Industrial Internet of Things-Enabled Fault Detection and Diagnosis in Measurement and Control Systems: A Bibliometric Analysis and Systematic Mapping Review

Bui Thanh Lam

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam

Corresponding author: Bui Thanh Lam; Email: lamt78789@gmail.com; ORCID: https://orcid.org/0009-0009-4575-0757

Article type	Review Article
Running title	IIoT-enabled fault detection and diagnosis
Word/illustration count	5,081 words; 7 figures; 6 tables

Abstract

Industrial Internet of Things (IIoT) architectures are shifting fault detection and diagnosis (FDD) from centralized supervisory computers toward smart instruments, edge controllers and distributed cyber-physical systems, yet evidence remains fragmented across metrology, networking, embedded computing and machine learning. This bibliometric analysis and systematic mapping review maps the field's structure and examines the measurement-to-action chain. A PRISMA-ScR-informed protocol retrieved 447 OpenAlex records published in 2010-2025. After normalized-title deduplication, deterministic engineering-relevance screening, retraction checks, Crossref DOI verification and direct source/year auditing against the July 2026 official Scopus source list, 192 journal articles and reviews formed the frozen analytic corpus. The source list was used for source/year parity, not document-level Scopus coverage. All 192 DOI-title pairs matched Crossref. Output rose from five papers in 2018 to 55 in 2025 (40.9% compound annual growth). The corpus spans 97 sources, 742 disambiguated author identifiers and 54 countries. Bradford zones showed a compact core but not a strict geometric source sequence; a discrete Lotka model yielded an exponent of 4.16 with a bootstrap goodness-of-fit p value of 0.306. A three-community co-citation network had moderate modularity (Q = 0.267). Explicit edge-cloud architecture occurred in 86.3% of 2023-2025 title/abstract records, whereas explicit decision or closed-loop control language occurred in 8.5%. Because 34 records lacked abstracts, reporting frequencies are lower-bound indicators rather than study-quality estimates. The review concludes that progress depends on traceable measurement uncertainty, grouped cross-machine validation, calibrated abstention, bounded end-to-end latency, semantic interoperability and safely governed supervisory action.

Keywords: industrial Internet of Things; fault detection and diagnosis; condition monitoring; measurement uncertainty; edge intelligence; bibliometric analysis

1. Introduction

Industrial fault detection and diagnosis converts measured evidence into statements about whether an asset or process is abnormal, where the abnormality resides, why it occurred and what action should follow. Classical installations are hierarchical: transducers feed programmable logic controller (PLC) or remote terminal unit (RTU) inputs; deterministic control executes locally; supervisory control and data acquisition (SCADA) or distributed control systems archive trends and alarms; and engineering workstations support diagnosis. This structure remains valuable because bounded scan times, validated function blocks and controlled configuration suit availability- and safety-critical operation. It is less well suited to heterogeneous fleets, high-rate waveform data, remote assets and models that learn across sites.

IIoT architectures extend this hierarchy with addressable smart instruments, industrial Ethernet and wireless links, edge/fog computation, scalable storage, digital twins and fleet analytics.[1-6] Connectivity is not the engineering endpoint. A defensible FDD system is a measurement-to-action loop: a sensor produces a traceable, time-stamped signal; acquisition preserves bandwidth and quality; communications preserve order, semantics and freshness; edge or cloud software makes a calibrated inference; and a human, maintenance planner or supervisory controller acts before the condition becomes unsafe.

FDD predates IIoT. Canonical reviews distinguish quantitative model-based, qualitative model-based and process-history methods, while machinery and process-monitoring reviews connect acquisition, signal processing, diagnosis and remaining useful life.[7-13] Later syntheses emphasize deep learning and machine health, predictive-maintenance algorithms, deployment barriers, Industry 4.0 taxonomies, knowledge-based diagnosis, industrial cyber-physical FDD and AI-oriented bibliometrics.[14-24] These reviews are valuable, but they do not jointly audit bibliometric structure, the hardware-network-firmware-algorithm stack, measurement uncertainty, real-time execution, cyber-physical ambiguity and control closure.

This review addresses four questions:

- RQ1. How have publication output, source concentration, author productivity, geographic collaboration and the co-citation base evolved?
- RQ2. Which sensing, networking, embedded-computing and algorithmic themes structure IIoT-enabled FDD, and how have they changed?
- RQ3. How has the architecture evolved from PLC/RTU-centred monitoring to smart transmitters, edge controllers and digital twins?
- RQ4. Which metrological, timing, interoperability, security, generalization and closed-loop validation problems should define the next research agenda?

The contribution is not the visual mapping alone. Bibliometric structure is used to direct a standards-informed engineering synthesis, and numerical claims are intended to remain traceable to a frozen record-level audit trail. The systematic component is a structured engineering evidence synthesis, not an intervention-effectiveness meta-analysis: no pooled effect estimate or study-level risk-of-bias grade is attempted.

Table 1. Position of the present review relative to representative syntheses.

Review stream	Primary emphasis	Main strength	Gap addressed here
Classical process/machinery FDD	Models, signals, diagnosis and prognostics	Durable engineering taxonomy	Limited distributed architecture and embedded constraints
Machine/deep-learning reviews	Features and diagnostic accuracy	Algorithmic breadth	Sparse metrological, timing and control-closure analysis
Predictive-maintenance reviews	Applications, decisions and barriers	Links diagnosis to asset management	IIoT stack and uncertainty are not jointly quantified
IIoT/CPS reviews	Knowledge-based or cross-domain diagnosis	System context and method families	No integrated bibliometrics plus four-layer assurance audit
Present review	Traceable measurement-to-action chain	Bibliometrics plus technical synthesis	Hardware, timing, firmware, networks, algorithms, security and safe action

2. Research methodology

2.1. Review design and reporting safeguards

The design combines bibliometric performance analysis, science mapping and a structured technical synthesis. PRISMA 2020 informed identification and flow reporting, PRISMA-ScR informed scope charting and PRISMA-S informed reporting of databases, dates, fields and limits.[25-27] AMSTAR 2 was developed for reviews of healthcare interventions and was therefore not used to issue a formal confidence grade for this engineering review.[28] Its safeguards - an explicit protocol, justified exclusions, duplicate assessment, source-bias discussion and disclosure - were used only as a methodological cross-check.

Table 2. Reporting and appraisal crosswalk. Residual limitations are stated rather than presented as compliance.

Framework	Safeguard applied	Residual limitation
PRISMA 2020	Eligibility rules, record-level decisions, exclusion reasons and flow diagram	One author-supervised screening stream; no independent duplicate screener
PRISMA-ScR	Explicit concept, context, evidence types, charted technical domains and gap synthesis	Not a full-text effectiveness review or meta-analysis
PRISMA-S	Syntax, field mapping, retrieval date, limits, deduplication and update instructions	Licensed Scopus/WoS record exports were unavailable
AMSTAR 2 safeguards	Protocol, provenance, retraction checks, source sensitivity, limitations and disclosures	Formal AMSTAR 2 rating is inapplicable; protocol was not prospectively registered
Technical evidence audit	Reproducible title/abstract flags for deployment, validation, uncertainty, timing and action	Metadata absence is not evidence of study-level non-reporting

2.2. Search strategy and database sensitivity

The open audit layer uses OpenAlex because record-level metadata and cited-reference links can be redistributed and reproduced.[29] The query combined three concept blocks: IIoT/industrial cyber-physical systems; FDD/condition monitoring/predictive maintenance/prognostics/remaining useful life; and physical measurement/control. It was applied to titles and abstracts, limited to English-

language journal articles and reviews, and restricted to complete publication years 2010-2025. Retrieval occurred on 21 August 2026; citation counts are therefore a dated snapshot. Exact strings are supplied in Supplementary Methods S1.

OpenAlex was not treated as an equivalent substitute for Scopus or Web of Science (WoS). Comparative work reports competitive reference coverage but database-dependent metadata completeness.[30] The Scopus and WoS sensitivity strings use title and abstract fields only, avoiding an invalid comparison with Scopus TITLE-ABS-KEY or WoS Topic Search. Licensed record exports were unavailable, so record-level overlap and field-completeness estimates remain pending rather than being inferred.

