POLYPUBLIE
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

**Titre:** Title: Human-centered AI for industry 5.0 (HUMAI5.0) : design framework and case studies

**Auteurs:** Authors: Mario Passalacqua, Garrick Cabour, Robert Pellerin, Pierre-Majorique Léger, & Philippe Doyon-Poulin

**Date:** 2024

**Type:** Chapitre de livre / Book Section

**Référence:** Citation: Passalacqua, M., Cabour, G., Pellerin, R., Léger, P.-M., & Doyon-Poulin, P. (2024). Human-centered AI for industry 5.0 (HUMAI5.0) : design framework and case studies. In Human-centered AI : a multidisciplinary perspective for policy-makers, auditors and users (pp. 260-274). https://doi.org/10.1201/9781003320791-27

**Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

**URL de PolyPublie:** PolyPublie URL: https://publications.polymtl.ca/57479/

**Version:** Version finale avant publication / Accepted version
Révisé par les pairs / Refereed

**Conditions d'utilisation:** Terms of Use: Tous droits réservés / All rights reserved

**Document publié chez l'éditeur officiel**
Document issued by the official publisher

**Titre du livre:** Book Title: Human-centered AI : a multidisciplinary perspective for policy-makers, auditors and users

**Maison d'édition:** Publisher: CRC Press

**URL officiel:** Official URL: https://doi.org/10.1201/9781003320791-27

**Mention légale:** Legal notice: This is an Accepted Manuscript of a book chapter published by Routledge/CRC Press in Human-centered AI : a multidisciplinary perspective for policy-makers, auditors and users on 2024-03-22, available online: https://doi.org/10.1201/9781003320791-27

**Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal**
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

# Human-Centered AI for Industry 5.0 (HUMAI5.0): Design framework and Case Studies

Mario Passalacqua[1], Garrick Cabour[1], Robert Pellerin[1], Pierre-Majorique Léger[2], Philippe Doyon-Poulin[1]

1 Polytechnique Montréal, Department of Mathematics and Industrial Engineering, Montreal, Canada

2 HEC Montréal, Department of Information Technologies, Montreal, Canada

## 0. Abstract

The fourth industrial revolution (Industry 4.0) is characterized by strategies aimed towards process, product, and service improvement through technology interconnectivity, decision-making speed, and automation capacity. These strategies, in which artificial intelligence (AI) plays a central role, emphasize technological advancement to drastically improve performance-related factors in operational environments. However, their development follows a technocentric approach neglecting their profound impacts on human work. For the anticipated benefits of technological advancement to materialize, it is necessary to consider social and organizational factors (e.g., trust calibration, work redesign) in addition to technical factors. In this chapter, we present HUMan-centred AI for Industry 5.0 (HUMAI5.0) as an evolution of the current industrial revolution. It is an innovative 6-step framework to support the development of AI systems that promote human long-term well-being, engagement, and system performance within complex socio-technical environments. HUMAI5.0 builds upon established human factors and ergonomics practices, organizational psychology, and our own experiences working with AI-based technologies. We applied this framework to two use cases in industrial context: error detection in manufacturing and quality control in aircraft maintenance. We were able to develop future work environments that leveraged the best abilities of both parties (human and AI) in a way respectful of the workers. Artifacts of this framework were also helpful to offer a common ground to various stakeholders to understand the potential impacts AI has in their workplace and how best to address it.

# 1. Introduction

Integrating artificial intelligence (AI) in the workplace has created many challenges and opportunities for human work. Increased human-automation collaboration is expected on physical or cognitive tasks. Several disciplines have echoed the fact that these new technologies automate part of the work steps in collaboration with human operators rather than replacing entire professions, such as Information Technologies (Seeber et al., 2020), economics (Frey & Osborne, 2017), work psychology (Parker & Grote, 2022) and human factors & ergonomics (Mueller et al., 2021). The area of Industry 4.0 (I4.0) is at the forefront of the digitalisation of human work wherein AI plays a central role. I4.0 intends to increase production system capabilities in terms of productivity, repeatability, flexibility, real-time monitoring and process standardization (Zheng et al., 2021). This is done by integrating a set of digital, robotic and automated technologies into production (Kadir et al., 2019) and combining different digital solutions together (Zheng et al., 2021). The latest technological advances in I4.0 have increased the capabilities of machines in performing complex, cognitive tasks (Xiong et al., 2022). However, the development of I4.0 technologies follow a technocentric approach (Michael Sony & Subhash Naik, 2020) i.e., it focuses on technology development first and then looks for potential applications (Carayannis et al., 2022). Bibliometric analyses quantified the technocentric directions of I4.0. A recent literature review noted that out of a sample of 4885 studies with a search strategy that included the terms *Industry 4.0* and *Human Factors*, 4849 studies focused on technical factors and 36 on human factors (Passalacqua et al., 2022). This top-down approach often neglects the contextual factors that govern work systems and their potential integration into situated operational practices (Loup-Escande, 2022).

Consequently, increased attention and principles for more effective and safer human-AI collaboration have been proposed by academics, governments, and industrial groups (Kadir et al., 2019; Maddikunta et al., 2022; Neumann et al., 2021). The European Commission launched the Industry 5.0 (I5.0) stream to operationalize human-centricity, sustainability, and resiliency principles to align new technological development towards social goods. New sub-fields emerged in response to the potential harm created by adopting a techno-centered design approach, considering only the technical dimension of AI. First, several studies underlined the potential benefits that I5.0 could bring by relying on potential application cases and by formalising avenues for future research (Jiao et al., 2020; Maddikunta et al., 2022; Seeber et al., 2020); Second a series of frameworks or design principles were proposed to integrate humans and technologies in collaborative work settings (Dubey et al., 2020; Mueller et al., 2021) or to consider human and organizational factors in digital transformation (Liao et al., 2020; Peeters et al., 2020). However, while these efforts have provided initial coarse-grained guidance, these general principles do not consistently translate into everyday design practices or final AI-based products. Design teams require systematic methods that provide fine-grained guidance to design new solutions that promote long-term human well-being, engagement, and system performance within complex socio-technical environments (Roth et al., 2018).

This chapter presents HUMan-centred AI for Industry 5.0 (HUMAI5.0), an innovative 6-steps framework that supports design teams in adopting a human-centered approach for designing and implementing AI-fused systems in the workplace. We apply this framework to two use cases in the I5.0 context. This chapters adopts a human factors' perspective wherein human-centered AI focuses on the collaboration between humans and AI systems in the workplace. The rest of the chapter is organized as follows: Section 2 delves into the literature from which the framework is derived; Section 3 presents the framework; and Section 4 applies it to two use cases.

# 2. Related literature

## 2.1 Existing design methods and frameworks

### 2.1.1 Technology-centred design for AI and limits

AI-infused solutions often follow a techno-centric roadmap where figures of merit estimate how mature the hardware and software components are (Sony & Naik, 2020). When the solution reaches a satisfying level of technological readiness, development teams seek areas and problems where the device could be applied. While this top-down approach narrows the problem space by focusing only on technical details, it neglects the contextual and human factors that govern work systems (Spies et al., 2020). Moreover, techno-centrism often accompanies a desire to automate as many functions as possible to reduce human intervention in industrial processes, a source of uncertainty and failure in this philosophy (Millot et al., 2015). However, AI cannot, at present, fully acquire and reproduce the tacit knowledge, socio-cognitive processes, and "a*dvanced know-how rooted in decades of institutional experience*" (Miller & Feigh, 2019)*;* p.2). This technology could however support or automate specific portion of a task in human-technology collaboration scenarios (Xiong et al., 2022). Complementary design methods that pay equal attention to human, organizational and technical factors in design processes have emerged, which will be discussed in the following sub-section.

### 2.1.2 Complementary design and analysis methods to address techno-centrism shortcomings

Implementing AI solutions in organizations introduces simultaneous social and organizational challenges in addition to technical ones (Kadir et al., 2019; Maddikunta et al., 2022; Neumann et al., 2021). Several authors advocated to adopt human-centered design approaches to overcome these challenges and ensure a successful transition to digitalized work situations (Neumann et al., 2022; (Oppl & Stary, 2019). Human-centered design approaches place end-users at the center of product design (Rapp, 2021). Contrary to technocentric approaches, they are part of an inductive philosophy that considers users' characteristics and their working environment to *"inform and guide the product development process"* (Rapp, 2021; p.1). This sub-section will explore the most relevant in the context of I5.0 whereas Table 1 summarizes their applicability to HUMAI5.0.

