Practical Checklist for Responsible AI Integration in Workplace Tasks

In ChatGPT We Trust

Abstract

Artificial Intelligence (AI) nowadays is becoming more common in workplace and industrial settings, changing how tasks are organized and completed. Although AI can help improve efficiency and productivity, its use also brings challenges, such as possible workflow disruption, ethical concerns, impacts on workers' well-being, and risk of losing human skills and expertise. Many organizations still do not have clear or practical ways to decide whether a task should be supported by AI, partly automated, or fully handled by AI systems. This project focuses on developing a human-centered checklist to support AI task delegation decisions in workplace environments. The checklist was first created based on delegation criteria identified through a scoping review on human-AI collaboration and workplace automation. It was improved through several rounds of co-design activities, including peer reviews, storyboard scenario discussions, stakeholder interviews, and feedback sessions with both managers and operational workers. During the design process, the checklist gradually developed from a simple list of criteria into a more practical decision-making tool. Across five versions, the checklist was refined in terms of structure, wording, response format, and practical usability. The final version combines binary criteria with a five-point Likert scale, incorporates a final Decision Logic layer to aggregate outcomes, and includes design requirements under each item, making the checklist easier to use as a structured guide for AI task delegation decisions. The activities also showed the importance of factors such as workflow compatibility, human supervision, organizational trust, ethical transparency, data protection, operational feasibility, and the protection of workers' skills when introducing AI into workplaces. This co-design process specifically aimed to address the lack of practical tools that can help companies make AI task delegation decisions in real workplace settings.The resulting checklist aims to provide companies and decision-makers with a structured framework to facilitate more balanced, transparent and human-centered AI adoption strategies. While designed to be structurally transferable across sectors, its contextual usability was specifically validated through a worked use case in industrial design (parametric 3D modeling). Early feedback suggests that this tool may support organizations in evaluating AI integration more critically while preserving the role and well-being of workers in professional environments. Future work will involve testing and improving the checklist in different industries and workplace situations.

1 Introduction

In recent years, artificial intelligence (AI) systems have experienced exponential growth, with applications ranging from education and healthcare to finance and engineering. In the latter sector, the impact of AI is particularly profound, as it intervenes in every phase—from design to production—assisting or, in some cases, entirely replacing human labor. Within the specific context of rapidly evolving professional sectors , the introduction of automated tools promises to optimize workflows, yet it raises a crucial question: to what extent do workers truly benefit from such innovations? Although AI is often perceived as an extremely advanced technology, its practical implementation is not without obstacles. On the contrary, it forces a confrontation with technical and ethical criticalities that fuel an ongoing debate regarding the efficacy and sustainability of these integrations within the industrial fabric. While modeling software is technologically cutting-edge and AI integration could further enhance its performance, a paradox frequently emerges: the development of these solutions tends to overlook the end-user. This creates a deep disconnect between the tool and those who operate it daily, resulting in a gap between technical innovation and real-world operations. In particular, the impact of automation on the worker—both in terms of cognitive load and the evolution of their professional role—is often underestimated. To bridge this gap and foster a conscious technological transition, this study aims to investigate the internal dynamics within company workplaces. By administering a check-list targeted at both managers and workers, the research analyzes the perception of AI integration and its related impact on operational flows. The ultimate goal is to outline guidelines that facilitate a balanced adoption of the technology—one capable of leveraging the potential of AI while simultaneously safeguarding the role and well-being of workers within the production ecosystem.

2 Related Work

Integrating AI systems into real-world task delegation is a complex choice characterised by the constant conflict between ethical and social restrictions and performance advantages [3, 6, 11]. In recent years, a variety of studies and research efforts have developed frameworks suitable for minimising this conflict of values [3, 6]. Most current studies converge on identifying fundamental values such as fairness, accountability, transparency, and privacy [6]. However, practitioners and organisational actors who are supposed to apply these frameworks in real-world scenarios often face a practical gap between abstract ethical principles and concrete organisational implementation [4, 12, 14]. Ethical constraints are frequently presented without a clear set of instructions for applying them to complex, co-dependent, and organisation-influenced tasks [4, 14]. It is also notable that, while software tools for inspecting model behaviour, bias, and fairness-related outcomes have been developed [9, 10], common metrics for comparing different studies remain contested [8], and tools capable of evaluating broader social and ethical effects are still limited [11]. In current studies, checklists are a widely adopted tool to address the matter under study [1, 2, 7]. However, they are also affected by the issues already cited above. Often, checklists fall under one of two categories: ethically principle-structured [5, 7, 13] or development lifecycle stage-structured [1, 2]. Across many checklist-based approaches, complex ethical questions may be reduced to simplified prompts or "yes or no" questions [2, 5, 7]. Furthermore, an issue that belongs to the vast majority of checklists developed is the limited, if not complete lack of, input from the practitioners who are supposed to apply the frameworks developed [4, 14]. This issue can also cause a lack of consideration for the sociocultural, organisational, and workflow alignment required to render the checklist functional for real-world use [4, 14].