A source-level coverage audit used the official Scopus source-title workbook dated July 2026. OpenAlex source_is_core was retained as descriptive metadata but was not used as an eligibility gate. Among 234 engineering-relevant, journal and non-retracted candidates, 193 had source/year parity; 29 unmatched sources, four papers outside a source's coverage window and eight serial proceedings were excluded. One source/year-matched no-DOI record was excluded because Crossref document provenance was unavailable. This procedure checks source provenance, not equivalence between an OpenAlex query and a Scopus document export.

2.3. Eligibility, deduplication and integrity screening

The broad query returned 447 records. Normalized titles identified seven duplicate manifestations, leaving 440 records for title/abstract screening. A deterministic centrality rule retained records whose titles made fault, diagnosis, anomaly, failure, degradation, health monitoring, tool wear, predictive maintenance, prognostics or remaining useful life central, provided metadata also supported a physical measurement/control context. Identity/access-control, cloud/container availability and pure cyber-security or communications-link/network diagnostics were excluded when no physical-asset or measurement-fault contribution was present. Title centrality was chosen to increase precision but can omit relevant studies whose FDD objective is visible only in the abstract.

Of 440 screened records, 194 failed engineering relevance and 246 underwent provenance assessment. Exclusions comprised 10 non-journal manifestations, two retracted records, 41 official Scopus source/year or serial-proceedings failures and one unverified no-DOI record, leaving 192 records (Figure 1). All 192 included DOI-title pairs resolved and matched Crossref. Relative to v2.0.1, 179 records are common, 13 source/year-verified records were added and four out-of-scope records were removed; headline directions were stable. Excluding no-DOI material may underrepresent regional or older engineering work.

Table 3. Operational inclusion and exclusion criteria.

Dimension	Include	Exclude
Phenomenon	Detection, isolation, diagnosis, anomaly/failure monitoring, prognostics, predictive maintenance or RUL is title-central	Terminology is peripheral to the title or unrelated to asset health

Dimension	Include	Exclude
Industrial context	IIoT, industrial CPS/CPPS, smart manufacturing or measured industrial asset/process	Consumer IoT, business, healthcare or smart city without industrial automation
Measurement/control	Physical sensor, signal, instrument, PLC/RTU/SCADA, controller, actuator, machine or process	Identity/access, cloud/container, cyber or network diagnosis without a physical-fault consequence
Provenance	English journal article/review, 2010-2025, DOI verified, eligible source/year	Duplicate, proceeding, repository, retraction, unverifiable or out-of-coverage source

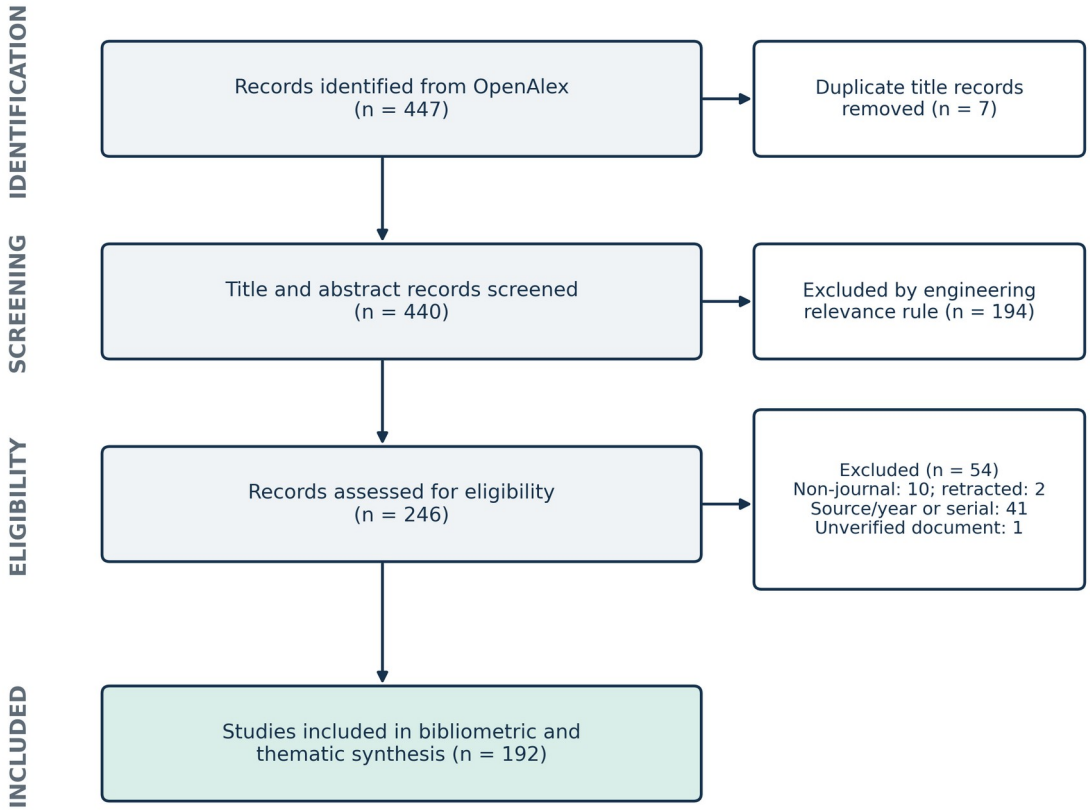

Figure 1. ***PRISMA-ScR-informed flow for the frozen metadata corpus. Commercial database document exports were not used; the source/year audit used the official Scopus source list.***

2.4. Bibliometric calculations and evidence charting

Annual output, snapshot citations, source productivity, author productivity and country participation were calculated from record matrices. Compound annual growth was calculated from 2018, the first year with at least five included records:

$$\text{CAGR} = (\text{N\_2025}/\text{N\_2018})\hat{}(1/7) - 1 = (55/5)\hat{}(1/7) - 1 = 0.4085. \quad (1)$$

Bradford zones were constructed by ranking sources and accumulating approximately one third of the 192 records per zone.[31] This tests concentration descriptively; it does not establish the geometric 1:n:n^2 sequence. A log-least-squares multiplier and relative root-mean-square error (RRMSE) evaluated that additional claim. Author productivity was fitted as a discrete power law:

$$p(x \mid \alpha) = x\hat{}(-\alpha)/\zeta(\alpha, 1), \quad x = 1, 2, \ldots \quad (2)$$

Maximum likelihood, the Kolmogorov-Smirnov (KS) distance and 1,000 refitted parametric-bootstrap samples were used.[32] Co-word analysis used a prespecified controlled engineering lexicon rather than uncurated keywords. Twenty-three phrases occurring in at least six records were grouped into FDD task, architecture, algorithm and measurement/asset families. Country collaboration counted country presence per record; an international paper contained at least two country codes. Reference co-citation used the 24 most locally shared topic-relevant references. The full weighted matrix was partitioned with deterministic greedy modularity. Export routines targeted bibliometrix/Biblioshiny and VOSviewer-compatible files.[33,34]

The four-layer technical charting frame described in Section 4 was applied to titles and abstracts. Seven evidence domains were flagged: physical/deployment validation; grouped or external validation; measurement traceability/uncertainty; calibration/OOD/abstention; timing/resource/energy; control or maintenance action; and cyber/physical ambiguity. Frequencies measure explicit metadata reporting only; they are not full-text risk-of-bias scores.

3. Bibliometric performance and science mapping

3.1. Growth, citations and source concentration

The first eligible record appeared in 2015. Annual output rose from five papers in 2018 to 55 in 2025, a 40.9% CAGR under Equation 1 (Figure 2). The corpus had accumulated 6,631 citations by retrieval (mean 34.5; median 18). The mean-median gap and dated snapshot preclude treating citation count as an invariant measure of quality.