### *Systemic methods of human work analysis to inform design*

**Cognitive Work Analysis (CWA)** is a five-phase framework that guides the design and evaluation of socio-technical systems (Vicente, 1999). CWA allows to gain a deep understanding of the constraints of the work domain, the system, the task, and the operator. CWA's first phase, the work domain analysis (WDA), is of particular importance for this chapter. WDA gives task-independent design information, allowing for better flexibility and adaptability to novelty or change. WDA identifies constraints in the work environment throughout five levels of hierarchy (abstraction hierarchy), ranging from most abstract to most concrete: functional purpose (end-goal of the system), abstract function (physics principles), generalized function (processes), physical function (equipment), and physical form (location and appearance) (Rasmussen, 1985). WDA is useful to map the means-end relationship that link to end-goal of the system to its physical implementation.

**Knowledge elicitation** is a family of methods and frameworks aimed at inferring the mental models and tacit knowledge of end-users (Shadbolt & Smart, 2015). Relying on advanced interviews and observation techniques, interviewers rely on timely presented prompts (e.g., pictures, documents, field notes, videotape of expert realizing an action) and systematic questions to guide the subject matter expert in the recall of situation-specific elements. Users can also be asked to think aloud while concurrently performing a task ("please tell me what you are doing and why"). The best-known methods in this category are Knowledge Audit (Gourova, 2009), Critical Decision Method (Klein et al., 1989) and Cognitive Task Analysis (Crandall et al., 2006).

### *Integrative design frameworks for Industry 5.0*

**Systems Framework and Analysis Methodology.** Neumann and his colleagues (2021) proposed a five-step framework to systematically consider human factors in designing and implementing I4.0 systems. They proposed that this framework paints a comprehensive portrait of the socio-technical changes resulting from I4.0-related technology, which allows for specific design and implementation recommendations to avoid dysfunctional Human-Machine Interactions (HMI).

**Motivational Design for Human-Machine Interaction.** Szalma (2014) mobilized self-determination theory's conceptualization of motivation for the design of anthropocentric interfaces or systems. The premise of motivational design is the support of employee autonomy, competence, and relatedness in order to build socio-technical interactions that favor long-term work engagement, performance, and well-being (Szalma, 2009, 2014). He puts forward the idea that considering only hedonic and utilitarian goals is not sufficient for employee performance and well-being. Rather, motivational requirements must be considered when designing or implementing socio-technical systems (Szalma, 2014). Autonomy refers to the perception of being in control of one's own decisions and actions, competence refers to the perception of being capable and effective in one's work, while relatedness refers to the perception of positive/meaningful connections and relationships in the work environment.

**Interdependent analysis** is a framework developed by Johnson et al. (2014, 2017, 2018) to understand and design how human and automated systems can effectively team up to achieve a joint activity together. Instead of focusing on designing the machine architecture in traditional automation design, the method supports exploring systematically the design space of joint human-AI work. The goal is to identify the *interdependent relationships* – when an entity (human or artificial) needs support to complete the objectives of a given task – to design the joint work. Three distinct phases are required to deploy the method:

1) Model the joint activity: define the tasks, subtasks, and capabilities required to achieve the operational objectives
2) Assess team member's capacity to perform and capacity to assist for each task and subtask in the joint activity model
3) Identify all possible human-automation workflows: establish all potential pathways to complete the set of tasks and the associated teamwork requirements

## 2.2 Limitations of the existing frameworks

Each of the presented frameworks has their strengths and limitations. Neumann et al's (2021) present a systematic methodology to properly address human factors in a I4.0 setting. It explains, in detail, how to apply each step of the framework in an I4.0 context. However, its proposed analysis of the technology and the human does not properly capture the complexity of the work domain in complex socio-technical systems. It also mentions the importance of examining the changing psychosocial environment (perception of the social environment), with no concrete recommendations.

The motivational design framework for HMI (Szalma, 2009, 2014) is meant to be an addition to other human factors frameworks instead of being a standalone framework. It provides a way to gain a deeper understanding of systems' psychosocial dynamics compared to Neumann’s et al.’s (2021) framework, which will result in design recommendations to enhance the HMI in socio-technical systems. In addition, it analyzes operator cognitive and affective traits. This analysis, absent from other HMI frameworks, can provide valuable insight into the operators’ cognition at a deeper level of abstraction, ultimately resulting in more specific implementation recommendations. Although thorough, its scope of analysis is mostly limited to psychosocial interactions within a system. Using it in conjunction with other frameworks would thus be ideal.

Work analysis methods, such as CWA or knowledge elicitation frameworks, provide an in-depth understanding of the current state of knowledge and the work domain. However, transitioning from in-depth analysis of the current system to designing the future work situations is not appropriately supported by current standalone methodologies. Additionally, the work domain analysis does not account for the psychosocial characteristics of systems. It should therefore be used with the other frameworks.

Finally, Interdependent analysis is the most comprehensive method that addresses the above drawbacks. As a formative tool, the interdependent analysis framework does not include a summative evaluation step for the proposed design. Its analytical lens is too coarse to grasp the specificities of complex socio-technical work systems, which suggests the need to complement it with other frameworks.

# 3. HUMan-centred AI for Industry 5.0 (HUMAI5.0) Framework

The presented frameworks and methods, found through a review of the literature, each have their benefits and limitations. Each seem to address, to varying depth, different components of the human-AI interaction. The proposed framework aims to mobilize the advantages of each to mitigate the drawbacks of others, creating an integrative framework to comprehensively understand the human-AI interaction.

