Electronic Markets (2025) 35:14
https://doi.org/10.1007/s12525-025-00766-y

RESEARCH PAPER

Automating quality control through an expert system

Giorgio Scarton[1] · Marco Formentini[1] · Pietro Romano[2]

Received: 23 July 2024 / Accepted: 29 January 2025

Abstract

In this article, we present findings from an interventional study conducted within a small enterprise in northern Italy, focused on automating quality control in press-in operation for the production of reduction gearboxes. Guided by Organizational Information Processing Theory, we developed an expert system to automate quality control and facilitate early fault detection. This novel approach enhances quality control within this production stage and could potentially impact other levels of the supply chain. We contribute to the theory by providing a revised version of the Organizational Information Processing Theory framework which integrates technological advancements and variability of the task over time as critical factors affecting information processing, and shows the iterative nature of the digitalization process in SMEs. Operationally, the solution increases defect identification from 6% at end-of-line to 15% through step-by-step checks. It provides a cost-effective, practical example of AI-driven quality control, advocating for data-driven decision-making demonstrating a scalable pathway for SMEs to adopt AI with limited resources.

Keywords Automation · Artificial intelligence · Quality control · Expert system · Digital supply chain · Industry 4.0

JEL Classification O300

Introduction

The supply chain digitalization literature highlights the key role played by digital technologies in reshaping supply chain business process and in turn improving performance (Perano et al., 2023). Recently, the combination of technologies such as AI and IoT has demonstrated to be able to impact supply chains in different ways (Culot et al., 2024). For example, Goodarzian et al. (2023) propose a vaccine supply chain model that integrates IoT and AI to optimize the distribution, allocation, and location of vaccination and treatment centers. Moreover, in the context of predictive maintenance, Ayvaz and Alpay (2021) describe the integration of IoT sensors in manufacturing equipment to enable real-time monitoring of critical parameters. AI is then used to analyze this data to identify patterns and anomalies that indicate potential failures preventing disruption of the supply chain related to delays in suppliers production (Brintrup et al., 2020). In the context of automotive quality control, Cardellicchio et al. (2024) show how lightweight machine learning (ML) models can be used to detect anomalies and damage on the weld surface, allowing human operators to intervene promptly, reducing costs and delays in the production line, and ensuring that only high-quality products are shipped.

The automation of quality control represents an important area of investigation in the supply chain digitalization literature, with an increasing attention paid to the integration of AI-based solutions (Cardellicchio et al., 2024; Escobar & Morales-Menendez, 2018; Peres et al., 2019). Recently, Culot et al. (2024) reviewed the literature on AI adoption in supply chain management, discussing the benefits provided by AI and integrated technologies such as IoT for quality control and automation in production processes. Moreover,

Responsible Editor: Andrea Patrucco

✉ Giorgio Scarton
giorgio.scarton@unitn.it

Marco Formentini
marco.formentini@unitn.it

Pietro Romano
pietro.romano@uniud.it

[1] Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy

[2] Polytechnic Department of Engineering and Architecture (DPIA), University of Udine, Udine, Italy

Perano et al. (2023) underlined the contribution offered by AI in improving product/service quality and in turn increase revenues for operations planning.

Indeed, quality control is pivotal in realizing Industry 4.0's (I4.0) potential, driving advancements in manufacturing, and impacting the entire supply chain (Escobar & Morales-Menendez, 2018). In fact, traditional statistical techniques commonly adopted in quality control provide limited system-level insights, especially with stage interdependencies (Arinez et al., 2020). By contrast, automated quality control using AI and IoT has the potential to enable real-time issue identification, reducing errors across supply chain stages and optimizing operations from suppliers to buyers (Cruz Guerrero et al., 2019; Hamm et al., 2023).

However, implementing advanced I4.0 technologies requires expertise and substantial financial investments (Rüßmann et al., 2015), resources often lacking in small and medium-sized enterprises (SMEs) (Hansen & Bøgh, 2021). This limitation often restricts their implementation to a subset of technologies such as Cloud Computing and IoT (Moeuf et al., 2018). Moreover, the lack of a comprehensive data management strategy often results in the collection of data that are not immediately applicable for advanced solutions, thus limiting the extraction of value and in turn requiring further data analysis. Fitting the information processing needs with corresponding capacities has demonstrated to be challenging for SMEs, which are still in the process of understanding how to effectively implement AI in their processes. While pathways for digitalization of SMEs are discussed in the literature (Battistoni et al., 2023), there is a lack of tailored guidelines for AI implementation, which presents several specificities (Dubey et al., 2020; Fosso Wamba, 2022). This leads us to draw the following research questions:

> *RQ1: To what extent is it possible to develop an automated quality control solution in a real manufacturing setting?*
> *RQ2: What are the implications both at the organizational and theoretical level of such implementation?*

To address these questions, we conducted an interventional study (Coughlan & Coghlan, 2002) in a SME based in northern Italy which acts as system integrator. Guided by Organizational Information processing Theory (OIPT) we developed an expert system (ES) for automated quality control and early fault detection based on historical data produced by the assembly of reduction gearboxes. The research project was conducted over one year, from September 2021 to September 2022, involving iterative cycles of data gathering, development, implementation, and validation (see Fig. 1). Data were collected through unstructured interviews, internal documents, and historical database analysis, revealing insights that identified critical issues in the company's processes. The study was organized into two loops, each comprising two iterations, facilitating close collaboration with the company team and real-world validation of solutions.

Our work contributes to the theory by proposing a revised OIPT framework, based on empirical evidence, that illustrates how digitalization in context of quality control automation allows to handle uncertainty derived by high production variability. In turn, this enhances information processing capacities, while also influencing information processing requirements in a cyclical way. At the supply

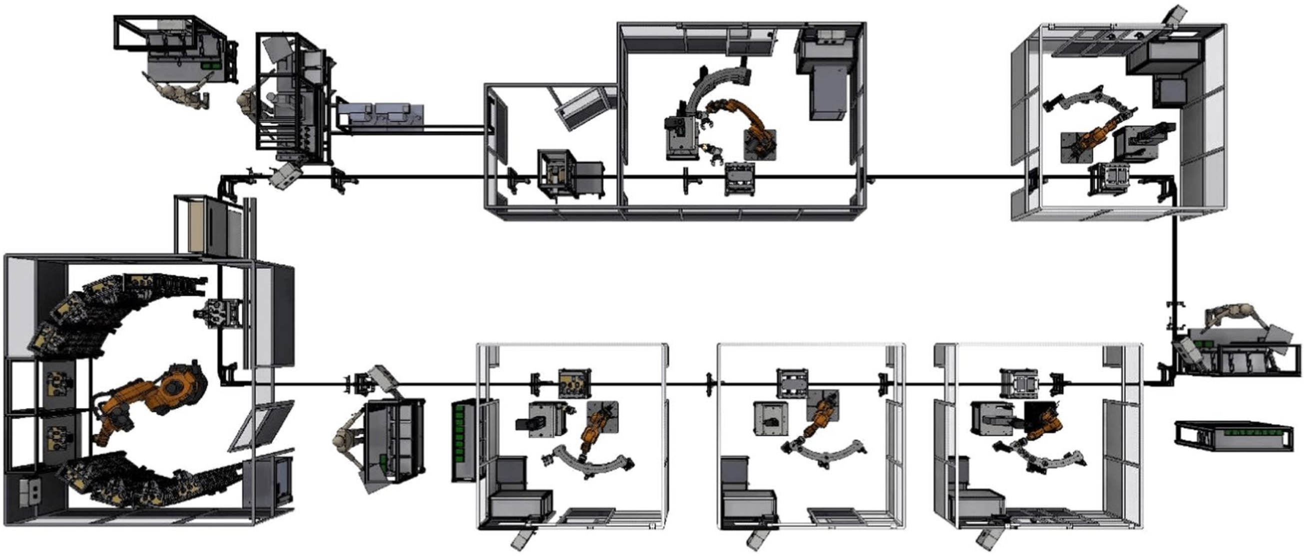


Fig. 1 3D representation of the plant for the assembly of planetary reduction gears (company source)

chain level, our ES facilitates the early identification of raw material quality deviations enhancing overall efficiency and reliability, ensuring higher product quality and more streamlined operations from suppliers to manufacturers and ultimately to buyers.

From an industrial perspective, we present a cost-effective and explainable solution designed to automate quality control within the context of an SME, characterized by limited financial investments and low AI expertise (Hansen & Bøgh, 2021). The notable outcomes achieved by the company include the capacity to identify issues throughout the assembly process, rather than solely at its conclusion. The system also enables pinpointing the specific component responsible for the fault, integrates automated data flow as part of a broader digitalization initiative, and facilitates the exchange of AI knowledge between researchers and the company. In addition, we outline a practical implementation pathway for SMEs, drawing on the work of Battistoni et al. (2023), to facilitate the adoption of advanced quality control systems while addressing common challenges in digitalization.

The reminder of the paper is structured as follows: first, we provide a theoretical background and framing of the work. Next, we describe the research design and the context of the intervention. Then, we describe the development processes and the results, followed by a discussion of their implications. Finally, we present our conclusions about the implementations, including limitations and future developments.

Theoretical background and framing

According to OIPT, an organization as a social system which has to manage work-related uncertainty (Galbraith, 1974; Tushman & Nadler, 1978). For this reason, the organization needs to facilitate collection, gathering and processing of information related to its environment, its performance and quality of outputs (Tushman & Nadler, 1978). OIPT has been used to explain how organizations can effectively manage and utilize data to improve decision-making and operational performance across various contexts. Benzidia et al. (2021) apply OIPT to investigate how AI impacts environmental performance in hospitals. Their study highlights the role of information technology in facilitating efficient data processing, enabling organizations to achieve specific strategic objectives. Expanding the application of OIPT, Dubey et al. (2019) examine how Big Data Analytics Capability (BDAC) fosters trust and collaborative performance between civil and military organizations during disaster relief operations. Their findings emphasize that BDAC enhances information processing capacity, reduces uncertainty, and promotes swift trust and effective collaboration, particularly in complex and uncertain environments. In another study, Srinivasan and Swink (2018) explore the relationship between supply chain visibility, analytics capability, and organizational flexibility. Using OIPT as a theoretical foundation, they demonstrate that enhanced information visibility and lateral organizational relationships significantly improve information processing capacity, ultimately leading to better operational performance. Their research highlights the interconnectedness of information processing, organizational structure, and performance outcomes.

