

Inteligência artificial generativa para otimização da análise de causa raiz: um estudo de caso na manufatura de equipamentos 5G

Generative artificial intelligence for root cause analysis optimization: a case study in 5G equipment manufacturing

Holden Offmenn¹
Ítala Lorena de Lima Ferreira²
Daniely Dantas Lobato³
Jhônatas Cardoso Luz⁴
Luana Dalla Rosa de Carvalho⁵
Laís Miranda Olímpio⁶
André Luiz Samistraro Santin⁷

Resumo

A complexidade da manufatura de equipamentos de telecomunicações de última milha, como a família OptiXstar, exige diagnósticos de falhas extremamente ágeis para manter a competitividade global. Este artigo apresenta os resultados da implementação da plataforma Cyber Nexus, que utiliza Modelos de Linguagem de Grande Escala (LLMs) para automatizar a Análise de Causa Raiz (RCA) no Polo Industrial de Manaus. O objetivo foi elevar o patamar de maturidade em Indústria 4.0 do processo, migrando de uma visibilidade reativa para uma transparência preditiva. A metodologia consistiu num estudo de caso experimental com validação em ambiente relevante (TRL 6), envolvendo engenheiros de qualidade e dados reais de produção. Os resultados demonstraram uma redução drástica no ciclo de RCA, de uma média de 15 dias para menos de 48 horas, com uma taxa de precisão na detecção de anomalias superior a 80%. A solução permitiu a correlação em tempo real entre variáveis de calibração de rádio frequência e logs de falhas, estabelecendo as bases para a manutenção preditiva e a manufatura verde através da redução de retrabalhos.

Palavras-chave: Indústria 4.0. Análise de Causa Raiz. Manufatura Inteligente. Eficiência Operacional. LLM.

Abstract

The complexity involved in manufacturing last-mile telecommunications equipment, such as the OptiXstar family, requires extremely agile fault diagnosis in order to maintain global competitiveness. This article presents the results of implementing the Cyber Nexus platform, which uses Large

¹ Especialista em Inteligência Artificial do Instituto Conecthus. e-mail: holden.offmenn@conecthus.org.br

² Cientista de Dados do Instituto Conecthus. e-mail: itala.ferreira@conecthus.org.br

³ Cientista de Dados do Instituto Conecthus. e-mail: daniely.lobato@conecthus.org.br

⁴ Desenvolvedor Back-End do Instituto Conecthus. e-mail: jhonatas.luz@conecthus.org.br

⁵ Analista de Testes do Instituto Conecthus. e-mail: luana.carvalho@conecthus.org.br

⁶ Gerente de Negócios da Huawei. e-mail: lais.olimpio@huawei.com

⁷ Gerente de Pesquisa e Desenvolvimento da Huawei. e-mail: andre.santin@huawei.com

Language Models (LLMs) to automate Root Cause Analysis (RCA) at the Manaus Industrial Pole. The objective was to raise the process's Industry 4.0 maturity level by shifting from reactive visibility to predictive transparency. The methodology consisted of an experimental case study with validation in a relevant environment (TRL 6), involving quality engineers and real production data. The results demonstrated a drastic reduction in the RCA cycle, from an average of 15 days to less than 48 hours, with an anomaly detection accuracy rate above 80%. The solution enabled real-time correlation between radio frequency calibration variables and failure logs, establishing the foundation for predictive maintenance and green manufacturing through the reduction of rework.

Keywords: Industry 4.0. Root Cause Analysis. Smart Manufacturing. Operational Efficiency. LLM.

1 INTRODUÇÃO

A quarta revolução industrial consolidou a necessidade de uma integração profunda entre sistemas físicos e digitais para otimizar a competitividade no setor de manufatura de alta tecnologia. Segundo Schuh et al. (2020), o sucesso da transformação digital não reside apenas na coleta massiva de dados, mas na capacidade das organizações de evoluírem através de estágios de maturidade que permitam a compreensão clara dos processos e a antecipação de falhas. No entanto, muitas indústrias ainda enfrentam barreiras significativas na transição entre os níveis de conectividade e visibilidade total, resultando em uma subutilização de ativos digitais e em processos de análise frequentemente reativos.

