

Applying AI in Supporting Additive Manufacturing Machine Maintenance

An Information Architecture

*Luiz F. C. S. Durão,
Florian Schmitt*,
Júlio César R. A. Paz,
Denilton Donizetti Silva,
Eduardo Zancul and
Eckhard Kirchner*

Based on recent Artificial Intelligence advancements, this paper addresses implementing an AI-based maintenance strategy within a controlled production environment. The architecture employs real-time machine monitoring. Data is first processed locally for early fault detection and later cloud-based to support predictive maintenance functions. A Large Language Model trained with domain-specific knowledge provides operators with basic instructions to manage repetitive faults. The architecture is implemented and tested as a source of experimental data.

Introduction

With the industry's expansion on applying big data and machine learning, digitalized manufacturing has generated increasing volumes of real-time data containing valuable information about machine and process conditions [1]. This data can be processed to implement predictive maintenance and optimize operations, improving productivity and ensuring workplace safety [2–4].

The application of Artificial Intelligence (AI) to support human-operated

machine maintenance in manufacturing is an emerging field with the potential to transform the industry. AI technologies, such as Large Language Models (LLMs), have enabled more effective interactions between humans and machines, playing an essential role in their operation, maintenance, and training [5–7]. LLMs, like GPT-3 [8] and GPT-4 [9], are recent technologies that allow more natural and efficient interaction between humans and machines. These AIs can assist in operation, maintenance, and training, offering personalized and contextualized support

to operators. However, significant data protection and integrity challenges remain, especially in industrial settings.

This paper discusses an AI-assisted information architecture focused on supporting the maintenance of an additive manufacturing machine in a learning factory environment. The proposed model uses sensors for real-time monitoring and a gateway for local and cloud processing, facilitating fault detection and supporting predictive maintenance, assisted by an LLM trained with additive manufacturing contextual information. The study addresses the implementation of a secure architecture, combining AI and VPN networks to maintain data privacy and integrity in industrial environments. By integrating emerging technologies, such as intelligent virtual assistants and secure networks, this approach seeks to validate a robust architecture that is practical and innovative with applications for the smart manufacturing industry.

The implemented framework is characterized by remarkable applicability and adaptability across various manufacturing machines by adjusting to and incorporating each machine type's spe-

* Corresponding author

Florian Schmitt, M. Sc.; Technical University Darmstadt, Institute for Product Development and Machine Elements (pmd); Otto-Berndt-Straße 2, 64287 Darmstadt; Phone: +49 (0) 6151 16-21170, E-Mail: office@pmd.tu-darmstadt.de

Co-authors

Luís F. C. S. Durão, Dr.; University of São Paulo, Production Engineering Department
Júlio César R. A. Paz; University of São Paulo, Production Engineering Department
Denilton Donizetti Silva; University of São Paulo, Production Engineering Department
Prof. Dr. Eduardo Zancul; University of São Paulo, Production Engineering Department
Eckhard Kirchner, Prof. Dr.-Ing.; Institute for Product Development and Machine Elements (pmd)

Note

This article is peer reviewed by the members of the ZWF Special Issue Advisory Board