In this study, we adopt OIPT as an appropriate lens to investigate the use of AI and IoT in the automated quality control process. To build on this foundation, we now introduce three core concepts that form the pillars of OIPT: information processing requirements, information processing capacities, and the fit between the two (Tushman & Nadler, 1978).

Information processing requirements in modern quality control

Zhu et al. (2018) defined information processing requirements as the amount of information required for decision making in the context of uncertainty. Uncertainty has been defined by Tushman and Nadler as the difference between information requirements and information possessed, and can be internal or external to the organization.

With the introduction of Industry 4.0 technologies, manufacturing has seen significant advancements in quality, pushing current benchmarks to very high levels (Escobar et al., 2021). In the context of our intervention, i.e., the production of robot precision gearboxes, the error tolerance is very low due to the high precision required for their operations (Li et al., 2024). In this production process, automated assembly lines helps minimize errors but perfect replicability over time is not granted (Li & Li, 2023; Wang et al., 2015), introducing a form of internal uncertainty.

This pushed information processing requirements to an higher level. Traditional end-of-line (EOL) quality control methods, which focus on inspecting finished products, do not provide enough information to endure the proper identification of defects, often missing opportunities for early detection and prevention of defects. Indeed, Arinez et al. (2020) and Turetskyy et al. (2020) emphasize the importance of early-stage quality evaluations, which can identify deviations early in the production process by leveraging a greater amount of data collected along the whole production process. This early detection, as highlighted by Arinez et al. (2020) and Meiners et al. (2021), not only prevents further processing of defective parts, but also reduces rework costs.

Regarding external uncertainty, it is mainly related to the supply chain level. As already described by Zhu et al. (2018), supplier products' quality is unpredictable since it is related to capacities not under the control of the organization. This means that, within supply chain, there is an increasing need for information to handle such uncertainty.

Building on these insights, this study aim to demonstrate how sequential quality evaluation better aligns with current information processing requirements, leading to significant improvements in both product quality and cost-efficiency along the whole supply chain.

Information processing capacities: impact of AI and IoT

Information processing capacities refer instead to the tools and methods employed in decision-making processes. According to Tushman and Nadler (1978), these capacities are influenced by the structure of the organizational unit, which can be either organismic or mechanistic. An organismic structure is particularly suited to managing high levels of uncertainty, as in quality control, due to its informal communication, flexibility, and adaptability to changing conditions (Srinivasan & Swink, 2018). However, such flexibility often leads to higher communication costs and slower decision-making because of the lack of standardized procedures.

To address these challenges, balancing flexibility with formalization through well-defined rules and procedures is essential. Galbraith (1974) proposed coordination and control mechanisms tailored to varying levels of task complexity and uncertainty. For high-complexity tasks like quality control, formal information systems are particularly effective. These systems enable structured data collection, analysis, and dissemination, ensuring consistency in insights and streamlined communication across organizational units; Battistoni et al. (2023) underline the need for formalization also in the context of manufacturing SMEs in the digitalization journey.

As discussed earlier, digitalization has increased both quality and quantity of information required, making traditional statistical methods for quality control inefficient. AI-based algorithms leveraging time-series data have been proposed (Carbery et al., 2018; Peres et al., 2019) to increase processing capacity. For instance, Wang et al. (2019) implemented quality control of work-in-progress products, while Iglesias et al. (2018) used computer vision with 3D cameras to detect imperfections in slate slabs. Despite the advantages of ML, challenges such as the low interpretability of results (Hamm et al., 2023; Wanner et al., 2022) and the need for high-quality data still persist.

To improve interpretability, two approaches has been proposed. On the one side, intervene on the algorithm to increase its interpretability (Guidotti et al., 2019). For example, Senoner et al. (2022) applied explainable AI in quality control in a semiconductor company. On the other side, integrating expert knowledge into AI models can improve both their interpretability and practical relevance. While data-driven approaches have proven effective, the expertise of human remains essential to contextualize AI outputs, although presenting several challenges (Arinez et al., 2020; Bousdekis et al., 2023). For this reason, we aim to explore methods for synthesizing expert insights with data-driven approaches, to enhance model performance and facilitate a more intuitive interpretation of results.

Regarding data sources, obtaining real-world datasets is challenging, costly and time-consuming (Cardellicchio et al., 2024; Tao et al., 2018). Consequently, several studies relied on simulated data and laboratory-based experiments, which lack the complexity and unpredictability of real production environment. This limits the scalability and generalizability of AI models. Therefore, we focus on implementing AI-based quality control solutions directly in real production settings.

Fit between information processing requirements and capacities

The third pillar of OIPT is the fit between an organization's information processing capacities and requirements posed by its operational environment. This principle stresses the need to design processes and structures to avoid inefficiencies that result from either underperformance, due to inadequate data handling, or overperformance, which leads to wasted resources. In our intervention, we initially observed a misalignment between the organization's information processing capacities and the requirements driven by its operational complexity.

As previously discussed, various sources of uncertainty exist within the production process, significantly increasing the need for advanced information processing to capture, analyze, and act promptly. However, the current reliance on EOL testing and statistical quality control methods is insufficient. While these tools are useful for post-production assessments, they lack the real-time data integration and proactive monitoring needed to manage uncertainty effectively across the entire production cycle. This results in a reactive, rather than proactive, approach to quality control, leaving the organization prone to inefficiencies and errors.

To address this, our intervention aimed to enhance information processing by incorporating dynamic and real-time data management solutions, aligning capacities with the complex requirements of the production environment. By improving this fit, the organization will improve management of uncertainty and optimize overall performance.

A key aspect of this process is the collaboration between academia and industry, which is often overlooked. Integrating academic research with industry needs is essential to address real-world challenges effectively. Studies by Hamm et al. (2023), Inoue et al. (2023), and Sharma et al. (2022) demonstrated the value of collaborative approaches such as action research. We aim to leverage these collaborative research frameworks to ensure that the AI-driven quality control systems developed are aligned with industry needs and directly applicable in operational settings.

Table 1 compares previous contributions in the literature, identifies key gaps in AI-driven automated quality control research in manufacturing, and positions the present study by demonstrating how it can address these gaps.

Research methodology and design

To address the aforementioned gaps, we conducted an intervention study (Coughlan & Coghlan, 2002) in collaboration with a system integrator company. The intervention was derived from an action research process, which aimed to solve a real organizational problem while also contributing to scientific knowledge (Shani & Coghlan, 2021). This methodology was chosen for the empirical aspect of our research project since it enables us to address research objectives within their natural setting (Schmidberger et al., 2009) by gathering firsthand information and creating a new system. Adopting it also meant that our findings could be directly implemented by the company, providing practical results more quickly. Additionally, action research has been successfully used in the context of other studies focusing on AI (Liu et al., 2022) and business (Athanasopoulou & De Reuver, 2020; Inoue et al., 2023).

Process and data description

The process under analysis involves the assembly of precision planetary reduction gears for robotics. These reducers are produced through a highly coordinated sequence of sub-assemblies, where pairs of semi-processed components are pressed together. The production plant, depicted in Fig. 1, has a semi-automatic design, combining automatic and manual stations connected via a conveyor belt. Spanning over 40 m the production line includes six automatic cells and four manual cells. Each automatic cell comprises a robotic arm and an automatic press. During the assembly, the arm picks up pieces from the conveyor, places them into the press, and performs the necessary tooling; the press then merges the components together.

In this plant, each gearbox is characterized by five key parameters, resulting in around 250 unique configurations identified by a Configuration IDentifier (CID). These configurations involve around 2500 different components, making it possible for the manufacturer to be required to produce a gearbox with a configuration never seen before and for which historical data exist.

The automatic cells rely on pressing machines equipped with an encoder and a load cell. The encoder captures the

Table 1 Key topics, literature gaps, and positioning of the present study in AI-driven automated quality control in manufacturing

Topic	Previous contributions	Identified gaps	Positioning of the present study
Sequential quality evaluation	Traditional EOL quality control methods (Arinez et al., 2020; Turetskyy et al. 2020)	Limited emphasis on early diagnosis and defect prevention	Our study focuses on sequential quality evaluation and demonstrate how this can lead to product quality improvement and cost reduction
Use of expert knowledge	Data-driven approaches only (Arinez et al., 2020; Bousdekis et al., 2023) ML-based approaches (Carbery et al., 2018; Peres et al., 2019)	Difficulty integrating domain knowledge into AI models and interpreting model results	Our study explores how expert knowledge can be effectively used for model training, and result interpretation
Real-world case implementations	Laboratory case studies and simulations (Cardellicchio et al., 2024; Tao et al., 2018)	Lack of scalable and generalizable implementations in real production environments	Our study highlights the implementation in a real-world case and discuss the challenges and lessons learned during the implementation process
University-industry collaboration	Collaborative research methods (Hamm et al., 2023; Inoue et al., 2023; Shani & Coghlan, 2021; Sharma et al., 2022)	Lack of collaboration on AI-specific applications	Our study focuses on using collaborative research methods such as action research to develop AI solutions tailored to industry needs

displacement position of the press head, while the load cell measures the force applied at each recorded displacement. These measurements generate a series of displacement-force tuples, initially collected by the human machine interface. The data is stored as CSV files in the local storage of the machine and can be accessed remotely through a secure connection.

Figure 2 presents a 2D Cartesian representation of the joining operation, with displacement (in millimeters) along the x-axis and force values (in kilonewtons) along the y-axis. The represented operation can be segmented into three main segments:

- Approach phase (70–125 mm, light grey): the press head advanced without applying force as it moves toward the components.
- Joining phase (125–160 mm, black): the press head engages the components, applying increasing force until reaching the peak, marking the end of the press stroke.
- Return phase (160–15 mm, dark grey): the press head retracts, with the force rapidly decreasing as the press head returns to its initial position.

A critical issue lies in the poor quality of the collected data. Unpredictable errors often leads to duplicated records or the splitting of data into multiple files. As data recording is managed directly through the robot's proprietary software, acting as a black box for the system integrator, there is no way to improve the data collection process. This misalignment between information processing requirements and capacities represents a key challenge to be addressed to successfully implement the proposed quality control mechanism.

