IADIS International Journal on Computer Science and Information Systems
Vol. 21, No. 2, pp. 103-125
ISSN: 1646-3692

DESIGNING COLLABORATIVE ROBOTICS FOR AUTOMOTIVE DEFECT DETECTION AND REPAIR: A HUMAN-CENTERED APPROACH

Günter Alce[1], Björn Fischer[1,6], Fatih Çiftçi[2], Dilek Tükel[3,4], Daniele Fontanelli[5] and Susanne Frennert[1]
[1]*Lund University, Box 118, 221 00 Lund, Sweden*
[2]*Tofaş, İstanbul Cad. No: 574, 16110 Bursa, Türkiye*
[3]*Altinay Robot Technologies, AYOSB Mah. 2.Sanayi Cad. No:22 Tuzla/İstanbul, TR-34953, Türkiye*
[4]*Doğuş University, Nato Yolu Cad. 265, 34776 İstanbul, Türkiye*
[5]*University of Trento, via Sommarive 9, 38123 Povo, Trento, Italy*
[6]*School of Humanities, Arts & Social Sciences, Singapore University of Technology & Design, Singapore*

ABSTRACT

The increasing deployment of collaborative robots (cobots) in industrial settings requires a nuanced understanding of their impact on workers. Unlike traditional automation that replaces human tasks, cobots are designed to complement and enhance human capabilities. This study reports a human-centered investigation within the MAGICIAN project, which develops cobot-supported defect detection and repair for automotive manufacturing. Fieldwork at Tofaş in Türkiye combined structured observations of eight workers across Bodyshop and Press settings, 31 stakeholder interviews, and unweighted NASA-TLX assessments with 27 workers. We detail how observational records, stakeholder accounts, and workload patterns were triangulated and translated into traceable requirements. Results identify high temporal and effort demands, distinct workload profiles between production areas, reliance on tacit visual and tactile expertise, and the importance of predictability, worker control, and integration with existing reporting practices. This work contributes to an empirically grounded requirements-elicitation framework, and a transparent, confidence-aware cobot architecture that preserves human oversight while reducing repetitive physical and cognitive strain.

KEYWORDS

Human-Computer Interaction, Human-Robot Collaboration, Human-Centered Design

1. INTRODUCTION

The integration of collaborative robots (cobots) into industrial environments is transforming how we think about automation (Alves et al., 2023). In Industry 4.0, the focus was on digitalization, data analysis, and automation, which led to significant efficiency gains (Alves et al., 2023). Despite these advances, the technology-centric focus of Industry 4.0 has often been criticized for neglecting the human factor, resulting in increased system complexity, procedural inertia, and new skill gaps in the workforce (Nguyen et al., 2022). In response to these critiques, Industry 5.0 has emerged as a more human-centered approach. The core of Industry 5.0 is not to replace humans, but to create collaborative environments where the precision and strength of machines enhance human creativity, critical thinking, and adaptability (Whitley, 2025). Therefore, unlike traditional robotic systems that operate in isolation or replace human labor, cobots are designed to work with humans, augmenting their capabilities rather than substituting them. This shift brings new challenges and opportunities, particularly in sectors like automotive manufacturing, where defect detection and repair still rely heavily on human expertise and sensory input (Villani et al., 2018; Krüger et al., 2009).

Maintaining high production quality under tight cycle constraints demands a balance between automation and human judgment. While advances in computer vision and AI have improved the capacity to detect surface anomalies, the nuanced tactile and visual assessments performed by skilled workers remain difficult to replicate algorithmically (Segura et al., 2025). Moreover, introducing automation into such workflows carries cognitive and organizational implications, including changes in worker roles, increased monitoring demands, and potential reductions in perceived autonomy (Segura et al., 2025). Despite advances in computer vision and shared autonomy, few studies report field-validated cobot architectures that align with workers' sensory workflows and address trust, role boundaries, and documentation burdens in defect detection and repair. We address this gap by conducting a multi-method field study in an operational automotive setting, deriving empirically grounded design guidelines and a reference architecture for confidence-aware cobot collaboration in defect detection and minor repair, and validating the implications with workers and stakeholders through iterative feedback cycles.

This paper has three goals: (1) to introduce the MAGICIAN project and its human-centered approach to an adaptive cobot system for defect detection and repair in real-world manufacturing; (2) to present the field studies conducted at Tofaş, including observations, interviews, and workload assessments, and to show how they informed user requirements; and (3) to translate the field findings into personas, workflow models, traceable requirements, and a proposed system architecture.

What we add to the field is a human-centered approach to understand how collaborative inspection can be designed and deployed without replacing workers. We ground this in a case study at a major automotive factory in Türkiye, then derive robotics-oriented requirements that preserve operator control while reducing repetitive strain. Our research contributions:

1. Empirical findings from observations, interviews, and workload data about defect detection.
2. Identifying key challenges, including cognitive workload, task adaptation, and trust in automation, that influence worker acceptance of automated systems.
3. Requirement synthesis in which field findings are interpreted through personas and current-state user journeys and translated into traceable human-centered requirements

4. A proposed cobot system overview, including future user journeys and a modular, confidence-aware architecture, designed to enhance human oversight rather than replace worker judgment.

The following section presents relevant related work, then the multi-method approach, which incorporated the field studies, followed by system architecture, discussion, and conclusions.

2. RELATED WORK

The field of human-robot collaboration has evolved significantly in recent years, moving from isolated robotic automation to tightly integrated systems in which humans and machines share tasks in real time. This development has broadened the scope of research to include fields such as industrial robotics, human factors, cognitive science, and interaction design.

2.1 Human-in-the-Loop Robotics and Adaptive Autonomy

Human-in-the-loop (HITL) approaches have gained traction as a strategy to ensure that automation enhances, rather than displaces, human decision-making. Adaptive autonomy frameworks allow control to dynamically shift between humans and robots based on context, uncertainty, and operator input. For instance, Dragan et al. (2013) proposed shared autonomy systems where robots infer user intent to provide seamless assistance. Similarly, Nikolaidis et al. (2017) demonstrated that trust-sensitive policies improve collaborative fluency in mixed-initiative settings.

In industrial contexts, real-time arbitration between human and robot control is particularly relevant for tasks involving perception and judgment under variability, such as quality inspection. Researchers like Wang et al. (2023) introduced trust-based role allocation methods that modulate robot behavior based on operator confidence, an approach that inspires our work on confidence-aware cobot collaboration.

Similarly, Igbinedion and Karaman (2024) proposed an interactive navigation framework where robots request human intervention when uncertainty is high. These approaches highlight the potential for adaptive human-robot collaboration in industrial settings.

2.2 Human Factors in Industrial Automation

Understanding the cognitive and physical demands placed on workers is essential when introducing automation into skilled manual workflows. Studies have shown that augmented feedback can help workers better adapt to non-anthropomorphic robot behavior, particularly in inspection or control tasks (Edraki et al., 2023). Cognitive workload, perceived autonomy, and predictability are consistently identified as key determinants of user acceptance and performance (Parasuraman and Riley, 1997).

In manufacturing environments, Krüger et al. (2009) emphasized the importance of designing robotic systems that fit seamlessly into established routines and role hierarchies. Building on this work, we combined field observations and workload assessments with interviews involving production line workers, managers, and developers, ensuring that the proposed cobot system respects workers' tacit knowledge and work culture.

