ENGENHARIAS, TECNOLOGIA, GESTÃO E TURISMO
ENGINEERING, TECHNOLOGY, MANAGEMENT AND TOURISM
INGENIERÍA, TECNOLOGÍA, ADMINISTRACIÓN Y TURISMO

millenium

Millenium, 2(Edição Especial Nº24)

pt

ARQUITETURA SENSOR-AGNÓSTICA PARA MONITORIZAÇÃO INDUSTRIAL: ESQUEMA CANÓNICO, DESCOBERTA AUTOMÁTICA E EXTENSIBILIDADE

SENSOR-AGNOSTIC ARCHITECTURE FOR INDUSTRIAL MONITORING: CANONICAL SCHEMA, AUTOMATIC DISCOVERY AND EXTENSIBILITY

ARQUITECTURA AGNÓSTICA AL SENSOR PARA LA MONITORIZACIÓN INDUSTRIAL: ESQUEMA CANÓNICO, DESCUBRIMIENTO AUTOMÁTICO Y EXTENSIBILIDAD

Nuno Santos[1]
Ricardo Correia[1] *https://orcid.org/0000-0003-3767-4538*

[1] Universidade Politécnica de Viseu, Viseu, Portugal

Nuno Santos – nunogoncalobsantos@gmail.com | Ricardo Correia – ricardocorreia@estgv.ipv.pt

CC BY

Autor Correspondente:
Nuno Santos
Rua Aquilino Ribeiro
3500 - 228 - Viseu - Portugal
nunogoncalobsantos@gmail.com

RECEBIDO: 16 de julho de 2026
REVISTO: 23 de julho de 2026
ACEITE: 10 de setembro 2026
PUBLICADO: 17 de novembro de 2026

DOI: https://doi.org/10.29352/mill0224e.48333

RESUMO

Introdução: A integração de sensores heterogéneos em plataformas IIoT confronta a engenharia com o desafio da normalização semântica, dado que diferentes fabricantes implicam nomenclaturas, unidades e protocolos distintos.
Objetivo: Apresentar e validar quantitativamente uma arquitetura sensor-agnóstica para resolução da heterogeneidade semântica em plataformas IIoT industriais.
Métodos: A arquitetura assenta em três mecanismos: (1) um esquema canónico InfluxDB com nomenclatura de campos estrita em seis domínios industriais; (2) uma biblioteca partilhada de *parsers* com interface uniforme, comum ao pipeline de produção e ao simulador de telemetria; e (3) um registo de sensores JSON para descoberta automática de dispositivos. A validação realizou-se numa unidade industrial com três sensores heterogéneos (Modbus RTU e I2C), complementada por simulação com 16 tipos de sensores em três instalações virtuais.
Resultados: A integração de novos tipos de sensores requer apenas 36–51 linhas de código, sem alterações nos serviços existentes. Registou-se o processamento de 216 mensagens/minuto, latência média de 107 ms por ativo e escalabilidade linear até 300 ativos simultâneos. A validação incluiu 313 testes automatizados com 93% de cobertura, deteção de anomalias por Z-Score e Média Móvel, e correlação vibração/energia por tipo de material em *deployment* CNC.
Conclusão: A convergência entre esquema canónico, biblioteca partilhada e registo de sensores constitui uma resposta eficaz e verificável à heterogeneidade semântica, garantindo extensibilidade com impacto nulo no código base dos serviços existentes.

Palavras-chave: IIoT; esquema canónico; biblioteca partilhada; descoberta automática; MQTT

ABSTRACT

Introduction: The integration of heterogeneous sensors into IIoT platforms confronts engineers with a semantic normalization challenge, as data from different manufacturers are exposed with distinct nomenclatures, units, and protocols.
Objective: Present and quantitatively validate a sensor-agnostic architecture for addressing semantic heterogeneity in industrial IIoT platforms..
Methods: The architecture comprises: (1) a canonical InfluxDB schema with strict nomenclature across six industrial domains; (2) a shared parser library with a uniform interface, used by both the production pipeline and the telemetry simulator; and (3) a JSON sensor registry for automatic device discovery. Validation was conducted at a real industrial manufacturing unit with three heterogeneous sensors (Modbus RTU and I2C), supplemented by a simulation covering 16 sensor types across three virtual sites.
Results: Integration of new sensor types requires 36–51 lines of code with zero changes to existing services. Data processing reached 216 messages/minute, with average latency of 107 ms per asset and linear scalability up to 300 simultaneous assets. The suite of 313 automated tests achieved 93% instruction coverage. The platform includes statistical anomaly detection (Z-Score and Moving Average) and vibration/energy correlation by material type, operational in an industrial CNC deployment.
Conclusion: The combination of a canonical schema, shared library, and sensor registry constitutes an effective and verifiable solution to the semantic heterogeneity problem in industrial IIoT platforms, ensuring extensibility with zero impact on existing services.

Keywords: IIoT; canonical schema; shared library; automatic discovery; MQTT

RESUMEN

Introducción: La integración de sensores heterogéneos en plataformas IIoT plantea un desafío de normalización semántica por la diversidad de nomenclaturas, unidades y protocolos de cada fabricante.
Objetivo: Presentar y validar cuantitativamente una arquitectura independiente del tipo de sensor para resolver la heterogeneidad semántica en plataformas IIoT industriales.
Métodos: La arquitectura comprende: (1) un esquema canónico InfluxDB con nomenclatura estricta en seis domínios industriales; (2) una biblioteca de parsers compartida con interfaz uniforme, común al pipeline de producción y al simulador de telemetría; y (3) un registro JSON de sensores para descubrimiento automático de dispositivos. La validación se realizó en una unidad industrial de manufactura con tres sensores heterogéneos (Modbus RTU e I2C), complementada con simulación de 16 tipos de sensores en tres instalaciones virtuales.
Resultados: La integración de nuevos tipos de sensores requiere 36–51 líneas de código sin modificaciones en los servicios existentes. Se procesaron 216 mensajes/minuto con latência media de 107 ms por activo y escalabilidad lineal hasta 300 activos simultáneos. El sistema superó 313 pruebas automatizadas con 93% de cobertura de instrucciones. La plataforma incluye detección de anomalías estadísticas (Z-Score y Media Móvil) y correlación vibración/energía por tipo de material en despliegue CNC industrial.
Conclusión: La convergencia entre el esquema canónico, la biblioteca compartida y el registro de sensores constituye una respuesta eficaz y verificable a la heterogeneidad semántica, garantizando extensibilidad con impacto nulo en los servicios existentes.