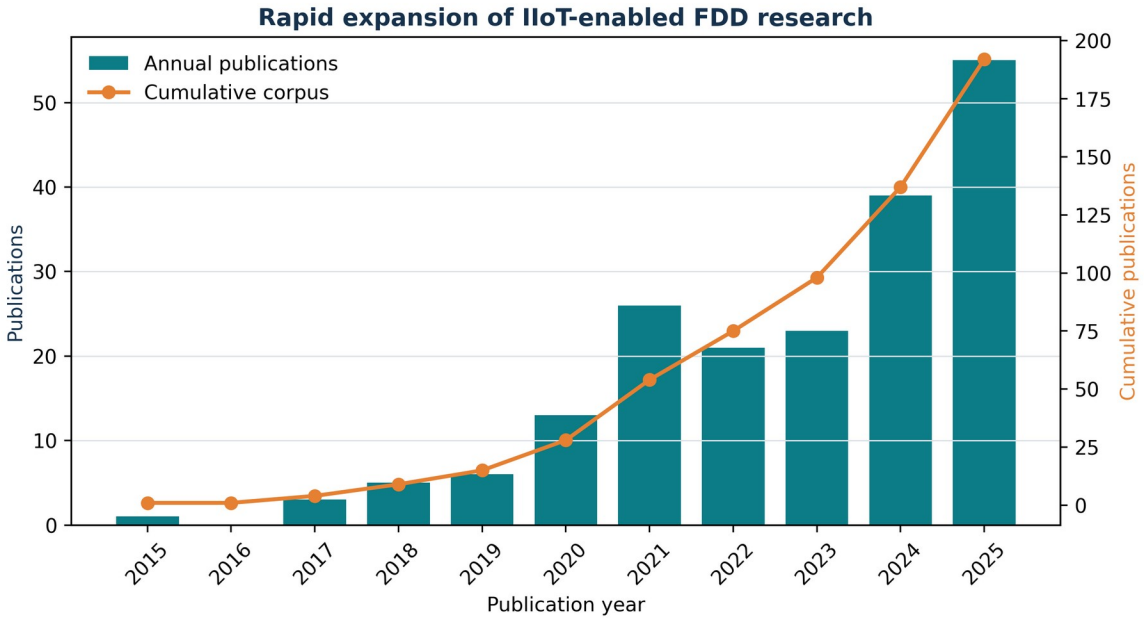

Figure 2. ***Annual and cumulative publication output in the 192-record primary corpus.***

The 192 records occupied 97 sources. IEEE Internet of Things Journal led with 32 papers, followed by Sensors (20), IEEE Transactions on Industrial Informatics (9) and Applied Sciences (6). Bradford zones contained 67, 64 and 61 papers across 4, 32 and 61 sources. Article balance was tight (coefficient of variation 3.8%), but source multipliers were 8.00 and 1.91. The best geometric multiplier was 4.51, predicting 4, 18.0 and 81.3 sources; observed zones were 4, 32 and 61 (RRMSE 31.7%). The defensible interpretation is a compact core and long tail, not confirmation of a strict Bradford law.

Authorship was fragmented: 742 disambiguated OpenAlex author identifiers produced 807 author-paper appearances. Counts were 689 authors with one paper, 44 with two, six with three and three with four. Identifiers were complete for 187 of 192 records. The discrete Lotka fit gave alpha = 4.156, KS = 0.00396 and bootstrap p = 0.306. The model is not rejected at 0.05, but its narrow observed support (one to four papers) and steep exponent indicate high transience; it does not establish a universal law.

Table 4. Performance and integrity indicators for the frozen primary corpus.

Indicator	Value	Indicator	Value
Included records	192	Crossref-verified DOI pairs	192 (100%)
Publication years	2015-2025	Snapshot citations	6,631
Sources	97	Bradford zones (sources)	4 / 32 / 61
Unique author IDs	742	One-paper authors	92.9%
Countries	54	International papers	60 (31.3%)
Lotka alpha / p	4.156 / 0.306	Co-citation Q / communities	0.267 / 3

3.2. Countries and collaboration

Fifty-four countries were represented. China led with 71 papers, followed by India (29), the United States (14), Australia, Canada and Spain (10 each), and the United Kingdom (8). Sixty papers involved more than one country: 31.3% of all 192 records or 32.8% of the 183 records with country data. China-USA was the strongest bilateral link (nine papers); Australia-China, Canada-China and China-UK each had five. Nine records lacked country metadata, so individual links and rankings remain sensitivity-limited.

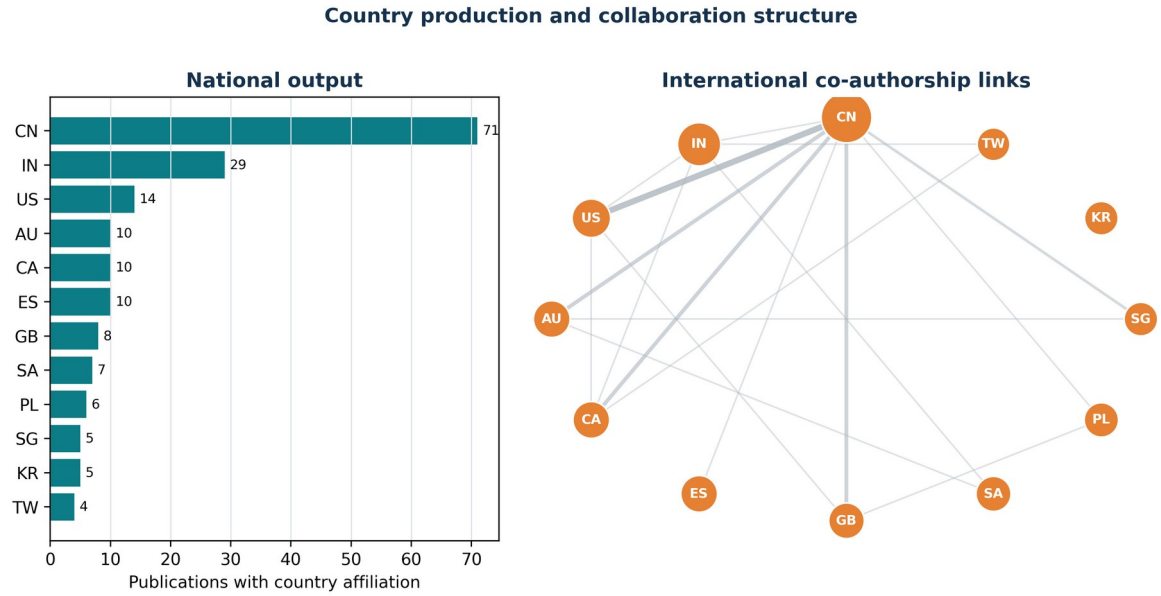

Figure 3. ***Country production and international co-authorship. Country codes follow ISO alpha-2 notation; edge width denotes co-authored document count.***

3.3. Co-word, thematic evolution and co-citation

The controlled term network (Figure 4) places predictive maintenance (73 records), fault diagnosis (58), fault detection (48) and condition monitoring (35) at the task core. Machine learning (34), deep learning (28), LSTM (23), CNN (17) and transformers (10) form the main algorithm branch. Edge computing (25), cloud computing (14), industrial CPS (12) and digital twins (11) form the architecture branch. Bearings and vibration (26 each), process systems (14), motors (13), sensor fusion (11) and machining/tool wear (9) anchor the map physically.