No Polo Industrial de Manaus, a fabricação de equipamentos de telecomunicações de última geração, como os rádios 5G e sistemas Mesh da linha OptiXstar da Huawei, exemplifica essa complexidade operacional. O processo de Análise de Causa Raiz (RCA), fundamental para a garantia da qualidade e eficiência da linha, opera atualmente em um cenário de silos informacionais onde dados de calibração do sistema QualitySys e logs de produção estão dispersos em bases heterogêneas. Esta fragmentação impõe um ciclo de diagnóstico moroso, que consome entre 4 e 15 dias para ser concluído, impactando diretamente a agilidade da cadeia de suprimentos e a produtividade da planta.

A literatura acadêmica aponta que a Inteligência Artificial Generativa, impulsionada por Modelos de Linguagem de Grande Escala (LLMs), oferece uma nova fronteira para a automação desses diagnósticos industriais complexos. Conforme Wang et al. (2024), essas tecnologias permitem a interpretação semântica de logs técnicos e a correlação automatizada de variáveis, transformando tarefas manuais exaustivas em fluxos de *insights* acionáveis em tempo real. A adoção de técnicas como o *Retrieval-Augmented Generation* (RAG) possibilita

que esses modelos fundamentem suas conclusões em evidências reais de produção, aumentando drasticamente a confiabilidade das predições de falha (MDPI, 2024).

A implementação da plataforma Cyber Nexus buscou endereçar especificamente esse gargalo, visando elevar a maturidade do processo produtivo do Nível 2/3 (Conectividade e Visibilidade Parcial) para o Nível 4 (Transparência) e estabelecer as bases para o Nível 5 (Capacidade Preditiva) do índice Acatech. O desafio central consistiu em conectar bases protegidas por múltiplas camadas de criptografia para permitir que a inteligência artificial analisasse o ciclo de vida dos dados de forma holística, superando a latência imposta por métodos tradicionais de integração manual e extração em lote.

Este artigo descreve o desenvolvimento e a validação experimental desta solução, que alcançou o nível de maturidade tecnológica TRL 6 através de testes realizados em ambiente relevante. Os resultados demonstraram que a orquestração de agentes inteligentes e LLMs pode reduzir o tempo de resposta para a RCA em mais de 85%, estabilizando o ciclo de análise em menos de 48 horas e mantendo uma precisão superior a 80% na detecção de anomalias. Através deste estudo de caso, demonstra-se como a IA Generativa atua como catalisadora da eficiência e da manufatura verde ao reduzir drasticamente o retrabalho e o desperdício de recursos industriais.

2 FUNDAMENTAÇÃO TEÓRICA OU REVISÃO DA LITERATURA

A sustentação científica deste estudo baseia-se na convergência entre os modelos de maturidade digital, a aplicação de modelos de linguagem em sistemas industriais e a busca pela eficiência operacional através da manufatura inteligente.

2.1 O Índice de Maturidade da Indústria 4.0 (Acatech)

A jornada de transformação digital nas organizações é frequentemente guiada por modelos que definem estágios evolutivos de competência técnica. O índice de maturidade proposto pela Acatech estabelece seis estágios: informatização, conectividade, visibilidade, transparência, capacidade preditiva e adaptabilidade (SCHUH et al., 2020). Enquanto a conectividade e a visibilidade garantem que os dados estejam disponíveis e centralizados, o estágio de transparência é alcançado quando a organização consegue entender o "porquê" dos eventos ocorrerem, através da integração semântica de informações. O projeto Cyber Nexus visa consolidar este nível de transparência, fornecendo as ferramentas necessárias para que a

análise de dados complexos revele as causas fundamentais de desvios de qualidade (SCHUH et al., 2020).