2.3 Human-Centered Design and Participatory Methods

A growing body of literature supports the integration of participatory design methods in robotic system development. Research by Frennert and Östlund (2014) and Leong et al. (2016) highlights how co-design with workers leads to more acceptable and usable systems, particularly when addressing tasks involving high sensory demands or contextual judgment.

Within HCI, personas, user journeys, and requirements analysis can serve as complementary representations between empirical inquiry and system design. Personas synthesize recurring goals, behaviors, competencies, and concerns into evidence-based representations that help design teams maintain a shared understanding of different user groups (Pruitt and Grudin, 2003). User journey mapping complements personas by locating users' actions, decisions, handovers, and difficulties within a temporal workflow, thereby revealing where support or redesign may be needed (Endmann and Keßner, 2016). Requirements analysis subsequently translates knowledge about users, tasks, and contexts of use into explicit design criteria and system responses (Maguire and Bevan, 2002). This progression is particularly important in HRI because the introduction of robotic functions may redistribute agency, coordination, and safety responsibilities across workers, managers, developers, and safety specialists. In the present study, personas and current-state journeys are therefore used as intermediate analytical representations, while the resulting requirements provide a traceable basis for the proposed collaborative workflows and system architecture.

The GiraffPlus project (Coradeschi et al., 2013), which integrated long-term health monitoring and robotic telepresence into elderly homes, is a noteworthy example of how iterative, field-based development can inform robust, real-world systems. In line with this, the ongoing European project adopts a similar philosophy, emphasizing deployment-driven design, transparency, and user empowerment as critical dimensions of success.

2.4 AI for Industrial Perception and Assistance

Advances in computer vision and machine learning have enabled real-time defect detection systems with high classification accuracy (Rahman et al., 2019; Zhou et al., 2023), yet challenges remain in explainability, edge deployment, and context adaptation. Systems like DeepInspect (Kumbhar et al., 2023) and industrial AI platforms for surface inspection show promise but still require careful integration with human workflows to ensure trust and robustness. Our cobot framework complements such systems by allowing human oversight and providing real-time feedback loops for shared control.

3. METHOD: FIELD STUDY AND DESIGN SYNTHESIS

To understand the complex interplay between humans and automation in defect detection and repair workflows, we employed a multi-method research strategy grounded in human-centered design principles. The study was conducted at a major automotive factory in Türkiye, a real-world manufacturing setting characterized by high production volume, tight cycle times, and well-established work routines.

Our methodological approach was designed to capture both the physical realities of the factory floor and the lived experiences of three key stakeholder groups: 1) production line workers, 2) managers, and 3) system developers. We combined qualitative and quantitative methods, enabling triangulation across multiple data sources, including field observations, structured interviews, and a questionnaire measuring the perceived workload.

The study protocol was reviewed and approved by the institutional ethics board at Koç University under protocol number: 2023.431.IRB3.184, and all participants provided informed consent.

3.1 Field Observations

Field observations were conducted across multiple sections of the factory, including the Bodyshop and the Quality Control Press areas, to examine production line workers' tasks in real time. The observations focused on work processes, interactions with materials, and how workflow was maintained throughout a work cycle. To capture variations in work practices and interaction patterns, data collection was carried out across multiple shifts. Site access and scheduling were coordinated through Tofaş, a partner in the MAGICIAN project, and approved by the factory's HR department. The observations were conducted by Günter Alce, an experienced interaction design researcher. The factory operates in three shifts: 08:00-16:00, 16:00-00:00, and 00:00-08:00. The following sections present in more detail the observations conducted on work roles and tasks in Bodyshop defect detection, Bodyshop repair, Quality Control Press defect detection, and Quality Control Press repair areas.

3.1.1 Work Roles and Tasks - Bodyshop Defect Detection Area

The observations were conducted through direct oversight of the production line workers' work environment. The observer was always at a safe distance to minimize interference while ensuring detailed observation. The structured notes template captured station details, station layout, work roles, number of personnel, time-stamped task sequences, tool usage, body postures, communication events, and tasks performed.

A total of eight individuals participated in the observations, including three females. While the number of observed participants is limited, the goal of the field study was not statistical generalization but analytic generalization, focusing on identifying recurring patterns in work practices and interaction structures across stations and shifts. The study included observations of two production line workers for approximately three hours during the morning shift (1st shift) at the Bodyshop and two production line workers for about 1.5 hours in the afternoon shift (2nd shift) at the same location. Additionally, two production line workers were observed for approximately two hours at the Press in the morning, and another two production line workers for about one hour at the Press in the afternoon.

The production line workers were responsible for inspecting the vehicle's body structure in the Bodyshop, focusing on detecting defects in the sheet metal. Their inspection process combined visual examination with manual manipulation to identify any irregularities. A mirror was used to check ceiling details that were otherwise difficult to see.

One production line worker was responsible for inspecting the vehicle's hood and the front side areas near the wheels. This production line worker also recorded all detected defects, which were verbally communicated by the other two production line workers during the inspection.

The documentation was carried out on paper and digitally, ensuring accurate tracking of any identified issues.

Each team member had a defined role, contributing to a structured, streamlined process. Inspection tasks relied heavily on the production line workers' senses, particularly vision and touch, to identify potential defects. The workers' ability to detect irregularities through experience and trained perception was considered crucial for maintaining high-quality standards in the production process.

Each production line worker carried a marking pen and a stone, which were essential tools in the defect detection process. If a production line worker was uncertain about a potential defect, they lightly sanded the area with the stone to reveal any irregularities more clearly. Once a defect was confirmed, it was marked and coded using the pen. The defect was then verbally announced to ensure that all team members were aware and that it was properly documented and addressed.

The vehicle remained stationary for approximately 50 to 60 seconds during the inspection. If the production line workers completed their tasks ahead of time, the worker responsible for inspecting the hood manually advanced the vehicle to optimize workflow efficiency. Additionally, a safety system was in place to prevent accidental vehicle movement. If someone stood in front of or behind the vehicle, a designated area with a rubber mat and an integrated sensor detected their presence, preventing the vehicle from moving forward.

During the inspection, doors were also checked, including interior and exterior metal surfaces. The process was carried out with complete focus. No unnecessary conversation occurred, and workers only reported their findings. A logbook followed each vehicle, ensuring that all defects and actions were properly recorded.

When a new vehicle model was introduced, a female production line worker initially handled the stamping process. However, it was later clarified that this was a natural task reassignment, meaning that responsibilities were informally redistributed within the team based on perceived workload and experience, without managerial intervention. The worker responsible for inspecting the hood had identified and conducted a more thorough inspection, which required additional assistance, leading to the role adjustment.

From an outsider's perspective, defect reports often seemed mumbled. They could be announced at any moment, raising the question of how the report-responsible production line worker managed to hear and remember them. When asked about this, the response was that experience played a crucial role. Production line workers immediately recognized the type of defect based on the way it was reported, and depending on who announced it, they instinctively knew its location. The workers mainly used codes when they reported, and many of these defects were recurring, making the processing of the information more efficient.

3.1.2 Work Roles and Tasks - Bodyshop Repair Area

Three production line workers were responsible for repairing the defects identified in the previous inspection stages. Unlike the inspection process, no women participated in this stage. According to line managers, current role distributions in this section are influenced by a combination of task demands, internal allocation practices, and worker preferences.