Palabras Clave: IIoT; esquema canónico; biblioteca compartida; descubrimiento automático; MQTT.

DOI: https://doi.org/10.29352/mill0224e.48333

INTRODUÇÃO

A transição para a Indústria 4.0 está a reconfigurar o paradigma de produção industrial, impulsionada pela proliferação de dispositivos conectados no chão de fábrica. Em (Xu et al., 2023), documenta-se esta transformação nas indústrias de manufatura, energia e transportes, identificando-se a heterogeneidade dos dispositivos como um dos principais desafios técnicos para plataformas de dados industriais unificadas. Esta realidade é particularmente crítica nas Pequenas e Médias Empresas (PME): num inquérito a 271 PMEs do setor industrial no Reino Unido, (Masood & Sonntag, 2020) identificam constrangimentos financeiros e de conhecimento técnico como principais barreiras à adoção da Indústria 4.0, com a integração de sistemas *legados* entre os desafios mais reportados — barreiras que (Müller et al., 2024) confirmam permanecer prevalentes em PMEs industriais, nas dimensões tecnológica, organizacional e de integração em cadeia de valor.

Na prática, cada fabricante de sensores expõe os seus dados com nomenclaturas proprietárias — por exemplo, "v1" ou "voltage_L1" para a mesma grandeza física —, resultando em pipelines frágeis e modificações em múltiplos serviços por cada nova integração. Em (Rahman & Hussain, 2020), enquadra-se este fenómeno como heterogeneidade semântica, identificando-se a ausência de normalização como principal barreira à interoperabilidade multi-fabricante — problema ativo na investigação recente, onde (You et al., 2024) propõem mecanismos de identificação semântica para dispositivos IoT heterogéneos. Sistemas de monitorização energética baseados em computação *edge* demonstram a viabilidade de stacks *open-source* com MQTT e InfluxDB (Mirani et al., 2024), mas constroem pipelines de sensores específicos sem abstração reutilizável. De igual modo, investigações com ESP32 como *gateway* IoT e sensores Modbus confirmam a viabilidade técnica (He et al., 2024; Boonmeeruk et al., 2024; Eldho K J, 2025), contudo, não resolvem a heterogeneidade semântica multi-fabricante. Plataformas genéricas (ThingsBoard, Node-RED) oferecem extensibilidade via plugins (Monios et al., 2024) mas não garantem consistência de dados entre serviços.

Neste contexto, o presente artigo propõe e valida uma plataforma de monitorização IIoT sensor-agnóstica, capaz de integrar dispositivos heterogéneos de forma automática, sem necessidade de reengenharia dos serviços existentes a cada nova integração. A plataforma suporta a aquisição, normalização e análise de dados em tempo real provenientes de múltiplos tipos de sensores industriais, independentemente do fabricante ou protocolo de comunicação utilizado. A solução é validada quantitativamente num cenário industrial real, com sensores de três domínios distintos (energia, vibração e ambiente) e protocolos heterogéneos (Modbus RTU e I2C).

As contribuições específicas são:

-(C1) Definição de um esquema canónico InfluxDB: estabelece-se uma nomenclatura de campos estrita organizada em seis domínios industriais (energy, vibration, environment, proximity, flow e cnc), constituindo o contrato de dados único da plataforma.

-(C2) Implementação de uma biblioteca partilhada de *parsers*: utiliza-se uma interface uniforme (parse/getAlerts/THRESHOLDS) comum ao pipeline de produção e ao simulador de telemetria, garantindo que a normalização é testada com dados sintéticos antes da implementação física.

-(C3) Desenvolvimento de um sensor *registry* JSON: introduz-se um mecanismo de descoberta automática que elimina a configuração manual na *dashboard* e no serviço de análise, reduzindo a integração de novos sensores à criação de um único ficheiro *parser*.

-(C4) Validação quantitativa da extensibilidade: demonstra-se a capacidade de integração com zero linhas de código alteradas nos serviços existentes, apresentando evidências de LOC, *throughput*, latência e cobertura de testes.

As contribuições C1–C3 constituem artefactos de engenharia — um esquema de dados, uma biblioteca e um mecanismo de registo — cuja novidade reside na combinação e não nos mecanismos individuais. A contribuição científica do trabalho (C4) reside na validação quantitativa reprodutível desta combinação: a demonstração, com métricas verificáveis de LOC, *throughput*, latência e cobertura de testes, de que a convergência entre esquema canónico, parsers partilhados e discovery automático elimina a necessidade de reengenharia por sensor.

1. ENQUADRAMENTO TEÓRICO

Na literatura recente, diversas abordagens para a monitorização industrial apresentam limitações na abstração de dados: sistemas de monitorização energética por edge computing com MQTT e InfluxDB (Mirani et al., 2024) e de monitorização em edifícios (Verde Romero et al., 2024) constroem pipelines específicos ao sensor, sem *parsers* partilhados nem descoberta automática; sistemas SCADA open-source com ESP32 e Node-RED (He et al., 2024) e gateways Modbus-MQTT baseados em ESP32 (Boonmeeruk et al., 2024) confirmam a viabilidade de hardware de baixo custo sem resolver a heterogeneidade semântica. Quanto ao protocolo, (Eldho K J, 2025) confirma que o MQTT dominava mais de 85% das comunicações IoT em 2023.