2.2 Inteligência Artificial Generativa e Análise de Causa Raiz (RCA)

A Análise de Causa Raiz em sistemas ciber-físicos industriais é uma tarefa intensiva em conhecimento que, tradicionalmente, exige que especialistas interpretem logs e relatórios técnicos de forma manual. A emergência dos Modelos de Linguagem de Grande Escala (LLMs) permite que essa análise seja tratada através de inferência lógica e processamento de linguagem natural (WANG et al., 2024). De acordo com Wang et al. (2024), a aplicação de LLMs na RCA permite a construção de grafos de conhecimento dinâmicos que correlacionam anomalias detectadas em sensores com padrões históricos de falha. Esta abordagem reduz drasticamente o tempo necessário para diagnósticos precisos, especialmente em setores de hardware sensível, como o de telecomunicações, onde os logs de produção são extensos e heterogêneos (MDPI, 2024).

2.3 Eficiência Operacional e Smart Manufacturing

A manufatura inteligente (*Smart Manufacturing*) busca a otimização contínua de recursos e a redução de desperdícios através da integração de inteligência artificial (ZHOU et al., 2018). O impacto dessa integração é mensurado diretamente na Eficiência Global dos Equipamentos (OEE), ao minimizar o tempo médio de diagnóstico e reparo (TAO et al., 2019). Além do ganho de produtividade, a agilidade na identificação de falhas contribui para a "manufatura verde", visto que a redução de retrabalhos e de lotes defeituosos diminui diretamente o consumo de energia e o descarte de componentes eletrônicos sensíveis. A capacidade preditiva derivada da análise automatizada permite que a indústria atue proativamente, consolidando um modelo de produção mais sustentável e competitivo (ZHOU et al., 2018).

3 METODOLOGIA

O percurso metodológico deste estudo caracteriza-se como uma pesquisa aplicada com abordagem de estudo de caso experimental, conduzida nas instalações da Foxconn em Manaus, unidade responsável pela produção de ativos de rede da Huawei. A investigação buscou elevar a solução tecnológica do patamar de prova de conceito para a validação em ambiente relevante, atingindo o nível de maturidade tecnológica TRL 6.

3.1 Design da Pesquisa e Cenário Experimental

A pesquisa foi estruturada para analisar o fluxo de informações da família de produtos OptiXstar, cujos diagnósticos de falha exigiam a correlação de variáveis de três fontes principais: o banco de dados Oracle (QualitySys), contendo parâmetros de calibração; os Servidores de Logs, acessados via protocolos SSH para rastreabilidade de hardware; e os sistemas AICC/PLM, que armazenam o histórico de pós-venda. O cenário experimental consistiu na replicação dessas bases em um ambiente de produção controlado para permitir a execução de testes de estresse e análise de latência sem interferir na linha de montagem ativa.

3.2 Implementação do Motor Algorítmico e Ciclo Agêntico

A solução técnica foi desenvolvida utilizando uma arquitetura de microsserviços em Python, orquestrada pelo framework FastAPI e containerizada em Docker para garantir a portabilidade no ambiente Ubuntu Server. O núcleo da inteligência baseou-se no conceito de "Agentic Loop", implementado através da biblioteca LangChain, que permite ao sistema decidir de forma autônoma a sequência de ferramentas (*tools*) necessárias para responder a uma consulta técnica complexa.