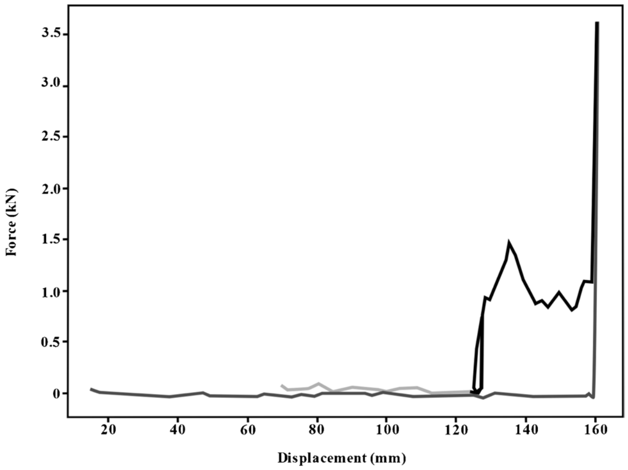


Fig. 2 Sample plot generated by raw data related to one joining operation

Research process

The research project spanned one year, from September 2021 to September 2022. Weekly meetings with the company facilitated information gathering and validation with the chief technical officer (CTO) and the project manager (PM) responsible for the implementation of the AI solution. As depicted in Fig. 3, the research followed iterative cycles of data collection, development, implementation, and validation of the solution. During this period, the research team contributed with their expertise in process engineering and AI, while the company provided in-depth technical knowledge of the plant and product.

The study was structured in two main loops, each comprising two iterative cycles, as depicted in Fig. 3. The first loop focused on the development of the ES, driven by the emergence of new information processing requirements. This phase involved two iterative cycles to refine and address these requirements. The second loop, derived by the information processing requirement emerged through the introduction of AI, centered on the development of an IoT device for data collection. Similarly, this phase included two cycles aimed at aligning information processing capacities with evolving information processing requirements. This approach allowed us to gain comprehensive insights by closely collaborating with the company team and validating our solutions in a real-world scenario.

To inform the development of the solution, multiple data sources were analyzed. The process began with the review of internal documents and 3D models of the plant (see Fig. 1) to understand the production process and identify data collection constraints. Additionally, an unstructured interview with the project manager provided further insights into the plant's highly optimized assembly process, highlighting limited potential for efficiency gains and redirecting the focus toward enhancing quality control.

System reports and discussions revealed that quality control was only conducted at the end of the production line, with no root cause data available for failures. Based on this findings, the project prioritized improving the quality control stage. As the project progressed, expert knowledge on process performance from the company was integrated into the system. Collaboratively, the research team and the CTO documented assembly failure causes and identified problematic production steps. The joint effort not only enabled the company to explore advanced technologies in quality control and AI but also enhanced the research team's understanding of common quality control challenges in manufacturing.

One year after the project concluded, additional unstructured interviews with the project manager and the CTO were conducted to evaluated the solution's impact. This interview approach allowed for flexibility in capturing emerging themes and insights, consistent with the iterative nature of

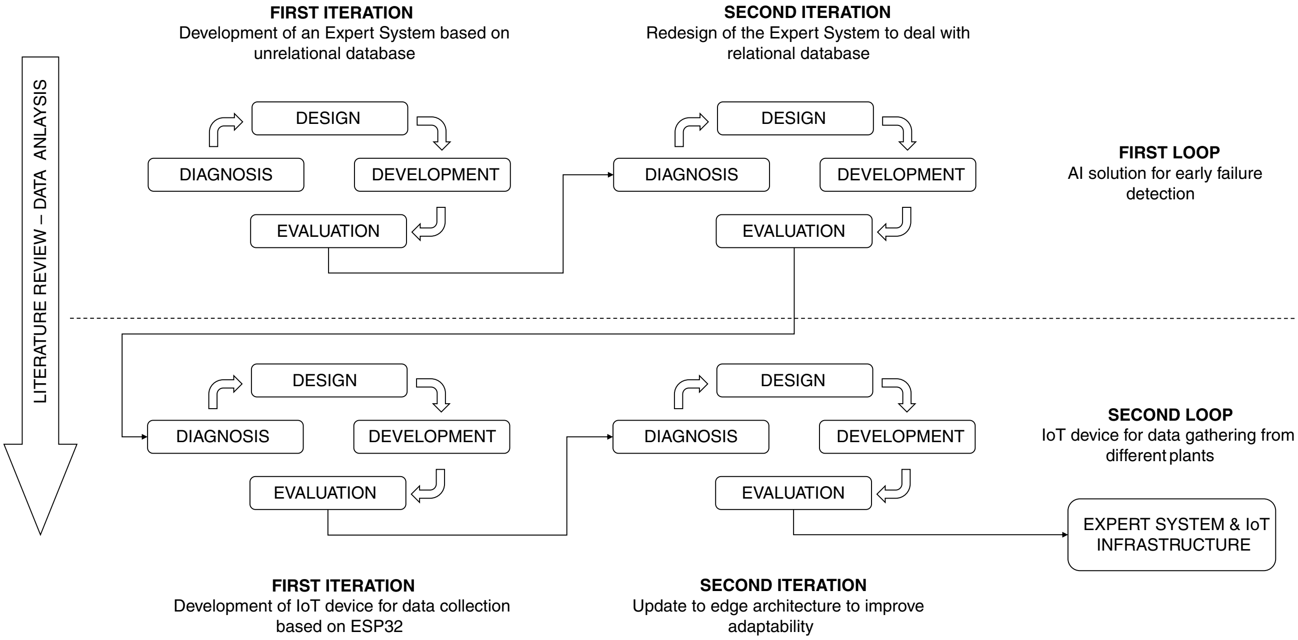


Fig. 3 The action research cyclical process adopted in the study

action research. Through open and focused coding, six thematic codes were developed and reviewed collaboratively, with results validated through participant feedback.

As per Reason and Bradbury's (2001) emphasis on evaluation action research based on "quality" rather than "validity," the intervention was designed to meet the four criteria proposed by Levin (2003). First, our research produced a *workable solution*, now implemented by the company. Second, our project addressed a *real-life problem* by improving quality control. Third, the collaborative approach fostered a shared understanding through joint interpretation of information, resulting in the construction of a *joint meaning*. Finally, the research team active *participation* in all stages of development reflected the close interaction between academia and industry throughout the project.

Development and implementation of the solutions

First loop: ES development and implementation

The analysis of the press-in process is central to this intervention. In traditional methods, quality evaluation relies on force–displacement curves and techniques such as threshold, window, or envelope methods to detect production failures by comparing assembly values with predefined boundaries (Meiners et al., 2021). However, these approaches are not suitable for our objective, as they lack the adaptability required for processes with high variability, as outlined by Tushman and Nadler (1978). In our case, the production process is highly personalized, and product variability is significant, necessitating an adaptive and learning-driven approach. This led us to identify AI as the most suitable technology to address these challenges.

With the decreasing costs of hardware and the increasing accessibility of data-driven techniques, AI has gained substantial attention in industrial applications. Given AI's ability to learn from data, it aligns well with the needs of our study. However, despite the growing interest in ML for monitoring press-in curves, there are still limited applications of these methods in this area (Meiners et al., 2021). For example, Cruz Guerrero et al. (2019) applied ML techniques such as k-nearest neighbors, naïve Bayes, and decision trees to assess press-in quality based on setup parameters like temperature, materials, and surface roughness. Similarly, Fang et al. (2018) utilized a 1D-convolutional neural network to automatically detect key points in press-in curves. These methods, however, rely on manually labeled datasets, which are not always readily available. Meiners et al. (2021) also explored ML models like random forests and naïve Bayes to predict quality from process curves, using data from multiple press-in machines. While these studies showed promise, they share common limitations related to the dependence on labeled data and feature engineering, which may not always be feasible in real-world industrial settings. Furthermore, ML techniques often sacrifice interpretability, which is crucial for diagnosing problems in complex industrial processes.

In contrast, ES offer a valuable alternative in such contexts, particularly where interpretability and rule-based diagnostics are essential. ES rely on explicit rules to capture human expertise (Bohez & Thieravarut, 1997; Hansen & Rieger, 2019), making them well-suited for quality control and fault detection, especially when labeled data or large datasets are unavailable. The interpretability of ES allows for a deeper understanding of the underlying process logic, which is critical in applications like fault detection. Given these advantages, we chose to implement an ES in our intervention.

Throughout the development process, we adopted an iterative approach, holding weekly group discussions and brainstorming sessions with the research team to refine and update the algorithm. These sessions allowed for valuable contributions from all team members and ensured that everyone had a full understanding of the project. In parallel, we maintained a collaborative relationship with the company through weekly roundtable meetings involving managers and technicians. During these meetings, the research team presented progress and asked clarifications on technical aspects. Results were shared with the company for validation, which led to important feedback that helped refine the system. The process was structured in two main iterations, and at the end of the second iteration, we delivered a fully functional version of the system to the company, along with preliminary results for further validation.

Model selection

To develop the ES, it was necessary to identify an appropriate model to represent the process data and facilitate effective evaluation. Given the broad range of products involved and the infeasibility of manual labeling or access to substantial support data, we adopted a methodology based on Lotter's (2006) approach. Lotter utilized a combination of thresholds and windows to assess the quality of press-in operations through analysis of 2D force–displacement curves, a methodology that remains prevalent in classical evaluation systems (Meiners et al., 2021).

The process windows technique involves segmenting the force–displacement curve into three distinct windows corresponding to specific phases of the joining process. These include the threading window, which detects when the joining partners begin to tilt and is marked by a force peak, as illustrated by the first dotted rectangle in Fig. 4; the transition window, which captures the core joining operation and is highlighted by the dark grey rectangle in Fig. 4; and the block window, which evaluates block dimensions and force, represented by the second dotted rectangle in Fig. 4.