O hardware ESP32 é validado para o desenvolvimento de sistemas IoT de baixo custo em (Hercog et al., 2023). Complementarmente, em (Monios et al., 2024) comparam-se quatro plataformas IoT *open-source* (ThingsBoard, Node-RED, Grafana+InfluxDB e OpenRemote), confirmando que nenhuma integra nativamente, sem esforço adicional, os três mecanismos em simultâneo: suporte Modbus, *parsers* extensíveis por tipo de sensor e descoberta automática (***Tabela 1***). A literatura confirma que as iniciativas de *retrofitting* de dispositivos legados para IIoT permanecem fragmentadas, focando-se em soluções técnicas isoladas sem *workflows* padronizados ou critérios de avaliação reproduzíveis (Boareto et al., 2026). A normalização semântica aqui proposta situa-se num espaço mais amplo de esforços de interoperabilidade industrial. Em (Busboom, 2024), documenta-se como o protocolo OPC-UA endereça este problema ao nível do modelo de informação; (Schmidt et al., 2023) demonstram a interoperabilidade da *Asset Administration Shell* (AAS) com outras especificações de *Digital Twin*; e (Hazra et al., 2023) sistematizam o papel das ontologias industriais na estruturação do conhecimento de domínio para além do simples mapeamento de campos. A arquitetura proposta adota uma estratégia deliberadamente mais leve do que estas abordagens, mais adequada a implementações de baixo custo, mas sacrificando a expressividade semântica e a interoperabilidade entre organizações que aquelas normas garantem.

DOI: https://doi.org/10.29352/mill0224e.48333

Tabela 1 - Comparação de Plataformas IIoT Open-Source

Plataforma	Modbus Nativo	Parser Extensível	Discovery Automático	Análise Integrada
ThingsBoard	Não[1]	Parcial[2]	Não	Média
Node-RED	Parcial[3]	Sim	Não	Não
Grafana + InfluxDB	Parcial[4]	Parcial	Não	Via plugins
OpenRemote	Sim	Não	Não	Média
Esta proposta	Sim	Sim	Sim	Sim

Note: [1] ThingsBoard requer o addon "IoT Gateway" para Modbus, sem interface uniforme por tipo de sensor; [2] e [3] Node-RED e Grafana+InfluxDB suportam Modbus apenas via plugins comunitários (node-red-contrib-modbus; Telegraf inputs.modbus), não integrados no core das plataformas.

2. MÉTODOS

Esta secção descreve a arquitetura da plataforma proposta e os mecanismos de implementação que suportam a extensibilidade sensor-agnóstica.

2.1 Arquitetura da Solução

A plataforma articula-se numa arquitetura em três camadas (Figura 1)*: Edge Tier*, *Platform Tier* e *Business Tier*. Adotou-se a separação de responsabilidades como o princípio central do projeto: o *edge* executa exclusivamente a aquisição e publicação de dados brutos, enquanto a plataforma assegura a normalização e o armazenamento, cabendo à camada de negócio a apresentação e atuação. Todos os serviços das camadas *Platform* e *Business* foram contentorizados com Docker e orquestrados via docker-compose.yml, abordagem reconhecida como facilitadora da portabilidade e reprodutibilidade em *deployments* computacionais de *edge* industrial (Chen & Zhou, 2021), garantindo a consistência do ambiente de execução entre instalações.

O *Edge Tier* é concretizado através do microcontrolador M5Stack CoreS3 (ESP32-S3), que gere três barramentos: UART1/UART2 via módulos M5Stack RS-485 Unit para sensores Modbus RTU (energia e vibração), e I2C na porta Grove A para o sensor ambiental, evitando contenção de barramento em *polling* simultâneo. O *Platform Tier* integra o broker Eclipse Mosquitto (MQTT/TLS), o serviço data-pipeline em Node.js e a base de dados InfluxDB 2.8, eficiente em séries temporais IoT face a bases NoSQL genéricas (Tripathi et al., 2023). Por fim, o *Business Tier* compõe-se de uma *dashboard* em Node.js/Express, com renderização *server-side* via EJS e comunicação em tempo real via Socket.IO, e de um serviço de análise baseado em FastAPI/Python.