Diferente de sistemas de automação rígidos, a metodologia empregada integrou a estratégia *Human-in-the-Loop* para mitigar o risco de "iterações infinitas" da IA. Definiu-se um limite mandatório de três iterações para cada consulta; caso o agente não alcançasse uma conclusão com alta confiança, o sistema realizava um *fallback*, solicitando informações adicionais ao engenheiro de qualidade. Essa abordagem assegurou a confiabilidade dos *insights* gerados, evitando diagnósticos errôneos baseados em dados ambíguos

3.3 Coleta de Dados e Integração Semântica

Para superar o desafio das múltiplas camadas de criptografia, a coleta de dados não seguiu o modelo tradicional de extração em massa. Em vez disso, utilizou-se o paradigma de *Retrieval-Augmented Generation* (RAG) descentralizado. Agentes especializados foram programados para realizar consultas pontuais e converter resultados heterogêneos em vetores semânticos interpretáveis pelos modelos Maritaca AI e DeepSeek. Esse processo permitiu que a plataforma "entendesse" a relação entre um desvio de milivolts em um teste de radiofrequência e uma falha lógica registrada nos logs de software, algo que anteriormente exigia horas de análise humana cross-referenciada.

3.4 Procedimentos de Validação e Métricas de Performance

A validação da eficácia operacional foi realizada através de Testes de Aceitação do Usuário (UAT) com uma amostra de engenheiros de dados e analistas de qualidade da Huawei e Foxconn. As métricas de avaliação foram fundamentadas no Índice de Maturidade da Indústria 4.0 da Acatech, focando na transição para os estágios de Transparência e Capacidade Preditiva (SCHUH et al., 2020). Os indicadores quantitativos incluíram:

- **Tempo de Ciclo de RCA:** Mensuração do tempo total entre a detecção da anomalia e o diagnóstico final, comparando o processo anterior (manual) com o automatizado.
- **Latência de Resposta:** Tempo decorrido para a consolidação de dados cruzados entre bases C7 (criptografia máxima).
- **Acurácia do Diagnóstico:** Percentual de acertos da IA na identificação da causa raiz em comparação com o parecer de especialistas humanos, com meta estabelecida em 80%.

A metodologia encerrou-se com a documentação técnica "as-is" e a transferência de conhecimento para as equipes de operação assistida, garantindo que os ganhos de eficiência fossem sustentáveis após o encerramento do projeto de P&D

4 RESULTADOS E DISCUSSÕES OU ANÁLISE DOS DADOS

A validação da plataforma Cyber Nexus em ambiente relevante (TRL 6) forneceu evidências concretas sobre o impacto da inteligência artificial generativa na otimização de processos industriais complexos. Os resultados foram analisados sob a ótica da performance técnica e da transformação da maturidade operacional.

4.1 Otimização do Tempo de Resposta e Latência de Dados

A implementação da arquitetura descentralizada de agentes permitiu uma superação significativa dos gargalos de latência impostos pelas múltiplas camadas de criptografia. A transição do modelo tradicional de extração em lote (*batch*) para a leitura via *stream* resultou em uma redução de 90% na latência de acesso aos metadados industriais. Essa eficiência possibilitou que a plataforma consolidasse informações de bases heterogêneas, como o QualitySys (Oracle) e logs de servidores Linux, em tempo quase real.

Observou-se que consultas técnicas complexas, que antes exigiam intervenções manuais demoradas para a extração e cruzamento de dados, passaram a ser retornadas em um tempo médio inferior a 12 segundos. Esse ganho de performance técnica é um indicador

fundamental da viabilidade de sistemas baseados em LLMs para a operação em "borda" (*edge*), onde a velocidade de processamento é crítica para a manutenção do fluxo produtivo. A redução da latência não apenas acelera a consulta, mas democratiza o acesso à informação estratégica para as equipes de engenharia no chão de fábrica.

4.2 Redução do Ciclo de Análise de Causa Raiz (RCA)

O impacto mais direto na eficiência operacional da planta foi a drástica redução no ciclo total da Análise de Causa Raiz (RCA). Antes da introdução do Cyber Nexus, o diagnóstico de falhas em produtos como a família OptiXstar exigia entre 4 e 15 dias de análise técnica manual devido à fragmentação informacional. Com a automação da correlação de dados via agentes inteligentes, esse tempo foi estabilizado em menos de 48 horas, representando uma melhoria superior a 85% na agilidade do diagnóstico.