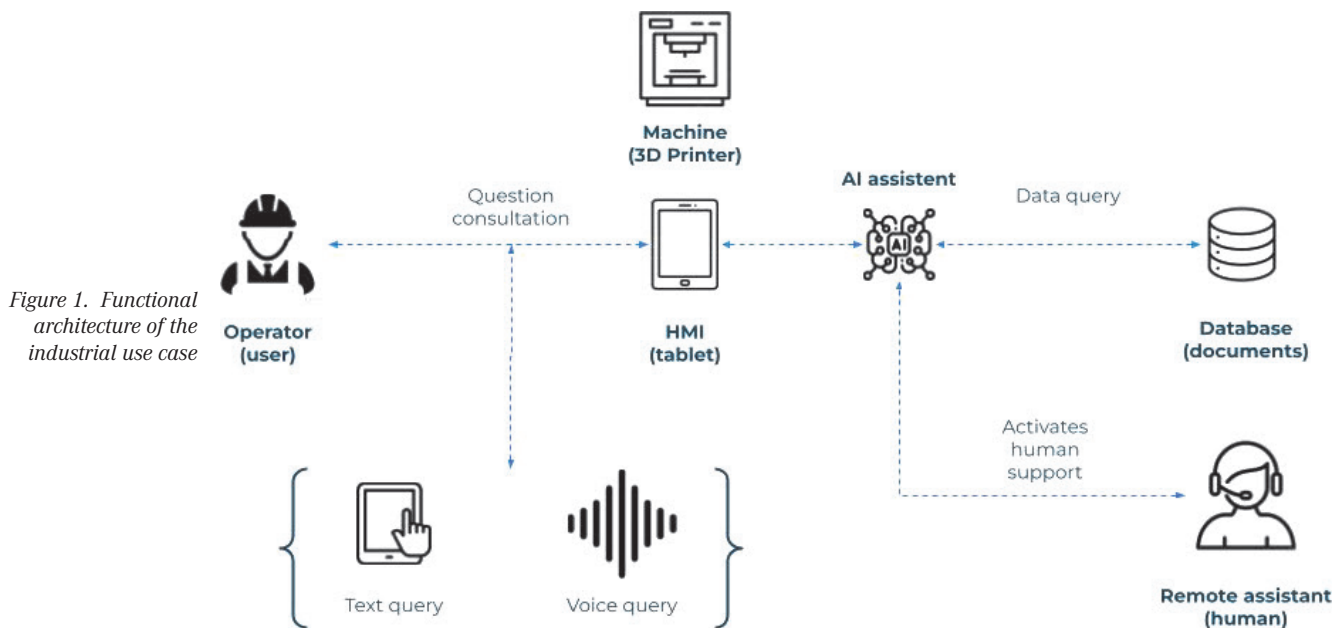


Figure 1. Functional architecture of the industrial use case

cific properties and processes. For instance, in the case of a CNC milling machine, the framework can be tailored to monitor spindle speed, tool wear, and precision tolerances, while in a robotic assembly line, it could focus on motor torque, alignment accuracy, and cycle times. This adaptability ensures that the framework can optimize secure manufacturing strategies by integrating machine-specific metrics into its monitoring and predictive maintenance algorithms, generally regardless of the origin of the specific optimization problem.

By leveraging secure data extraction methods through a VPD network and gateway, the framework ensures robust data security while enabling real-time data collection from multiple machine components and additionally integrated sensors. For example, in a smart factory producing automotive components, the AI could analyze vibration data and thermal signatures to detect anomalies, reducing risks of unplanned downtime. This secure and adaptable approach fosters resilience, enhances efficiency, and supports sustainable manufacturing operations across industries.

Architecture

The primary objective of the technological architecture is illustrated in Figure 1, high-

lighting several essential characteristics of assistive AI applications for machine operation. Traditionally, the Human-Machine Interface (HMI) is implemented as a screen-based device due to the operational conditions of machinery [10]. However, one of the critical innovations introduced by natural language AI is the voice communication capability, which further simplifies interaction within industrial HMIs [11]. Another essential feature is the AI's interpretative capability regarding a database (typically cloud-based) that can be continually updated, constraining the AI's response sample field and reducing the risk of erroneous outputs. Additionally, the architecture supports the AI's ability to execute varied interpretations and com-

mands, such as initiating human technical assistance when the user requests.

Current architectures for assistive Artificial Intelligence (AI) applications often adopt the model represented in Schematic A of Figure 2. In this architectural model, a web-based service sends requests via APIs to an AI engine hosted in a cloud computing environment. The web-based service is responsible for the interface and communication management, and the AI engine operates within cloud infrastructures, facilitating scalability and remote access [2–4]. However, this configuration presents significant challenges when applied to industrial networks, particularly concerning data security, latency, and reliability [12].

AI-Based Maintenance Strategies for Platforms

AI-based maintenance strategies for platforms, specifically an FDM printer, is designed for an exemplary manufacturing machine to enhance operational efficiency, reduce machine downtime, and enhance maintenance quality. By adapting to various machine properties and processes in the form of expert knowledge and real-time sensor data, the platform can be adapted to multiple manufacturing machines, exceeding the exemplary application in the additive manufacturing context. It facilitates maintenance by providing expert knowledge to operators through a human-machine interface. The AI framework securely extracts data from the manufacturing line via a VPD Network gateway, performs cloud-based analysis, and proactively informs operators of deviations to prevent downtime.