2.1 Theoretical Background: Scoping Review Summary

As said before, this checklist was created following a prior scoping review conducted by our team. There we analysed a total of 30 papers, exploring the critical factors that determine whether occupational tasks should be delegated to artificial intelligence. The core of our findings rests on a systemic framework divided into four distinct thematic areas, demonstrating that AI integration represents a reconfiguration of work rather than total human replacement. First, looking at organizational and labor market impacts, AI primarily automates routine tasks. This shifts human workers toward higher cognitive and emotional responsibilities but also contributes to job polarization by hollowing out mid-level positions. Second, the human-AI relationship and worker skills are paramount. Effective AI training and supervision rely heavily on the worker's initial expertise. However, prolonged task delegation can lead to severe skill decay, leaving workers unable to handle anomalies, alongside potential psychological strain due to isolation. Third, regarding AI advantages and task characteristics, the technology excels in standardized, rule-based environments but faces a hard limit when tasks require genuine creativity or emotional intelligence. Because AI models often operate as opaque "black boxes," continuous human oversight remains vital to catch hidden biases and ensure precision in high-stakes environments. Finally, social and ethical governance dictates that technical feasibility must be paired with strict accountability. Organizations must manage heavy data and privacy risks, define clear responsibility structures, and build trust to overcome public skepticism. Before integrating AI, organizations must thoroughly assess whether a task fits the system's capabilities and ensure human oversight remains in place for critical operations. Leaders must weigh clear operational advantages against the risk of eroding core human skills, while actively planning for workforce upskilling. Finally, strict data privacy and regulatory compliance are non-negotiable foundations for sustainable AI adoption that augments, rather than undermines, professional judgment. To translate these principles into action, we developed the following checklist to evaluate whether a task should be supported by AI, partly automated, or fully handled by AI systems.

3 Co-Design Procedure and Iterative Analysis

To translate our initial theoretical framework into a validated, practical organizational tool, we employed a participatory co-design approach. Rather than presenting our initial criteria as a finished product, we subjected the tool to repeated cycles of empirical feedback, using physical prototyping, scenario simulations, and expert interviews. This chapter details our data collection methods, the analytical coding process, and the specific item-level iterations that transformed the initial theoretical list (Checklist V1) into a fully operational decision-making framework (Checklist V5).

3.1 Data Collection

Our data collection was structured across three distinct activities, engaging progressively more specialized stakeholders to evaluate the checklist's usability, clarity, and organizational relevance.

3.1.1 Activity A: Initial Peer Review. To evaluate the preliminary usability of the tool, we conducted a first round of peer review with a group of 5 to 6 university colleagues drawn from various academic courses. The original criteria, derived directly from our scoping review, were physically printed on four A3-sized papers, corresponding to the initial four macro-themes. During this 30-minute session, participants interacted directly with the tangible prototype, utilizing post-it notes to highlight strengths, flag structural weaknesses, and point out redundancies. This physical interaction allowed the research team to gather immediate, unguided reactions without overly constraining the participants' cognitive processes.

3.1.2 Activity B: Storyboarding Simulation. To assess the contextual usability of the refined prototype, we engaged a subsequent group of 3 new reviewers, distinct from the first group to avoid anchoring bias. In this 30-minute session, we presented participants with a physically hand-drawn storyboard illustrating a concrete corporate scenario: an industrial and commercial designer tasked with developing a new parametric 3D model for a client, using an AI tool to augment the workflow. After describing and showing the physical storyboard, participants were asked to role-play the scenario and envision applying the checklist to assess the adoption of the AI tool. This simulation probed the practical flow of the items and anticipated potential client reactions to AI disclosure.