Essa aceleração no ciclo de RCA permitiu que a gestão industrial identificasse gargalos de qualidade de forma quase imediata, evitando que desvios de calibração em componentes de rádio frequência 5G se propagassem por lotes inteiros de produção. A redução do tempo de resposta reflete-se diretamente na diminuição do Tempo Médio de Reparo (MTTR) e no aumento da Eficiência Global dos Equipamentos (OEE), consolidando um modelo de manufatura ágil e responsiva às demandas do mercado global.

4.3 Acurácia e Confiabilidade do Diagnóstico com LLMs

A precisão dos diagnósticos gerados pelo motor algorítmico, utilizando os modelos Maritaca AI e DeepSeek, superou as expectativas iniciais do projeto. Nos Testes de Aceitação do Usuário (UAT), a plataforma demonstrou uma taxa de acerto na identificação de anomalias e padrões de falha superior a 80%. A capacidade de interpretação semântica dos modelos permitiu correlacionar logs de software abstratos com variáveis físicas de hardware de forma mais precisa que as análises manuais retrospectivas.

Para garantir a confiabilidade, a metodologia de *Human-in-the-Loop* foi fundamental, especialmente em casos de ambiguidade nos dados de entrada. O sistema foi configurado para solicitar validação humana sempre que a confiança do agente caísse abaixo de um patamar crítico, garantindo que a decisão final permanecesse sob supervisão técnica qualificada. Essa simbiose entre inteligência artificial e expertise humana assegurou que a plataforma não apenas gerasse *insights* rápidos, mas que fossem tecnicamente robustos e acionáveis.

4.4 Evolução da Maturidade em Indústria 4.0 e Manufatura Verde

A implementação do projeto marcou uma evolução significativa nos níveis de maturidade digital da unidade fabril segundo os critérios da Acatech. O processo de gestão de dados industriais migrou de um estado de conectividade fragmentada (Nível 2/3) para o estágio de Transparência plena (Nível 4), onde a organização passou a compreender as causas fundamentais dos eventos em tempo real. Além disso, a validação de algoritmos de correlação preparou o terreno para o Nível 5 (Capacidade Preditiva), permitindo antecipar falhas antes que estas resultem em interrupções na linha.

Este avanço tecnológico possui um impacto direto no conceito de "Manufatura Verde". Através da análise inteligente e ágil, houve uma redução drástica na necessidade de retrabalhos físicos e na geração de resíduos eletrônicos provenientes de lotes defeituosos. A otimização dos recursos industriais e a diminuição do desperdício de insumos sensíveis reforçam o compromisso com a sustentabilidade regional, demonstrando que a alta tecnologia pode ser um motor para o desenvolvimento econômico e ambiental equilibrado na Amazônia Ocidental.

5 CONCLUSÃO/CONSIDERAÇÕES FINAIS

A implementação do projeto Cyber Nexus demonstrou que a integração de Modelos de Linguagem de Grande Escala (LLMs) em sistemas de manufatura de alta complexidade é uma estratégia eficaz para superar a fragmentação de dados em silos protegidos. O estudo comprovou que a utilização de uma arquitetura de agentes inteligentes permite transformar processos de diagnóstico anteriormente manuais e morosos em fluxos de trabalho automatizados e de alta performance. Ao levar a inteligência para a borda da rede fabril, a solução não apenas resolveu o problema da latência de acesso, como também garantiu que a análise técnica fosse realizada com base em evidências em tempo real, preservando a soberania e o sigilo dos dados industriais.

Os ganhos operacionais quantificados durante a fase de validação em ambiente relevante (TRL 6) revelaram-se disruptivos para o padrão atual da planta fabril. A redução do ciclo de Análise de Causa Raiz (RCA) de uma média de 15 dias para menos de 48 horas representa uma melhoria superior a 85% na agilidade de resposta a falhas de produção. Aliado a uma taxa de precisão diagnóstica superior a 80%, este resultado valida a robustez do motor algorítmico e a sua capacidade de fornecer recomendações acionáveis que impactam diretamente na Eficiência Global dos Equipamentos (OEE) e na competitividade da unidade em Manaus.