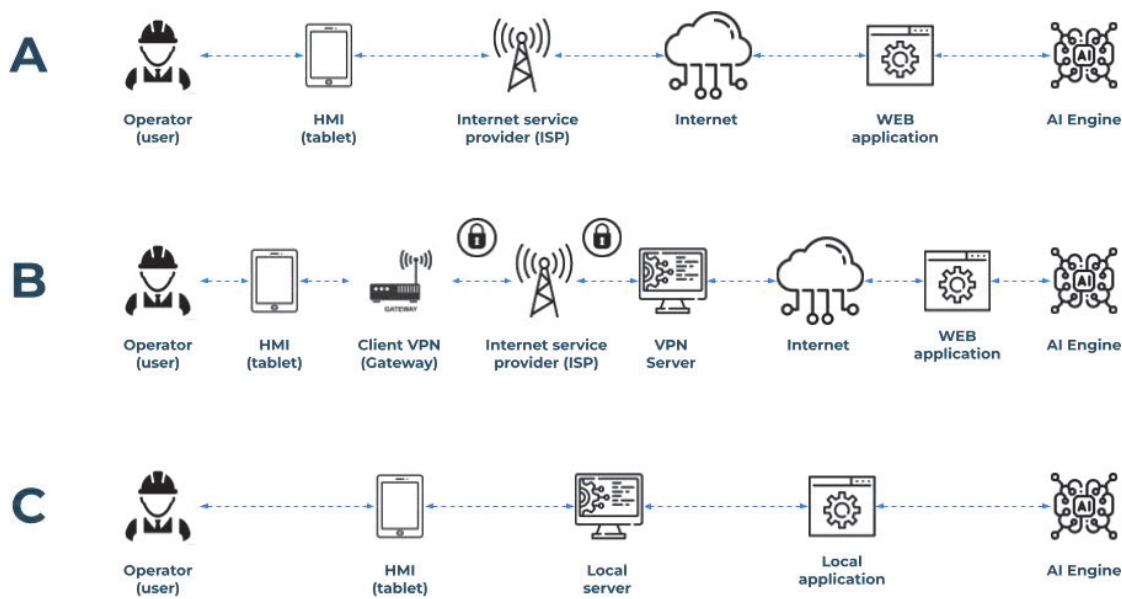


Figure 2. Schematic model architectures

One of the main challenges to overcome is data security. Protecting sensitive information is crucial in industrial environments, as transmitting data over the Internet to cloud services can expose the system to vulnerabilities and cyber-attacks. Additionally, recurring operational costs and dependence on cloud infrastructure can pose significant obstacles to large-scale adoption of this architecture in the industrial sector.

To address these issues, VPNs are necessary to ensure the security and privacy of transmitted data. Architectures B and C, illustrated in Figure 2, present alternative solutions. In Architecture B, a VPN creates a secure communication tunnel between system components, allowing the AI engine to remain in the cloud while protecting the entire system. In Architecture C, implementing a local LLM eliminates the need for external cloud

communication, enhances security, and reduces latency, though it limits response capacity and application scalability. This configuration enables assistive chats to operate more efficiently and securely in industrial environments.

Table 1 compares the different architectures discussed, highlighting their main characteristics and suitability for various industrial scenarios.

Implementation

The implementation described was evaluated within the controlled environment of the Factory of the Future Lab, a Learning Factory at the University of São Paulo (USP). In this context, latency was not a critical parameter. Thus, the experimental focus was assessing the feasibility of different application scenarios and their scalability potential. Consequently, Ar-

chitecture B (Figure 2) was selected for deployment due to its balanced approach to security, cost efficiency, and operational scalability. The implemented system architecture is illustrated in Figure 3.

The architecture is currently being implemented on a Fused Deposition Modeling (FDM) Additive Manufacturing machine, which serves as a source of experimental data and a demonstrator of the proposed maintenance strategy. Given that FDM processes are characterized by high variability across several parameters, finding an optimal configuration presents a complex and challenging scenario [13].

The technology was implemented by integrating a tablet as the main interface, a gateway equipped with a 5G sim card, and OpenAI's AI engine. A voltage and current sensor were attached to the additive manufacturing machine, allowing the AI system to detect and report electrical issues in real time by comparing them with the training dataset. The tablet is the primary interaction between the human operator and the assistive system. It provides a Human-Machine Interface for monitoring and controlling the machine. By interacting with the assistive chat, operators can receive immediate notifications about electrical failures or power interruptions and obtain guidance on operational and maintenance procedures.