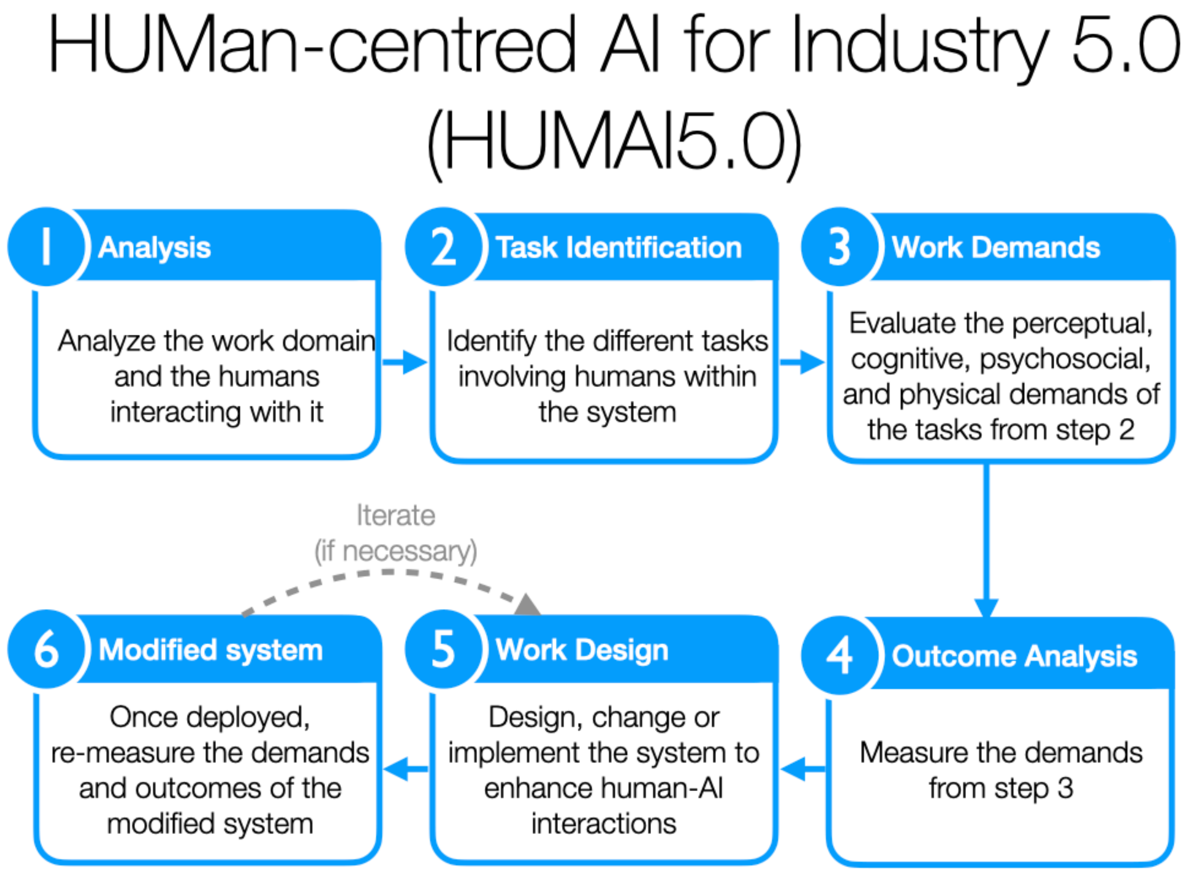


*Figure 1: HUMAI5.0 six-step framework*

**Step 1 – Analysis.** This step involves analyzing the work domain and the humans interacting with it. For this step, we recommend using CWA's work domain analysis (Vicente, 1999) to understand the system's constraints. Other candidates' methods to complete these steps could be People, Activity, Context, Technology (PACT) or Activity Analysis (St-Vincent et al., 2014). More specifically, the WDA identifies constraints in the work environment throughout five levels of hierarchy i.e., abstraction hierarchy (Rasmussen, 1985). We also recommend using step 2 of Neumann et al.'s (2021), which identifies the human roles that interact with the system. Lastly, we recommend using the first step of the motivation design framework, which involves measuring users' cognitive and affective traits. Using these methods offers a more complete understanding of the system's environment.

**Step 2 – Task identification.** This step entails identifying the different tasks involving humans within the system. If applicable, changes in task or job characteristics involving humans within the system should be identified. This refers to examine what tasks or job characteristics are added or removed for all humans involved with the system (e.g., additional sub-system monitoring, reduced physical workload). This is analogous to step 3 in Neumann et al.'s model.

**Step 3 – Work demand analysis.** This step necessitates evaluating the perceptual, cognitive, psychosocial, and physical demands of the tasks identified through the previous step (step 4 in Neumann et al.'s framework). Additionally, we recommend identifying the environment and task features that affect (positively or negatively) operators' feelings of autonomy, competence, and relatedness (e.g., monotony, isolation, decision latitude). For example, task repetition and monotony can hinder these needs, while supervisory support and skill developing programs can facilitate them. This refers to step 2 of the motivational design framework. Lastly, we recommend using the available literature to understand how the identified work demands affect (directly and indirectly) the success criteria.

**Step 4 – Outcome analysis.** For this step, the perceptual, cognitive, psychosocial, and physical demands should be quantitatively or qualitatively measured using various tools (e.g., questionnaires, neurophysiology). Additionally, the relevant outcomes should be measured (e.g., performance, mental and physical health). This step is analogous to step 4 in the motivational design framework and step 5 in Neumann et al.'s framework.

**Step 5 – Work design**. This step involves concretely applying the findings from the previous steps. This is done by designing, changing, or implementing a socio-technical system or interface, in ways that enhance the interaction between humans and other agents within a socio-technical system. When possible, the organizational context should also be tweaked. This step is equivalent to step 5 of the motivational design framework.

**Step 6 – Outcome analysis of new or modified socio-technical system.** Once the system is deployed, this step entails re-measuring the work demands and outcome measures. This is necessary to validate the application of the analyses (step 5). When necessary, the system, interface, and organizational context should be refined or redesigned.

# 4. Case studies

In this section, we illustrate the application of HUMAI5.0 involving the (re)design of human work in collaboration with AI agents in two domains: error detection in manufacturing and quality control in aircraft maintenance.

## a. Error detection in manufacturing

### Context

Within a snowshoe manufacturing plant, the final operator on the assembly line is tasked with error- and defect-detection before the product is sent to the customer. In essence, this final operator is in charge of product quality control. Non-defective snowshoes are put on one side of the workstation, while defective ones are put on the other. Figure 2 shows the workstation used by the operators. Currently, the quality control is being done manually, i.e., without support from any technological system. To aid the operator, the intention is to implement an artificial intelligence-based error and defect-detection system (AIEDS). This AIEDS will use computer vision to detect possible errors and defects within assembled snowshoes and presents its results to the operator on a computer screen at their workstation.

*Figure 2: Workstation used by operators*

Before proceeding with the implementation, HUMAI5.0 was applied to gain a complete understanding of the human-AI interaction. Two versions of the AIEDS were tested and compared to the current method of functioning (No AIEDS). In version A, the AIEDS instructs the operator whether an issue is present. It has a perfect reliability (100%) in terms of error/defect detection and the AIEDS decides whether to send or not the snowshoe to the customer. In Version B, AIEDS has an imperfect reliability (83%) and the decision-making authority is shifted towards the operator who decides to send or not the product to the customer.

## Framework application

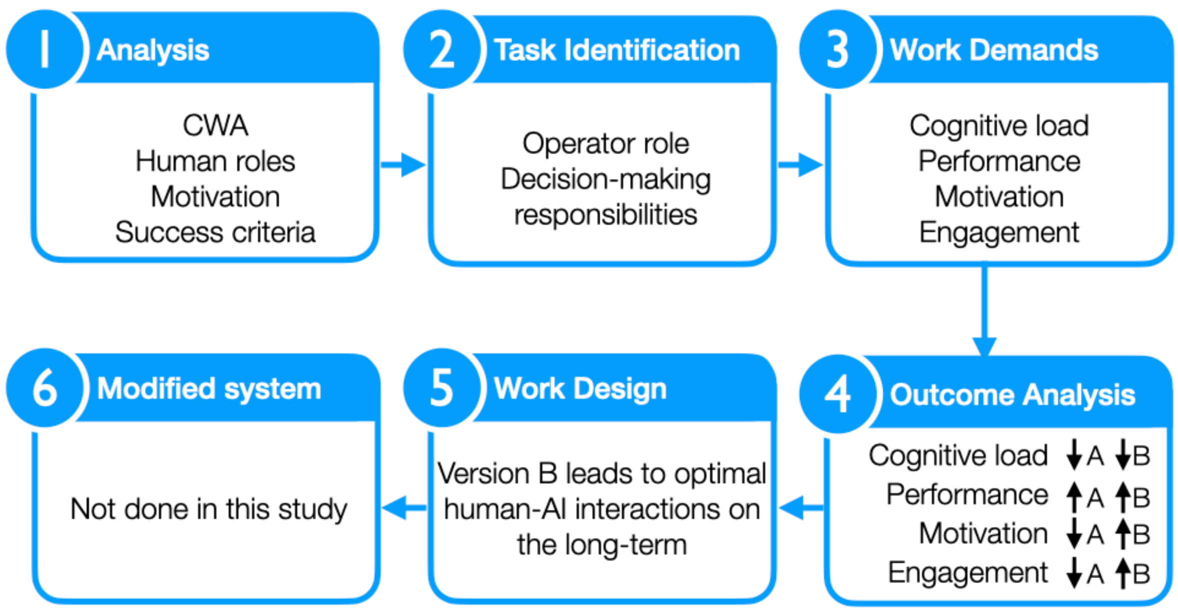


*Figure 3: Framework application to case study #1*

### *Step 1 – Analysis*

CWA's work domain analysis was applied. It allowed us to first delineate the purpose of the system, i.e., customer satisfaction through the delivery of defect-free snowshoes. Then, the domain values were derived (system performance, task engagement, balanced mental workload, balanced stress levels, etc.), followed by the domain functions (mental demand, motivation, image mapping capabilities, depth and geometry understanding, error complexity, training, experience, etc.). Most importantly, we were able to map out the interdependencies between values and functions through means-ends relationship. WDA mapped out that error complexity, training, and experience affect mental demand, which affects system performance. WDA also showed that image mapping capabilities, depth and geometry understanding, and error complexity affected system error-detection performance. Lastly, we identified the physical components of the work system and their functions (e.g., cameras for 3D computer vision).

Step 2 of Neumann et al. (2021) was also applied. The main humans in this system are the final operators on the assembly. We also identified the engineers who must initially calibrate and maintain the AIEDS for proper functioning.