All production line workers wore protective gear, including safety glasses, gloves, long pants, long-sleeved shirts, and arm protection. However, no hearing protection was used in this area.

The repair process involved a variety of tools, including hammers, grinders, screwdrivers, nut tighteners, and metal plates for bending. Before addressing the reported defects, production

line workers quickly checked using their hands and eyes to verify the issue. Repairs were performed within the main workflow, with at least two empty stations left between sections to create space for the repair team. In cases where more significant defects were detected, the vehicle could be moved to the side for more extensive work. Production line workers also had to assess tool wear during the process. For example, they relied on experience to determine when the sandpaper was worn and replaced it “on the fly” without disrupting their workflow.

3.1.3 Work Roles and Tasks - Quality Control Press Defect Detection (Larger Parts)

In this area, the production line workers are positioned along a large press line where stamped metal parts, such as door panels and other large body components, move along the conveyor to their stations. Two production line workers perform inspections, holding sandpaper in one hand and running their other hand over the surface to detect defects. One production line worker rotates the part while inspecting.

There are six control points, and after inspection, the larger parts are hung on racks. For larger parts, two production line workers are required for handling, whereas smaller parts can be managed by one person.

Every 200th part is examined more thoroughly, which means that one dedicated worker is responsible for conducting thorough quality control, counting all holes, checking if they are correctly placed, and conducting careful inspections that focus on detecting damages along the edges and other critical areas. The tools used include a stone, a marking pen, and sandpaper. These detailed inspections take approximately 10 minutes per part.

3.1.4 Work Roles and Tasks - Quality Control Press Defect Detection (Smaller Parts)

In this area, the production line workers are positioned along a smaller press line where small stamped metal parts move along the conveyor to their stations. The inspection team consists of one female and two male production line workers. A camera captures images of the parts as part of the quality control process.

Two production line workers handle the control process, with one responsible for hanging the parts. Additional control is performed every 200th part, like the larger parts, with the inspector moving between two identical machines.

No tools are used in this section. Once a batch is completed, a forklift picks up the finished parts. Production line workers visually inspect the parts as a final check.

Extra inspections focus on counting holes, which takes approximately three minutes per part. A proposed improvement is to use the MAGICIAN cobot's computer vision module to automate hole counting, which could be applied to both large and small parts.

3.1.5 Work Roles and Tasks - Quality Control Press Repair Area

Two male production line workers are responsible for repairing defects found in the press stage.

The work environment is calm, with multiple sanding machines hanging in place, allowing production line workers to access them from any position easily. The defects are often repetitive, meaning the same type of issue can occur across multiple parts in a batch.

Various tools are used in the repair process, including hammers, support stones, planers, sandpaper, sanding machines, and manual inspection using fingers or gloves to detect irregularities.

Production line workers wear protective equipment, including gloves, long sleeves, arm guards, and safety glasses. In the Press areas, hearing protection is mandatory due to high noise levels.

Different grades of sandpaper are used depending on the task. There are 6-7 different sanding machines in operation, spread across four workstations to maximize efficiency. If an adjustment is needed at one station, the work can be transferred to the next available table.

Key insights from observations include:

- **Collaborative yet specialized roles**: Workers operated in well-defined teams, each focusing on specific vehicle components. This specialization was considered important for efficiency and reducing redundancy, allowing for rapid defect detection and correction.
- **Visual and tactile inspections**: A combination of sensory inputs ensured accurate defect detection, emphasizing the importance of human expertise in quality control. Workers used their hands and eyes in a coordinated manner to identify even minute irregularities in the vehicle structure.
- **Adaptive workflows**: Workers adjusted their inspection strategies based on defect type and location, showcasing the flexibility of human-driven quality control. The ability to adapt to varying defect patterns and respond dynamically was a frequently observed characteristic among workers.

These findings suggest that automation should be designed to enhance rather than replace existing inspection workflows, enabling workers to maintain control over quality assessments. Systems should support human strengths, such as sensory perception and adaptive problem-solving, rather than imposing rigid automated processes. There are clear opportunities where a cobot could assist by taking over inspection tasks that require workers to bend, stretch, or reach into ergonomically challenging positions, thereby reducing physical strain. Additionally, the system could incorporate audio capture to automatically record verbal defect reports and integrate them with existing quality tracking systems.

3.2 Structured Interviews

Participants were recruited through HR and line managers using purposive sampling to cover production, management, and development roles. Within each group, convenience scheduling was applied across shifts while seeking variation in tenure and station type. Structured interviews were conducted to examine how production workers, developers, and managers perceived automation and robotics, including their experiences, expectations, and attitudes toward these technologies in their respective roles.

A total of 31 individuals were interviewed (five females): twelve production line workers (three females), nine managers and ten developers (two females, five of the developers were from Altinay). The overall average age of the participants was $M = 34.8$ ($SD = 6.08$).

The average age and standard deviation for each group were calculated:

- Production line workers: Mean age = 34.5, $SD = 5.13$
- Managers: Mean age = 40.2, $SD = 3.77$
- Developers: Mean age = 30.2, $SD = 5.03$

Prior to the interviews, a set of structured questions was prepared. These questions were reviewed and approved by HR at a major automotive factory in Türkiye to ensure relevance and appropriateness.

The interview guide contained role-specific questions for production workers, managers, and developers. Shared topics included current tasks, prior experience of automation, expected benefits and risks, trust, safety, training, job change, and desired forms of system feedback. Role-specific prompts addressed sensory inspection and reporting for workers, investment and implementation decisions for managers, and integration, reliability, cybersecurity, and maintainability for developers. Interviews were conducted in person, documented with participant consent, and organized by stakeholder group to preserve differences in role and decision authority.

3.2.1 Interview Procedure

The interviews were conducted in a comfortable setting equipped with air conditioning to ensure a relaxed environment for the participants. Prior to each interview, participants signed a consent form, confirming their voluntary participation and understanding of the interview process.

3.2.2 Production Line Workers

Interviews with production line workers revealed a complex relationship between automation and job satisfaction. While robots were seen as beneficial for reducing physical strain and improving efficiency, concerns about job security and skill erosion persisted. Workers emphasized the importance of maintaining autonomy over task allocation and expressed a preference for automation that supports rather than dictates workflow decisions.

Key themes from worker interviews:

- **Workplace transformation**: Improvements in lighting, ventilation, and ergonomics were linked to automation, enhancing worker comfort and productivity. One worker stated: *We did not have fans before, it was hot. Now we have fans, better lighting, and robots take care of some heavy jobs.*
- **Efficiency vs. autonomy**: While robots improve consistency and defect detection accuracy, workers preferred retaining decision-making control over quality assessments. One worker said: *The system does it all, click, click, click. But if there's a quality problem, they still come back to me.*
- **Job security concerns**: Many workers feared job displacement as automation adoption increased, underscoring the need for transparent communication about workforce transitions and upskilling opportunities. One worker stated: *Robots came in, and there were reductions. Improvements were made, but a lot of workers left.*

3.2.3 Managerial Perspectives

Managers acknowledged the cost-driven nature of automation adoption, emphasizing the need to balance efficiency gains with workforce stability. AI-driven defect detection was identified as a key area for future investment, but concerns about system reliability and worker acceptance were noted.