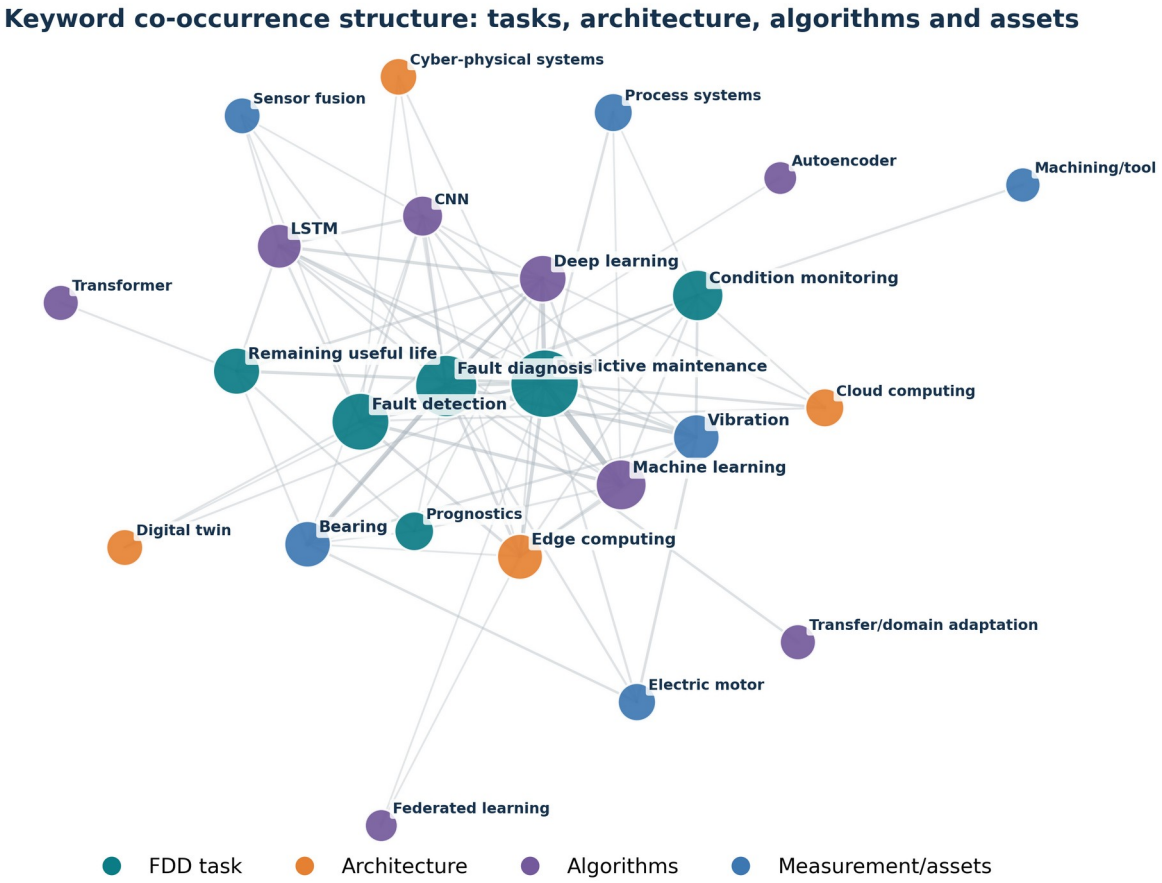

Figure 4. ***Binary record-level co-occurrence among controlled engineering terms occurring in at least six records. Node area denotes record frequency; colour denotes the prespecified analytical family.***

The thematic heat map (Figure 5) indicates that explicit edge-cloud architecture increased from 73.3% of 2015-2019 records to 86.3% of 2023-2025 records. Data-driven diagnosis/RUL increased from 20.0% to 59.8%, and trustworthy distributed-intelligence language from 20.0% to 37.6%. Explicit maintenance-decision or closed-loop control language was present in 8.5% of 2023-2025 records, compared with 16.7% in 2020-2022. These are title/abstract reporting patterns, not evidence that the corresponding methods were absent from full texts.

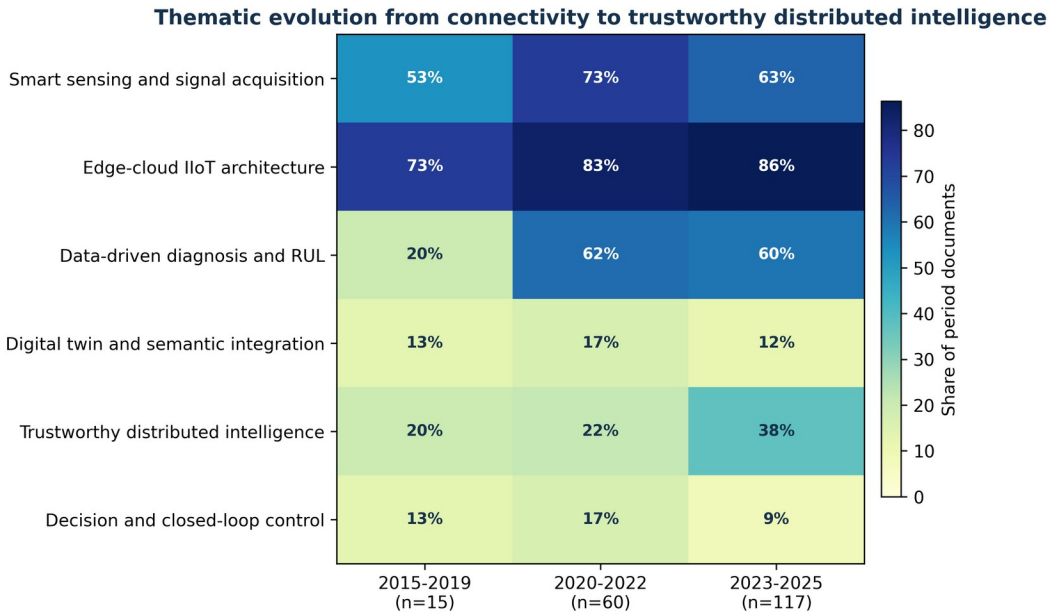

Figure 5. ***Share of records in each period containing each cross-cutting theme. Themes overlap; columns do not sum to 100%.***

The co-citation graph comprised 24 nodes, 127 non-zero links and three communities with Q = 0.267 (Figure 6). This is moderate rather than strong separation: IIoT architecture, predictive maintenance, machine health and learning methods share a connected knowledge base. Community colours should not be interpreted as sharply isolated schools. Referenced-work metadata were available for 177 of 192 records. Full node metadata, edge weights and the controlled lexicon accompany the submission.