We then applied the user analysis portion of step 1 in the Motivational Design for HMI. We assessed operators' personality traits (trait engagement) using the general causality orientations scale (Deci & Ryan, 1985; Ryan & Deci, 2008). This questionnaire allowed to control for predispositions to experience a certain level of motivation or engagement, which could impact our assessment of the effect of the AIEDS type on the various metrics.

Finally, measurable success criteria were selected. System performance was measured using (1) time taken to verify 30 snowshoes and (2) error detection performance (percentage of snowshoes correctly categorised). Other important metrics derived from the work domain analysis, such as operator mental demand (workload), motivation, and engagement were measured using questionnaires, and physiological measures when possible. The questionnaires and physiological measures were chosen based on a search of the literature.

### *Step 2 - Task identification*

We identified how the work task is being modified by the introduction of the AIEDS. Using version A, the operator will no longer check for error/defects in snowshoes. They will simply look at the computer screen to determine whether to put the snowshoe on one side or the other. Essentially, the operator's work becomes automatized without any decision-making responsibilities, and their role is limited to intervening if the system breakdowns or exceptions occur. Using version B, the operator will continue checking for defects, since the AIEDS has imperfect reliability. The operator will take on a supervisory role, verifying if the AIEDS's recommendation is correct. The operator has the final say regarding the presence of a defect.

### *Step 3 – Work demand analysis*

In this step, we examined how the AIEDS may impact perceptual, cognitive, psychosocial, and physical demands. Through the literature, we also evaluated the possible effect of these changing demands on our success criteria. We expected no changes in physical demands. We expected that both versions of the AIEDS would lead to a reduction in the operator's cognitive load since the operator must not longer manually detect errors. Version A, in which automatization is highest, was expected to lead to the greatest reduction in cognitive load. We expected both AIEDS to increase operators' performance, with version A leading to the best performance. However, version A leaves operators with very little decisional freedom and high monotony. Therefore, we expected version A to lead to lowered perceived autonomy, intrinsic motivation, and task engagement. On the other hand, version B provides operators with full decisional latitude i.e., agreeing or overriding the AIEDS. Therefore, we expected version B to lead to improved perceived autonomy, competence, intrinsic motivation, and engagement.

### *Step 4 – Outcome analysis*

To evaluate the impacts AIEDS has on the operator's task, we set up a pilot study in which 9 engineering students participated. Participants were randomly assigned to one of three conditions (Manual, AIEDS version A, or AIEDS version B). Participants were trained to complete the task and interact with the AIEDS by detecting errors/defects. After this training task, participants completed the experimental task, which consisted of sorting 30 snowshoes with 6 of them having a defect.

Performance time and error detection - Both versions of the AIEDS led to better performance time and error detection rate than manual execution of the task. However, no differences were observed when comparing versions A and B.

Mental demand (cognitive workload) - Cognitive workload was the lowest for version A of the AIEDS. No differences were detected between version B and manual execution. When interpreting the mean values according to guidelines presented by the authors of the scale (Hart & Staveland, 1988), both versions of the AIEDS resulted in *moderate* mental demand, while manual execution led to *somewhat high* demand.

Perceived autonomy and competence - Version B of the AIEDS led to the greatest perceived autonomy. No differences were found between version A and manual execution. No differences were observed for perceived competence

Intrinsic motivation - We found that version B of the AIEDS led to the greatest intrinsic motivation. No differences were observed between version A and manual execution.

Task engagement - Perceived engagement was higher for version B of the AIEDS than version A. No differences were observed between version B and manual execution, or between version A and manual execution. For physiological task engagement, version B was the highest, followed by manual execution, followed by version A.

### *Step 5 – Work design*

Results from the outcome analysis indicate important differences between the three conditions. Both versions of the AIEDS led to a significant increase in performance, with no differences being observed when comparing versions A and B. However, version B led to better outcomes in terms of perceived autonomy, intrinsic motivation, and task engagement compared to version A, all while maintaining an acceptable cognitive workload. These outcomes are directly related to long-term performance, well-being, technology acceptance, employee retention, and organisational citizenship behaviour (Deci et al., 2017;

Van den Broeck et al., 2021). The outcome analysis therefore indicates that version B of the AIEDS will lead to the most optimal human-AI interaction in the long term and should thus be implemented.

### Conclusion

Applying the steps to this case context allowed us to consider all aspects of the human-AI relationship. We systematically compared three methods of completing the work task to determine which would lead to the best outcomes for both the operator and the organisation. While version A of the AIEDS was fully automatized and thus restricted operator autonomy, version B gave operators a greater decisional latitude by allowing them to have the final say over the AIEDS' recommendation. Partially automating the task let operators take on a more supervisory role, which was perceived as being more meaningful and gratifying, leading to more positive short-term outcomes than version A. In turn, these positive short-term outcomes will develop into positive long-term outcomes for both operators and organisations.

## b. Quality control of aircraft maintenance

### Context

This real-world case study deals with integrating human factors in a research project to design a Cyber-Physical System (CPS) to automate aircraft components' inspection. Industrial inspection is a labor- and knowledge-intensive task carried out by skilled operators. Their objective is to ensure that aircraft components are free of harmful defects. Operators carefully examine each component for potential anomalies, and when one is detected, they evaluate and compare it to several decision criteria. A machine vision-based system was envisioned to support detecting and diagnosing the defects on service-run components such as gas turbines, discs, shafts, and fan blades to aid the inspectors in this tedious task. Starting from a proof-of-concept (Technology Readiness Level 3), the technology development plan was to reach Technology Readiness Level 6, i.e., where the technology functionalities are demonstrated in a relevant (operational) environment. HUMAI5.0 was applied to incorporate human factors input into a project that follows a technology-centric design pathway (see Figure 5). In addition, previous implementation failures due to a mismatch between human and automated diagnostics have motivated project leaders to understand how inspectors perform an inspection in detail. The framework allowed us to understand, model, and evaluate how an inspection is performed manually. From this understanding, requirements were derived to inform and guide the engineering process while considering future joint human-AI inspection workflows.

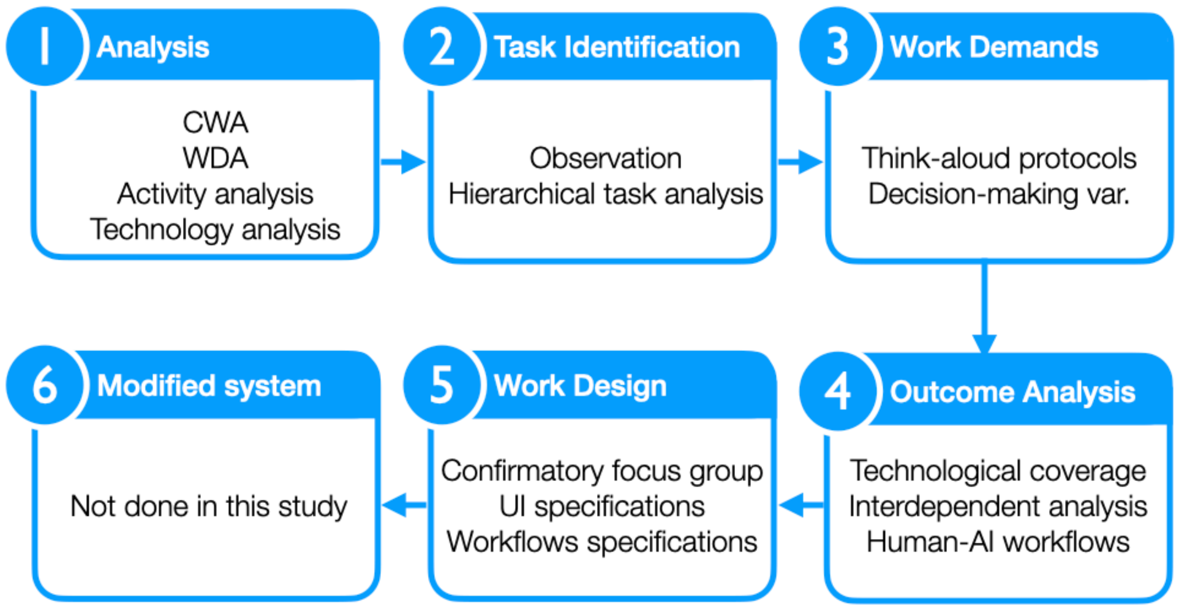


*Figure 4: Framework application to case study #2*

Due to space constraints, this case study's details are presented in this book's online supplementary material. Readers are also invited to consult previous publications on this topic (Cabour, 2023; Cabour, Ledoux, et al., 2022; Cabour, Morales-Forero, et al., 2022).