No que tange à maturidade digital, o projeto logrou êxito em conduzir o processo de gestão de dados para o Estágio 4 (Transparência) do índice Acatech, estabelecendo simultaneamente as competências técnicas para o alcance da Capacidade Preditiva (Estágio 5). A plataforma permitiu que a organização deixasse de apenas observar eventos isolados para compreender as correlações semânticas profundas entre variáveis de calibração e falhas lógicas. Este avanço é fundamental para a consolidação de uma cultura orientada a dados, onde a transparência operacional serve como base para a inovação contínua e a otimização de recursos.

Adicionalmente, o Cyber Nexus consolidou-se como um facilitador para a "Manufatura Verde" dentro do Polo Industrial de Manaus. Ao acelerar a identificação de causas raízes, o sistema reduziu drasticamente a ocorrência de lotes defeituosos e a necessidade de retrabalhos físicos, o que diminui o desperdício de insumos eletrônicos sensíveis e o consumo energético associado a ciclos de produção ineficientes. Esta dimensão sustentável reforça o papel da tecnologia de ponta não apenas como um vetor de lucro, mas como uma ferramenta de responsabilidade ambiental e social no ecossistema da Amazônia Ocidental.

Para trabalhos futuros, recomenda-se a expansão das capacidades do sistema para o suporte à decisão autônoma (Nível 6 de maturidade), integrando mecanismos de ajuste fino de parâmetros de produção em tempo real através de sistemas ciber-físicos (CPS). Sugere-se também o desenvolvimento de modelos de linguagem ainda mais especializados em hardware de rádio frequência para aumentar a acurácia em diagnósticos de falhas de sinal complexas. Em suma, o Cyber Nexus encerra esta etapa de P&D como uma solução validada e escalável, pronta para ser replicada em outros setores industriais que enfrentam desafios similares de complexidade e segurança de dados na era da Indústria 4.0.

REFERÊNCIAS

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF, ago. 2018.

DOCKER. **Docker Documentation.** 2025. Disponível em: <https://docs.docker.com/>. Acesso em: 10 mar. 2026.

FASTAPI. **FastAPI Documentation.** 2025. Disponível em: <https://fastapi.tiangolo.com/>. Acesso em: 10 mar. 2026.

LANGCHAIN. **LangChain Documentation**. 2025. Disponível em: <https://python.langchain.com/>. Acesso em: 10 mar. 2026.

MARITACA AI. **Maritaca AI: Modelos de Linguagem para o Contexto Brasileiro**. 2025. Disponível em: <https://www.maritaca.ai/>. Acesso em: 12 mar. 2026.

MDPI. **RCEGen: A Generative Approach for Automated Root Cause Analysis Using Large Language Models (LLMs)**. MDPI Electronics, v. 4, n. 4, p. 29, 2024. Disponível em: <https://www.mdpi.com/2674-113X/4/4/29>. Acesso em: 14 mar. 2026.

SCHUH, G. et al. **Industrie 4.0 Maturity Index. Managing the Digital Transformation of Companies**. Acatech Study, 2020. Disponível em: <https://www.acatech.de/publikation/industrie-4-0-maturity-index-update-2020/>. Acesso em: 15 mar. 2026.

TAO, F. et al. **Digital Twin-driven smart manufacturing**. Academic Press, 2019.

WANG, J. et al. **LLM-driven root cause analysis for industrial cyber-physical systems**. Journal of Manufacturing Systems, 2024.

ZHOU, J. et al. **Smart Manufacturing: Current Status and Future Perspectives**. Engineering, v. 4, n. 3, p. 408-420, 2018.