Based on Schematic B (Figure 2), the implemented architecture uses a Virtual Pri-

Characteristic	Architecture A	Architecture B	Architecture C
Scalability	High	Medium	Low
Security	Low	Medium	Medium
Latency	–	–	Low
Operational Costs	On demand	On demand	Low
Infrastructure Dependency	Low	Low	Local dependent

Table 1. Comparison of alternative architectures

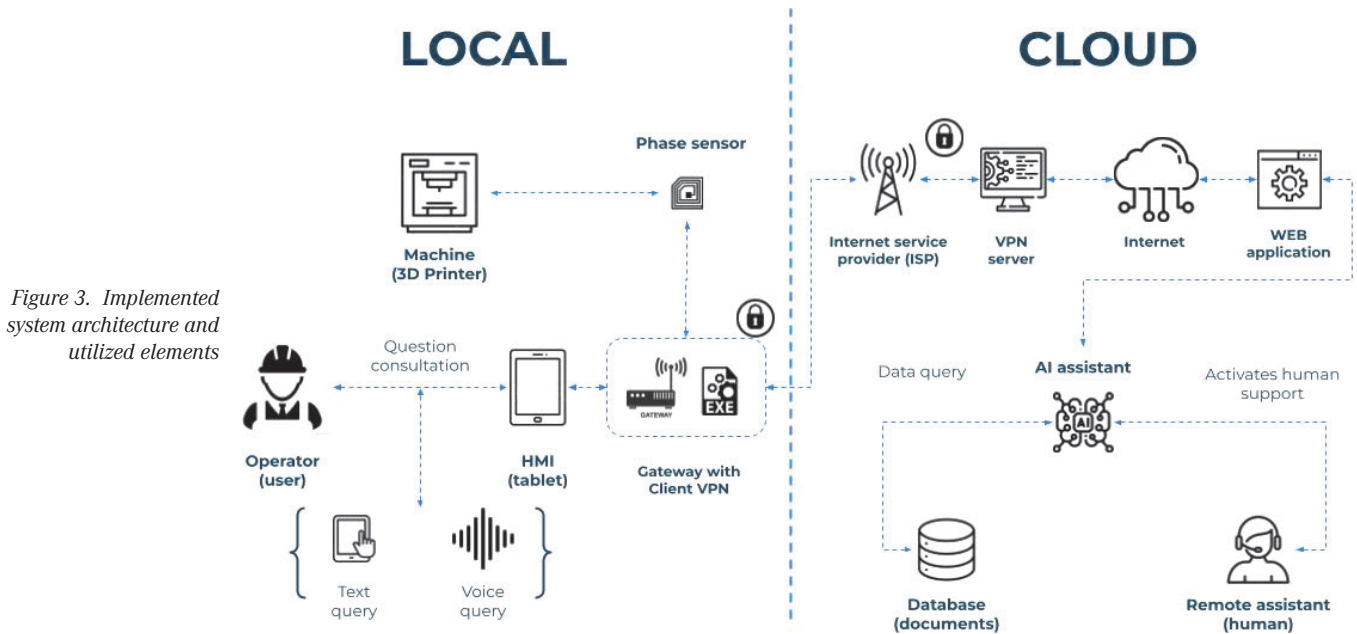


Figure 3. Implemented system architecture and utilized elements

vate Network (VPN) to ensure secure communication between local devices and the cloud-based AI engine. The gateway collects data from the voltage and current sensors and transmits it through the VPN to the AI system in the cloud. This configuration allows the AI to analyze real-time data and promptly alert the operator about detected anomalies, enhancing operational efficiency and equipment safety. The VPN addresses security concerns associated with industrial networks, creating a secure tunnel for transmitting sensitive data [14]. Additionally, by maintaining the AI engine in the cloud and ensuring secure communication, the system leverages cloud ser-

vices' computational power and scalability without compromising data integrity.