3.1.3 Activity C: Expert Stakeholder Interviews. The final and most critical phase involved semi-structured interviews with seven industry professionals. To prevent organizational blind spots and capture divergent needs, the sample was deliberately stratified into two clusters: an Operational group (an Electronic Designer, an Electronics Engineer/PhD student, and a Product Development Manager) and a Strategic-Managerial group (a Vice President, a Commercial Director, a Product Manager, and a Company Consultant). Interviews lasted between 20 and 30 minutes on average. To accommodate geographical and logistical constraints, the VP and the Commercial Director were interviewed in person, while the remaining stakeholders were engaged remotely via video calls or structured questionnaires. To ensure participants felt comfortable sharing candid organizational critiques, no audio recordings were made; instead, researchers took comprehensive live notes during the sessions.

3.2 Data Analysis

From the collected material, post-it notes during Activity A, observations made while storyboarding, also interview records taken in real time during Activity C, a pattern began to form. What emerged was shaped through inductive coding, letting themes rise without preset labels. Each piece of insight carried weight, not because it fit a plan, but because it repeated across settings. Meaning built slowly, pulled together by repetition rather than assumption. The process trusted what appeared, instead of forcing structure too soon. For stronger analysis and less personal influence, comments went into one joint Google Doc. Separate from each other, three people examined the original responses, linking what participants said to exact checklist points. Where changes in wording or layout came up, like combining entries or adjusting criteria, all the three people reviewed their results together. If views differed on applying a suggestion, agreement followed whichever interpretation had more support. Decisions stood when at least two out of three aligned. Should agreement fail between the three coders, input from the broader research group guided the final phrasing. With each modification logged, decisions followed only where data led.

3.3 Iterative Refinement: Traceable Change Logs

As illustrated in our procedural workflow (Figure 1), the checklist evolved through substantive structural and functional iterations. The following subsections document the explicit mapping between participant feedback and the resulting design changes, proving the empirical foundation of the tool.

3.3.1 The Transition from V1 to V2: Consolidation and Refinement. The first major evolution of the checklist (V1 to V2) was primarily driven by workshop feedback and storyboard activities, focusing on streamlining the tool and removing redundancies. Several criteria were merged to improve cohesion; for instance, separate items regarding technical execution (V1.1.6 and V1.3.1) were unified into a single “Technical Feasibility” statement within the first section. Similarly, the standalone point regarding human supervision (V1.2.5) was absorbed, leaving only the specific requirement to preserve workers’ skills for safe oversight. To further declutter the checklist, several items (V1.2.7, V1.2.9, V1.3.3, and V1.4.4) were entirely removed after being deemed superfluous, irrelevant, or simply too difficult to evaluate in a practical setting. Finally, a general refinement of the vocabulary was applied across all remaining statements to ensure better clarity.

3.3.2 The Transition from V2 to V3: Flattening for Feedback. The shift from V2 to V3, guided by group brainstorming sessions, was characterized by structural deconstruction to transform the checklist into a more effective feedback-gathering tool. The categorized subheadings were completely removed, flattening the document into a continuous list. To maximize the value of the testing phase, specific questions tailored to different stakeholders were added at the end of the document to collect targeted feedback. Content-wise, a notable change was the splitting of the previously merged technical feasibility item (V2.1.5); it was separated back into two distinct questions to independently assess whether the task is physically augmentable/automatable and whether the AI model is fully capable of tackling it.

3.3.3 The Transition from V3 to V4: Pragmatism and Risk Mitigation. The V3 to V4 transition marks the tool’s maturation into a pragmatic, business-ready asset, heavily influenced by direct stakeholder feedback. Structurally, order was restored: questions were once again grouped under specific categories with descriptive titles and tags. Additionally, explicit “Yes/No” labels were added to specific statements to clarify the expected response format. This iteration also introduced critical new criteria focused on real-world risks. An “Intellectual Property and Asset Protection” point was added to assess legal risks associated with stolen or copyrighted training data. Operational efficiency was addressed by adding a requirement that the AI must not make the workflow more laborious and expensive than manual human work. Finally, a fundamental safety measure was introduced to the Team section, requiring that manual procedures be kept active so workers can safely intervene if the AI makes mistakes or stops working.

