”). This semi-autonomous loop preserves human oversight for safety-critical or ethically sensitive decisions while automating routine optimization.

3.4 Transition from Isolated Systems to Integrated Cyber-Physical Ecosystems

Isolated turbine SCADA systems are giving way to integrated ecosystems linking turbines, substations, maintenance logistics, insurers, and grid operators through shared semantic models. KGs serve as the unifying substrate, representing not only technical entities but also contractual obligations, spare-part inventories, and regulatory constraints. DTs synchronized via 5G or satellite links enable real-time what-if simulation at farm scale, supporting coordinated wake steering that respects individual turbine health states. An original extension of the transformation framework is the notion of “knowledge-infused resilience”: semantic links permit rapid analogical reasoning when novel fault signatures appear, allowing the system to retrieve similar historical cases or physics-based simulations even when training data are sparse.

4. Applications of Artificial Intelligence in Intelligent Engineering Systems

4.1 Application in Intelligent Fault Diagnosis and Predictive Maintenance

The KG-driven approach constitutes the flagship application for OWT PdM and FD. Construction begins with systematic elicitation of entities and relations from FMEA, OEM manuals, and historical work orders, augmented by simulation-generated fault and degradation trajectories (e.g., via AMESim drivetrain models) [1]. The resulting maintenance KG encodes multi-label relationships among fault modes, severity indicators, sensor signatures, and feasible interventions. TRGCN layers then perform joint node classification (fault mode and degradation state) and link prediction (optimal maintenance action) on time-evolving graphs derived from SCADA and vibration streams.

Empirical performance on gearbox subsystems demonstrates practical viability: average fault-mode accuracy of 94.06%, degradation-state accuracy of 97.10%, and maintenance-action prediction ranging from 90.9% (replacements) to 93.4% (inspections) [1]. Complementary data–knowledge fusion methods combine real-time monitoring streams with KG reasoning to accelerate fault-type inference in offshore settings [6]. Digital twins extend these capabilities by maintaining high-fidelity physics-based replicas synchronized with live sensor data, enabling RUL estimation via hybrid physics–data models and deep architectures (CNNs, LSTMs, Transformers) [3]. Computer-vision pipelines applied to drone or fixed-camera imagery of blades detect early erosion or cracks; detected anomalies are instantiated as new KG entities, triggering diagnostic queries that fuse visual evidence with vibration and SCADA context.

Edge computing complements the architecture: lightweight inference models running on turbine controllers or nearby gateways execute real-time anomaly detection and protective logic, forwarding compact embeddings or flagged events to the central KG for fleet-level reasoning. This hybrid edge–cloud design satisfies the stringent latency requirements of safety-critical protection while preserving the explanatory power of the graph. Critical appraisal reveals limitations: KG construction demands substantial domain expertise and remains vulnerable to incomplete ontologies when rare or novel faults occur; offshore sensor drift (corrosion, biofouling) can degrade input quality; and high false-positive rates still translate into expensive vessel call-outs. Nevertheless, the explicit relational structure markedly improves explainability relative to opaque deep-learning baselines, facilitating regulatory acceptance and operator trust [2].

4.2 Application in Intelligent Process Optimization and Quality Control

Beyond fault management, AI contributes to process-level optimization. Health-aware control policies adjust individual turbine setpoints (pitch, torque) or farm-level wake steering to maximize energy capture while respecting remaining useful life constraints encoded in the KG or DT. Reinforcement-learning agents trained on high-fidelity simulators learn policies that trade short-term power against long-term component degradation under stochastic wind and wave conditions. In manufacturing and installation phases, in-line computer-vision inspection of composite layups or weld quality, coupled with KG-tracked component provenance, supports zero-defect production and traceability—foundational for subsequent PdM accuracy.

4.3 Application in Intelligent Human-Machine Collaboration and Operational Efficiency Enhancement

Remote operations centers increasingly rely on DT visualizations whose underlying KG supplies causally annotated decision support. Augmented-reality interfaces overlay repair procedures, torque specifications, and safety constraints directly onto the technician's field of view, reducing cognitive load and error rates during limited weather windows. Multi-agent coordination—where each turbine's local DT/KG negotiates with neighbors and central logistics—optimizes vessel routing, spare-part positioning, and personnel scheduling. These applications collectively compress the O&M cost curve while elevating human roles from reactive firefighting to strategic oversight and exception handling.