## Concluding remarks

The introduction of AI technology in complex socio-technical environments changes the operator's role and responsibilities in the workplace. In this chapter, we showed that adopting a human-centric approach to the design of human-AI collaboration in the industry leverages the best abilities of both parties to develop a work environment respectful to the long-term well-being of human workers and beneficial to the organization. To this end, we presented the HUMAI5.0 framework to analyze the work domain systematically, (re)design the human-AI interactions, and evaluate their outcomes. Indeed, the framework emphasizes the need for performing ethnographic studies to grasp the operational realities in which the AI-infused technology will be deployed. This emphasis resonates with other frameworks and methods cited in this chapter that rely on user-centered design for technology development projects, e.g., CWA. The strength of HUMAI5.0 is to integrate these existing methods used in silo into an integrative design approach that enables a shared understanding among stakeholders of AI's impacts in their workplace and how best to address it. For example, what tasks are being fully automated or partially automated? How can we suitably redesign work for human-AI collaboration? Since the integration of AI changes work practices, these methods help to envision and design future work practices during technology development cycles (and not post-implementation) and thus avoid falling classical automation issues, such as task-technology mismatches (Loup-Escande, 2022; Spies et al., 2020). Case study 2 illustrates this by methodically designing effective human-AI workflows and user interfaces that align with the work demands. Case study 1 delves into the connection between work performance and motivation, exploring the more profound meaning people ascribe to their work.

The rise of AI has been a topic of much discussion in recent years, with many experts predicting a bleak future of mass unemployment (Employment & Skills, 2014; Forum, 2016). However, this chapter takes a more nuanced approach, acknowledging that while AI will undoubtedly transform the nature of work, it may not necessarily lead to widespread job loss. Instead, AI is more likely to automate specific tasks that humans previously carried out, while new tasks will emerge from the collaboration between humans and AI. For example, AI can monitor the human state, such as detecting signs of fatigue or stress, while humans can validate the output of AI systems for critical decisions. This transformation of work will require a more holistic view of the human-machine system. Rather than viewing humans and machines as separate entities, it's essential to consider them as a whole. HUMAI5.0 can help bridge this gap by enabling a more cohesive approach to design effective and humane joint human-AI work.

The chapter illustrated the successful application of HUMAI5.0 in two cognitively demanding domains, i.e., error detection in manufacturing and quality control in aircraft maintenance. This also demonstrates the versatility of the approach by applying it to both experimental and naturalistic contexts. We believe the framework could be applied beyond industrial environments to support the shift from techno-centric to more human-centric design in AI projects (e.g., social-environmental projects). Applying the framework to other domains would create opportunities to investigate how it supports decision-makers in considering human and organizational factors in AI projects. Additionally, applying our framework to larger-scale scientific experiments examining human-AI interaction in the workplace could help further refine it.

# References

Cabour, G. (2023). Reconcevoir le travail dans des logiques de collaboration humain-technologie à l'ère du digital: Le cas de l'inspection industrielle [Thèse de doctorat]. Polytechnique Montréal.

Cabour, G., Ledoux, É., & Bassetto, S. (2022). Aligning Work Analysis and Modeling with the Engineering Goals of a Cyber-Physical-Social System for Industrial Inspection. Applied Ergonomics. In-Press.

Cabour, G., Morales-Forero, A., Ledoux, É., & Bassetto, S. (2022). An explanation space to align user studies with the technical development of Explainable AI. AI & SOCIETY. https://doi.org/10.1007/s00146-022-01536-6

Carayannis, E. G., Dezi, L., Gregori, G., & Calo, E. (2022). Smart Environments and Techno-centric and Human-Centric Innovations for Industry and Society 5.0: A Quintuple Helix Innovation System View Towards Smart, Sustainable, and Inclusive Solutions. Journal of the Knowledge Economy, 13(2), 926–955. https://doi.org/10.1007/s13132-021-00763-4

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). Working Minds: A Practitioner's Guide to Cognitive Task Analysis. MIT Press.

Deci, E. L., Olafsen, A. H., & Ryan, R. M. (2017). Self-determination theory in work organizations: The state of a science. *Annual review of organizational psychology and organizational behavior*, *4*, 19-43.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, *19*(2), 109-134.

Dubey, A., Abhinav, K., Jain, S., Arora, V., & Puttaveerana, A. (2020). HACO: A Framework for Developing Human-AI Teaming. Proceedings of the 13th Innovations in Software Engineering Conference on Formerly Known as India Software Engineering Conference, 1–9. https://doi.org/10.1145/3385032.3385044

Employment, U. C. f., & Skills. (2014). The future of work: jobs and skills in 2030. *Evidence Report*, *84*.

Forum, W. E. (2016). The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution. In: World Economic Forum Geneva.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114(C), 254–280.

Gourova, E. (2009). Knowledge audit concepts, processes and practice. 6(12).

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). Elsevier.

Jiao, J. (Roger), Zhou, F., Gebraeel, N. Z., & Duffy, V. (2020). Towards augmenting cyber-physical-human collaborative cognition for human-automation interaction in complex manufacturing and operational environments. International Journal of Production Research, 0(0), 1–23. https://doi.org/10.1080/00207543.2020.1722324

Johnson, M., Bradshaw, J. M., & Feltovich, P. J. (2017). Tomorrow's Human–Machine Design Tools: From Levels of Automation to Interdependencies: Journal of Cognitive Engineering and Decision Making. https://doi.org/10.1177/1555343417736462

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, M. B., & Sierhuis, M. (2014). Coactive Design: Designing Support for Interdependence in Joint Activity. J. Hum.-Robot Interact., 3(1), 43–69. https://doi.org/10.5898/JHRI.3.1.Johnson

Johnson, M., Vignati, M., & Duran, D. (2018). Understanding human-autonomy teaming through interdependence analysis. 1–20.

Kadir, B. A., Broberg, O., & Conceição, C. S. da. (2019). Current research and future perspectives on human factors and ergonomics in Industry 4.0. Computers & Industrial Engineering, 137, 106004. https://doi.org/10.1016/j.cie.2019.106004

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics, 19(3), 462–472. https://doi.org/10.1109/21.31053

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing Design Practices for Explainable AI User Experiences. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–15). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376590

Loup-Escande, E. (2022). Concevoir des technologies émergentes acceptables: Complémentarité des approches expérimentale, écologique et prospective. Activités, 19–1. https://doi.org/10.4000/activites.7345

Maddikunta, P. K. R., Pham, Q.-V., B, P., Deepa, N., Dev, K., Gadekallu, T. R., Ruby, R., & Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications. Journal of Industrial Information Integration, 26, 100257. https://doi.org/10.1016/j.jii.2021.100257

Michael Sony & Subhash Naik. (2020). Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model. Technology in Society. https://doi.org/10.1016/j.techsoc.2020.101248

Miller, M. J., & Feigh, K. M. (2019). Addressing the envisioned world problem: A case study in human spaceflight operations. Design Science, 5. https://doi.org/10.1017/dsj.2019.2

Millot, P., Pacaux-Lemoine, M.-P., & Trentesaux, D. (2015, October 26). Une conception anthropo-centrée pour casser le mythe de l'«Humain Magique» en Génie Industriel ?

Morineau, T., & Flach, J. M. (2019). The heuristic version of cognitive work analysis: a first application to medical emergency situations. *Applied ergonomics*, *79*, 98-106.