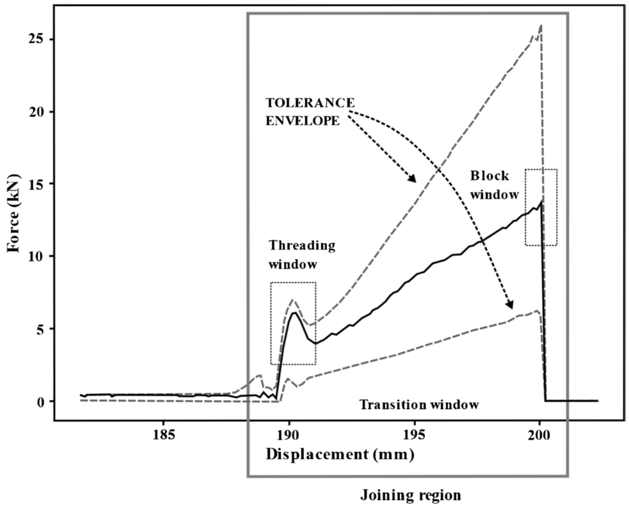


Fig. 4 A press-in process curve from our target process. In the region in which the joining takes place, we can see highlighted its main windows and the tolerance envelope (Lotter, 2006)

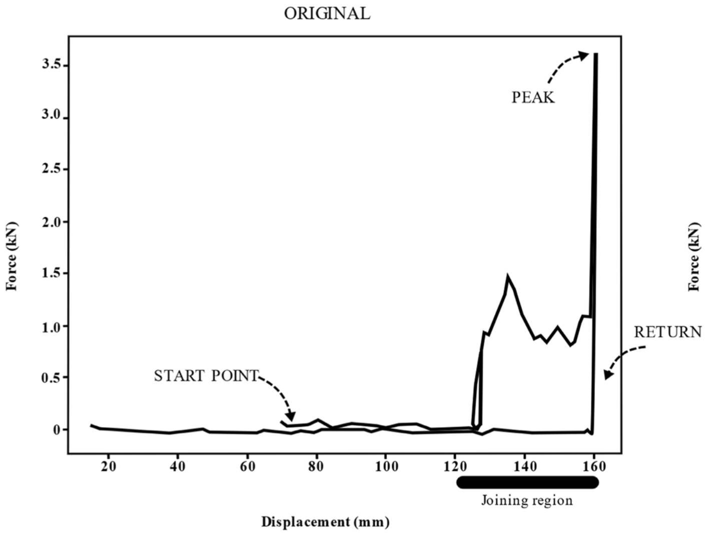
In contrast, the envelope technique defines a tolerance band around the curve, derived from successful assemblies, to "envelop" the acceptable range. This band is represented by the two dashed lines in Fig. 4. By comparing real-time process data to these predefined thresholds, windows, and tolerance envelopes, the ES can accurately assess the quality of press-in operations.

Development

Our solution works in three stages: pre-processing, training and evaluation. During pre-processing a sequence of operations are executed to reduce required computational resources and improve standardization. Observing Fig. 5a, we can notice that the force values recorded after the peak are always close to zero and so can be discarded. Figure 5 shows the difference between the original curve (a) and the one used in the following analysis (b) where the return phase has been dropped. We can note that the amount of data to handle is drastically reduced while the informative content required for the analysis is maintained.

At this point, envelope technique (Lotter, 2006) is implied to identify anomalous curves. From now on computations are executed by combining curves with the same CID. Given that the Maximum Displacement (MD) among curves with the same CID must be the same, we can compute it as in (1) where MD_{CID} is the target MD for the CID, n is the number of curves with the same CID and $D_i\left[\max\left(F_i\right)\right]$ is the displacement corresponding to the maximum force of the i-th curve.

$$MD_{CID} = \frac{\sum_{i=1}^{n} D_i\left[\max\left(F_i\right)\right]}{n} \tag{1}$$



(a)

(b)

Fig. 5 A sample curve before (**a**) and after (**b**) pre-processing

The projection of the windows on the x-axis is then computed as in (2) where W_{CID} it the interval of the projection of the window on the x-axis, $\sigma_{MD_{CID}}$ is the standard deviation of MD_{CID} and α is a coefficient to set the sensitivity of the windows. For further detail about the pre-processing, refer to Appendix A.

$$W_{CID} = MD_{CID} \pm \alpha \times \sigma_{MD_{CID}} \tag{2}$$

Although this technique may generate a flat curve after the force peak, this does not affect the final result of the quality of the output since most of the curve will have the same behavior, leading the mean value to approach 0.

$$F_b(d) = \begin{cases} \mu_{F(d)_{CID}} \pm \sigma_{F(d)_{CID}}, & \text{if } CP_{CID} \leq d \leq MD_{CID} + \alpha \times \sigma_{MD_{CID}} \\ \text{undefined}, & \text{otherwise} \end{cases} \tag{4}$$

Once curves belonging to the same CID are comparable, mean and standard deviation are computed for each point of the SV to be able to use envelope technique (Lotter, 2006). Through these value it is possible to generate both the target curve, which is the curve generated by the mean points, and upper and lower bound of the tolerance envelope, computed as in (4) where $F_b(d)$ is the boundary value at displacement d, $\mu_{F(d)_{CID}}$ is the mean of the force of the curves with the same CID at point d and $\sigma_{F(d)_{CID}}$ is the corresponding standard deviation (see Fig. 6). See Appendix A for a detailed description of the process.

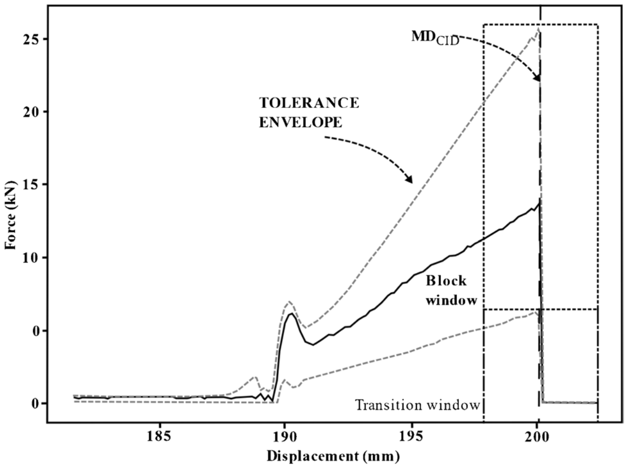


Fig. 6 Plot of a sample target curve with the parameters generated during the pre-processing and training steps

Once training is completed, the system can evaluate new force–displacement curves. To ensure comparability with the target curve, each new curve must first be resampled in the same manner as during the training phase. For each point on the curve, the system determines whether it lies within the predefined tolerance envelope. If a point falls outside the envelope, an out-of-bounds counter is incremented automatically.

The evaluation results in one of three flags being assigned to the curve: good, warning, or fault. These flags are determined based on the value of the out-of-bounds counter. Two manually adjustable thresholds are set to control the sensitivity of the evaluation process. These thresholds represent the maximum number of out-of-bounds points that can be tolerated for each classification. The first threshold distinguishes between good and warning, while the second differentiates warning from fault. Additionally, the system provides a visual representation of the evaluated curve alongside the tolerance boundaries, enabling users to perform a visual validation of the results. A detailed explanation of the evaluation process is available in Appendix A.

By iteratively performing training and evaluation steps, the algorithm progressively refines the boundary values. Curves flagged as warning or fault are excluded from subsequent training cycles, effectively narrowing the tolerance envelope and reducing oscillations in the target curve.

Deployment and results

In the initial phase, it took five weeks to deploy the first functional prototype of the algorithm. This version utilized a non-relational database, which facilitated the management of the unstructured data provided. However, this approach proved inefficient in terms of computational resource usage. Consequently, in collaboration with the company, we decided to transition to a relational database, necessitating a partial redesign of the code to accommodate the new data structure.

The adoption of a relational database introduced the need for a standardized data storage structure, representing a new information processing requirement. This transition marked the beginning of the second iteration of the intervention. During the redesign process, which spanned an additional five weeks, we not only adapted the system to the new database structure but also gained deeper insights into the manufacturing process. This enhanced understanding allowed us to improve the data pre-processing step, thereby increasing the efficiency and reliability of the algorithm.

After completing the development phase, we tested the ES on the provided historical data, which comprised approximately 175,000 joining operations conducted over 2 years, from September 2019 to September 2021. Unfortunately, precise fault identification records were unavailable; only aggregated statistics were provided. As a result, we could not fully validate the algorithm's fault-detection capacities.

Table 2 highlights our primary achievement, which is the ability to evaluate production quality at every stage of the process, rather than exclusively at the end of the production line. This milestone was achieved by implementing a modified version of the adaptive improvement principle proposed by Turetskyy et al. (2020), as illustrated in Fig. 7.

According to this principle, tolerances are initially defined based on experience or empirical data, and the system is configured accordingly. Each assembly step is then assessed in real-time. If a problem is detected, the functionality of the entire reducer is evaluated, and the results are presented to a decision-maker. This approach enables informed decisions on whether to halt production to prevent material waste or rework, or to proceed if the issue is deemed manageable.

We also conducted an ablation study to identify the most relevant features for evaluating the curves, with detailed experimental results provided in Appendix B. The algorithm proved particularly effective at detecting faults related to component non-conformity and improper alignment during the joining phase, which can have significant downstream implications. On one hand, these faults may result in an unbalanced final product, reducing operational precision. On the other hand, they may damage assembly elements, leading to machine stoppages and financial losses for the customer company.

Table 2 Summary of results achieved, based on the historical information

Problem	Initial conditions	After implementation
Defect identification time	Only at end of line	At each assembly step
Defect identification rate	~6% of final products	~15% of steps identified
Defect identification efficiency	~6% of final products	~8% of final products
Hidden defect before shipping	Not identified	~1% of final products

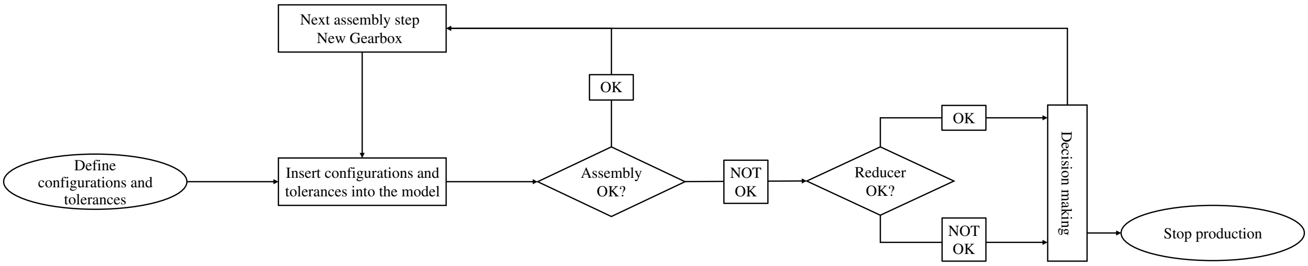


Fig. 7 Modified version of adaptive improvement principle by Turetskyy et al. (2020)

Our system enables halting production immediately upon detecting a defect, thereby avoiding costly reworking processes. This approach allows workers to focus on value-added activities and conserves materials, particularly since many components cannot be reused after assembly. Furthermore, the system can identify faults that would otherwise go unnoticed during manual quality inspections. For example, an unbalanced assembly might not cause immediate failure but could lead to operational issues over time.