Insights from manager interviews:

- **Strategic automation adoption**: Cost reduction was a primary driver, with automation decisions balancing labor costs and efficiency gains.
- **Data-driven quality control**: AI and sensor-based defect detection were seen as promising but required further refinement to align with production realities.
- **Workforce impact**: Managers emphasized upskilling initiatives to mitigate worker displacement and support collaboration between humans and machines.

3.2.4 Developer Insights

Developers highlighted the rapid evolution of AI technologies and the challenges of integrating new systems into established workflows. Ethical considerations, particularly related to job displacement, were acknowledged but often secondary to efficiency gains.

Key developer perspectives:

- **Technology adaptation challenges**: Rapid technological advancements necessitated continuous learning and adaptation.
- **Human-centered AI**: Developers stressed the importance of designing AI tools that complement human decision-making rather than replacing it.
- **Risk mitigation**: Safety and reliability remained critical concerns, requiring robust fail-safes and transparency in AI decision-making.

3.3 NASA-TLX Workload Analysis

In addition to the observations and interviews, workers from the Bodyshop and Press completed the NASA Task Load Index (NASA-TLX) questionnaire (NASA, 2020). The instrument assesses perceived workload across six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration. The optional pairwise weighting procedure was not used because it may reduce experimental validity (Hart, 2006). We therefore report unweighted ratings to describe the factors contributing to workload in the two production areas.

A total of 27 participants took part in the study. The average age of the participants was $M = 35.8$ years, with a standard deviation of $SD = 5.76$ years. The group consisted of 21 males and six females. Participants were distributed across two workplace sections: 10 participants worked in the Press section, while 17 participants were from the Bodyshop section. The average age of the participants from the Press section was $M = 34.3$, with a standard deviation of $SD = 6.68$, while the average age of the participants from the Bodyshop section was $M = 36.7$, with a standard deviation of $SD = 5.14$. There were three females from each section. The NASA TLX scores are illustrated in Figure 1a, the mean NASA TLX score was $M = 41.9$, $SD = 6.51$.

The three data collections had different participant groups and purposes. Observations involved eight production workers, four in the Bodyshop and four in the Press. Interviews involved 31 people across roles: 12 production workers, nine managers, and ten developers. The NASA-TLX questionnaire was completed by 27 production workers, 17 in the Bodyshop and ten in the Press. The available study records do not establish which individuals participated in more than one data collection, or how the interviewed workers were distributed between the two production areas. Accordingly, these counts should not be added to obtain a total number of unique participants, and the findings are triangulated across methods at the group and workflow level rather than linked person by person.

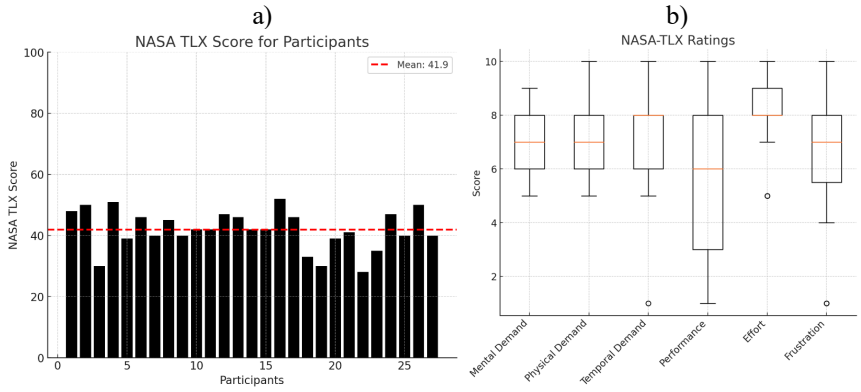


Figure 1. a) Individual unweighted NASA-TLX total scores (M = 41.9, SD = 6.51);
b) Distribution of NASA-TLX ratings across all six subscales

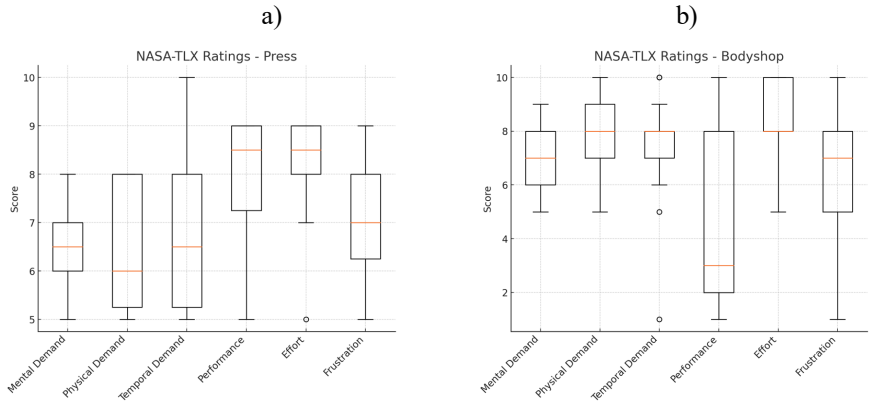


Figure 2. a) NASA-TLX subscale ratings for participants in the Press section;
b) NASA-TLX subscale ratings for participants in the Bodyshop section

The Temporal Demand and Effort sub-scores had the largest median MD = 8.0. The performance sub-score had the lowest value, MD = 6.0. While the Mental Demand, Physical Demand, and Frustration had MD = 7.0. These values are considered to be high values, thus indicating a high perceived workload.

In the Press section, performance and effort had the highest median ratings (MD = 8.5; Figure 2a), followed by frustration (MD = 7.0). Mental demand and temporal demand each had a median of MD = 6.5, while physical demand had the lowest median rating (MD = 6.0).

In the Bodyshop, physical demand, temporal demand, and effort had the highest median ratings (MD = 8.0; Figure 2b). Mental demand and frustration each had a median of MD = 7.0, while performance had the lowest median rating (MD = 3.0).

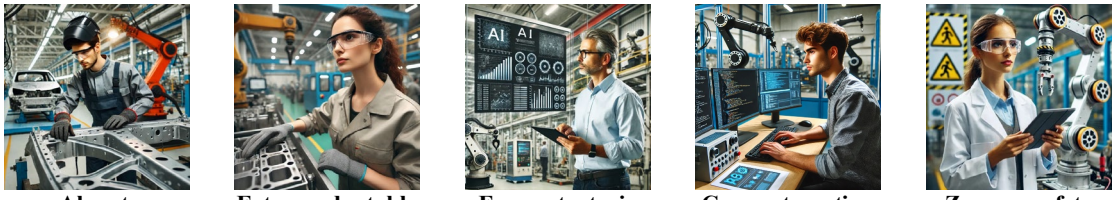
The subscale profiles indicate greater physical and temporal demands in the Bodyshop than in the Press. The largest difference concerned performance, with a median of MD = 3.0 in the Bodyshop and MD = 8.5 in the Press. This contrast may reflect differences in perceived task performance, but it may also reflect how respondents interpreted the performance scale. We did

not conduct follow-up debriefs to establish how individual workers understood its direction. We therefore report the performance ratings descriptively and do not treat the difference as evidence of a difference in actual task performance. The comparatively high Bodyshop ratings for physical demand, temporal demand, and effort are consistent with the observed postures, manual inspection activities, and strict production rhythm.