Mueller, S. T., Veinott, E. S., Hoffman, R. R., Klein, G., Alam, L., Mamun, T., & Clancey, W. J. (2021). Principles of Explanation in Human-AI Systems. ArXiv:2102.04972 [Cs]. http://arxiv.org/abs/2102.04972

Neumann, W. P., Winkelhaus, S., Grosse, E. H., & Glock, C. H. (2021). Industry 4.0 and the human factor – A systems framework and analysis methodology for successful development. International Journal of Production Economics, 233, 107992. https://doi.org/10.1016/j.ijpe.2020.107992

Oppl, S., & Stary, C. (2019). Designing Digital Work: Concepts and Methods for Human-centered Digitization. Springer Nature. https://doi.org/10.1007/978-3-030-12259-1

Parker, S. K., & Grote, G. (2022). Automation, Algorithms, and Beyond: Why Work Design Matters More Than Ever in a Digital World. Applied Psychology, 71(4), 1171–1204. https://doi.org/10.1111/apps.12241

Passalacqua, M., Pellerin, R., Doyon-Poulin, P., Boasen, J., Léger, P-M. (2022, July). Human-Centred AI in the Age of Industry 5.0: A Systematic Review Protocol. HCI International 2022 Conference, Montréal, QC, Canada.

Peeters, M. M. M., van Diggelen, J., van den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., & Raaijmakers, S. (2020). Hybrid collective intelligence in a human–AI society. AI & SOCIETY. https://doi.org/10.1007/s00146-020-01005-y

Rapp, A. (2021). In Search for Design Elements: A New Perspective for Employing Ethnography in Human-Computer Interaction Design Research. International Journal of Human–Computer Interaction, 37(8), 783–802. https://doi.org/10.1080/10447318.2020.1843296

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on systems, man, and cybernetics*(2), 234-243.

Roth, E., Depass, B., Harter, J., Scott, R., & Wampler, J. (2018). Beyond Levels of Automation: Developing More Detailed Guidance for Human Automation Interaction Design: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. https://doi.org/10.1177/1541931218621035

Ryan, R. M., & Deci, E. L. (2008). Self-determination theory and the role of basic psychological needs in personality and the organization of behavior. In *Handbook of personality: Theory and research, 3rd ed.* (pp. 654-678). The Guilford Press.

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G.-J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. Information & Management, 57(2), 103174. https://doi.org/10.1016/j.im.2019.103174

Shadbolt, N., & Smart, P. R. (2015). Knowledge Elicitation: Methods, Tools and Techniques. In J. R. Wilson & S. Sharples (Eds.), Evaluation of Human Work (pp. 163–200). CRC Press. https://eprints.soton.ac.uk/359638/

Spies, R., Grobbelaar, S., & Botha, A. (2020). A Scoping Review of the Application of the Task-Technology Fit Theory. LNCS-12066(Part I), 397. https://doi.org/10.1007/978-3-030-44999-5_33

St-Vincent, M., Vézina, N., Bellemare, M., Denis, D., Ledoux, É., & Imbeau, D. (2011). *L'intervention en ergonomie*. Éditions Multimondes.

Szalma, J. L. (2009). Individual differences in human–technology interaction: incorporating variation in human characteristics into human factors and ergonomics research and design. *Theoretical Issues in Ergonomics Science*, *10*(5), 381-397.

Szalma, J. L. (2014). On the application of motivation theory to human factors/ergonomics: Motivational design principles for human–technology interaction. *Human factors*, *56*(8), 1453-1471.

St-Vincent, Vézina, Bellemare, Denis, Ledoux, & Imbeau. (2014). Ergonomic Intervention. Institut de recherche Robert-Sauvé en santé et en sécurité du travail.

Van den Broeck, A., Howard, J. L., Van Vaerenbergh, Y., Leroy, H., & Gagné, M. (2021). Beyond intrinsic and extrinsic motivation: A meta-analysis on self-determination theory's multidimensional conceptualization of work motivation. *Organizational Psychology Review*, *11*(3), 240-273.

Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. CRC Press.

Xiong, W., Fan, H., Ma, L., & Wang, C. (2022). Challenges of human—Machine collaboration in risky decision-making. Frontiers of Engineering Management, 9(1), 89–103. https://doi.org/10.1007/s42524-021-0182-0

Zheng, T., Ardolino, M., Bacchetti, A., & Perona, M. (2021). The applications of Industry 4.0 technologies in manufacturing context: A systematic literature review. International Journal of Production Research, 59(6), 1922–1954. https://doi.org/10.1080/00207543.2020.1824085

# Supplementary material

## Quality control of aircraft maintenance

### Context

This real-world case study deals with integrating human factors in a research project to design a Cyber-Physical System (CPS) to automate aircraft components' inspection. Industrial inspection is a labor- and knowledge-intensive task carried out by skilled operators. Their objective is to ensure that aircraft components are free of harmful defects. Operators carefully examine each component for potential anomalies, and when one is detected, they evaluate and compare it to several decision criteria. A machine vision-based system was envisioned to support detecting and diagnosing the defects on service-run components such as gas turbines, discs, shafts, and fan blades to aid the inspectors in this tedious task. Starting from a proof-of-concept (Technology Readiness Level 3), the technology development plan was to reach Technology Readiness Level 6, i.e., where the technology functionalities are demonstrated in a relevant (operational) environment. HUMAI5.0 was applied to incorporate human factors input into a project that follows a technology-centric design pathway (see Figure 5). In addition, previous implementation failures due to a mismatch between human and automated diagnostics have motivated project leaders to understand how inspectors perform the inspection in detail. The framework allowed us to understand, model, and evaluate how an inspection is performed manually. From this understanding, requirements were derived to inform and guide the engineering process, with respect to future joint human-AI inspection workflows.

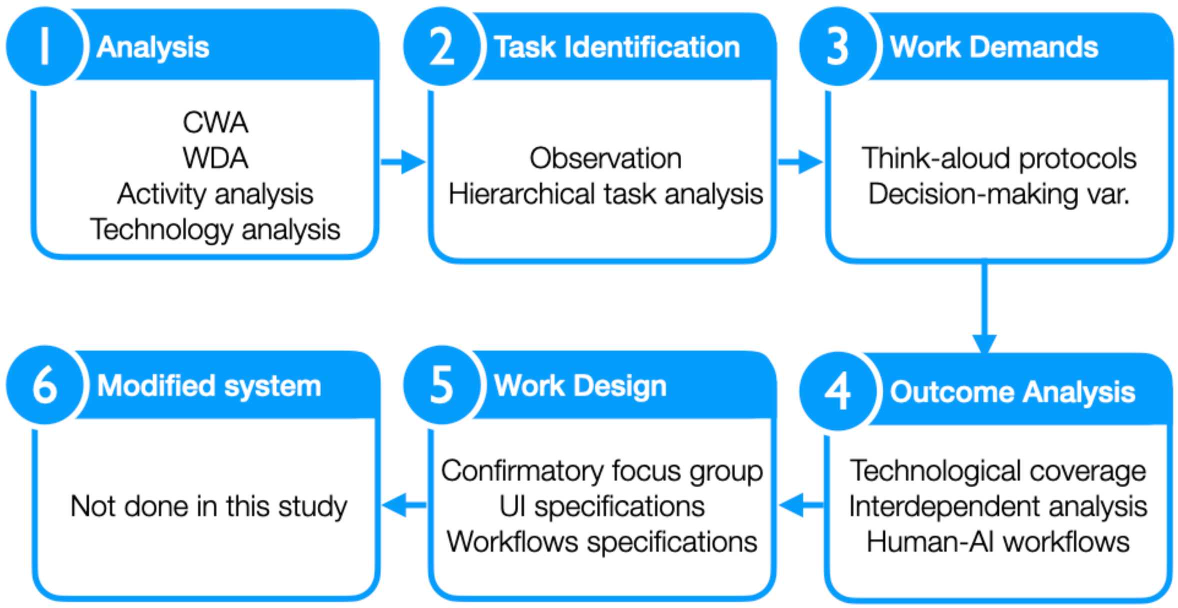


*Figure 5: Framework application to case study #2*

## Results

**Step 1 – Analysis.** Understanding the work domain was conducted by combining and adapting several methods. The first method employed was WDA, formulated by Vicente (1999). This method helped to understand the work system objectives and core functions, the overhaul process, and the upstream/downstream activities. Data were collected through open-ended questions with domain experts. To complete this top-down understanding, we used a second set of methods: Activity Analysis (Vincent et al., 2014) and heuristic Cognitive Work Analysis (Morineau et al., 2019). These methods were used to derive the domain constraints and link them to the operators' behavior. For example, strict airworthiness directives and institutional rules lead to conservative behavior from inspectors who are more likely to reject aircraft components with near-tolerance defects. Finally, a technology analysis was conducted to understand developers' technology development intentions and work packages. This step was as important as the domain analyses in that it aligned user studies with technology milestones and established the foundation to integrate human factors into design engineering.