To validate the operational impact of our solution, we conducted unstructured interviews with the project manager and the CEO one year after implementation. Table 3 summarizes the findings from these validation interviews. It includes the key codes identified during the analysis, representative quotes from both managers, and concise descriptions of the concepts, offering deeper insights into the practical benefits of the system.

Second loop: IoT device

As outlined earlier, the development of the AI solution necessitated the establishment of a structured approach for input data. While this significantly enhanced computational efficiency and minimized data redundancy, it also introduced limitations to the generalizability of the solution, as the data must be collected and stored in a specific format. To address this issue and expand the applicability of our solution, we initiated a new development phase aimed at creating a plug-and-play device. This device is designed to be easily installed across a variety of plants and preconfigured to collect and store data in alignment with the requirements of our ES.

This phase consisted of two key iterations. In the first iteration, we selected an ESP32 microcontroller equipped with an Ethernet module to collect and store the necessary data. This choice offered a compact and cost-effective solution while maintaining data integrity. In the second iteration, we enhanced the system by incorporating a programmable router to serve as a bridge between the plant's hardware and the network. Additionally, we utilized a virtual machine to act as a centralized server for data collection and storage.

To collect and transmit data to a remote storage, we developed an IoT device based on an ESP32 board, drawing inspiration from Aghenta et al. (2020). This device supports both MQTT and HTTP protocols, enabling it to utilize classical internet communication alongside specialized IoT protocols (Arnold et al., 2022). The ESP32 board is easily programmable and debuggable through its USB interface and the Arduino IDE, offering flexibility and ease of use. The flow of information among the devices in the infrastructure is depicted in Fig. 8a.

The ESP32 board interfaces with the plant using the HTTP protocol to request data from the sensors. After collecting the data, the board parses it and transmits it to the database via a proprietary API call (see Fig. 8a). While this architecture provided a practical starting point, it has two significant drawbacks. First, it introduces a single point of failure, as any disruption to the board results in the loss of all collected data. Second, it lacks universal compatibility with all machine brands due to the potential unavailability of drivers required for this architecture.

Edge infrastructure implementation

At this stage of the project, our primary goal was to enhance the adaptability of our solution, aligning with the cross-platform compatibility principles proposed by Spaeth and Niederhöfer (2022). After discussing with company technicians, we decided to modify the system architecture, as depicted in Fig. 8b. This updated design, inspired by the approach outlined by Kao et al. (2018), integrates an edge device as a bridge between the plant and the network, complemented by a server for data collection and storage.

While this architecture increases network traffic, as shown in Fig. 8b, it offers several critical advantages. Firstly, connecting the plant to the company network is simplified, requiring only a physical or wireless access point, which facilitates straightforward installation across various plant configurations. Secondly, the server-based data collection approach allows for the deployment of a wide array of drivers, enabling compatibility with different environments and machine brands. This enhanced adaptability significantly broadens the potential applications of the system, making it more versatile for diverse industrial contexts.

Deployment and results

We successfully developed a functional prototype of a data collection system capable of adapting to any type of machinery by simply modifying the collection script. This flexibility enables seamless integration into the ecosystem of various companies, ensuring compatibility with diverse industrial environments. The collected data can be directly utilized within the company's ecosystem and is readily leveraged by the ES. Furthermore, when integrated with other cloud services, as illustrated in Fig. 9, the IoT device becomes a robust tool for supporting business intelligence and AI systems, fully aligning with Industry 4.0 principles.

The prototype was initially tested in a controlled laboratory environment designed to replicate the structure of the plant where the first iteration was implemented. The IoT device was connected to an IO-Link Master equipped with a photo cell to

Table 3 Summary of the results of the validation interviews, grouped according to the 6 identified codes and correlated with quotes. Quotes marked with (PM) are from the project manager, once with (CEO) are from the CEO of the company. The results demonstrate the effectiveness of the proposed solution by highlighting its ability to enable early fault detection, enhance incoming material quality control, optimize production logic, and integrate seamlessly into ongoing digitalization efforts. Additionally, the findings validate the system's role in improving operational efficiency, facilitating technological transfer, and enabling integration with external processes

Result codes	Main quotes	Description
Early fault detection	"The proposed expert system is able to notify a fault in the assembly process in near real-time, allowing us to identify the problem and halt production before assembly is completed." (PM) "In this way, it is possible to save both time and resources related to reworking." (CTO)	The ES enables near real-time fault notifications during the assembly process, allowing for quick identification of issues and halting production before completion. This capacity helps save time and resources associated with rework
Intensification of incoming material quality control	"We can now proactively address quality issues before they become significant." (PM) "The introduction of the expert system allows to identify trends in faults by analyzing historical data clustered by component code." (CTO)	The new solution enhances the quality control process for incoming materials by analyzing historical data, allowing for the identification of trends and non-compliances that contribute to faulty assembly behaviors
Update of production logics	"This shift in logic helps optimize the supply chain and reduce costs." (PM) "Through the analysis of historical results, management is able to identify components to in- or out-source based on their performances during the assembly." (CTO)	Historical data analysis enables management to refine production strategies, deciding whether to insource or outsource components based on their performance metrics, ultimately enhancing operational efficiency
Extension of ongoing digitalization project	"This is a significant step toward fully integrating our systems." (PM) "The developed architecture has been introduced into the ecosystem developed by the company during their digitalization process." (CTO)	The integration of the new architecture within the existing digital ecosystem automates data flows, enhancing the efficiency of data collection, processing, and application back into the assembly process, furthering the company's digital transformation efforts
Integration with external processes	"By linking these processes, it is possible to enhance overall operational performance." (PM) "Indications provided by our system are suitable to identify faulty components which may be involved in other processes such as rectification of other forms of reworking." (CTO)	The system's ability to identify faulty components facilitates integration with external processes, allowing for rectification and better handling of rework, thereby improving overall operational efficiency
Technological transfer	"This transfer of knowledge empowers our team and sets us apart in the market." (PM) "Thanks to the chosen methodology, it has been possible to transfer AI and data management knowledge from the research group our staff." (CTO)	The chosen methodology has enabled the effective transfer of artificial intelligence and data management expertise from the research group to company staff, allowing the company to leverage this knowledge in its projects and gain a competitive edge in the industry

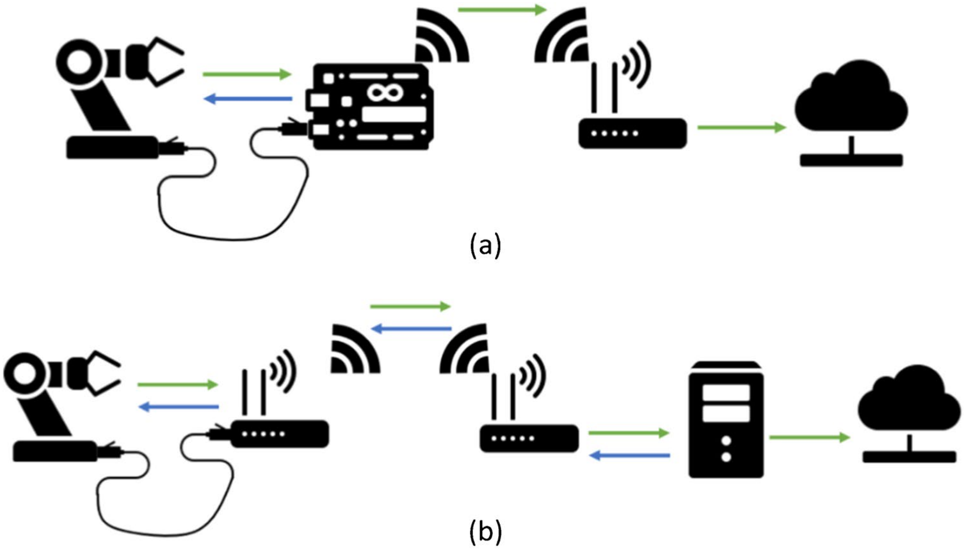


Fig. 8 **a** IoT architecture with programmable board and **b** IoT architecture with edge infrastructure

generate data at irregular intervals. Through this new setup, it became possible to remotely access process information in near real-time. Once data were generated, they were immediately transmitted to the cloud database, ensuring they were accessible to the AI algorithm. This architecture also enabled other applications, such as a plant dashboard or BI tools, to directly access the data, providing real-time insights to workers and managers.

Following the initial development, a modified version of the solution was deployed in a real-world scenario to address a company's requirement for a system capable of collecting data from their CNC lathe. This implementation was essential for maintaining an operational history and monitoring the lathe's real-time state. By making a few adjustments to accommodate the plant's specific hardware, the IoT device was successfully installed and tested in the production environment.

Discussion of finding

Organizational and theoretical implications

For what concern our research question "What are the implications both at the organizational and theoretical level of such implementation?", based on our observations we propose a revised version of OIPT framework (Tushman & Nadler, 1978) in the context of automated quality control, as in Fig. 10.