Given the exploratory nature of the study and the limited sample size, the NASA-TLX data were analyzed descriptively rather than inferentially. The results should therefore be interpreted as indicative workload patterns rather than statistically generalizable effects.

3.4 Personas and Their Design Value

Personas were developed to translate recurring patterns from the observations and interviews into accessible representations of the people who will use, supervise, develop, or be affected by the proposed automation (Figure 3). Rather than describing individual participants, the personas synthesize shared goals, frustrations, motivations, and needs across production workers, managers, developers, and safety specialists. Their design value is to keep heterogeneous stakeholder perspectives visible when requirements and system functions are discussed, and to reduce the risk that technical decisions are made in isolation from everyday work practices.



Ahmet, experienced worker | **Fatma, adaptable newcomer** | **Emre, strategic leader** | **Can, automation developer** | **Zeynep, safety engineer**

Figure 3. Evidence-based personas representing production workers, management, and development roles. All the images are generated using Generative AI

Ahmet represents experienced Bodyshop workers whose goals are defect-free production, personal safety, and preservation of quality standards. His expertise is embodied in tactile examination, learned defect codes, and rapid interpretation of weak signals. His central concern is that automation could erode both job security and the recognition attached to craftsmanship. For this persona, assistance must expose uncertainty, accept correction, and make expert contribution visible rather than presenting machine output as final.

Fatma represents a newer Press worker who is motivated by learning, career progression, and mastery of manual and automated inspection. Her frustration is the lack of formal preparation for advanced automation. She needs situated guidance and opportunities to progress from routine operation to interpretation and troubleshooting. Emre represents operational management, balancing productivity, cost, quality, employee development, and change management. He requires traceable performance information and training plans that connect technical implementation to workforce transition.

Can represents developers who prioritize scalable and accurate automation but may have limited access to the tacit details of line work. He needs direct, structured worker feedback and field-based validation. Zeynep represents safety engineering, with responsibility for monitoring,

fail-safe behavior, and alignment between rapid technical change and safety standards. She needs real-time performance data and collaboration across production, development, and safety roles. Together, the personas reveal that the same feature can have different meanings: automation confidence is a model metric for Can, an override cue for Ahmet, a learning opportunity for Fatma, a governance signal for Emre, and a safety input for Zeynep.

Used in this way, the personas support three connected activities: interpreting field evidence, comparing design alternatives, and reviewing whether the resulting requirements serve all affected roles. They are therefore analytical and communication tools, not claims that five fictional profiles represent the full diversity of the workforce.

3.5 Current-State User Journeys

This section presents user journey visualizations by creating a workflow chart from the workers' and managers' perspectives. This section aims to understand the current defect detection workflow for designing effective and user-centered AI-driven solutions. The current worker journey begins with a visual and tactile inspection of the vehicle body. Workers use their hands and a stone or abrasive tool to reveal subtle surface irregularities, then decide whether the body is clean. When a defect is confirmed, it is marked on the vehicle, communicated through concise defect codes, recorded on paper and in digital systems, and, when required, routed to a repair station for re-evaluation and correction. Figure 4 makes visible three points at which human expertise is indispensable: detecting weak signals, classifying uncertain findings, and deciding whether escalation is required. It also shows where workload accumulates through repeated scanning, time pressure, simultaneous communication, and duplicate documentation.

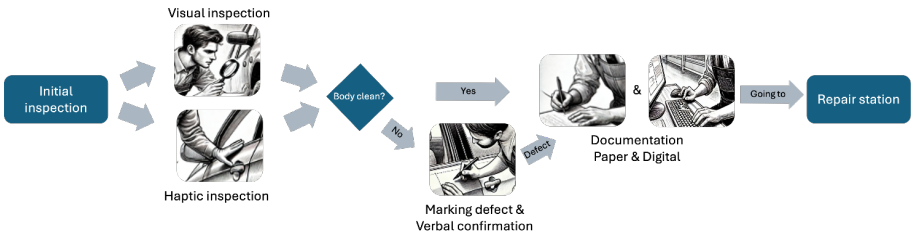


Figure 4. Defect detection workflow before MAGICIAN

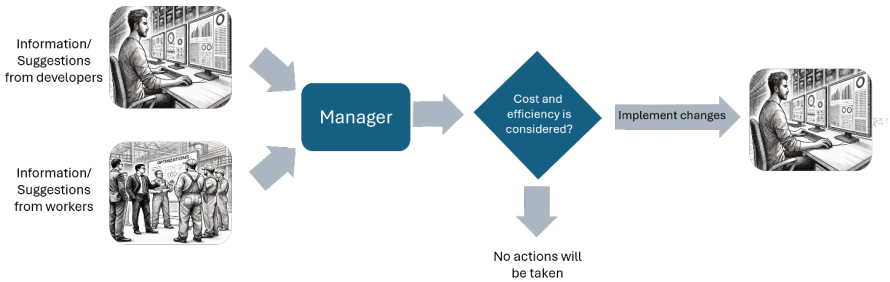


Figure 5. Current managerial workflow for implementing or optimizing production technology

The current managerial journey complements the worker journey. Managers gather production reports and feedback from workers, engineers, and team leaders, assess whether a proposed change meets cost, quality, safety, and efficiency objectives, and either approve implementation or request further refinement. After implementation, the team iteratively reviews performance data and operational feedback. Figure 5 shows that decisions about automation already depend on both quantitative information and situated expertise, but that cost-benefit assessment, cross-role communication, and follow-up can delay or constrain change.

3.6 Requirements Synthesis

Requirements synthesis converted empirical findings into implementable design directions while retaining their evidential basis. Each row in Table 1 begins with an observed or reported work condition, articulates the associated human or organizational need, states a reader-facing requirement, and identifies a corresponding system response. This structure shows where a proposed feature is grounded in field evidence and where later prototyping and validation are still required.

Table 1. From field evidence to human-centered requirements and proposed system responses

Req.	Evidence	Finding or need	Requirement statement	System response
R1	Observation, interviews	Inspection is sensory, time-critical, and embedded in an established station rhythm.	Assist within the existing station and augment visual or tactile inspection without relocating or replacing workers.	Workflow-integrated inspection with retained visual and tactile operator verification.
R2	Observation	Defects are announced through concise codes and recorded on paper and digitally.	Integrate with current logging practices and avoid duplicate data entry.	Shared defect record and interface integration.
R3	NASA-TLX, observation	Temporal demand and effort were high, especially in Bodyshop.	Use AI-supported detection to reduce repetitive analysis while retaining complex judgment for workers.	Automated screening and human review of uncertain cases.
R4	NASA-TLX, interviews	High-tempo coordination creates cognitive strain and frustration.	Provide real-time guidance and decision support without adding monitoring burden.	Task status, alerts, and prioritized guidance.
R5	Interviews, observation	Workers want automation that can be adapted to task and operator needs.	Allow adjustment of response speed, precision, and level of autonomy.	Configurable assistance and shared-control settings.
R6	Interviews	Trust depends on understanding what the system is doing and why.	Communicate ongoing tasks, alerts, confidence, and reasons for changes.	Transparent interface with confidence and rationale.