5. Challenges and Prospects of Artificial Intelligence Empowering Intelligent Engineering Systems

5.1 Current Main Challenges

Several interrelated challenges impede large-scale deployment. Data heterogeneity and quality remain acute: SCADA protocols differ across OEMs, vibration sampling rates vary, and marine-induced sensor degradation introduces non-stationary noise. Class imbalance is inherent—catastrophic faults are rare, yet their consequences are severe—necessitating sophisticated semi-supervised or simulation-augmented learning [5]. KG scalability for farms comprising hundreds of turbines strains both storage and real-time inference; incremental graph updates and approximate reasoning techniques are active research areas. Cybersecurity constitutes a systemic risk: connected CPS expose attack surfaces that could manipulate sensor feeds or KG content, with potentially cascading physical consequences. False-positive costs in offshore logistics are punishing; even modest rates can negate predictive savings. Finally, standardization gaps persist regarding data formats, ontology alignment, model validation, and liability allocation when AI recommendations influence safety-critical decisions.

5.2 Prospects for Future Development Trends

Prospects are nonetheless substantial. Neuro-symbolic integration—combining large foundation models with KGs—promises automated population and querying of maintenance graphs from unstructured maintenance reports and design documentation, mitigating the ontology bottleneck [2]. Physics-informed embeddings that embed governing equations (structural dynamics, fluid–structure interaction) directly into graph neural networks will improve generalization under distribution shift. Edge–cloud–satellite hierarchies leveraging 6G and non-terrestrial networks will enhance resilience of remote farms. Multi-agent cognitive CPS architectures, in which autonomous agents negotiate maintenance priorities and energy-market participation while maintaining traceable rationales, represent a natural evolution of the IES transformation. Pilot deployments combining drone-based computer vision, onboard edge inference, and central KG reasoning are already demonstrating closed-loop autonomy potential. An original forward-looking insight is that the convergence of KGs, DTs, and foundation models heralds "cognitive CPS" in renewable energy—systems capable not only of prediction but of causal explanation, counterfactual reasoning, and proactive self-configuration—thereby fostering the trust, auditability, and regulatory alignment required for autonomous operation at scale.

6. Conclusion

Knowledge graph-driven approaches to predictive maintenance and fault diagnosis constitute a technically mature and theoretically grounded instantiation of AI empowerment within intelligent engineering systems for offshore wind turbines. By furnishing an explicit semantic substrate that integrates heterogeneous data, domain knowledge, and physics-based models, these methods deliver measurable gains in diagnostic accuracy, maintenance-action prediction, and operational explainability while exemplifying the broader transitions from reactive to proactive, isolated to ecosystem-integrated, and manual to semi-autonomous modes of engineering system operation. Concrete realizations—simulation-augmented TRGCN frameworks achieving >94% fault classification and hybrid DT–AI pipelines for RUL estimation—illustrate both the promise and the remaining hurdles of data quality, scalability, and cybersecurity in harsh marine environments. Future progress hinges on continued advances in neuro-symbolic architectures, standardized semantic interoperability, and human-centered validation protocols. As offshore wind scales globally, KG-augmented cognitive CPS offer a principled pathway toward higher availability, lower lifecycle costs, and ultimately more sustainable and resilient renewable energy infrastructure.

REFERENCES

[1] Kim, H. S., & Lee, J. (2023). Maintenance knowledge graph and time-series relational graph convolution neural networks for enhancing predictive maintenance of wind turbine gearbox. Knowledge-Based Systems, 338, 115523. https://doi.org/10.1016/j.knosys.2023.115523

[2] Jiang, R., Fang, F., Rodríguez-Andina, J. J., Song, Z., Liu, J., Chen, Y., & Wang, H. (2023). Artificial intelligence in wind turbine fault diagnosis: A systematic knowledge mapping and trend analysis. Renewable and Sustainable Energy Reviews, 226(Part D), 116321. https://doi.org/10.1016/j.rser.2025.116321

[3] Maldonado-Correa, J., Cuenca-Granda, J., Torres-Cabrera, J., et al. (2023). Digital Twin Technology in Wind Turbine Condition Monitoring, Predictive Maintenance, and RUL Estimation: A Systematic Literature Review. Energies, 19(6), 1477. https://doi.org/10.3390/en19061477

[4] Kandemir, E., et al. (2024). Predictive digital twin for wind energy systems. Energy Informatics. https://doi.org/10.1186/s42162-024-00373-9

[5] Barnabei, V. F., et al. (2025). Semi-Supervised Deep Learning Framework for Predictive Maintenance of Horizontal Axis Wind Turbines. Wind, 10(3), 14. https://doi.org/10.3390/wind10030014

[6] Xu, Q., et al. (2023). Knowledge and Data Fusion-driven for Offshore Wind Turbine Fault Diagnosis. Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023).

[7] Chen, X., et al. (2023). A concept for human-cyber-physical systems of future wind turbines. Technical University of Denmark.

[8] Yeman, R. (2023). Systems Engineering applied to Artificial intelligence in context of Cyber-Physical Systems. Systems Engineering Research Center.