Tushman and Nadler (1978, p. 616) proposed that "as work related uncertainty increase, so does the need for increased amounts of information, and thus the need for increased information processing capacity". This proposition described a unidirectional influence, where information processing requirements, driven by uncertainty, affect information processing capacities. While confirming this effect,

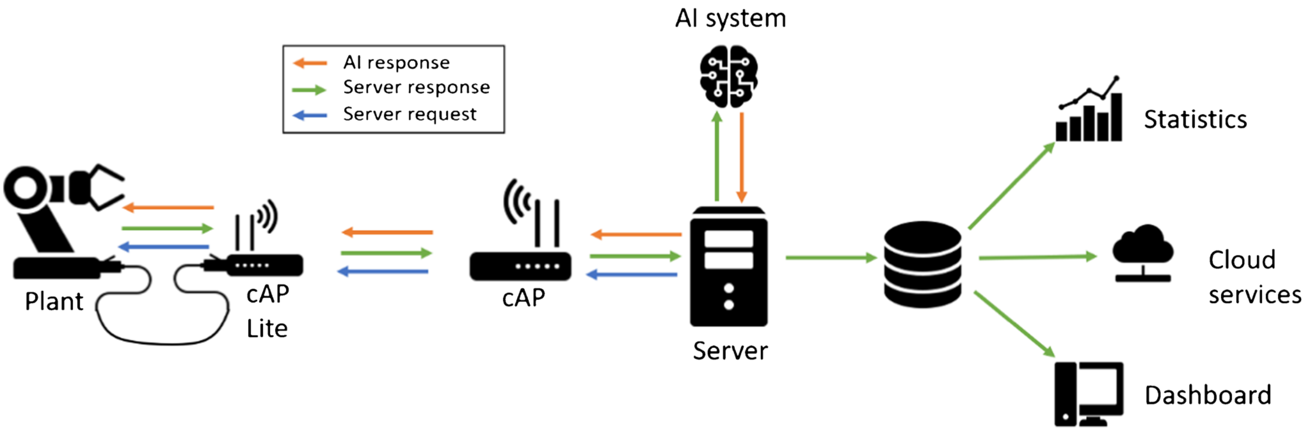


Fig. 9 Architectural design of the IoT device integrated with AI and other cloud services provided by the system integrator company

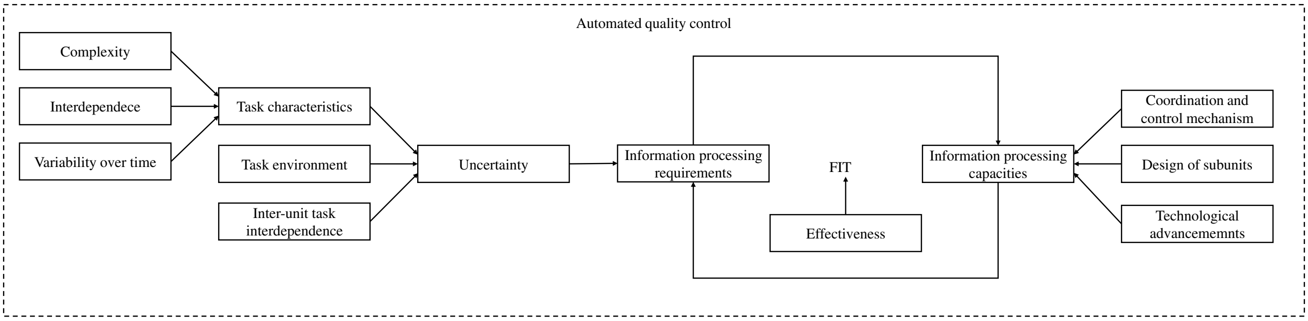


Fig. 10 Revised version of OIPT framework by Tushman and Nadler (1978), incorporating technological advancements and task variability over time as factors influencing information processing. The framework highlights the iterative and interactive relationship between information processing requirements and capacities, where increasing requirements drive the adoption of new technologies, which in turn generate additional information demands for their effective operation

our findings reveal a bidirectional relationship between these constructs. As illustrated in Fig. 10, we propose a feedback loop in which information processing requirements and capacities continuously influence each other. During our AR, the introduction of the ES to address the original need for greater information processing capacities, generated new information requirements, particularly the demand for structured and detailed data. This, in turn, led to the development of enhanced information processing capacities through the implementation of IoT to expand data collection. The iterative nature of this dynamic aligns with literature, which describes the AI implementation process as inherently iterative (Neumann et al., 2024), requiring adaptation to emerging business goals and technical needs (Uren & Edwards, 2023). Based on these insights, we propose the following statement:

> *S1 Information processing capacities and information processing requirements influence each other in a feedback loop*

Our statement also integrates another proposition by Tushman and Nadler (1978, p. 619): "organizations will be more effective when there is a match between information processing requirements facing the organization and information processing capacity of the organization's structure." While focusing on the importance of achieving this fit, we emphasize the need for a reciprocal interaction between the two constructs. AI has been shown to be an augmenting technology rather than purely an automating one (Colombo et al., 2023; Revilla et al., 2023). This means that AI collaborates with humans rather than replacing them, fostering a dynamic partnership where human input refines the outputs of the technology, which, in turn, informs human decisions. Such collaboration is achievable only through successive iterations, where both human and technological components adapt and evolve together. We observed a similar interaction during the development of the proposed ES. Technicians provided key inputs for the system's design, which then offered new insights into the production process. Additionally, the newly developed system architecture created a need for IoT infrastructure. This infrastructure, while supporting the ES, also empowered technicians with an enhanced means of data collection, opening opportunities for diverse applications beyond the immediate scope of the system.

This mutual influence between technology and the organization highlights the need for flexibility on both sides. As noted, AI implementation processes must adapt to emerging business goals (Uren & Edwards, 2023), while organizations must address new technical requirements that may reshape their processes, as observed in our action research. Organizational flexibility has proven essential for managing uncertainty (Srinivasan & Swink, 2018). In our intervention, this flexibility was evident in the shift from EOL quality control to a fundamentally different paradigm, accommodating new information processing requirements and fully leveraging the capacities provided by the ES.

Recalling Tushman and Nadler (1978, p. 621) proposition that "if organization (or subunits) face different conditions over time, more effective units will adapt their structure to meet the changed information processing requirements," we observe that variability over time introduces an intrinsic source of uncertainty in the quality control task. This variability has to be accounted for defining information processing requirements, underscoring the critical need for flexibility in both technology and organizational structures to adapt to evolving conditions. Based on this analysis we can draw the following statement:

> *S2 In the context of high information requirements, flexibility is crucial to handle various sources of uncertainty*

Additionally, Srinivasan and Swink (2018), in alignment with OIPT principles, emphasize that implementing a vertical information system enhances efficient data processing, facilitates rapid decision-making, and reduces communication channel overload (Galbraith, 1974). These systems ensure that critical information flows effectively to decision-makers while minimizing unnecessary complexity in communication processes.

In line with these principles, we observed that the implementation of the IoT architecture significantly improved communication by providing detailed, real-time information about the production process at each stage. This capacity offered a comprehensive view of the production line and allowed management to monitor operations with unprecedented granularity. However, the increased flow of information introduced a potential risk of overwhelming operators and decision-makers with excessive data, which could lead to redundancy, inefficiencies, and slower response times. To address this challenge, the ES was designed with an intelligent filtering mechanism. By prioritizing critical information and issuing alerts only when operator intervention was necessary, the ES ensured that operators could focus on resolving key issues without being distracted by irrelevant data. The case highlights the delicate balance required when integrating advanced technologies like IoT and AI into production environments. While these technologies have the potential to drastically enhance information processing capacities, their successful implementation depends on the coordination between process control mechanisms and communication strategies. Without such coordination, the influx of data can overwhelm human operators and reduce the overall effectiveness of the system. We can then derive a third statement:

> *S3 Technological advancements impact information processing capacities, but has to be coordinated with the process control mechanisms and communication strategies*

Guidelines for AI implementation in SMEs

Our intervention provided valuable insights into the implementation of AI in SMEs, contributing empirical evidence regarding the specific challenges and characteristics of such setting. Battistoni et al. (2023) proposed a hierarchical framework for the implementation of digital technologies in SMEs, emphasizing the particular constraints these companies face, such as limited financial availability and human resources. While addressing our research question “To what extent is it possible to develop an automated quality control solution in a real manufacturing setting?,” we integrate this framework with the findings of our action research. By doing so, we examine the critical pitfalls encountered during the implementation process, discuss the strategies employed to address these challenges, and offer practical guidelines for SMEs seeking to integrate AI into their operations.

Battistoni et al. (2023) provided a valuable framework to guide digital transformation in manufacturing SMEs. This framework builds on a hierarchical four-layered model (i.e., sensor, integration, intelligent, and response) clustering digital technologies and discussing adoption pathways for SMEs. We believe it is an appropriate tool to review our results.

Sensor layer: Data collection and quality

As a first step, it was crucial to align the quality control paradigm with the inherent uncertainty of the manufacturing process. Highly uncertain production environments require a shift from traditional EOL quality control to step-by-step evaluation. Such a paradigm shift, as observed in our study, enables the real-time assessment of process quality rather than focusing exclusively on the final product. This approach allows for immediate identification of defects during assembly, reducing the need for rework and minimizing resource waste (Arinez et al., 2020; Turetskyy et al., 2020). Additionally, this method generates valuable data about the quality of individual components, enabling manufacturers to detect supplier-related trends and collaboratively address quality issues, thereby improving overall supply chain performance (Zhu et al., 2018).

The introduction of step-by-step quality control, while addressing the initial need for early defect detection and prevention, increased the volume and complexity of data to be managed. This necessitated investment in sensor technologies and robust data management systems, as discussed in the “sensor” layer by Battistoni et al. (2023). Our findings underline the importance of ensuring high-quality data collection, as the accuracy and relevance of the data directly affect the performance of AI systems. In our intervention, we observed that the limited availability of data posed challenges for the ES in reliably identifying assembly failures. To mitigate this issue, expert knowledge was integrated into the system, enhancing its interpretability and supporting decision-making processes. Nevertheless, this approach proved insufficient to fully align information processing capacities with requirements, leading to the implementation of IoT devices to enhance data collection.

The integration of IoT technologies facilitated real-time, cost-effective data acquisition, showcasing their potential for SMEs with limited resources (Matt et al., 2020). The implementation costs of this technology were notably low, and it integrated seamlessly with existing workflows. By utilizing boards such as the ESP32, the company was able to deploy IoT infrastructure that could easily connect to a wide range of manufacturing systems. This flexible solution not only proved effective in enhancing data collection for the AI system but also provided the company with a straightforward approach to improve its overall data strategy.