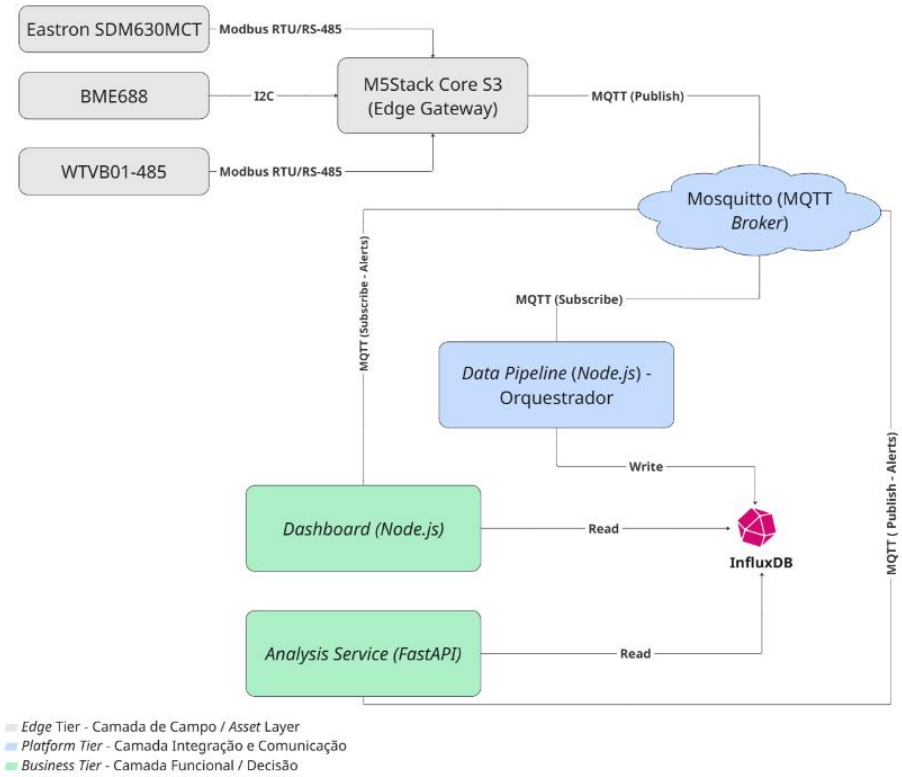


Figura 1 - Arquitetura da Plataforma

DOI: https://doi.org/10.29352/mill0224e.48333

2.2 Implementação

O *firmware* do M5Stack CoreS3 publica cada leitura de sensor num tópico MQTT hierárquico (site/cluster/sensor) sob a forma de um envelope JSON canónico. Este envelope transporta o contexto completo do dispositivo — identidade, localização e dados brutos — numa única mensagem, sendo o campo sensor_type a chave utilizada pelo pipeline para a seleção do *parser* adequado. Os valores brutos, com a nomenclatura proprietária do fabricante, são mantidos no campo data sem normalização no *edge*. Esta decisão de projeto permite centralizar toda a normalização no pipeline, assegurando que a lógica de tratamento de dados é testável e auditável.

O esquema canónico define a *measurement telemetry* no InfluxDB como o único ponto de escrita autorizado pelo data-pipeline. Os sensores são identificados através de *tags* (como sensor_id, domain e asset_id), enquanto os *fields* obedecem a uma nomenclatura estrita organizada por domínio industrial. Esta estrutura abrange as áreas de: energy (ex.: tensões, correntes, potências e THD), vibration (valores RMS e aceleração), environment (parâmetros atmosféricos e IAQ), proximity, flow e cnc (incluindo velocidade do spindle e vibração em maquinação).

A normalização semântica é implementada pela biblioteca shared/parsers, desenvolvida como um módulo Node.js partilhado entre o pipeline de produção e o simulador de telemetria. Cada ficheiro *parser* expõe uma interface uniforme (parse, getAlerts e THRESHOLDS), integrando limiares baseados em normas industriais, tais como (ISO, 1995) para vibração e (CENELEC, 2010) para tensão de linha (400 V ± 10%, independentemente da existência de condutor neutro), e os algoritmos BSEC da Bosch para qualidade do ar (IAQ). O parâmetro nominal_voltage permite configurar os limiares de tensão para diferentes configurações de alimentação, incluindo redes monofásicas (230 V) e trifásicas (400 V), independentemente da existência de condutor neutro, com o mesmo *parser*. Durante a inicialização, todos os *parsers* são carregados seguindo o princípio *fail-fast*, garantindo que eventuais anomalias de configuração são detetadas no arranque do sistema.

O serviço data-pipeline subscreve o tópico MQTT global e executa três etapas distintas por mensagem: (1) validação do contrato e rejeição de *payloads* malformados; (2) normalização semântica através de um mapeamento *hash map* com complexidade O(1) por mensagem; e (3) escrita no InfluxDB com a correta distinção entre *tags* e *fields*. A fiabilidade da comunicação é assegurada pelo nível de serviço QoS 1 (at least once) — modalidade que garante pelo menos uma entrega por mensagem em contextos de instabilidade de rede (Pawar et al., 2024) — enquanto a idempotência da escrita é garantida pela utilização de *timestamps* e etiquetas de identificação idênticas.

O registo de sensores separa a identidade imutável do dispositivo da apresentação configurável sem alteração de código. Na *dashboard*, durante a fase de inicialização, o servidor procede à leitura do registo, sendo os componentes de visualização instanciados dinamicamente de acordo com a chave de layout de cada sensor ativo. Simultaneamente, no serviço de análise (FastAPI), executa-se um ciclo autónomo a cada 60 segundos que deteta os ativos em funcionamento através de consultas ao InfluxDB. Este mecanismo permite o ajuste automático do conjunto monitorizado sem que seja necessário reiniciar o sistema, garantindo a continuidade da operação.

2.3 Metodologia de Avaliação

A avaliação quantitativa da solução incide em quatro dimensões fundamentais: (1) o esforço de integração, medido em linhas de código (LOC) por componente; (2) o *throughput* e a latência do ciclo analítico perante uma carga incremental de 3 a 300 ativos; (3) a extensibilidade, aferida pelo número de LOC alterados nos serviços preexistentes; e (4) a qualidade do sistema, assegurada pela cobertura de testes automatizados. Os ensaios de desempenho (*benchmarks*) foram executados em Docker, com 16 vCPUs e 32 GB de RAM. Como caso de estudo, utilizou-se um *deployment* industrial real.

A seleção dos três domínios de sensor (energia, vibração e ambiente) não foi arbitrária: representam as categorias de grandezas mais recorrentes na monitorização de ativos industriais rotativos — condição elétrica, condição mecânica e contexto ambiental — e correspondem, cada uma, a um protocolo de comunicação distinto (Modbus RTU a duas velocidades diferentes e I2C), maximizando a heterogeneidade de integração testada com um número reduzido de dispositivos físicos. A fresadora CNC foi selecionada como caso de estudo por reunir simultaneamente as três dimensões — consumo energético, vibração de maquinação e condições ambientais da oficina — num único ativo, permitindo testar a correlação inter-domínio (vibração/energia por material) que uma instalação mono-domínio não permitiria demonstrar.

3. RESULTADOS

A arquitetura proposta foi validada num cenário industrial real, especificamente numa fresadora CNC numa unidade industrial, onde se integraram três sensores de domínios e protocolos distintos (Tabela 2). A monitorização da corrente elétrica trifásica, com estimativa da potência ativa associada, foi assegurada pelo sensor Eastron SDM630MCT via Modbus RTU (9600 baud, 2 stop bits); a captura de vibração triaxial foi realizada pelo WitMotion WTVB01-485 (115200 baud); e as condições ambientais (temperatura, humidade, pressão e IAQ) foram monitorizadas pelo sensor Bosch BME688 via I2C.