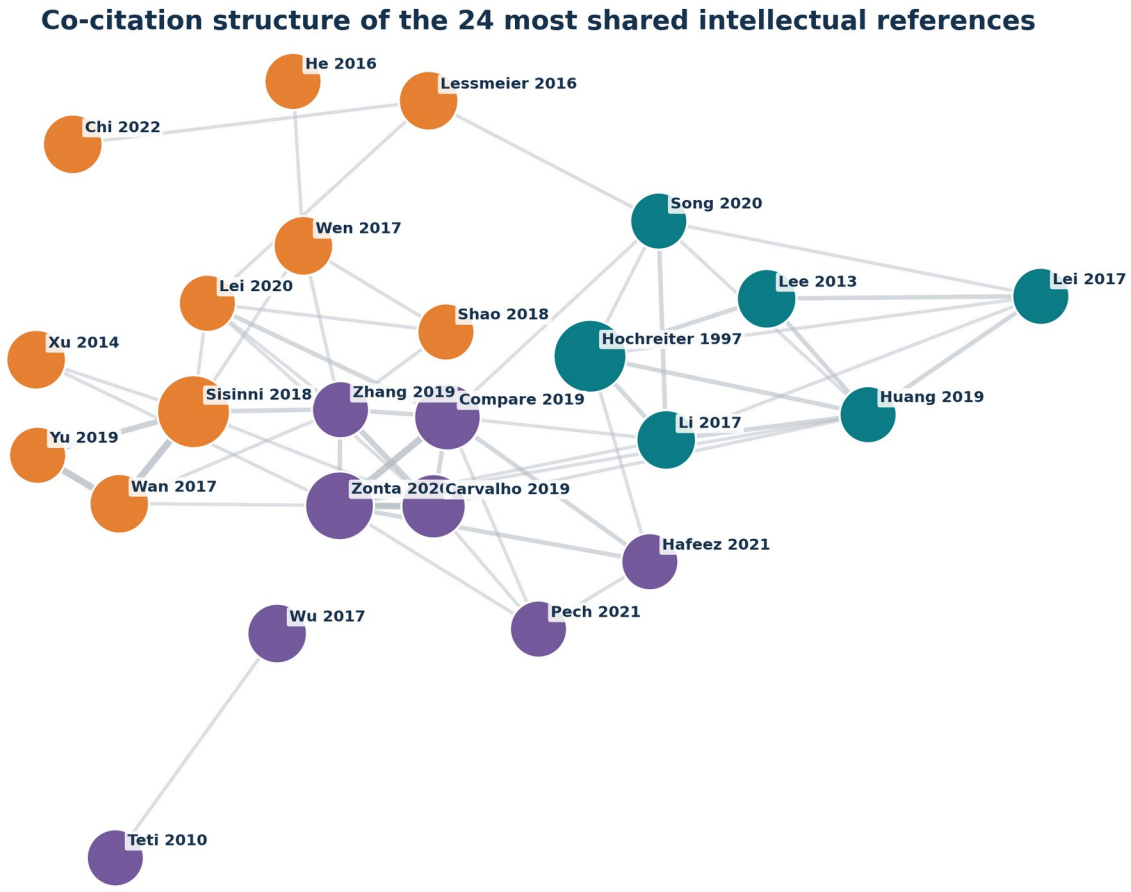

Figure 6. ***Weighted co-citation network of 24 locally shared topic-relevant references. Labels show first author and year; node area denotes local citation count; colours denote greedy-modularity communities.***

The metadata audit found explicit physical/deployment validation in 66 records (34.4%), grouped, external or domain-shift validation in 13 (6.8%), metrological traceability or uncertainty in three (1.6%), confidence calibration/OOD/abstention in three (1.6%), timing/resource/energy evidence in 70 (36.5%), and control or maintenance-action validation in 14 (7.3%). No title/abstract matched the conservative cyber/physical ambiguity pattern. Because 34 records lacked abstracts, these values diagnose reporting visibility and should not be presented as full-text prevalence.

4. Thematic and technological synthesis

4.1. Four-layer measurement-to-action stack

The literature is interpreted as four coupled layers (Figure 7). Each has a local engineering objective and a hidden failure that can invalidate downstream diagnostic accuracy. The framework separates what the bibliometric corpus explicitly reports from the minimum evidence that a deployable measurement and control system should provide.

Table 5. Technical audit of the four-layer IIoT-FDD stack.

Layer	Required evidence	Typical hidden failure	Deployment acceptance test
Measurement	Transfer function, mounting, calibration, bandwidth, ADC, clock and uncertainty	Alias, drift, cross-sensitivity, clipping or timestamp error appears as a plant fault	Traceable injection/simulation across temperature, load and mounting tolerance

Layer	Required evidence	Typical hidden failure	Deployment acceptance test
Networking	Unit, status, source timestamp, identity, information model and bounded freshness	Delay/loss, retained stale value or tag mismatch corrupts temporal features	Impairment test with PTP error, packet-delay variation, loss, reorder and failover
Firmware/edge	WCET, RTOS schedule, SRAM/Flash, watchdog, versioning and secure rollback	Average latency passes while a deadline, buffer or update fails in the worst case	Hardware-in-the-loop stress test at maximum acquisition/network load
Algorithms/action	Grouped validation, calibration, OOD policy, cost, authority and safe fallback	Leakage inflates accuracy; confident error triggers unsafe or costly action	Leave-one-site/regime-out test plus a supervisory closed-loop scenario

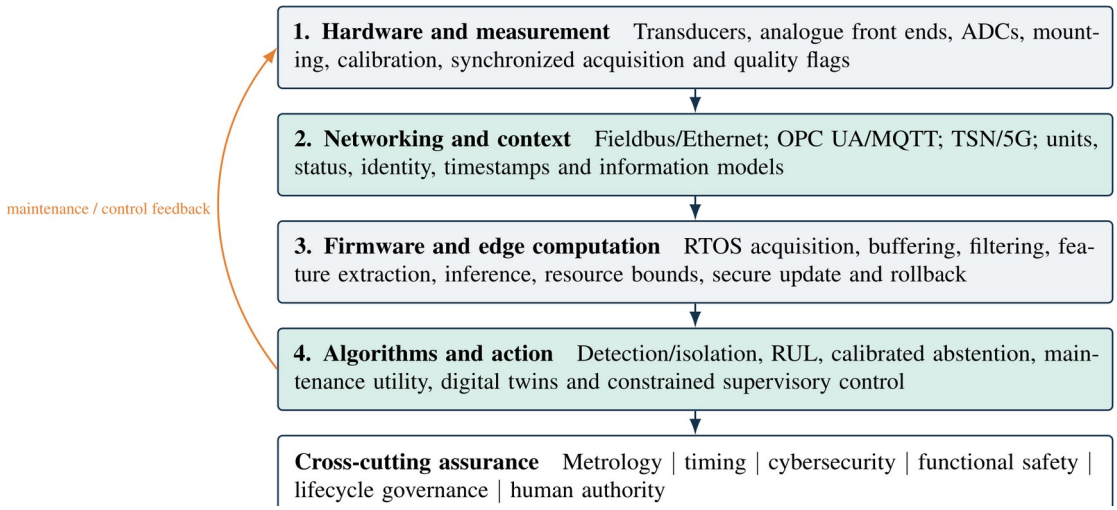

Figure 7. ***Layered IIoT-FDD architecture. Diagnosis becomes a control contribution only when uncertainty, latency, authority and fallback are bounded.***

4.2. Layer 1: metrological and sensing integrity

ISO 13374 structures condition-monitoring data processing, communication and presentation, but it does not remove the need to specify the measurement model.[35-38] For measurand y = f(x1,...,xp), the GUM combined standard uncertainty is:[39]

$$u_c^2(y) = \Sigma_i \Sigma_j (\partial f/\partial x_i)(\partial f/\partial x_j)u(x_i,x_j); \quad U = k u_c(y). \tag{3}$$

An FDD record should therefore bind each sample to calibration identity, unit, combined uncertainty, coverage factor, clock quality, sensor orientation/mounting and operating regime. Normalizing a public array does not remove uncertainty from transduction, mounting, gain, bandwidth or time. Minimum physical-layer reporting includes transducer location and attachment, analogue-front-end gain and saturation, temperature and cross-axis sensitivity, anti-alias response, ADC resolution and sampling-clock stability.

For an ideal N-bit ADC with full-scale range VFS, the quantization step is Delta = VFS/2^N and standard uncertainty is Delta/sqrt(12); the full-scale sine signal-to-quantization-noise ratio is approximately 6.02N + 1.76 dB. Sampling must include a guard band above the highest fault frequency because a real analogue filter has finite roll-off. Clock jitter sigma_t bounds sinusoidal jitter SNR as:

$$SNR_{jitter} \approx -20 \log_{10}(2\pi f_{in} \sigma_t). \tag{4}$$

Multi-sensor fusion improves observability but introduces correlated uncertainty, asynchronous sampling and sample-rate conversion. Resampled features should retain source timestamps and state anti-imaging/anti-alias filters and group delay. Drift, bias, clipping, packet loss and clock error can resemble plant or actuator faults; credible studies inject these measurement faults separately and jointly rather than placing them in one undifferentiated anomaly class.