Integration layer: System interoperability

The subsequent challenge was to ensure effective system integration, as highlighted in the "integration" layer of Battistoni et al.'s framework. AI implementation requires seamless interoperability between existing systems and the AI infrastructure, alongside a focus on human-AI collaboration (Colombo et al., 2023; Revilla et al., 2023). In tasks such as quality control, it is essential to design systems that augment human expertise by providing interpretable and actionable insights. In our study, the visual and comprehensible output of the ES allowed operators to make informed decisions, illustrating the importance of combining AI with domain expertise in contexts where interpretability is critical.

The integration of IoT and ES within a single system architecture has proven beneficial for quality control in SMEs. These technologies required minimal financial investment, offering a valuable solution for resource-constrained companies. While the relative simplicity of the ES, compared to more complex ML or deep learning models, provided actionable insights with minimal disruption to existing processes, the incorporation of IoT significantly expanded data processing capacities. This integration allowed the ES to be seamlessly incorporated into the existing plant infrastructure, enhancing overall system efficiency and effectiveness.

Intelligent layer: Iterative AI development

The iterative nature of AI development became apparent during our intervention. As noted in the "intelligent" layer, AI systems must be continuously refined based on real-world performance and evolving requirements (Neumann et al., 2024). In the first iteration, the implementation of the ES enhanced the organization's ability to process information but revealed data limitations that hindered its effectiveness. The subsequent iteration addressed these limitations by integrating IoT devices, which improved data collection capacities and aligned information processing capacities with requirements. This iterative approach underscores the importance of adopting a long-term perspective, where technological advancements are introduced incrementally and aligned with organizational needs over time (Uren & Edwards, 2023).

Response layer: Collaboration and scalability

Finally, our findings emphasize the value of collaboration, as captured in the "response" layer. SMEs often lack the internal resources and expertise required for AI implementation, making partnerships with external organizations essential. In our case, collaboration with the research team enabled the transfer of knowledge and access to cutting-edge technologies, while also allowing the manufacturing company to adopt AI solutions at a low cost.

The AI and IoT-based quality control system developed in our study offers promising potential for adaptation in sectors like the food industry. In food production, quality control is heavily reliant on monitoring critical parameters such as temperature, humidity, pH, and product integrity. By integrating IoT sensors that track these variables in real-time, the system could be adapted to provide continuous, step-by-step assessments throughout the production process, ensuring that any deviations from optimal conditions are detected early. Similar to its application in manufacturing, the system's ability to process and analyze data at each stage can prevent defects or contamination before they reach the final product, thus improving food safety and reducing waste. The use of low-cost, flexible IoT devices, like those implemented in our study, can be easily integrated into existing food production lines, making this solution highly accessible. Furthermore, incorporating expert knowledge to contextualize the data and refine decision-making would enhance the system's effectiveness in the in the case of processes with high variability or dependent on environmental conditions. This process highlights the potential for SMEs to leverage external expertise to overcome barriers to AI adoption and demonstrates that AI implementation is not solely the domain of large enterprises.

In summary, we highlight the following managerial implications of our study. The framework derived from our study integrates these observations into a cohesive pathway for SMEs seeking to implement AI. By addressing process uncertainty through step-by-step quality control, investing in robust data collection and management, ensuring system interoperability, adopting iterative development processes, and fostering external collaborations, SMEs can overcome barriers to AI implementation and enhance their competitiveness. Importantly, our findings highlight that the implementation of AI should be approached as a dynamic and evolving process, where each stage builds on prior advancements to achieve sustainable improvements in quality control and operational efficiency.

Conclusions

Our study contributes to the burgeoning literature on supply chain digitalization (Perano et al., 2023) by focusing on the joint adoption of AI and IoT in the context of automated quality control, a key supply chain process for manufacturing companies. Beyond the established impacts discussed in the literature (Culot et al., 2024), to the best of our knowledge, this article represents one of the first applications of AR in developing a tool that combines AI and IoT in the context of SMEs. This approach allowed us to generate context-specific knowledge while addressing the constraints typical of smaller companies, which provided opportunities to explore solutions that yield high-quality results without necessitating substantial financial investments. Unlike projects conducted in larger companies, the limited availability and quality of data in our study posed a significant challenge. To overcome this, we devise an innovative approach effectively combines AI and expert knowledge. This integration allowed us to enhance the information embedded in the data by filtering out irrelevant data points and excluding data associated with collection failures, identifiable by improper curve behavior or incomplete representation. The quality improvement at the manufacturing level may provide further positive impacts at the overall supply chain level, such as customer satisfaction, cost-efficiency, and prompt problem-solving with suppliers. Overall, the tailored approach we implemented in the SME context has yielded significant advantages and novel insights, laying the foundation for future applications in similar scenarios. In doing so, we provide guidelines for SMEs aiming to implement AI in their processes, drawing on empirical insights from our study, while also identifying potential challenges and offering practical strategies to address them.

The design and implementation of our solution address gaps in the existing literature that motivated our research questions. By extending the OIPT framework through observations made during the action research, we incorporate technological advancements and the variability over time. This integration highlights the specificities of AI implementation and emphasizes the iterative nature of the process.

This study, while offering valuable insights, has limitations related to its specific application context, the assembly of reduction gearboxes in a small enterprise. One notable limitation is the presence of low-quality data, which affects the robustness of the ES and can impair the accuracy of early fault detection, reducing the system's reliability in identifying defects. Addressing data quality constraints is critical for improving the system's precision and expanding its applicability to broader industrial contexts. Moreover, the system's reliance on specific physical measures relevant to gearbox assembly limits the system's direct adaptability to other industries. Future research should the scalability of the system in more contexts, such as multi-national or multi-plant companies, where intricate supply chains may challenge the solution's applicability.

Further validation could also involve testing the system's architecture with time-series data or other standardized behaviors to extend its utility in automated quality control processes. Future developments should prioritize case studies across diverse industrial settings to confirm the system's flexibility and scalability. Addressing these areas could enhance the ES's applicability, making it a more versatile solution for automated quality control across varied production environments.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12525-025-00766-y.

Funding Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement.

Data Availability The datasets generated and analyzed during the current study are not publicly available due to confidentiality agreements with the providing company. Access to the data is restricted as per the company's policy and cannot be shared.

Declarations

Conflict of Interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

Aghenta, L. O., Iqbal, T., Department of Electrical and Computer Engineering, Faculty of Engineering and Applied Science, Memorial University of Newfoundland (MUN), St. John's, NL A1B 3X5, Canada. (2019). Design and implementation of a low-cost, open source IoT-based SCADA system using ESP32 with OLED, ThingsBoard and MQTT protocol. *AIMS Electronics and Electrical Engineering, 4*(1), 57–86. https://doi.org/10.3934/ElectrEng.2020.1.57

Arinez, J. F., Chang, Q., Gao, R. X., Xu, C., & Zhang, J. (2020). Artificial intelligence in advanced manufacturing: Current status and future outlook. *Journal of Manufacturing Science and Engineering, 142*(11), 110804. https://doi.org/10.1115/1.4047855

Arnold, L., Jöhnk, J., Vogt, F., & Urbach, N. (2022). IIoT platforms' architectural features – a taxonomy and five prevalent archetypes. *Electronic Markets, 32*(2), 927–944. https://doi.org/10.1007/s12525-021-00520-0

Athanasopoulou, A., & De Reuver, M. (2020). How do business model tools facilitate business model exploration? *Evidence from Action Research. Electronic Markets, 30*(3), 495–508. https://doi.org/10.1007/s12525-020-00418-3

Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Systems with Applications, 173*, 114598. https://doi.org/10.1016/j.eswa.2021.114598

Battistoni, E., Gitto, S., Murgia, G., & Campisi, D. (2023). Adoption paths of digital transformation in manufacturing SME. *International Journal of Production Economics, 255*, 108675. https://doi.org/10.1016/j.ijpe.2022.108675

Benzidia, S., Makaoui, N., & Bentahar, O. (2021). The impact of big data analytics and artificial intelligence on green supply chain process integration and hospital environmental performance. *Technological Forecasting and Social Change, 165*, 120557. https://doi.org/10.1016/j.techfore.2020.120557

Bohez, E. L. J., & Thieravarut, M. (1997). Expert system for diagnosing computer numerically controlled machines: A case-study. *Computers in Industry, 32*(3), 233–248. https://doi.org/10.1016/S0166-3615(96)00077-2

Bousdekis, A., Lepenioti, K., Apostolou, D., & Mentzas, G. (2023). Data analytics in quality 4.0: Literature review and future research directions. *International Journal of Computer Integrated Manufacturing, 36*(5), 678–701. https://doi.org/10.1080/0951192X.2022.2128219

Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2020). Supply chain data analytics for predicting supplier disruptions: A case study in complex asset manufacturing. *International Journal of Production Research, 58*(11), 3330–3341. https://doi.org/10.17863/CAM.45246

Carbery, C. M., Woods, R., & Marshall, A. H. (2018). A Bayesian network based learning system for modelling faults in large-scale manufacturing. *IEEE International Conference on Industrial Technology (ICIT), 2018*, 1357–1362. https://doi.org/10.1109/ICIT.2018.8352377

Cardellicchio, A., Nitti, M., Patruno, C., Mosca, N., Di Summa, M., Stella, E., & Renò, V. (2024). Automatic quality control of aluminium parts welds based on 3D data and artificial intelligence. *Journal of Intelligent Manufacturing, 35*(4), 1629–1648. https://doi.org/10.1007/s10845-023-02124-1

Colombo, J., Boffelli, A., Kalchschmidt, M., & Legenvre, H. (2023). Navigating the socio-technical impacts of purchasing digitalisation: A multiple-case study. *Journal of Purchasing and Supply Management, 100849*. https://doi.org/10.1016/j.pursup.2023.100849

Coughlan, P., & Coghlan, D. (2002). Action research for operations management. *International Journal of Operations & Production Management, 22*(2), 220–240. https://doi.org/10.1108/01443570210417515

Cruz Guerrero, R., De Los Angeles Alonso Lavernia, M., & Simon Marmolejo, I. (2019). Prediction of press-fit quality via data mining techniques and artificial intelligence. *IEEE Access, 7*, 159599–159607. https://doi.org/10.1109/ACCESS.2019.2950642

Culot, G., Podrecca, M., & Nassimbeni, G. (2024). Artificial intelligence in supply chain management: A systematic literature review of empirical studies and research directions. *Computers in Industry, 162*, 104132. https://doi.org/10.1016/j.compind.2024.104132