No que respeita à monitorização de energia, os terminais de tensão do SDM630MCT não foram ligados fisicamente: a ligação elétrica ao quadro foi da responsabilidade da unidade industrial de acolhimento, que optou por não autorizar essa ligação. Nestas condições, apenas as correntes de fase (I1, I2, I3) foram medidas em tempo real através das pinças CT instaladas nos cabos de alimentação da fresadora; a potência ativa foi calculada a partir dessas correntes segundo $P = \sqrt{3} \times V_LL \times Iavg \times FP_nominal$, assumindo os valores nominais da rede industrial 3P3 para a tensão de linha (V_LL = 400 V) e para o fator de potência (FP_nominal = 0,90). Os valores de tensão, fator de potência e potência ativa publicados no esquema canónico

DOI: https://doi.org/10.29352/mill0224e.48333

correspondem, por isso, a uma estimativa baseada nestes nominais, e não a uma medição direta — o rigor desta estimativa está sujeito ao desvio real do fator de potência da carga face ao valor assumido, uma vez que o fator de potência varia com o regime de funcionamento do equipamento e não é uma propriedade fixa da rede elétrica.

A utilização destes três sensores, com nomenclaturas proprietárias distintas, validou a eficácia da normalização semântica da biblioteca shared/parsers: cada sensor publica os seus campos brutos, normalizados pelo pipeline para o esquema canónico. No contexto da maquinação CNC, cada sessão é associada a um session_id e ao material selecionado, viabilizando a correlação de perfis de vibração e energia por tipo de material, em conformidade com as normas de referência para vibração em maquinaria (Hassan et al., 2024; ISO, 1995), com padrões consistentes com os reportados em maquinagem CNC (Apostolou et al., 2024).

Durante o período de validação experimental, a infraestrutura processou a telemetria de 18 sensores ativos, distribuídos por três locais de teste. Com um intervalo de publicação MQTT configurado para 5 segundos por sensor, a taxa de ingestão total ascendeu a 216 mensagens/minuto (correspondendo a 3,6 msg/s). Assegurou-se a persistência de 1 080 campos/minuto no InfluxDB com uma taxa de descarte nula, facto que confirma a estabilidade operacional da *stack* tecnológica implementada. Este valor de 216 mensagens/minuto reflete a procura real do deployment (18 sensores × 5 s de intervalo), não a capacidade máxima da plataforma: os testes de escalabilidade da Tabela 3 confirmam margem de crescimento linear até 300 ativos simultâneos — mais de 16 vezes o volume atualmente em produção — sem degradação do tempo médio de processamento por ativo.

Tabela 2 - Sensores Integrados no Caso de Estudo CNC

Sensor	Protocolo	Domínio	Campos Canónicos (exemplos)
SDM630MCT (Eastron)	Modbus RTU, 9600 baud	energy	voltage_l1/2/3, active_power, power_factor, import_energy, thd_v1
WTVB01-485 (WitMotion)	Modbus RTU, 115200 baud	vibration	rms_x/y/z, rms_total, acceleration_x/y/z, temperature_internal
BME688 (Bosch)	I2C, 0x76	environment	temperature, humidity, pressure, iaq, gas_resistance

Nota: os campos voltage_l1/2/3, power_factor, active_power, import_energy e thd_v1 do SDM630MCT resultam de uma estimativa baseada em corrente medida e nos valores nominais da rede (ver secção 3, Resultados); apenas current_l1/2/3 correspondem à medição direta neste caso de estudo

Os testes de desempenho do motor analítico (FastAPI) indicaram um tempo de processamento médio de 107 ms por ativo, num *benchmark* com carga incremental de 3 a 300 ativos simultâneos (Tabela 3). Observou-se um crescimento estritamente linear, o que confirma a inexistência de *overhead* de coordenação entre ativos; cada unidade é processada de forma independente, tirando partido da descoberta automática *data-driven* via consultas ao InfluxDB. O *Service Level Objective* (SLO) de 30 s foi violado apenas quando atingido o patamar de 300 ativos simultâneos — volume que excede o cenário de validação deste estudo (18 ativos em operação). Esta restrição de desempenho é intrínseca ao modelo unidirecional (single-threaded) do ciclo de execução atual, delineando-se como trabalho futuro a paralelização dos processos por identificador de instalação (site_id).

Tabela 3 - Escalabilidade do Loop de Análise (discovery automático)

N.º Ativos	Tempo Total (s)	Tempo Médio/Ativo (s)	SLO 30 s	SLO 60 s
3	0,3	0,095	CUMPRIU	CUMPRIU
10	1,0	0,102	CUMPRIU	CUMPRIU
25	2,6	0,105	CUMPRIU	CUMPRIU
50	5,3	0,107	CUMPRIU	CUMPRIU
100	10,8	0,108	CUMPRIU	CUMPRIU
300	32,1	0,107	VIOLOU	CUMPRIU

Nota: SLO = Service Level Objective. Benchmark em Docker (16 vCPUs, 32 GB RAM)

A extensibilidade sem necessidade de reengenharia, contribuição central desta arquitetura, é quantificada na Tabela 4 para três sensores de domínios e complexidades distintas. A título de exemplo, a integração do caudalímetro ultrassónico KLINGER_ST (Modbus RTU) requereu exclusivamente a criação de um novo parser (51 linhas de código) e uma entrada no sensor registry, adicionada automaticamente a partir do momento em que existe telemetria no InfluxDB; os três serviços identificaram o novo ativo (Figura 2) no ciclo seguinte, sem reinicialização. À data da validação, o Sensor Registry contabilizava 22 sensores ativos, distribuídos por 6 domínios industriais e 3 locais distintos, abrangendo 16 tipologias de hardware.

Tabela 4 - LOC por Componente para Integração de Novos Sensores (Prova de Extensibilidade)

Sensor	Domínio	LOC Parser (shared/)	LOC data-pipeline	LOC dashboard	LOC analysis-service
KLINGER_ST	Flow	51	0	0	0
SD6501	Flow	49	0	0	0
SHT31	Environment	36	0	0	0

Nota: Zero linhas alteradas nos serviços existentes. O sensor registry (is_active: true) ativa o *discovery* automático na *dashboard* e no serviço de análise sem reinício.