3.3.4 The Transition from V4 to V5: Decision Logic and Scoring Integration. The final transition (V4 to V5), shaped by peer reviews and lectures, focused heavily on user experience, clarity, and depth of assessment. The lexicon across all items was simplified to ensure the checklist was accessible to a broader audience without losing technical accuracy. Responding to the Product Development Manager’s directive to “don’t overcomplicate... so as to have a clear output”, we acknowledged that a flat list of questions was insufficient for managerial triage. A major formatting update was also introduced for the evaluation mechanism: while binary constraints retained their Yes/No format, all qualitative, non-binary items were upgraded to a 5-point Likert scale to allow for more nuanced assessments. Furthermore, the checklist evolved from a simple evaluation tool into a generative guide by adding specific “design requirements” under every item, explaining how the development of the AI tool should actively tackle each criterion. Finally, the overall structure was polished by reordering the items within each category by order of importance and adding a brief explanatory preamble to guide the user through the document.

Overall evolution:

The journey from the initial V1 draft to the final V5 iteration highlights a profound shift from a theoretical screening questionnaire to a highly actionable managerial framework. Driven by continuous stakeholder feedback, the tool progressively abandoned idealistic assumptions to address concrete workplace realities. Early versions struggled with a rigid binary format, which was ultimately replaced by a dual-track system. Absolute constraints like legal compliance became strict "Yes/No" gates, while complex socio-technical issues transitioned to a five-point Likert scale, allowing decision-makers to effectively evaluate the nuanced gray areas of AI integration. Operationally, the framework shifted toward strict business pragmatism. It now demands empirical proof of AI capabilities rather than blind trust, assesses tangible threats like intellectual property infringement, and rigorously calculates operational overhead to ensure that supervising the AI does not cost more than performing the work manually. Simultaneously, the concept of job displacement was reframed. Rather than viewing AI solely as a threat to employment, the final iterations treat human expertise as a critical safety net. By requiring companies to maintain manual fallback procedures, preserving core worker skills became a fundamental operational necessity to guard against AI hallucinations or system failures. Finally, the checklist evolved from a passive evaluative hurdle into a proactive blueprint. With the addition of specific design guidelines beneath each criterion, the framework now actively steers the development process, ensuring that ethical constraints and operational efficiencies are built directly into the AI workflow from the start.

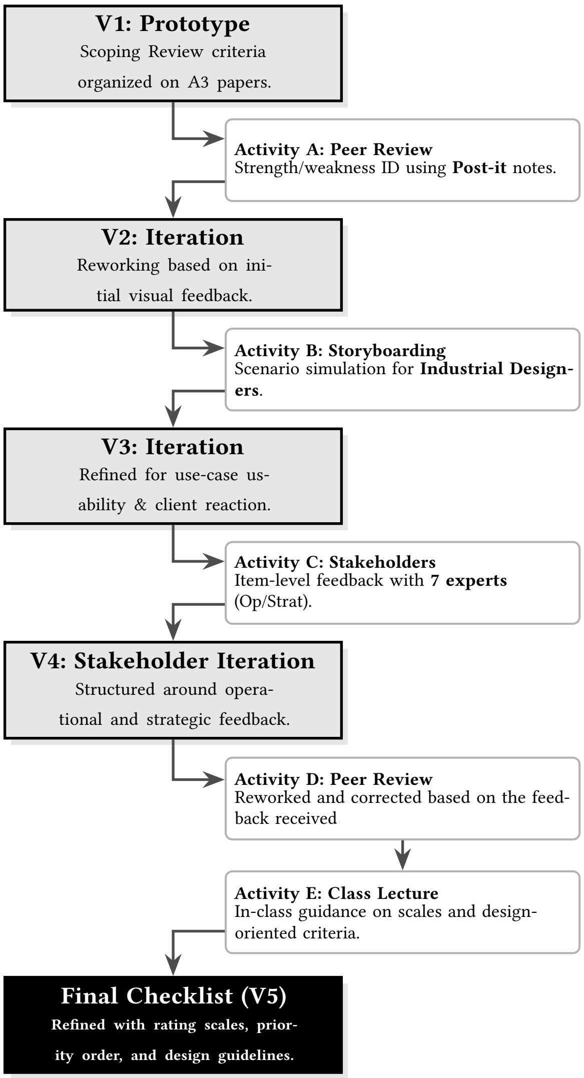

Figure 1: The iterative co-design process from V1 to the final V5 checklist.