Dubey, R., Gunasekaran, A., Childe, S. J., Roubaud, D., Fosso Wamba, S., Giannakis, M., & Foropon, C. (2019). Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain. *International Journal of Production Economics, 210*, 120–136. https://doi.org/10.1016/j.ijpe.2019.01.023

Dubey, R., Gunasekaran, A., Childe, S. J., Bryde, D. J., Giannakis, M., Foropon, C., Roubaud, D., & Hazen, B. T. (2020). Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. *International Journal of Production Economics, 226*, 107599. https://doi.org/10.1016/j.ijpe.2019.107599

Escobar, C. A., & Morales-Menendez, R. (2018). Machine learning techniques for quality control in high conformance manufacturing environment. *Advances in Mechanical Engineering, 10*(2), 168781401875551. https://doi.org/10.1177/1687814018755519

Escobar, C. A., McGovern, M. E., & Morales-Menendez, R. (2021). Quality 4.0: A review of big data challenges in manufacturing. *Journal of Intelligent Manufacturing, 32*(8), 2319–2334. https://doi.org/10.1007/s10845-021-01765-4

Fang, S., Ding, G., & Chen, X. (2018). Detection of keypoint in press-fit curve based on convolutional neural network. *World Academy of Science, Engineering and Technology International Journal of Materials and Metallurgical Engineering, 12*(12), 650–654. https://doi.org/10.5281/zenodo.2363238

Fosso Wamba, S. (2022). Impact of artificial intelligence assimilation on firm performance: The mediating effects of organizational agility and customer agility. *International Journal of Information Management, 67*, 102544. https://doi.org/10.1016/j.ijinfomgt.2022.102544

Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces, 4*(3), 28–36. https://doi.org/10.1287/inte.4.3.28

Goodarzian, F., Navaei, A., Ehsani, B., Ghasemi, P., & Muñuzuri, J. (2023). Designing an integrated responsive-green-cold vaccine supply chain network using Internet-of-Things: Artificial intelligence-based solutions. *Annals of Operations Research, 328*(1), 531–575. https://doi.org/10.1007/s10479-022-04713-4

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2019). A survey of methods for explaining black box models. *ACM Computing Surveys, 51*(5), 1–42. https://doi.org/10.1145/3236009

Hamm, P., Klesel, M., Coberger, P., & Wittmann, H. F. (2023). Explanation matters: An experimental study on explainable AI. *Electronic Markets, 33*, 17. https://doi.org/10.1007/s12525-023-00640-9

Hansen, L. K., & Rieger, L. (2019). Interpretability in intelligent systems – a new concept? Explainable AI: Interpreting, explaining and visualizing deep learning. *Lecture Notes in Computer Science, 11700*. https://doi.org/10.1007/978-3-030-28954-6_3

Hansen, E. B., & Bøgh, S. (2021). Artificial intelligence and internet of things in small and medium-sized enterprises: A survey. *Journal of Manufacturing Systems, 58*, 362–372. https://doi.org/10.1016/j.jmsy.2020.08.009

Iglesias, C., Martínez, J., & Taboada, J. (2018). Automated vision system for quality inspection of slate slabs. *Computers in Industry, 99*, 119–129. https://doi.org/10.1016/j.compind.2018.03.030

Inoue, Y., Takenaka, T., Kasasaku, T., Tamegai, T., & Arai, R. (2023). How to design platform ecosystems by intrapreneurs: Implications from action design research on IoT-based platform. *Electronic Markets, 33*, 9. https://doi.org/10.1007/s12525-023-00618-7

Kao, K.-C., Chieng, W.-H., & Jeng, S.-L. (2018). Design and development of an IoT-based web application for an intelligent remote SCADA system. *IOP Conference Series: Materials Science and*

Engineering, 323, 012025. https://doi.org/10.1088/1757-899X/323/1/012025

Levin, M. (2003). Action research and the research community. *Concepts and Transformation, 8*(3), 275–280. https://doi.org/10.1075/cat.8.3.08lev

Li, C., & Li, L. (2023). An analytical framework for error propagation effects in multiprocess manufacturing. *Mathematical Problems in Engineering, 2023*(1), 2769757. https://doi.org/10.1155/2023/2769757

Li, T., Zhang, Y., Xu, J., Dai, Z., Yao, J., Luo, L., & Gong, L. (2024). Study on the establishment method of multi-dimensional chain model and precision assembly for robot precision reducers. *Journal of the Brazilian Society of Mechanical Sciences and Engineering, 46*(9), 530. https://doi.org/10.1007/s40430-024-05108-1

Liu, Y., Yang, M., & Guo, Z. (2022). Reinforcement learning based optimal decision making towards product lifecycle sustainability. *International Journal of Computer Integrated Manufacturing, 35*(10–11), 1269–1296. https://doi.org/10.1080/0951192X.2022.2025623

Lotter, B. (Ed.). (2006). *Montage in der industriellen Produktion: Ein Handbuch für die Praxis ; mit 16 Tabellen*. Springer. https://doi.org/10.1007/978-3-642-29061-9

Matt, D. T., Modrák, V., & Zsifkovits, H. (Eds.). (2020). *Industry 4.0 for SMEs: Challenges, opportunities and requirements*. Springer International Publishing. https://doi.org/10.1007/978-3-030-25425-4

Meiners, M., Mayr, A., & Franke, J. (2021). Process curve analysis with machine learning on the example of screw fastening and press-in processes. *Procedia CIRP, 97*, 166–171. https://doi.org/10.1016/j.procir.2020.05.220

Moeuf, A., Pellerin, R., Lamouri, S., Tamayo-Giraldo, S., & Barbaray, R. (2018). The industrial management of SMEs in the era of Industry 4.0. *International Journal of Production Research, 56*(3), 1118–1136. https://doi.org/10.1080/00207543.2017.1372647

Neumann, O., Guirguis, K., & Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: A comparative case study. *Public Management Review, 26*(1), 114–141. https://doi.org/10.1080/14719037.2022.2048685

Perano, M., Cammarano, A., Varriale, V., Del Regno, C., Michelino, F., & Caputo, M. (2023). Embracing supply chain digitalization and unphysicalization to enhance supply chain performance: A conceptual framework. *International Journal of Physical Distribution & Logistics Management, 53*(5/6), 628–659. https://doi.org/10.1108/IJPDLM-06-2022-0201

Peres, R. S., Barata, J., Leitao, P., & Garcia, G. (2019). Multistage quality control using machine learning in the automotive industry. *IEEE Access, 7*, 79908–79916. https://doi.org/10.1109/ACCESS.2019.2923405

Reason, P., & Bradbury, H. (2001). *Handbook of action research*. Sage. https://doi.org/10.4135/9781848607934

Revilla, E., Saenz, M. J., Seifert, M., & Ma, Y. (2023). Human–artificial intelligence collaboration in prediction: A field experiment in the retail industry. *Journal of Management Information Systems, 40*(4), 1071–1098. https://doi.org/10.1080/07421222.2023.2267317

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing. *Boston Consulting Group*. https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries. Accessed 15 Dec 2024.

Schmidberger, S., Bals, L., Hartmann, E., & Jahns, C. (2009). Ground handling services at European hub airports: Development of a performance measurement system for benchmarking. *International Journal of Production Economics, 117*(1), 104–116. https://doi.org/10.1016/j.ijpe.2008.10.006

Senoner, J., Netland, T., & Feuerriegel, S. (2022). Using explainable artificial intelligence to improve process quality: Evidence from semiconductor manufacturing. *Management Science, 68*(8), 5704–5723. https://doi.org/10.1287/mnsc.2021.4190

Shani, A. B. (Rami), & Coghlan, D. (2021). Action research in business and management: A reflective review. *Action Research, 19*(3), 518–541. https://doi.org/10.1177/1476750319852147

Sharma, G., Greco, A., Grewatsch, S., & Bansal, P. (Tima). (2022). Cocreating forward: How researchers and managers can address problems together. *Academy of Management Learning & Education, 21*(3), 350–368. https://doi.org/10.5465/amle.2021.0233

Spaeth, S., & Niederhöfer, S. (2022). Compatibility promotion between platforms: The role of open technology standards and giant platforms. *Electronic Markets, 32*(4), 1891–1915. https://doi.org/10.1007/s12525-022-00590-8

Srinivasan, R., & Swink, M. (2018). An investigation of visibility and flexibility as complements to supply chain analytics: An organizational information processing theory perspective. *Production and Operations Management, 27*(10), 1849–1867. https://doi.org/10.1111/poms.12746

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems, 48*, 157–169. https://doi.org/10.1016/j.jmsy.2018.01.006

Turetskyy, A., Wessel, J., Herrmann, C., & Thiede, S. (2020). Data-driven cyber-physical system for quality gates in lithium-ion battery cell manufacturing. *Procedia CIRP, 93*, 168–173. https://doi.org/10.1016/j.procir.2020.03.077

Tushman, M. L., & Nadler, D. A. (1978). Information processing as an integrating concept in organizational design. *The Academy of Management Review*, *3*(3), 613–624. JSTOR. https://doi.org/10.2307/257550

Uren, V., & Edwards, J. S. (2023). Technology readiness and the organizational journey towards AI adoption: An empirical study. *International Journal of Information Management, 68*, 102588. https://doi.org/10.1016/j.ijinfomgt.2022.102588

Wang, X., Liu, M., Ge, M., Ling, L., & Liu, C. (2015). Research on assembly quality adaptive control system for complex mechanical products assembly process under uncertainty. *Computers in Industry, 74*, 43–57. https://doi.org/10.1016/j.compind.2015.09.001

Wang, G., Ledwoch, A., Hasani, R. M., Grosu, R., & Brintrup, A. (2019). A generative neural network model for the quality prediction of work in progress products. *Applied Soft Computing, 85*, 105683. https://doi.org/10.1016/j.asoc.2019.105683

Wanner, J., Herm, L.-V., Heinrich, K., & Janiesch, C. (2022). The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study. *Electronic Markets, 32*(4), 2079–2102. https://doi.org/10.1007/s12525-022-00593-5

Zhu, S., Song, J., Hazen, B. T., Lee, K., & Cegielski, C. (2018). How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective. *International Journal of Physical Distribution & Logistics Management, 48*(1), 47–68. https://doi.org/10.1108/IJPDLM-11-2017-0341