4.3. Layer 2: networking, semantics and synchronized time

OPC UA provides typed variables, status codes, source/server timestamps, security services and Companion Specifications; its value is lost when deployment reduces it to vendor-specific string tags.[40] A portable diagnostic service should consume an information model or UANodeSet containing unit, range, asset/component relation, sampling configuration, quality and calibration reference. MQTT 5.0 offers lightweight publish/subscribe delivery with QoS 0, 1 and 2, but none guarantees a real-time deadline.[41] QoS 1 may duplicate a message, QoS 2 adds handshakes, and retained messages can be stale unless application freshness is checked.

$$\begin{aligned} &\mathrm{L\_e2e} = \mathrm{L\_acq} + \mathrm{L\_window} + \mathrm{L\_pre} + \mathrm{L\_infer} + \mathrm{L\_queue} \\ &\quad + \mathrm{L\_net} + \mathrm{L\_decide} + \mathrm{L\_act}; \quad \max(\mathrm{L\_e2e}) \le \mathrm{D\_FDD}. \end{aligned} \tag{5}$$

Average inference time covers only one term. IEEE 1588 PTP can support sub-microsecond synchronization under an engineered profile, but asymmetry, residence-time correction, grandmaster failover, oscillator holdover and packet-delay variation still require measurement.[42] TSN scheduled traffic can bound gate schedules and should be co-designed with feature-window deadlines.[43] The IMT-2020 URLLC target is a radio-interface evaluation requirement, not an end-to-end industrial control guarantee covering queueing, core networks, inference and actuation.[44]

4.4. Layer 3: edge execution, lifecycle security and communications integration

Edge FDD is a real-time embedded workload. For fixed-priority tasks, response-time analysis requires:

$$\begin{aligned} &\mathrm{R\_i} = \mathrm{C\_i} + \mathrm{B\_i} + \Sigma\_(\mathrm{j} \in \mathrm{hp(i)})\ \lceil \mathrm{R\_i}/\mathrm{T\_j} \rceil \mathrm{C\_j}, \\ &\qquad\qquad \mathrm{R\_i} \le \mathrm{D\_i}. \end{aligned} \tag{6}$$

Here C_i is a bounded worst-case execution time, B_i is blocking, T_j denotes higher-priority periods and D_i is the diagnostic deadline.[45] Measurement-based maxima are not WCET bounds unless path and interference coverage are justified. Cache, DMA, interrupt storms, multicore contention and network bursts belong in the test envelope. TinyML feasibility likewise requires explicit Flash, SRAM and energy envelopes:

$$\mathrm{M\_code} + \mathrm{M\_weights} + \mathrm{M\_constants} \le \mathrm{M\_Flash,max}. \tag{7}$$

$$\begin{aligned} &\mathrm{M\_activation,peak} + \mathrm{M\_buffers} + \mathrm{M\_stack} + \mathrm{M\_RTOS} \\ &\qquad\qquad \le \mathrm{M\_SRAM,max}. \end{aligned} \tag{8}$$

MLPerf Tiny illustrates why accuracy, latency and energy should be reported together.[46] Quantization and pruning should be evaluated for sensitivity, calibration, OOD rejection, WCET, memory peaks and energy per reliable decision, not accuracy alone.

Security is lifecycle evidence. IEC 62443 separates zones, conduits, asset-owner programmes and component requirements; NIST SP 800-193 frames protect-detect-recover firmware resilience.[47-50]

An edge node therefore needs verified boot, signed firmware and model packages, anti-rollback controls, atomic update, recovery, audit trails and a tested safe state appropriate to its risk. Network security cannot protect an unsigned model loaded below the protocol stack. Cyber and physical diagnoses should also be separable: sensor bias, spoofing, replay, denial of service, network loss and genuine degradation can produce similar residuals. A single anomaly score cannot support a safety or security case.

4.5. Layer 4: generalization, confidence and control closure

Random window-level train/test splits are invalid when windows from the same run, bearing, machine or operating regime share signatures. If g(z) identifies the asset, site, regime or degradation trajectory that generated sample z, evaluation should enforce:

$$\{g(z): z \in D\_train\} \cap \{g(z): z \in D\_test\} = \emptyset. \tag{9}$$

Hyperparameters should be selected by nested grouped validation. At minimum, studies should report leave-one-rig/machine-out and leave-one-regime-out results; multi-site claims require leave-one-site-out or prospective external-site evaluation.

Confidence must be validated separately from discrimination. Expected calibration error should be accompanied by reliability diagrams and a proper score because it depends on binning.[51] Split conformal prediction can provide marginal coverage under exchangeability, but the guarantee does not automatically survive regime shift.[52] OOD scores, drift monitors and a predeclared abstention threshold should therefore trigger a safe fallback, request another measurement or defer to an operator.[53] Report coverage, set size and selective risk under both nominal and shifted regimes.

Diagnosis ultimately closes a decision loop. A risk-weighted threshold should minimize:

$$J(\theta) = c\_FN\ P\_\theta(FN) + c\_FP\ P\_\theta(FP) + c\_d\ E\_\theta[t\_d] + c\_m\ E\_\theta[C\_maintenance]. \tag{10}$$

A diagnostic output may change a supervisory mode, constrain a set point or request shutdown only if confidence, freshness and latency gates pass and the post-switch state remains inside a verified safe set. Protection and interlock logic should remain on validated deterministic controllers unless the diagnostic component has a commensurate safety case.

4.6. From PLC/RTU hierarchies to smart instruments and edge controllers

The transition is evolutionary. Traditional PLC/RTU systems retain deterministic scans, modest attack surfaces, long support cycles and clear maintenance ownership. Their limitations are waveform bandwidth, vendor-specific context, weak fleet learning and delayed analysis. A smart transmitter can check range and drift, extract spectral features and publish quality metadata; an edge controller can fuse vibration, current, acoustic and temperature channels while disconnected; and a cloud layer can train across fleets and manage twins. Representative demonstrations include edge data reduction, edge-to-cloud sensing, distributed FDD, temporal-logic diagnosis, federated domain generalization and distributed twins.[54-59]

The architectural mistake is to replace validated protection with an opaque network dependency. A staged authority ladder is safer: offline shadow mode; advisory alarm; maintenance recommendation; supervisory constraint; and direct control only when justified. Each stage should

exercise rollback, clock/network loss, OOD input and contradictory sensor evidence. Hybrid FDD - physical residuals plus learned features, semantic component models and calibrated uncertainty - offers stronger failure detectability than either physics-free classifiers or brittle complete models.

5. Research gaps and testable trajectories

Table 6 converts the synthesis into six testable research trajectories. The first four address metrology, embedded/network execution and algorithmic validity; the final two connect security and control consequence.

Table 6. Research gaps, technical remedies and minimum validation protocols.

Gap	Research upgrade	Minimum acceptance evidence
1. Leakage and domain shift	Physics-guided self-supervision, domain generalization and explicit OOD abstention	Nested leave-one-machine/site/regime-out; calibration and risk-coverage under shift
2. Missing traceability	GUM-compatible uncertainty graph from sensor and clock through feature and decision	Calibration/mounting/temperature perturbation; uncertainty budget and detection limit
3. Unbounded edge timing	Joint acquisition-RTOS-model-network design with WCET and memory envelopes	Worst-case HIL load; deadline-miss rate, WCET argument, SRAM/Flash and energy/decision
4. Weak interoperability	ISO 13374 pipeline plus OPC UA information model and ISO 23247 twin	Transfer between two vendor stacks without manual unit/tag remapping
5. Fault/security ambiguity	Multi-hypothesis process/sensor/network/model/attack diagnosis and signed lifecycle	Coordinated physical fault, spoofing, delay, loss and malicious-update scenarios
6. Open-loop benchmarks	Cost-sensitive diagnosis linked to invariant degraded envelopes and human authority	Detection delay, false alarms/hour, constraint metrics, downtime and maintenance utility

Generalization studies should make asset, run and site identifiers mandatory grouping variables. Preprocessing must be refitted within every training fold, tuning nested by machine/site, and evaluation repeated after sensor replacement, load change and maintenance. Alongside AUROC or F1, report false alarms per operating hour, time to detection, proper calibration scores, conformal coverage/set size and selective risk.

A measurement digital thread should carry measurand, unit, calibration identifier/date, transfer function, bandwidth, mounting, uncertainty, clock state, preprocessing/model version and quality flags. Uncertainty may be propagated analytically for simple residuals or by Monte Carlo for nonlinear pipelines. The actionable comparison is whether better sensing reduces decision risk more than a larger model.