**Step 2 – Task identification.** The overarching goal of this step in the industrial inspection process was to fully comprehend and situate where industrial inspection fits within the larger overhaul process. To this end, we began with *in situ* observations of the process to create a clear and detailed process map as a hierarchical task analysis (Figure 3). This analysis broke down the process into smaller, manageable tasks, allowing us to understand better the specific roles and responsibilities of each step in the inspection process and the decision points. Industrial inspection consists of the five following phases:

1. Work preparation: bring the component and the necessary inspection tools to the workstation
2. Multi-sensorial search: scrutinize each component individually until one defect is detected
3. Diagnose the condition of the defect/component by comparing its characteristics to the decision-making criteria, i.e., domain-specific rules
4. Execute corrective actions resulting from the diagnosis to restore the part using different material polishing operations
5. Work completion: carry on computerized procedures

In several focus group meetings, this understanding was presented to the relevant design stakeholders (engineers, researchers, inspectors, and human factors specialists). It allowed the design team to determine that the CPS would be applied primarily to automate phase 2 (search) and serve as a decision aid for phase 3 (Diagnose). It also guided subsequent analyses to understand better the tacit knowledge and operational demand of phases 2 and 3 of the inspection process.

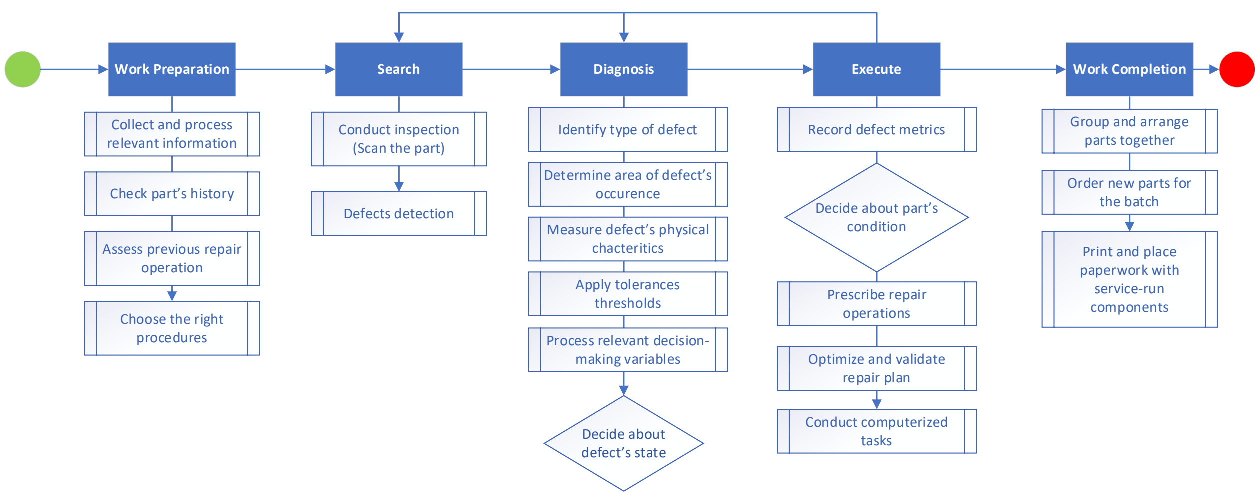


*Figure 6: Five steps model of inspection. From Cabour et al. (in preparation)*

**Step 3 – Work demand analysis.** This step aimed for a fine-grained understanding of operational demand in each of the 5 inspection phases (Figure 3) to ensure that our technology design supports inspection accordingly and avoids generating any mismatch between human and machine diagnoses. To achieve this, we employed various knowledge elicitation methods to grasp the two most crucial phases: search and diagnosis. One critical method was using think-aloud protocols, in which we asked the inspectors to verbalize their thoughts and actions while performing their tasks. This helped us understand the mental processes and strategies used during the inspection and identify the 36 decision-making variables that inspectors must process to diagnose the component's state (Table 3). Inspectors were also asked to range these variables in order of importance. Those variables can be specified in inspection aid documents or arise from operational constraints and experience (e.g., case-based reasoning). They relate to the defect's physical characteristics, the part's overall condition, and its history (e.g., previous polishing area) or the available course of action. From an automation design perspective, some variables are likely to disappear, especially the ones associated with human dimensional measurements (Table 3 - #9). However, the engineering team must assess technology capabilities in taking cognizance of all these variables to ensure that their human-AI system covers the core tasks and decision variables, which will be the focus of the next step. Additional details from Steps 1, 2, and 3 can be found in previous publications (Cabour, 2023; Cabour, Ledoux, et al., 2022; Cabour, Morales-Forero, et al., 2022).

*Table 1: Decision-making variables of inspection*

| # | Decision-making Variables | Techno. Coverage |
|---|---|---|
| 1 | Depth, width, length, circumference of a defect | Yes |
| 2 | Family of defect (defect classification) | Partially (Hybrid) |
| 2.1 | Harmful defect for airworthiness safety ("sharp" defect) | Non relevant anymore |
| 3 | Area of defect's occurrence | Partially (Hybrid) |
| 3.1 | Defect occurrence on critical area (attachment zone) | No |
| 3.2 | Defect occurrence on unregistered area ("grey area") | Partially (Hybrid) |
| 3.3.1 | Defect occurrence on repaired area | No |
| 3.3.2 | Defect occurrence adjacent to a repaired area | No |
| 3.4 | Defect in-between two areas | Yes |
| 3.4.1 | Diagnose each area with relevant tolerance | Yes |
| 3.4.2 | Apply the more restrictive criteria | Yes |
| 3.5 | Defect extension into an unauthorized area | No |
| 4 | Blending/dressing restrictions | No |
| 5 | Applicability of repair scheme | No |
| 6 | Repair scheme restrictions | No |
| 6.1 | Max applications reached for the scheme | No |
| 6.2 | Max applications reached for conjunctive scheme(s) | No |
| 7 | Additional instructions from local reference manuals | No |
| 8 | Optimization of the repair process | No |
| 8.1 | Defect removal by inspectors using sandpaper | Non relevant anymore |
| 8.2 | Defect removal by downstream repair operations | Non relevant anymore |
| 8.2.1 | Assigning defect to a nominal repair operation | No |
| 8.2.2 | Assigning defect to a non-nominal repair operation | No |
| 9 | Ambiguity in the process or No reasoning | Partially (Hybrid) |
| 9.1 | Reachable area(s) | Yes |
| 9.2 | Measurements tools wear | Non relevant anymore |
| 9.3 | Quality and perceptibility of a defect | Yes |
| 9.4 | Proximity with other defects | Yes |
| 9.5 | Physical characteristics borderline with tolerance threshold | Non relevant anymore |
| 9.6 | Unclear/non-existent instructions for defect's configuration | Partially (Hybrid) |
| 10 | Risk of bottle-necking the Material Review Board | No |
| 11 | Interdependencies with adjacent part(s) | No |
| 12 | Centralization of repair operations | Non relevant anymore |
| 13 | Number of batch parts affected | Non relevant anymore |
| 14 | Part's thickness | No |
| 14.1 | Maximal material removal allowed | No |
| 14.2 | Current material thickness | No |

**Step 4 – Outcome analysis.** The previous steps enabled us to understand the current work domain parameters systemically. However, transitioning from the current state (human inspection) to the future state (joint human-automation inspection) required additional design artifacts and methods. This step aims at linking how the industrial inspection will be performed in future joint human-automation workflows. We relied on two methods. The first consisted of calculating the number of functions that could be automated or semi-automated (technology coverage), given the current and anticipated automation capabilities. As a result, three tasks will be automatable (in green on Figure 7) and 5 partially automatable (in yellow on Figure 7), which means that the technology will co-execute these tasks in collaboration with inspectors. Regarding the decision-making variables, the CPS fully covers 7 of them and partially covers 5 of them (Table 3).