DOI: https://doi.org/10.29352/mill0224e.48333

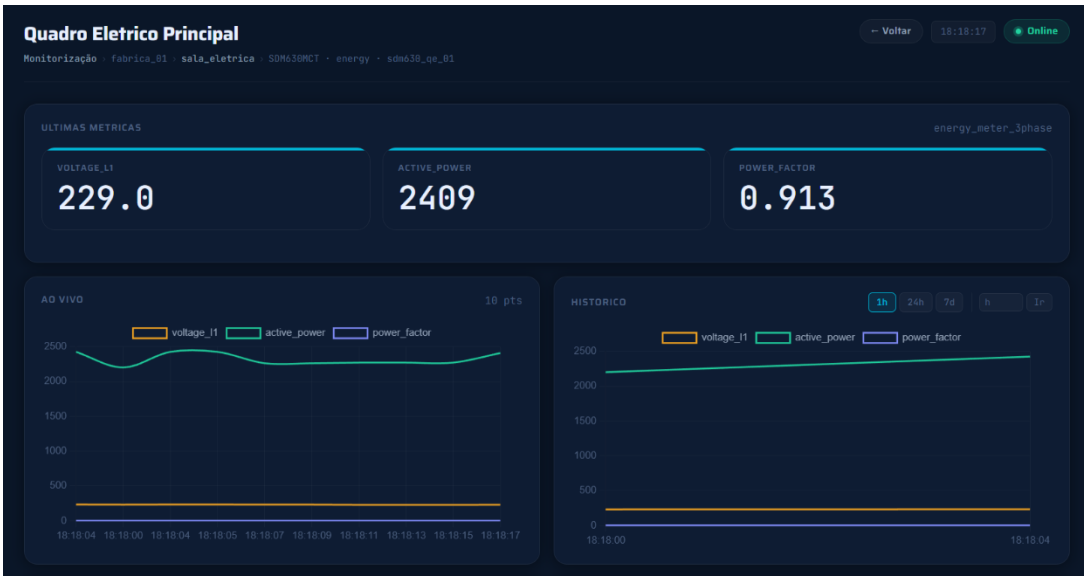


Figura 2 - Página de monitorização gerada automaticamente pela dashboard para um sensor adicionado ao sensor-registry.json com is_active: true. Nenhum serviço foi reiniciado nem alterado; a página surgiu no ciclo de descoberta seguinte (≤ 60 s)

A fiabilidade da plataforma é assegurada por 313 testes automatizados: 117 no serviço de análise (92% na lógica de deteção de anomalias), 95 no data-pipeline (93,67% de instruções; 95,89% de linhas) e 101 na dashboard. O simulador de telemetria — que partilha a biblioteca shared/parsers com o pipeline de produção — garante testes de integração com 16 tipos de sensor em três instalações virtuais, detetando inconsistências entre parser e esquema canónico antes do *deployment* físico.

A plataforma disponibiliza três interfaces: (i) uma dashboard web com visualização dinâmica por tipo de sensor, gráficos temporais e alertas em tempo real via Socket.IO; no contexto CNC, inclui vista de correlação vibração/energia por sessão de maquinagem e tipo de material processado (figura3); (ii) uma API REST documentada, que expõe deteção de anomalias por Z-Score e Média Móvel (bandas de desvio padrão) — métodos revistos em (Blázquez-García et al., 2022) e enquadrados no estado da arte de deteção de anomalias IoT (DeMedeiros et al., 2023); e (iii) um registo estruturado de mensagens rejeitadas (drop log), acessível na *dashboard*, que permite auditoria do pipeline.



Figura 3 - Correlação entre vibração (rms_total) e velocidade do spindle (spindle_speed) por tipo de material em maquinagem CNC, obtida a partir dos dados normalizados no esquema canónico. Cada série corresponde a um material processado

DOI: https://doi.org/10.29352/mill0224e.48333

4. DISCUSSÃO

Os resultados obtidos demonstram que a arquitetura proposta resolve eficazmente o problema da heterogeneidade semântica em plataformas IIoT industriais. A evidência quantitativa de zero LOC alterados nos serviços existentes para a integração de cada novo sensor revela-se uma métrica mais robusta do que as avaliações qualitativas de extensibilidade frequentemente encontradas na literatura.

A interface uniforme parse/getAlerts/THRESHOLDS, partilhada entre o pipeline e o simulador, constitui a contribuição central que distingue esta abordagem. Ao reutilizar-se a mesma biblioteca em ambiente de produção e em testes sintéticos, elimina-se a principal fonte de divergência entre ambientes — um problema recorrentemente documentado em implementações industriais reais.

A escalabilidade linear decorre diretamente do modelo de discovery automático: cada ativo é resolvido de forma independente a partir do sensor registry, sem estado partilhado nem sincronização entre ativos, pelo que o tempo total cresce proporcionalmente ao número de ativos. O sensor registry reduz o acoplamento por indireção: a dashboard e o serviço de análise não referenciam sensores individuais no código, apenas a estrutura do registry, pelo que a adição, remoção ou reconfiguração de um sensor não implica alterações nos consumidores dessa informação. Quanto à robustez, o carregamento fail-fast dos parsers deteta módulos malformados no arranque, e mensagens que violem o contrato de dados são desviadas para uma fila de erros em vez de processadas silenciosamente. A evolução do esquema apresenta, contudo, uma limitação atual: a introdução de um novo tipo de sensor requer o reinício do serviço de ingestão, e o registry não implementa versionamento explícito de esquemas de sensor — alterações retrocompatíveis não são distinguidas de alterações que quebram compatibilidade.