The role of the VPN in data security is fundamental in this context. A VPN protects sensitive data transmitted between the machine, the gateway, and the cloud-based AI engine against interception and unauthorized access. The VPN creates an encrypted tunnel for data transmission, guaranteeing the confidentiality and integrity of information [14]. This is especially critical in industrial environments, where data security is essential to prevent cyber threats and protect intellectual property. By implementing the VPN, the system architecture meets the neces-

sary security requirements to operate reliably and securely within an industrial setting. Figure 4 shows the solution implementation in the selected machine.

The technology was integrated and tested using the OpenAI AI engine to process received data and generate query responses. Initially, we extracted the content from the technical manual and reformulated significant portions of the material, organizing it into a structured dataset in FAQs (frequently asked questions). During this process, we emphasized critical information on operational procedures, such as startup and shutdown, and maintenance guidelines, in-



Figure 4. Photos of the local installation

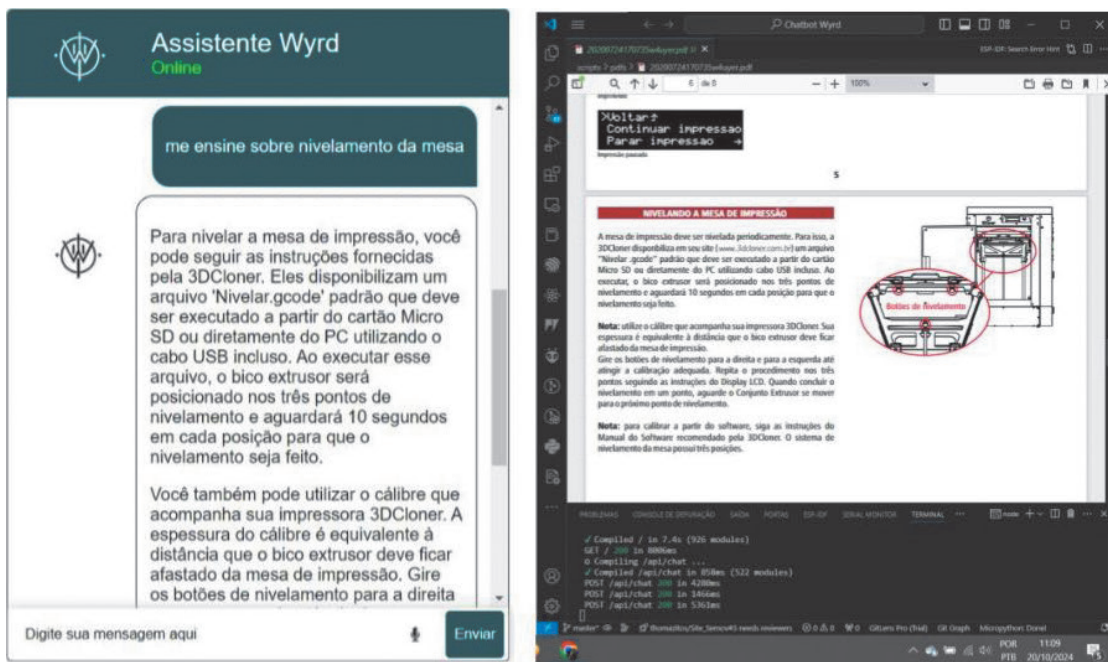


Figure 5. AI-assisted chat. The left figure represents a question made by the user and the answer given by the system (The right figure illustrates the training method proposed by the system)

cluding troubleshooting for operational failures or interruptions during functioning, such as a printing stoppage.

Subsequently, we consolidated these two documents into a specialized database to enhance efficient responses to technical queries. This approach resulted in the ChatGPT model delivering more precise and contextually accurate responses, effectively addressing specific technical demands. Various query scenarios were conducted during troubleshooting to validate the system's ability to identify relevant content within technical materials such as manuals and datasheets. Figure 5 shows the output of the chat.

Results and Conclusions

Implementing a technical assistant copilot for operating and maintaining an additive manufacturing machine and utilizing a VPN for secure data transmission proved practical and innovative. By integrating a gateway equipped with a 5G sim card and connected to a voltage and current sensor, the virtual assistant can monitor the machine's status in real time, identifying electrical issues and providing operational support. The combination of the tablet interface, the gateway with a sensor, and advanced AI constitutes an innovative solution that

enhances human-machine interaction within the industrial context while validating an architecture that offers elevated levels of security and reliability.