Req.	Evidence	Finding or need	Requirement statement	System response
R7	Observation, literature	Unexpected robot motion can disrupt fluent work and reduce comfort.	Use predictable movement patterns and notify workers of deviations.	Legible motion, timing cues, and safe fallback behavior.
R8	Interviews	Workers want influence over how automation changes their work.	Include workers in introduction, refinement, and validation of system functions.	Iterative co-design and feedback loop.
R9	Interviews, field study	Hands-on learning is preferred and roles will require new skills.	Provide situated guidance and progression toward troubleshooting competence.	Step-by-step work support and training mode.
R10	Interviews, observation	Safety needs and confidence vary across users and tasks.	Support adaptive safety zones and customizable safety alerts.	User-aware safety configuration within certified limits.

The synthesis exposes two complementary flows. The first is operational: detection, confidence estimation, operator confirmation, logging, and repair allocation. The second is organizational: worker participation, configuration, training, and iterative review. Treating both flows as part of the system boundary avoids reducing human-centered design to interface features alone and prepares the transition from the current-state journeys in Section 3.5 to the proposed journeys in Section 4.3.

4. SYSTEM OVERVIEW FOR DEFECT DETECTION AND REPAIR

To operationalize the principles identified through our field study, we propose a modular system architecture for human-in-the-loop defect detection and repair using collaborative robots (cobots). Our system employs two Doosan H2515 cobots configured for complementary tasks: defect identification and defect repair, working alongside human production line workers.

4.1 System Overview

The proposed architecture combines tactile sensing, computer vision, and AI for predictive and adaptive defect detection. One cobot is equipped with the camera system from the ongoing European project, using a SONY XCG-CP510 global shutter polarization 5.1-megapixel CMOS sensor with a GigE interface, together with a tactile sensor (e.g., GelSight) for car body panel inspection. It follows a pre-planned trajectory and adapts inspection granularity based on defect probability maps, which are continuously updated through an AI learning loop informed by historical patterns and real-time data. The AI learning loop is based on supervised machine learning models trained on historical defect datasets, periodically updated using new inspection data.

Human workers collaborate closely with the inspection cobot, using their visual and tactile expertise to supplement the automated sensing process (R1, R3). A separate computer vision-based human tracking system supports safe human-robot collaboration by dynamically adjusting robot behavior and maintaining appropriate safety zones (R7, R10).

When a defect is detected, the system records its location and classification in a shared defect record intended to integrate with existing documentation practices (R2). A second cobot is deployed to address the defect using a grinding tool for minor surface corrections. Current hardware limitations restrict the repair cobot from performing high-impact operations (e.g., hammering), but minor surface polishing and grinding are supported.

The current contribution focuses on system requirements and interaction design rather than full technical implementation. Moreover, as the system is currently under development, quantitative performance metrics such as detection accuracy, task completion time, and repair success rate are not yet available. These metrics are planned as primary evaluation criteria in the next project phase.

4.2 AI-Assisted Defect Localization

Defect localization is enhanced through the use of a machine learning model trained on historical factory data to predict high-probability defect zones. When a previously detected defect recurs in the same region, the inspection cobot prioritizes that area. This model enables proactive inspection strategies, reducing the overall cycle time and improving consistency. The prioritization is intended to support workers by focusing attention on likely defect areas, while uncertain or context-dependent cases remain subject to operator review (R3, R4, R6).

4.3 Proposed User Journeys with MAGICIAN

4.3.1 Worker Journey

The proposed worker journey retains the current human decision structure while adding cobot-supported scanning, automated recording, and AI-based suggestions (R1, R2). The inspection cobot scans designated areas and presents the detected location, classification, and confidence to the operator (R4, R6). High-confidence routine findings can be recorded consistently, while ambiguous cases are escalated for visual or tactile verification. The operator can confirm, correct, or reject the output before the item proceeds or is allocated for repair (R3, R6). Figure 6 therefore describes a complementary workflow: automation addresses repetition, awkward reach, and documentation burden, while the worker remains responsible for contextual judgment and exceptions.

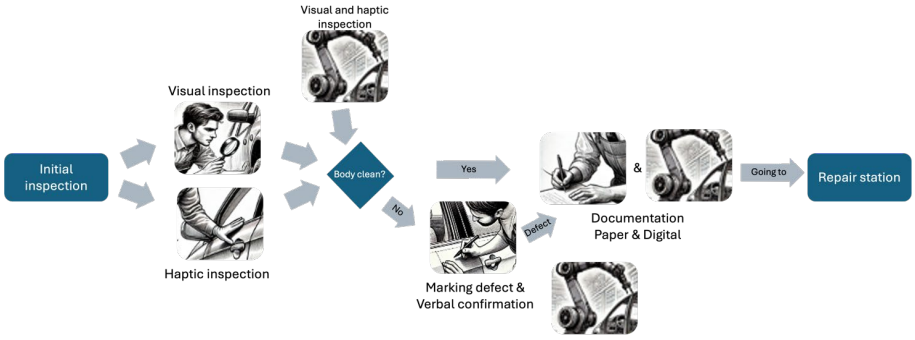


Figure 6. Proposed worker journey with the MAGICIAN cobot complementing human inspection

4.3.2 Manager Journey

The proposed managerial journey adds AI-supported performance alerts, consolidated defect records, and scenario-based cost-benefit information to the existing decision process (R2, R4). Managers still combine quantitative evidence with feedback from workers, developers, and safety specialists, but can identify recurring problems earlier, compare proposed adjustments, and monitor their effects after deployment (R5, R8). As shown in Figure 7, this is not automated managerial decision-making. Recommendations remain subject to human review, organizational priorities, and iterative feedback from the production environment.

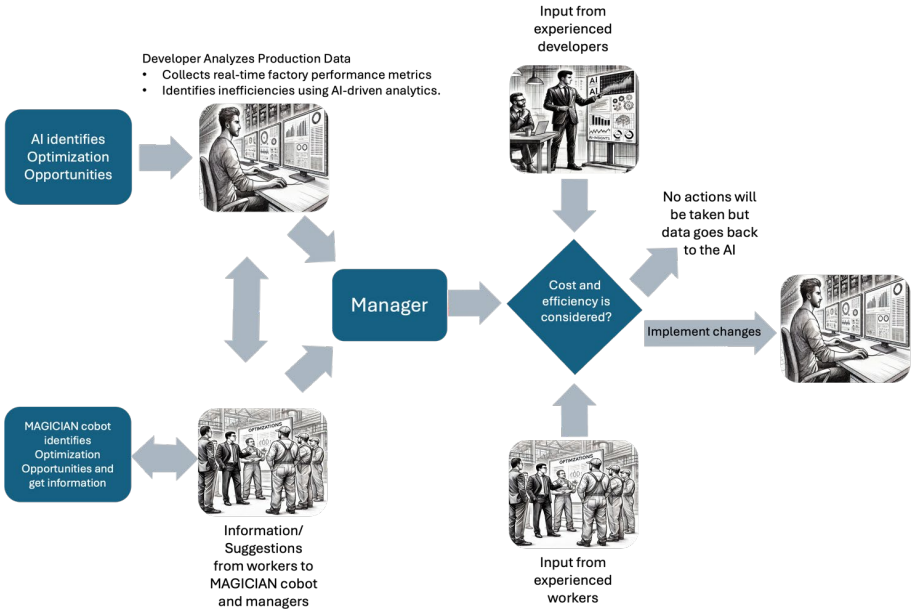


Figure 7. Proposed managerial journey for implementing or optimizing production technology with MAGICIAN

4.4 Human Oversight and Interaction

A human operator supervises the entire process and works side-by-side with the cobot during inspection. They may also verify or refine defect classification. The system provides real-time feedback via a screen-based interface, ensuring clarity and traceability of system behavior. A vision-based safety layer ensures robots dynamically yield space or pause activity to maintain human safety. This design supports a human-in-the-loop paradigm and aligns with findings from our field study emphasizing trust, skill preservation, and adaptability.