A escalabilidade linear observada até 300 ativos, com uma latência média de 107 ms por unidade, evidencia uma margem de crescimento superior a 16 vezes face aos 18 ativos do cenário de validação — indicador relevante para contextos industriais que operam tipicamente com um número reduzido de ativos monitorizados. Esta latência de processamento de 107 ms por ativo, embora superior aos 59 ms obtidos por (Boonmeeruk et al., 2024) para transações puras Modbus TCP/IP, justifica-se pelo *overhead* computacional da camada de normalização semântica e validação de contratos em tempo real executada no *Platform Tier*, um *trade-off* que troca desempenho bruto por flexibilidade de agnosticismo de hardware.

A comparação com plataformas de referência (ThingsBoard, Node-RED, Grafana+InfluxDB e OpenRemote) evidencia que a convergência entre *parsers* extensíveis, esquema canónico partilhado e descoberta automática nativa não se encontra disponível em nenhuma das soluções analisadas sem a necessidade de desenvolvimento adicional significativo.

Esta extensibilidade decorre de um agnosticismo a dois níveis: ao nível do *Platform/Business Tier*, os serviços correm em Docker sobre hardware convencional (≥ 2 vCPUs, ≥ 4 GB de RAM). Ao nível do *Edge Tier*, o agnosticismo é de contrato de dados: qualquer dispositivo que publique no tópico MQTT canónico com o envelope JSON definido é integrado sem alteração de código no pipeline, independentemente do microcontrolador utilizado. Os sistemas revistos na literatura (Boonmeeruk et al., 2024; Verde Romero et al., 2024; He et al., 2024) seriam, por isso, compatíveis sem alteração, bastando publicarem nesse formato de envelope; o firmware ESP32-S3 aqui descrito constitui uma implementação de referência do contrato, não um requisito de hardware.

A generalização dos resultados apresenta limitações inerentes ao contexto de validação. O deployment decorreu numa única máquina e numa única unidade industrial; a validação numa segunda empresa ou máquina de tipologia distinta permitiria distinguir efeitos específicos do ativo dos efeitos genéricos da arquitetura. A monitorização de energia constitui uma segunda limitação: por decisão da unidade industrial de acolhimento, responsável pela ligação elétrica ao quadro, os terminais de tensão do SDM630MCT não foram ligados; a potência ativa publicada resulta de corrente medida e de tensão e fator de potência assumidos nos nominais da rede (400 V, FP = 0,90), pelo que o erro relativo da estimativa iguala, em primeira aproximação, o desvio real do fator de potência face ao valor assumido — erro não quantificado neste estudo por ausência de medição direta de referência. Adicionalmente, os benchmarks foram executados com recursos fixos (16 vCPUs, 32 GB RAM), podendo diferir em hardware de menor capacidade típico de PME, e a comparação com as plataformas da Tabela 1 é qualitativa, sem métricas de desempenho diretamente comparáveis. Por fim, a violação do SLO de 30 s aos 300 ativos é atribuível ao modelo single-threaded do ciclo de análise, mitigável através da paralelização por site_id, proposta como trabalho futuro.

CONCLUSÃO

Este artigo apresentou e validou uma plataforma de monitorização IIoT sensor-agnóstica, implementada e testada numa unidade industrial de manufatura, respondendo ao problema de heterogeneidade semântica identificado na Introdução. A integração de sensores de corrente elétrica, vibração e ambiente — heterogéneos em protocolo e fabricante — foi alcançada sem alterações nos serviços existentes, a um custo médio inferior a 51 linhas de código por novo tipo de sensor (Tabela 4), com *throughput* e latência estáveis até 300 ativos simultâneos (Tabela 3). Esta extensibilidade resulta da convergência entre o esquema canónico do InfluxDB, a biblioteca partilhada de parsers e o sensor registry JSON, cuja combinação — e não os mecanismos isolados — constitui a resposta proposta ao problema de partida.

DOI: https://doi.org/10.29352/mill0224e.48333

A contribuição central — a interface uniforme parse/getAlerts/THRESHOLDS partilhada entre o pipeline de produção e o simulador de telemetria — assegura que a lógica de normalização seja validada com dados sintéticos realistas previamente ao *deployment* físico, eliminando a principal fonte de divergência entre ambientes de teste e produção.
Como trabalho futuro propõe-se:

- A integração do protocolo OPC-UA para reforçar a interoperabilidade com sistemas SCADA;
- A paralelização do ciclo de análise por site_id, visando superar o limite de 300 ativos mantendo o rigor do SLO;
- O desenvolvimento de modelos de manutenção preditiva baseados no histórico de dados acumulado; (4) A publicação da biblioteca shared/parsers como um pacote open-source (NPM), fomentando a reutilização e desenvolvimento pela comunidade IIoT.

AGRADECIMENTOS

Este trabalho foi desenvolvido no âmbito do projeto final de licenciatura em Engenharia Informática da Escola Superior de Tecnologia e Gestão de Viseu, Instituto Politécnico de Viseu. Expressa-se um agradecimento ao STAR Institute pela disponibilização do ambiente industrial e do apoio técnico necessários para a validação experimental da arquitetura proposta.

CONTRIBUIÇÃO DOS AUTORES

Conceptualização, N.S.; tratamento de dados, N.S.; investigação, N.S. e R.C.; metodologia, N.S.; administração do projeto, N.S. e R.C.; supervisão, R.C.; validação, N.S. e R.C.; visualização, N.S.; redação-preparação de rascunho original, N.S.; redação. Revisão e edição, N.S. e R.C.

CONFLITO DE INTERESSES

Os autores declaram não existir conflito de interesses.