A key innovation of this project was transforming the assistive chat from a passive helper into an active agent. The assistant anticipates the operators' needs and takes action by proactively notifying them of detected electrical issues, potentially reducing downtime and preventing possible failures. Moreover, a VPN ensures secure communication between devices and the virtual assistant, allowing private networks to be used within the industrial environment without compromising data security.

The results highlight the substantial potential of integrating emerging technologies, such as intelligent virtual assistants and secure networks, in industrial applications. This approach can be extended to other equipment and processes, fostering more effective human-machine interaction. Future research may further explore expanding the assistant's functionalities, incorporating additional sensor parameters, and optimizing human-machine interaction to improve operational efficiency and safety.

Literature

- Gao, R.X.; Wang, L.; Helu, M.; Teti, R.: Big Data Analytics for Smart Factories of

- the Future. CIRP Annals 69 (2020) 2, pp. 668–692
DOI:10.1016/j.cirp.2020.05.002
- Fett, M.; Turner, E.; Breimann, R.; Kirchner, E.: Extension of the System Boundary of the Digital Twin onto the Sensors of the Physical Twin through the Introduction of Redundant Soft Sensors. Forschung im Ingenieurwesen/Engineering Research 87 (2023) 1, pp. 479–488
DOI:10.1007/s10010-023-00653-y
- Becker-Dombrowsky, F. M.; Koplin, O. S.; Kirchner, E.: Individual Feature Selection of Rolling Bearing Impedance Signals for Early Failure Detection. Lubricants 11 (2023) 7
DOI:10.3390/lubricants11070304
- Bienefeld, C.; Becker-Dombrowsky, F. M.; Shatri, E.; Kirchner, E.: Investigation of Feature Engineering Methods for Domain-Knowledge-Assisted Bearing Fault Diagnosis. Entropy 25 (2023) 9
DOI:10.3390/e25091278
- Gao, R.X.; Krüger, J.; Merklein, M. et al.: Artificial Intelligence in Manufacturing: State of the Art, Perspectives, and Future Directions. CIRP Annals 73 (2024) 2, pp. 723–749
DOI:10.1016/j.cirp.2024.04.101
- Kusiak, A.: Generative Artificial Intelligence in Smart Manufacturing. Journal of Intelligent Manufacturing 36 (2024) 5
DOI:10.1007/s10845-024-02480-6
- Liu, A.; Lu, S.; Tao, F.; Anwer, N.: Integration of Data Science with Product Design towards Data-driven Design. CIRP Annals 73 (2024), pp. 509–532
DOI:10.1016/j.cirp.2024.06.003

8. Brown, T.B.; Mann, B.; Ryder, N. et al. Language Models are Few-Shot Learners. 2020 DOI:10.48550/arXiv.2005.14165
9. OpenAI, Achiam, J.; Adler, S. et al. (eds.): GPT-4 Technical Report. 2023 DOI:10.48550/arXiv.2303.08774
10. Romero, D.; Bernus, P.; Noran, O. et al.: The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation towards Human-Automation Symbiosis Work Systems. In: Nääs, I.; Vendrametto, O.; Mendes Reis, J. et al. (eds.): Advances in Production Management Systems. Initiatives for a Sustainable World. Springer, Cham 2016, pp. 677–686 DOI:10.1007/978-3-319-51133-7_80
11. Zhou, L.; Gao, J.; Li, D.; Shum, H. Y.: The Design and Implementation of Xiaoice, an Empathetic Social Chatbot. Computational Linguistics 46 (2020) 1, pp. 53–93 DOI:10.1162/COLLA_00368
12. Stouffer, K.; Pease, M.; Tang, C. et al.: Guide to Operational Technology (OT) Security. United States National Institute of Standards and Technology, NIST Special Publication 800–82r3, 2023 DOI:10.6028/NIST.SP.800-82r3
13. Durão, L. F. C. S.; Barkoczy, R.; Zancul, E. et al.: Optimizing Additive Manufacturing Parameters for the Fused Deposition Modeling Technology Using a Design of Experiments. Progress in Additive Manufacturing 4 (2019) 3, pp. 291–313 DOI:10.1007/s40964-019-00075-9
14. Stallings, W.: Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloudtext. Pearson International, Indianapolis 2015

Authors

Luiz Fernando Cardoso dos Santos Durão was born in 1992. He holds a Doctorate in Industrial Engineering from the University of São Paulo (2020), where he also earned his Master of Science (2017) and Bachelor's degree in Engineering (2014). His doctoral research focused on the concept and application of the „Closed-loop Digital Twin.“ His professional career includes various leadership and academic roles. Since 2023, he has served as the CEO of Delta Goal, a startup that uses AI to analyze football

players in São Paulo, Brazil. Additionally, he is a lecturer at Insper, São Paulo, where he has been teaching since 2019, covering topics such as innovation management and data science.