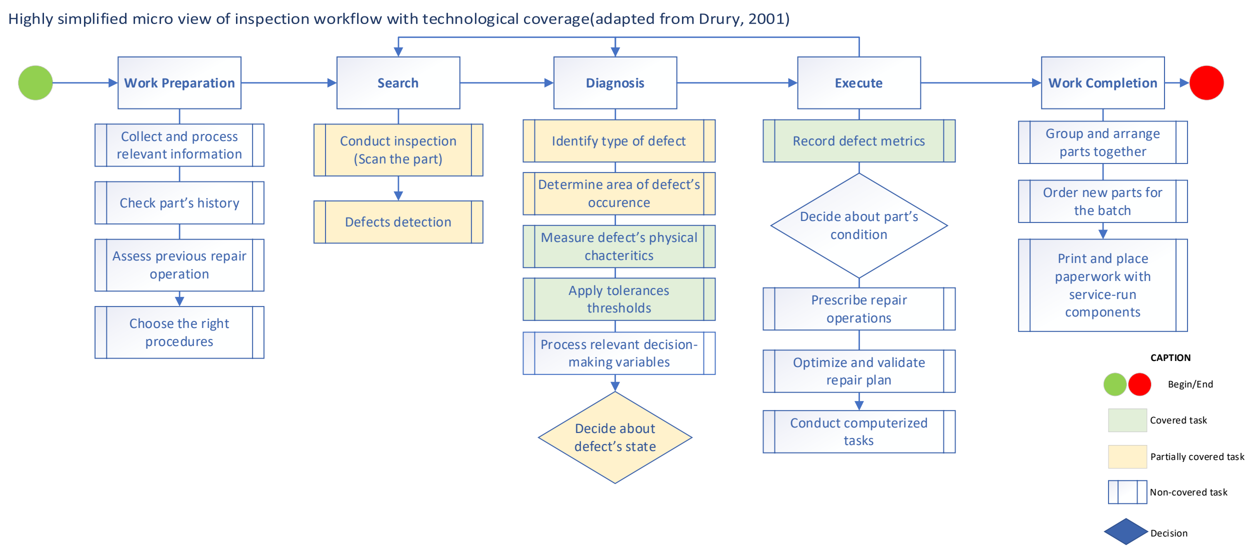


*Figure 7 - Technological coverage embedded in hierarchical task analysis*

The second method, called interdependent analysis (Johnson et al., 2018), was used to define how the human-automation system will conduct the work tasks as a team, given the work demands, the technological and human capabilities, and the interaction need. The focus is on the joint cognitive system rather than the human component (or the technology). An excerpt of the Interdependent Analysis Table (IAT) is given in Figure 8. The IAT formalizes every step required to effectuate a function. Section 1 helps designers to model the joint work in terms of tasks, sub-tasks, and required capacities/operational requirements. Section 2 helps identify the different teaming alternatives (human-assisted or automation-assisted in our case) to complete the work. Section 3 analyzes the potential workflows and pathways, e.g., to switch from one human task to an automated task, which can highlight automation boundaries and design flaws. Finally, section 4 lists possible human automation configurations and requirements to make it happen. The IAT revealed eight critical pathways currently unsupported due to technical or interactional shortcomings. These critical paths need to be addressed by corrective actions to integrate the system into production, whether at the level of technical development, organizational design, or training.

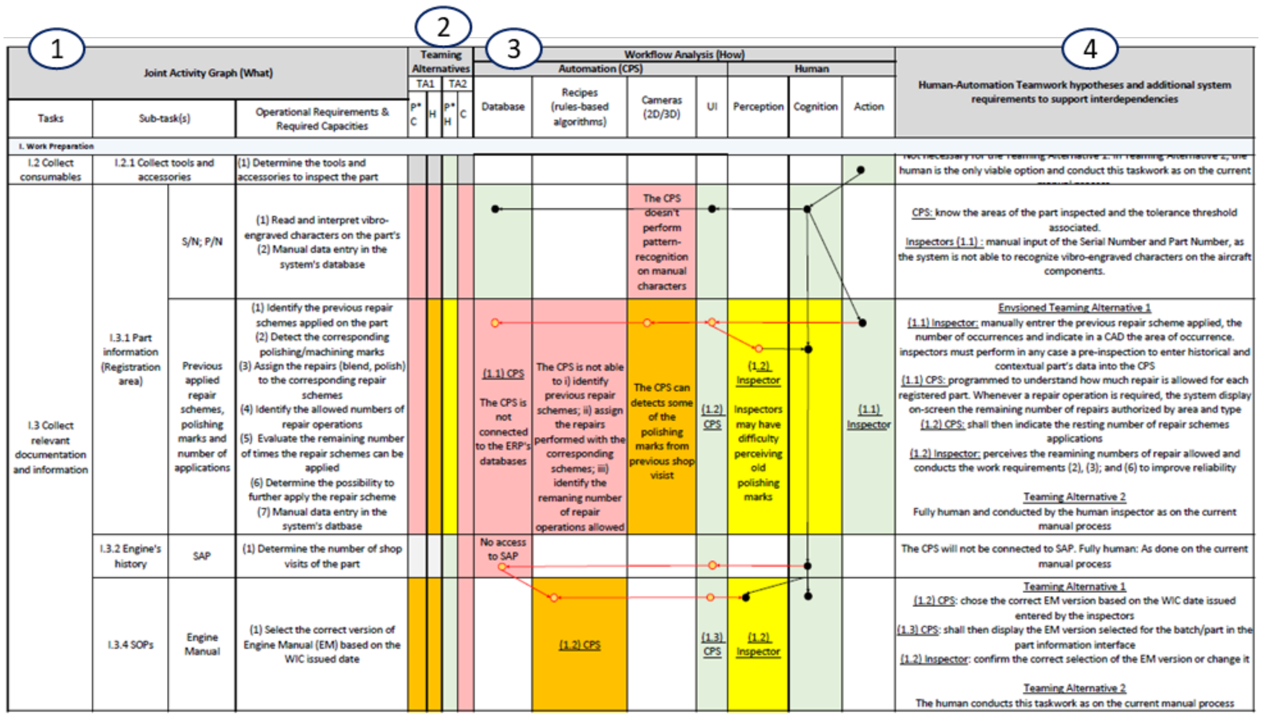


*Figure 8 - Extract from the IAT of the information-gathering task, including the upstream operations.*

**Step 5 – Work design.** The outcome analysis revealed several key areas for improvement in our joint human-automating teaming workflows. To address these issues, we conducted several confirmatory focus groups with design engineers to gather feedback and identify specific actions to be taken. We developed short-term and long-term solutions for each identified problem from these discussions. Additionally, we proposed specific changes to the user interface and human-automation workflows to improve overall usability and efficiency. These changes will be implemented in phase 6, with the short-term solutions being implemented immediately and the long-term solutions being implemented over time as resources become available.

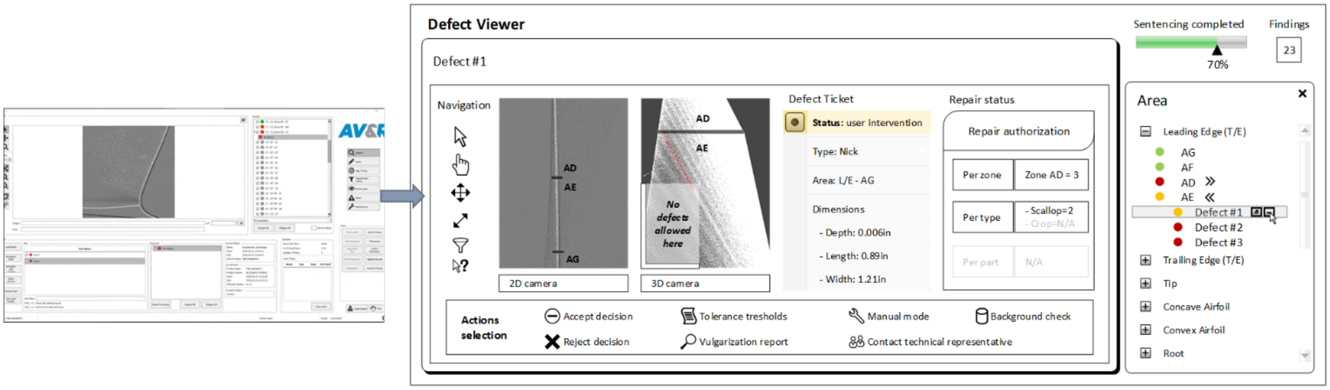


*Figure 9 - Screenshot of the current user interface (left) and recommendations using low-fidelity mockups (right).*