Deterministic edge/network co-design should derive the diagnostic deadline from fault dynamics and safe intervention time, then allocate budgets across Equation 5. RTOS response-time analysis, TSN gates and PTP quality should be tested together under maximum sampling, logging, update and network load. Link loss must exercise local autonomy, backlog discard, stale-message rejection and recovery without duplicate action.

Interoperable twins should deploy the same service on at least two vendor stacks and demonstrate transfer of units, timestamps, status and asset relations without hand-coded remapping. ISO 23247

supplies a manufacturing-twin framework.[60] A failure-seeking twin should generate boundary regimes, sensor/clock faults, cyber manipulations and counterfactual maintenance actions, while reporting its own measurement fidelity and uncertainty.

Joint safety-security studies need synchronized scenarios across physical, network and firmware layers: gradual drift, abrupt sensor bias, replay, timestamp spoofing, delay/loss, actuator degradation, model poisoning and failed signed update. Evaluation should report confusion between fault and attack classes, not attack detection alone. Finally, closed-loop work should optimize thresholds against plant-specific false-negative, false-positive, delay, inspection and downtime costs, and report constraint violations, stability/robustness margins, production loss and nuisance interventions under delayed or wrong diagnoses.

6. Limitations

Eight limitations bound the interpretation. First, a frozen OpenAlex retrieval and official Scopus source/year audit cannot replace document-level Scopus/WoS export overlap. Second, 34 records lacked abstracts, nine lacked country data, five lacked complete author identifiers and 15 lacked referenced works; affected results state denominators. Third, screening was deterministic and author-supervised but not independently duplicated by two human reviewers. Fourth, the protocol was not prospectively registered. Fifth, title centrality can miss abstract-central FDD. Sixth, English, journal, DOI and source-coverage filters underrepresent regional, conference and grey literature. Seventh, controlled lexicons and citation networks measure attention and reporting, not efficacy, safety or maturity. Eighth, the v3.0.0 correction changed the former 183-record boundary to 192; all 13 additions and four removals are disclosed and headline directions remained stable. The corpus, code and audits accompany the paper.

7. Conclusions

IIoT-enabled FDD is rapidly expanding, geographically distributed and author-fragmented. Its knowledge base connects predictive maintenance, fault diagnosis, edge/cloud computing, bearings/vibration and deep temporal models. Quantitative laws require restraint: Bradford concentration is visible but its geometric sequence is not; the Lotka fit is statistically plausible but supported over only four observed productivity values; and co-citation separation is moderate. The stronger engineering observation is that distributed inference is more visible in metadata than traceable uncertainty, deterministic execution and validated action, although title/abstract evidence cannot establish full-text absence.

Future work should be judged across the complete chain: what was measured and with what uncertainty; whether mounting, sampling and time were valid; how semantics and freshness survived the network; whether edge resources and deadlines were bounded; whether confidence failed safely under new machines and regimes; and how diagnosis changed maintenance or control without violating constraints. That evidence turns a connected classifier into an industrial measurement and control system.

Acknowledgements

[AUTHOR ACTION: Insert verified acknowledgements, including any non-author writing or editing assistance, or state 'Not applicable'.]

Author contributions (CRediT)

[AUTHOR ACTION: Bui Thanh Lam must confirm the final CRediT roles before submission.]

Statements and Declarations

Ethical considerations

Ethical approval was not required because this study analysed public bibliographic metadata and did not involve human participants, identifiable personal data, animals or biological material.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Declaration of conflicting interest

[AUTHOR ACTION: If accurate for Bui Thanh Lam, replace this text with: 'The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article'.]

Funding statement

The author received no financial support for the research, authorship, and/or publication of this article.

Data availability

The accompanying reproducibility release contains the frozen broad and analytic corpora, record-level screening and provenance decisions, DOI and source-coverage audits, version-sensitivity audit, exclusion logs, data dictionary, analysis code, numerical outputs, figure source files and VOSviewer-compatible networks. **[AUTHOR ACTION: Deposit this verified release in Zenodo/OSF or another durable repository and insert the versioned DOI here before submission.]** The release includes an MIT code licence, CC0 metadata boundary and SHA-256 checksums, consistent with FAIR principles.[61]

Generative AI disclosure

OpenAI Codex in ChatGPT Work was used for language revision, document restructuring, reference-style conversion and figure/layout preparation. It was not treated as an author or evidence source. The named human author is responsible for verifying every query, record

decision, calculation, citation, technical statement and declaration before submission, consistent with Sage and COPE policies.[62,63]

References

1. Xu LD, He W and Li S. Internet of Things in industries: a survey. IEEE Trans Ind Inform 2014; 10: 2233-2243. DOI: 10.1109/TII.2014.2300753.

2. Lee J, Bagheri B and Kao HA. A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manuf Lett 2015; 3: 18-23. DOI: 10.1016/j.mfglet.2014.12.001.

3. Monostori L, Kadar B, Bauernhansl T, et al. Cyber-physical systems in manufacturing. CIRP Ann 2016; 65: 621-641. DOI: 10.1016/j.cirp.2016.06.005.

4. Boyes H, Hallaq B, Cunningham J, et al. The industrial Internet of Things (IIoT): an analysis framework. Comput Ind 2018; 101: 1-12. DOI: 10.1016/j.compind.2018.04.015.

5. Sisinni E, Saifullah A, Han S, et al. Industrial Internet of Things: challenges, opportunities, and directions. IEEE Trans Ind Inform 2018; 14: 4724-4734. DOI: 10.1109/TII.2018.2852491.

6. Lu Y. Industry 4.0: a survey on technologies, applications and open research issues. J Ind Inf Integr 2017; 6: 1-10. DOI: 10.1016/j.jii.2017.04.005.

7. Venkatasubramanian V, Rengaswamy R, Yin K, et al. A review of process fault detection and diagnosis: Part I: quantitative model-based methods. Comput Chem Eng 2003; 27: 293-311. DOI: 10.1016/S0098-1354(02)00160-6.

8. Venkatasubramanian V, Rengaswamy R and Kavuri SN. A review of process fault detection and diagnosis: Part II: qualitative models and search strategies. Comput Chem Eng 2003; 27: 313-326. DOI: 10.1016/S0098-1354(02)00161-8.

9. Venkatasubramanian V, Rengaswamy R, Kavuri SN, et al. A review of process fault detection and diagnosis: Part III: process history based methods. Comput Chem Eng 2003; 27: 327-346. DOI: 10.1016/S0098-1354(02)00162-X.

10. Jardine AKS, Lin D and Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech Syst Signal Process 2006; 20: 1483-1510. DOI: 10.1016/j.ymssp.2005.09.012.

11. Lei Y, Li N, Guo L, et al. Machinery health prognostics: a systematic review from data acquisition to RUL prediction. Mech Syst Signal Process 2018; 104: 799-834. DOI: 10.1016/j.ymssp.2017.11.016.

12. Qin SJ. Survey on data-driven industrial process monitoring and diagnosis. Annu Rev Control 2012; 36: 220-234. DOI: 10.1016/j.arcontrol.2012.09.004.

13. Yin S, Ding SX, Xie X, et al. A review on basic data-driven approaches for industrial process monitoring. IEEE Trans Ind Electron 2014; 61: 6418-6428. DOI: 10.1109/TIE.2014.2301773.

14. Zhao R, Yan R, Chen Z, et al. Deep learning and its applications to machine health monitoring. Mech Syst Signal Process 2019; 115: 213-237. DOI: 10.1016/j.ymssp.2018.05.050.

15. Khan S and Yairi T. A review on the application of deep learning in system health management. Mech Syst Signal Process 2018; 107: 241-265. DOI: 10.1016/j.ymssp.2017.11.024.

16. Lei Y, Yang B, Jiang X, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap. Mech Syst Signal Process 2020; 138: 106587. DOI: 10.1016/j.ymssp.2019.106587.

17. Carvalho TP, Soares FAA, Vita R, et al. A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 2019; 137: 106024. DOI: 10.1016/j.cie.2019.106024.