Florian Schmitt, M.Sc., born in 1992, completed his Master of Science in General Mechatronics at TU Darmstadt in 2020, with part of his studies conducted at the University of São Paulo. He is pursuing a doctorate in methodological product development and modeling, focusing on quantitative working space models, at TU Darmstadt. Portions of his doctoral research are conducted within a collaborative project at the University of São Paulo, examining the effectiveness and acceptance of new modeling approaches. Since December 2020, he has been a research associate at the Institute for Product Development and Machine Elements at TU Darmstadt.

Júlio César R. A. Paz is a mechatronics engineering undergraduate student and a junior researcher at the Fábrica do Futuro learning factory at the University of São Paulo (USP).

Denilton Donizetti Silva is a technical sales executive at V2COM WEG in São Paulo, focusing on industry connectivity solutions. He has long collaborated on research initiatives at the Fábrica do Futuro learning factory at the University of São Paulo (USP).

Eduardo Zancul is an Associate Professor at the School of Engineering at the University of São Paulo (USP). He earned his Ph.D. in Industrial Engineering at USP. Before joining USP as a scholar, he worked as a management consultant. His main research interests are product development process management, advanced manufacturing, and design education.

Prof. Dr.-Ing. Eckhard Kirchner, born in 1969, completed his Mechanical Engineering and Applied Mechanics studies in 1995 at the Technical University (TH) of Darmstadt and the Norwegian Institute of Technology in Trondheim. He earned his doctorate (Dr.-Ing.) at TU Darmstadt in 1999. Following his industrial career in the gearbox and drivetrain development at Adam Opel AG, GM Powertrain, Schaeffler AG, and Siemens AG, he was appointed a Professor for Product Development and Machine Elements at TU Darmstadt in 2016.

His research focuses on sensory machine elements, the associated product development methods, and the design of components for additive manufacturing. Prof. Kirchner has

been a full member of the Scientific Society for Product Development (WiGeP) since 2019 and was named spokesperson for the Society for Machine Elements and Systems in 2022.

Abstract

Einsatz von KI zur Unterstützung der Wartung von additiven Fertigungsmaschinen – Eine Informationsarchitektur. Basierend auf den jüngsten Fortschritten im Bereich der Künstlichen Intelligenz adressiert dieser Beitrag die Implementierung einer KI-gestützten Wartungsstrategie in einer kontrollierten Produktionsumgebung. Die Architektur nutzt einen Echtzeitfähigen Überwachungsansatz. Zunächst werden die Daten lokal verarbeitet, um frühe Fehlererkennungen zu ermöglichen, und später in einer Cloud verarbeitet, um prädiktive Wartungsfunktionen zu unterstützen. Ein mit domänenspezifischem Wissen trainiertes Large Language Model stellt den Bedienern grundlegende Anweisungen zur Bewältigung wiederkehrender Fehler bereit. Die Architektur wird implementiert und getestet, um experimentelle Daten zu generieren.

Acknowledgements

The authors thank CAPES (process 88881.964206/2024-01) for supporting this research.

Keywords

Condition Monitoring, Predictive Maintenance, Human Machine Interaction, Data Security, Large Language Models

Schlüsselwörter

Zustandsüberwachung, Prädiktive Wartung, Mensch-Maschine-Interaktion, Datensicherheit, Large Language Models

Bibliography

DOI:10.1515/zwf-2024-0169
ZWF 120 (2025) Special Issue; page 208 – 213
Open Access. © 2025 bei den Autoren, publiziert von De Gruyter. 
Dieses Werk ist lizenziert unter der Creative Commons Namensnennung 4.0 International Lizenz.
ISSN 0947-0085 · e-ISSN 2511-0896