5. DISCUSSION

The multi-method study yielded a number of insights into the human factors, operational constraints, and design considerations relevant to collaborative defect detection and repair. This section integrates the findings across methods and considers their implications for the system development and deployment.

5.1 Observations

Our field observations revealed that workers perform inspection tasks using a highly developed blend of visual, tactile, and auditory cues. For example, in the Bodyshop, surface irregularities were detected not only by sight but also by gliding fingers across the sheet metal or using the stone tool to reveal any irregularities more clearly. These nuanced practices are not easily captured by conventional automation systems, underscoring the need for cobots that can defer to human judgment when ambiguity arises. However, certain inspection tasks that require workers to stretch, bend, or reach into awkward positions could be delegated to the MAGICIAN system, reducing physical strain while maintaining inspection quality.

Workflow rhythms were tightly coupled to the production cycle. Coordination occurred through minimal verbal communication and implicit role division. This calls for cobots that integrate into these routines without introducing latency or cognitive overhead, this is consistent with findings by Lasota et al. (2017).

5.2 Interviews

Across 31 interviews, workers expressed interest in automation that reduces repetitive strain but emphasized the importance of retaining control over defect classification. Common concerns included job displacement, over-reliance on opaque AI systems, and increased documentation burden. Managers and developers, in contrast, focused on efficiency and system integration, often underestimating the value of human oversight.

This divergence highlights the importance of transparent system behavior and explainability, both central to building trust in automation (Hancock et al., 2011; Ribeiro et al., 2016). Workers were more willing to engage with robotic systems when they could verify system decisions or intervene when needed, confirming the relevance of shared autonomy frameworks (Dragan et al., 2013).

5.3 NASA-TLX Questionnaire

The NASA-TLX ratings showed descriptive differences between the Bodyshop and Press sections. Physical demand, temporal demand, and effort had the highest median ratings in the Bodyshop (*MD* = 8.0), while performance had the lowest (*MD* = 3.0). In the Press, performance and effort had the highest median ratings (*MD* = 8.5). The contrast in performance ratings is difficult to interpret because participants' understanding of that scale was not checked in a follow-up debrief. We therefore draw no conclusion about differences in actual performance between the sections. The physical and temporal demand ratings, considered alongside the observations, provide clearer evidence of demanding work conditions in the Bodyshop.

5.4 Implications for Deployment

Our findings highlight that the deployment of cobots in automotive defect detection and repair cannot be reduced to technical efficiency alone. Human inspectors rely on a rich integration of sensory channels, such as vision, touch, and hearing, and complement this with experience-based intuition and tacit knowledge. These elements are not easily codified in algorithms but remain essential for accurate and confident decision-making.

From a deployment perspective, the lesson is that collaborative systems must be designed to preserve and extend these human strengths rather than substitute them. Fischer and Frennert (2025) similarly emphasize the importance of transparency, explainability, and role clarity in collaborative settings. Artificial intelligence can contribute by recognizing recurring defect patterns, monitoring ergonomically demanding or hard-to-reach areas, and providing consistent logging of verbal defect reports. However, human oversight and team-based collaboration are crucial for ensuring contextual judgment, flexibility, and accountability.

Equally important is acknowledging the social dimension of production work. Workers value not only their individual skills but also the shared sense of belonging, teamwork, and professional pride that comes from collective inspection and repair. Deployment strategies should therefore strengthen rather than undermine this sense of community. Transparent interfaces, opportunities for skills development, and shared autonomy mechanisms will help workers feel empowered rather than displaced.

In practice, this means deploying cobots as complementary partners: fast in repetitive or physically demanding tasks, while leaving space for human workers to apply judgment, adapt strategies, and sustain the collaborative rhythm of the production line. By balancing AI-driven pattern recognition with human sensory intelligence and valuing team cohesion, manufacturers may help balance both rapid production and sustainable acceptance of automation.

From our findings we derive three robotics-oriented imperatives: confidence-aware perception with explicit operator override, a transparent and traceable interface, and ergonomically targeted assistance that reduces awkward reaches while preserving worker control, in line with prior research on trust, training, and explainability in human-robot collaboration (Wang, 2023).

5.5 From Evidence to Architecture

The progression from field evidence to the proposed architecture was mediated through personas, current-state user journeys, and requirements synthesis. The personas consolidated

recurring goals, competencies, concerns, and responsibilities across worker, management, development, and safety roles, while the current-state journeys situated these role-specific needs within the temporal and organizational sequence of defect detection, reporting, decision-making, and repair. These representations made it possible to identify where assistance was needed, where automation might interrupt established coordination, and which human and organizational needs should be retained as explicit requirements. The resulting progression is therefore not a direct leap from interview statements to technology. It moves from empirical observations and stakeholder accounts, through role and workflow interpretation, to requirements and proposed system responses.

The experienced-worker persona and the current inspection journey illustrate this progression. Together, they indicate that automated detection must be introduced without displacing tactile verification, removing operator control, or delaying the established station rhythm. These needs are expressed through requirements concerning workflow compatibility, workload reduction, operator review, and transparency (R1, R3, and R6). Architecturally, they motivate the separation of perception, confidence estimation, human review, and action. A confidence threshold consequently functions as an interaction boundary rather than merely as a model parameter: uncertain cases are routed to a worker, who can confirm, correct, or reject the classification before the decision is logged and acted upon.

The management, developer, and safety personas reveal complementary organizational progression. Adoption depends not only on the robot but also on consolidated records, configurable behavior, predictable motion, role-appropriate training, safety adaptation, and opportunities for worker participation (R2, R5, and R7–R10). These needs are reflected in the proposed logging, configuration, safety, training, and feedback components. They should be evaluated together during demonstrator trials using task performance and perceived workload measures alongside observations of coordination, trust calibration, and perceived control. This traceability also supports later evaluation because each proposed component can be assessed against the human or organizational need that motivated it.

It is important to distinguish between what the study demonstrates empirically and what it proposes as future system design. The empirical contribution consists of the observed inspection and repair practices, the accounts provided in 31 interviews, and the descriptive NASA-TLX workload patterns. Together, these data support the identified needs concerning sensory expertise, time pressure, physical strain, worker control, understandable automation, compatible documentation, and participation in deployment. The personas, current-state journeys, and requirements synthesis are analytic design representations derived from this material. They have been reviewed as part of the project process, but they should not be interpreted as evidence that a particular technical configuration will produce the intended effects. By contrast, the confidence-aware detection workflow, modular architecture, adaptive safety configuration, consolidated interface, AI-supported prioritization, and proposed future journeys are design propositions informed by the field evidence and relevant literature. At the present stage, the complete system has not been evaluated in a production environment, and the study therefore does not establish improvements in detection accuracy, cycle time, workload, safety, or acceptance. These outcomes require evaluation after implementation. Separating findings from propositions prevents the architecture from being presented as a validated solution and clarifies the role of the present study: it establishes an empirically grounded design basis and a set of testable expectations for subsequent prototyping and field evaluation.