18. Zonta T, da Costa CA, da Rosa Righi R, et al. Predictive maintenance in the Industry 4.0: a systematic literature review. Comput Ind Eng 2020; 150: 106889. DOI: 10.1016/j.cie.2020.106889.

19. Zhang W, Yang D and Wang H. Data-driven methods for predictive maintenance of industrial equipment: a survey. IEEE Syst J 2019; 13: 2213-2227. DOI: 10.1109/JSYST.2019.2905565.

20. Compare M, Baraldi P and Zio E. Challenges to IoT-enabled predictive maintenance for Industry 4.0. IEEE Internet Things J 2020; 7: 4585-4597. DOI: 10.1109/JIOT.2019.2957029.

21. Angelopoulos A, Michailidis ET, Nomikos N, et al. Tackling faults in the Industry 4.0 era: a survey of machine-learning solutions and key aspects. Sensors 2020; 20: 109. DOI: 10.3390/s20010109.

22. Chi Y, Dong Y, Wang ZJ, et al. Knowledge-based fault diagnosis in industrial Internet of Things: a survey. IEEE Internet Things J 2022; 9: 12886-12900. DOI: 10.1109/JIOT.2022.3163606.

23. Dowdeswell B, Sinha R and MacDonell SG. Finding faults: a scoping study of fault diagnostics for industrial cyber-physical systems. J Syst Softw 2020; 168: 110638. DOI: 10.1016/j.jss.2020.110638.

24. Keleko AT, Kamsu-Foguem B, Ngouna RH, et al. Artificial intelligence and real-time predictive maintenance in Industry 4.0: a bibliometric analysis. AI Ethics 2022; 2: 553-577. DOI: 10.1007/s43681-021-00132-6.

25. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71. DOI: 10.1136/bmj.n71.

26. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018; 169: 467-473. DOI: 10.7326/M18-0850.

27. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021; 10: 39. DOI: 10.1186/s13643-020-01542-z.

28. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 358: j4008. DOI: 10.1136/bmj.j4008.

29. Priem J, Piwowar H and Orr R. OpenAlex: a fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv. Epub ahead of print 17 May 2022. DOI: 10.48550/arXiv.2205.01833.

30. Culbert JH, Hobert A, Jahn N, et al. Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. Scientometrics 2025; 130: 2475-2492. DOI: 10.1007/s11192-025-05293-3.

31. Bradford SC. Sources of information on specific subjects. Engineering 1934; 137: 85-86.

32. Lotka AJ. The frequency distribution of scientific productivity. J Wash Acad Sci 1926; 16: 317-323.

33. Aria M and Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr 2017; 11: 959-975. DOI: 10.1016/j.joi.2017.08.007.

34. Van Eck NJ and Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010; 84: 523-538. DOI: 10.1007/s11192-009-0146-3.

35. International Organization for Standardization. ISO 13374-1:2003: Condition monitoring and diagnostics of machines - data processing, communication and presentation - Part 1: General guidelines. Geneva: ISO, 2003.

36. International Organization for Standardization. ISO 13374-2:2007: Condition monitoring and diagnostics of machines - data processing, communication and presentation - Part 2: Data processing. Geneva: ISO, 2007.

37. International Organization for Standardization. ISO 13374-3:2012: Condition monitoring and diagnostics of machines - data processing, communication and presentation - Part 3: Communication. Geneva: ISO, 2012.

38. International Organization for Standardization. ISO 13374-4:2015: Condition monitoring and diagnostics of machines - data processing, communication and presentation - Part 4: Presentation. Geneva: ISO, 2015.

39. Joint Committee for Guides in Metrology. JCGM 100:2008: Evaluation of measurement data - guide to the expression of uncertainty in measurement. Sevres: BIPM, 2008.

40. OPC Foundation. OPC Unified Architecture Specification, Part 1: Overview and Concepts. Release 1.05. Scottsdale, AZ: OPC Foundation, 2023.

41. Banks A, Briggs E, Borgendale K, et al. MQTT Version 5.0. Burlington, MA: OASIS, 2019.

42. Institute of Electrical and Electronics Engineers. IEEE Std 1588-2019: Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems. New York: IEEE, 2020.

43. Institute of Electrical and Electronics Engineers. IEEE Std 802.1Qbv-2015: Enhancements for Scheduled Traffic. New York: IEEE, 2016.

44. International Telecommunication Union. ITU-R M.2410-0: Minimum requirements related to technical performance for IMT-2020 radio interface(s). Geneva: ITU, 2017.

45. Wilhelm R, Engblom J, Ermedahl A, et al. The worst-case execution-time problem - overview of methods and survey of tools. ACM Trans Embed Comput Syst 2008; 7: 1-53. DOI: 10.1145/1347375.1347389.

46. Banbury C, Reddi VJ, Torelli P, et al. MLPerf Tiny Benchmark. In: Proceedings of the NeurIPS Track on Datasets and Benchmarks, 2021.

47. International Electrotechnical Commission. IEC TS 62443-1-1:2009: Industrial communication networks - network and system security - Part 1-1: Terminology, concepts and models. Geneva: IEC, 2009.

48. International Electrotechnical Commission. IEC 62443-2-1:2024: Security for industrial automation and control systems - security program requirements for IACS asset owners. Geneva: IEC, 2024.

49. International Electrotechnical Commission. IEC 62443-4-2:2019: Security for industrial automation and control systems - technical security requirements for IACS components. Geneva: IEC, 2019.

50. Regenscheid A. NIST SP 800-193: Platform firmware resiliency guidelines. Gaithersburg, MD: National Institute of Standards and Technology, 2018.

51. Guo C, Pleiss G, Sun Y, et al. On calibration of modern neural networks. In: Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6-11 August 2017, pp.1321-1330. PMLR.

52. Angelopoulos AN and Bates S. Conformal prediction: a gentle introduction. Found Trends Mach Learn 2023; 16: 494-591. DOI: 10.1561/2200000101.

53. Hendrycks D and Gimpel K. A baseline for detecting misclassified and out-of-distribution examples in neural networks. In: International Conference on Learning Representations, Toulon, France, 24-26 April 2017.

54. Wang X, Lu S, Huang W, et al. Efficient data reduction at the edge of Industrial Internet of Things for PMSM bearing fault diagnosis. IEEE Trans Instrum Meas 2021; 70: 3508612. DOI: 10.1109/TIM.2021.3051668.

55. Li Z, Fei F and Zhang G. Edge-to-cloud IIoT for condition monitoring in manufacturing systems with ubiquitous smart sensors. Sensors 2022; 22: 5901. DOI: 10.3390/s22155901.

56. Marino R, Wisultschew C, Otero A, et al. A machine-learning-based distributed system for fault diagnosis with scalable detection quality in industrial IoT. IEEE Internet Things J 2021; 8: 4339-4352. DOI: 10.1109/JIOT.2020.3026211.

57. Chen G, Liu M and Kong Z. Temporal-logic-based semantic fault diagnosis with time-series data from industrial Internet of Things. IEEE Trans Ind Electron 2021; 68: 4393-4403. DOI: 10.1109/TIE.2020.2984976.

58. Zhao C and Shen W. Federated domain generalization: a secure and robust framework for intelligent fault diagnosis. IEEE Trans Ind Inform 2024; 20: 2662-2670. DOI: 10.1109/TII.2023.3296894.

59. Abdullahi I, Longo S and Samie M. Towards a distributed digital twin framework for predictive maintenance in Industrial Internet of Things. Sensors 2024; 24: 2663. DOI: 10.3390/s24082663.

60. International Organization for Standardization. ISO 23247-1:2021: Automation systems and integration - digital twin framework for manufacturing - Part 1: Overview and general principles. Geneva: ISO, 2021.

61. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016; 3: 160018. DOI: 10.1038/sdata.2016.18.

62. Sage Publishing. Artificial intelligence policy, https://www.sagepub.com/journals/publication-ethics-policies/artificial-intelligence-policy (2026, accessed 24 August 2026).

63. Committee on Publication Ethics. Authorship and AI tools: COPE position statement, https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools (2023, accessed 24 August 2026).