REFERÊNCIAS BIBLIOGRÁFICAS

Apostolou, G., Ntemi, M., Paraschos, S., Gialampoukidis, I., Rizzi, A., Vrochidis, S., & Kompatsiaris, I. (2024). Novel Framework for Quality Control in Vibration Monitoring of CNC Machining. *Sensors*, *24*(1), 307. https://doi.org/10.3390/s24010307

Blázquez-García, A., Conde, A., Mori, U., & Lozano, J. A. (2022). A Review on Outlier/Anomaly Detection in Time Series Data. In *ACM Computing Surveys* (Vol. 54). Association for Computing Machinery. https://doi.org/10.1145/3444690

Boareto, P. A., Assad, A., Loures, E. F. R., Deschamps, F., & Santos, E. A. P. (2026). Design and Implementation of Industrial Internet of Things for Legacy Field Devices: A Case Study. *IEEE Access*, *14*, 45882–45895. https://doi.org/10.1109/ACCESS.2026.3674418

Boonmeeruk, P., Palrat, P., & Wongsopanakul, K. (2024). Cost-Effective IIoT Gateway Development Using ESP32 for Industrial Applications. *Engineering Journal*, *28*, 93–108. https://doi.org/10.4186/ej.2024.28.10.93

Busboom, A. (2024). Automated generation of OPC UA information models — A review and outlook. *Journal of Industrial Information Integration*, *39*, 100602. https://doi.org/10.1016/j.jii.2024.100602

CENELEC. (2010). *EN 50160: Voltage characteristics of electricity supplied by public electricity networks*.

Chen, S., & Zhou, M. (2021). Evolving Container to Unikernel for Edge Computing and Applications in Process Industry. *Processes*, *9*(2), 351. https://doi.org/10.3390/pr9020351

DeMedeiros, K., Hendawi, A., & Alvarez, M. (2023). A Survey of AI-Based Anomaly Detection in IoT and Sensor Networks. *Sensors*, *23*(3), 1352. https://doi.org/10.3390/s23031352

Eldho K J. (2025). Optimizing Data Transfer Speed and the Performance Evaluation of MQTT in IoT: A Study on MQTT for Atmospheric Condition Monitoring. *Journal of Information Systems Engineering and Management*, *10*, 231–244. https://doi.org/10.52783/jisem.v10i39s.7144

Hassan, I. U., Panduru, K., & Walsh, J. (2024). An In-Depth Study of Vibration Sensors for Condition Monitoring. *Sensors*, *24*(3), 740. https://doi.org/10.3390/s24030740

Hazra, A., Adhikari, M., Amgoth, T., & Srirama, S. N. (2023). A Comprehensive Survey on Interoperability for IIoT: Taxonomy, Standards, and Future Directions. *ACM Computing Surveys*, *55*(1), 1–35. https://doi.org/10.1145/3485130

He, W., Baig, M. J. A., & Iqbal, M. T. (2024). An Open-Source Supervisory Control and Data Acquisition Architecture for Photovoltaic System Monitoring Using ESP32, Banana Pi M4, and Node-RED. *Energies*, *17*. https://doi.org/10.3390/en17102295

DOI: https://doi.org/10.29352/mill0224e.48333

Hercog, D., Lerher, T., Truntič, M., & Težak, O. (2023). Design and Implementation of ESP32-Based IoT Devices. *Sensors*, *23*(15), 6739. https://doi.org/10.3390/s23156739

ISO. (1995). *ISO-10816-1-1995*. https://www.iso.org/standard/18866.html

Masood, T., & Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs. *Computers in Industry*, *121*. https://doi.org/10.1016/j.compind.2020.103261

Mirani, A. A., Awasthi, A., O'Mahony, N., & Walsh, J. (2024). Industrial IoT-Based Energy Monitoring System: Using Data Processing at Edge. *Internet of Things*, *5*, 608–633. https://doi.org/10.3390/iot5040027

Monios, N., Peladarinos, N., Cheimaras, V., Papageorgas, P., & Piromalis, D. D. (2024). A Thorough Review and Comparison of Commercial and Open-Source IoT Platforms for Smart City Applications. In *Electronics (Switzerland)* (Vol. 13). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/electronics13081465

Müller, J. M., Islam, N., Kazantsev, N., Romanello, R., Olivera, G., Das, D., & Hamzeh, R. (2024). Barriers and Enablers for Industry 4.0 in SMEs: A Combined Integration Framework. *IEEE Transactions on Engineering Management*, 1–13. https://doi.org/10.1109/TEM.2024.3365771

Pawar, S., Jadhav, D. B., Lokhande, M., Raskar, P., & Patil, M. (2024). Evaluation of quality of service parameters for MQTT communication in IoT application by using deep neural network. *International Journal of Information Technology*, *16*(2), 1123–1136. https://doi.org/10.1007/s41870-023-01664-2

Rahman, H., & Hussain, M. I. (2020). A comprehensive survey on semantic interoperability for Internet of Things: State-of-the-art and research challenges. *Transactions on Emerging Telecommunications Technologies*, *31*. https://doi.org/10.1002/ett.3902

Schmidt, C., Volz, F., Stojanovic, L., & Sutschet, G. (2023). Increasing Interoperability between Digital Twin Standards and Specifications: Transformation of DTDL to AAS. *Sensors*, *23*(18), 7742. https://doi.org/10.3390/s23187742

Tripathi, P., Miraz, M. H., & Joshi, S. (2023). Comparative Analysis of MongoDB and InfluxDB for Time Series Data Management in IoT Environments: A Study on Performance, Scalability, and Concurrency. *2023 International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA)*, 39–42. https://shre.ink/GS40

Verde Romero, D. A., Villalvazo Laureano, E., Jiménez Betancourt, R. O., & Navarro Álvarez, E. (2024). An open source IoT edge-computing system for monitoring energy consumption in buildings. *Results in Engineering*, *21*. https://doi.org/10.1016/j.rineng.2024.101875

Xu, B., Guertler, M. R., & Sick, N. (2023). Analyzing Industry 4.0 Adoption Barriers of Small and Medium-sized Enterprises and Existing Support. *2023 IEEE Engineering Informatics, EI 2023*. https://shre.ink/GS4w

You, S., Li, X., Yu, H., & Chen, W. (2024). S-ID: Implementing Intelligent Interoperability for Heterogeneous IoT Identifiers Through a Semantic-Based Framework. *IEEE Access*, *12*, 1341–1359. https://doi.org/10.1109/ACCESS.2023.3347586