5.6 Limitations and Transferability

The study was conducted within one automotive manufacturing organization and focused on defect detection and repair activities in selected Bodyshop and Press settings. The findings are consequently shaped by local production routines, workforce organization, reporting practices, available equipment, and the MAGICIAN project objectives. The observational sample was small and intended to reveal situated practices rather than estimate their prevalence. Interview data represent participant accounts and may be influenced by role expectations, organizational priorities, and anticipated consequences of automation. The NASA-TLX analysis is descriptive, based on a limited cross-sectional sample, and cannot establish causal effects or represent workload across all shifts, tasks, or production conditions. In addition, the methods contribute different forms of evidence: observations capture enacted work, interviews capture experience and expectations, and NASA-TLX captures perceived workload. Their convergence strengthens the interpretation of recurring needs, but it does not eliminate uncertainty or make the results statistically generalizable.

Transferability is therefore best understood at the level of design reasoning rather than direct generalization of numerical results. The progression from field evidence to personas, workflows, requirements, and proposed system components can be applied in other industrial settings, provided that each setting is investigated independently. The specific personas, workload profiles, defect categories, station rhythms, and technical parameters should not be transferred without renewed fieldwork and stakeholder validation. Future studies should examine the requirements with larger and more diverse worker groups, including additional shifts and production areas, and compare sites with different levels of automation. Longitudinal evaluation is also needed to determine whether initial acceptance persists after workers gain experience with the system and whether assistance reduces workload rather than shifting it to monitoring, exception handling, or documentation.

5.7 Future Development

Further development will include the first iteration of our system, which will be followed by new studies and user tests. Efforts will focus on refining the AI model using a larger and more diverse defect dataset, integrating force-based feedback to better detect deeper surface anomalies, and enabling mobile base configurations to accommodate larger vehicle models. These enhancements aim to improve the system's coverage, accuracy, and responsiveness across a variety of production environments.

6. CONCLUSION

In conclusion, this study shows how field-based insights can be used to guide the design of collaborative robotics for automotive defect detection and repair. Observations and interviews showed that workers rely on sensory expertise and want to remain in control of the work. The NASA-TLX results also showed high workload, especially in terms of time pressure and effort. We used the fieldwork to develop personas, user journeys, and requirements for the proposed system architecture. At this stage, the architecture is a design proposal grounded in the fieldwork. It has not yet been validated in production. The next step is to implement and test the system in practice.

ACKNOWLEDGEMENT

This research was co-funded by the project MAGICIAN - "iMmersive leArninG for ImperfeCtion detectIon and repAir through human-robot interactioN" (European Union's Horizon Europe programme, under grant agreement ID 101120731), and the European Union. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

Alves, J., Lima, T. M., & Gaspar, P. D. (2023). Is Industry 5.0 a human-centred approach? A systematic review. *Processes, 11*(1), 193.

Coradeschi, S., et al. (2013). GiraffPlus: Combining social interaction and long-term monitoring for promoting independent living. *Proceedings of the International Conference on Human System Interactions* (pp. 578–585). IEEE.

Dragan, A. D., Lee, K. C., & Srinivasa, S. S. (2013). Legibility and predictability of robot motion. *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 301–308). IEEE.

Edraki, M., Maurice, P., & Sternad, D. (2023). Humans need augmented feedback to physically track non-biological robot movements. *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 9872–9878). IEEE.

Endmann, A., & Keßner, D. (2016). User journey mapping: A method in user experience design. *i-com, 15*(1), 105–110. https://doi.org/10.1515/icom-2016-0010

Fischer, B., & Frennert, S. (2025). Towards an experiential ethics of AI and robots: A review of empirical research on human encounters. *Technological Forecasting and Social Change, 219*, 124264.

Frennert, S., & Östlund, B. (2014). Seven matters of concern of social robots and older people. *International Journal of Social Robotics, 6*, 299–310.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human–robot interaction. *Human Factors, 53*(5), 517–527.

Hart, S. G. (2006). NASA-task load index (NASA-TLX): 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 904–908.

Igbinedion, I., & Karaman, S. (2024, May). Learning when to ask for help: Efficient interactive navigation via implicit uncertainty estimation. *2024 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 9593-9599). IEEE.

Kumbhar, A., Chougule, A., Lokhande, P., Navaghane, S., Burud, A., & Nimbalkar, S. (2023). DeepInspect: an AI-powered defect detection for manufacturing industries. *arXiv preprint arXiv:2311.03725*

Krüger, J., Lien, T. K., & Verl, A. (2009). Cooperation of human and machines in assembly lines. *CIRP Annals, 58*(2), 628–646.

Lasota, P. A., Fong, T., & Shah, J. A. (2017). A survey of methods for safe human–robot interaction. *Foundations and Trends in Robotics, 5*(4), 261–349.

Leong, T. W., & Johnston, B. (2016). Co-design and robots: A case study of a robot dog for aging people. *International Conference on Social Robotics* (pp. 702–711). Springer.

Maguire, M., & Bevan, N. (2002). User requirements analysis: A review of supporting methods. In J. Hammond, T. Gross, & J. Wesson (Eds.), *Usability: Gaining a Competitive Edge*, IFIP, Vol. 99 (pp. 133–148). Springer. https://doi.org/10.1007/978-0-387-35610-5_9

NASA. (2020). *NASA TLX*.

Nguyen Ngoc, H., Lasa, G., & Iriarte, I. (2022). Human-centred design in Industry 4.0: Case study review and opportunities for future research. *Journal of Intelligent Manufacturing, 33*(1), 35–76.

Nikolaidis, S., Nath, S., Procaccia, A. D., & Srinivasa, S. (2017). Game-theoretic modeling of human adaptation in human–robot collaboration. *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction* (pp. 323–331).

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253.

Petric, T., Gams, A., Zlajpah, L., Ude, A., & Morimoto, J. (2014). Online approach for altering robot behaviors based on human-in-the-loop coaching gestures. *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 4770–4776). IEEE.

Pruitt, J., & Grudin, J. (2003). Personas: Practice and theory. *Proceedings of the 2003 Conference on Designing for User Experiences* (pp. 1–15). ACM. https://doi.org/10.1145/997078.997089

Rahman, N. N. S. A., Saad, N. M., Abdullah, A. R., & Ahmat, N. (2019). A review of vision-based defect detection using image processing techniques for beverage manufacturing industry. *Jurnal Teknologi*, *81*(3).

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. *In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144).

Segura, P., Lobato-Calleros, O., Soria-Arguello, I., & Hernandez-Martínez, E. G. (2025). Work roles in human–robot collaborative systems: Effects on cognitive ergonomics for the manufacturing industry. *Applied Sciences, 15*(2), 744.

Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics, 55*, 248–266.

Wang, Q., Liu, D., Carmichael, M. G., & Lin, C.-T. (2023). Robot trust and self-confidence based role arbitration method for physical human–robot collaboration. *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 9896–9902). IEEE.

Whitley, K. (2025). *Industry 5.0 brings a shift to human-centered innovation in manufacturing.*

Zhou, Y., Yuan, M., Zhang, J., Ding, G., & Qin, S. (2023). Review of vision-based defect detection research and its perspectives for printed circuit boards. *Journal of Manufacturing Systems*, *70*, 557–578.