4 Stakeholders Involved

Table 1: Stakeholders involved in the co-design process

Managerial Roles	Worker Roles
Product Development Manager	Electronic Designer
Product Manager	Electronics Engineer & PhD student in AI studies
Company Consultant	
VP & Commercial Director	

5 Checklist Applied

To demonstrate the practical applicability and validity of the final decision making tool, we tested the checklist on the reference worked use case: the integration of an AI tool for parametric optimization in 3D modeling. The application of the five thematic clusters produced the following decision analysis:

5.0.1 I. Intended Uses. The optimization activity is governed by stable and systematic geometric rules and is computationally automatable. AI acts strictly as technical support (Workflow Framing), while the protection of brand identity is not delegated, as AI lacks the necessary creative and emotional nuances (Boundary Evaluation). Furthermore, objective metrics are established to quantify the output and distinguish intended design results from any algorithmic errors.

5.0.2 II. Harms. In industrial design, identifying legal risks regarding copyright is vital to ensure that training data does not derive from the misappropriation of competitors' projects (IP and Asset Protection). Additionally, the company must assess the potential impact on the designer's psychological well-being and build social acceptance of the tool through practical demonstrations of its real capabilities, avoiding the requirement of blind trust (Empirical Trust Building).

5.0.3 III. System and Data. Before implementation, it is necessary to confirm accessibility to a high-quality dataset and determine whether the costs of CAD file preparation outweigh the expected economic gains (Cost-Benefit of Data Prep). At a system level, the cloud must guarantee the security and privacy of trade secrets (Infrastructure Safeguards), while maintaining a level of technical transparency sufficient to justify the model's internal optimization choices to the designer (Contextual Explainability).

5.0.4 IV. Oversight. Since the validation of physical components often requires the final approval of a qualified professional (Regulatory Compliance), an analysis of operational overhead is imperative. It must be confirmed that human oversight requirements do not make the AI-based workflow more laborious ("double the effort") compared to purely manual modeling (Operational Overhead).

5.0.5 V. Team. Finally, implementation must be managed to prevent uncontrolled job loss (Socio-Economic Stability), evaluating whether AI will enhance or downgrade the long-term skills of the workforce and planning specific training programs. It is essential to preserve the designer's core expertise so that they remain capable of supervising the task safely (Safe Oversight Maintenance), maintaining the flexibility and manual practices necessary to intervene promptly in case of AI failures or geometric "hallucinations" (Resilience and Safeguards).

5.1 Final Outcome and Numerical Breakdown

To demonstrate the operational validity of Checklist V5, we aggregated the responses for the parametric 3D modeling scenario using the established Decision Logic rubric.
Absolute Blockers (Yes/No Items): All Passed. The scenario successfully cleared all critical gateways. Specifically, the task is computationally automatable, data preparation costs do not exceed economic gains, and the AI acts strictly as a supportive tool without overriding the designer's final sign-off.
Likert-Scale Scoring (1-5): For the remaining qualitative items, the scenario yielded the following breakdown (Max Score: 45):

- **I. Intended Uses (Score: 5/5):** Highly favorable, as the parametric output is perfectly measurable.
- **II. Harms (Score: 12/15):** Medium-High. Mental health risks are minimal, though minor concerns persist regarding social acceptability.
- **III. System and Data (Score: 3/5):** Medium. The opacity of generative AI models ("black box" effect) limits total transparency.
- **IV. Oversight:** (Category evaluated entirely via Yes/No blockers, all cleared).
- **V. Team (Score: 13/20):** Medium-Low. This category revealed the most vulnerabilities, scoring lower due to the high risk of skill decay over time and the immediate need for proactive upskilling to maintain genuine trust.

Total Aggregated Score: 33/45 (73%). **Final Verdict:** Partial Automation Recommended. Because all absolute blockers were successfully passed and the total score falls into the "Medium" tier (50%–80%), the AI tool can be safely adopted for parametric optimization. However, the integration must proceed with caution: the tool cannot run autonomously and human design supervisors must remain strictly in the loop to intercept geometric hallucinations and prevent long-term skill erosion.

6 Gaps and Future Work

Several methodological limitations emerged during the transition from the theoretical findings of the scoping review to the development of a practical and operational checklist. Chief among these is the narrow scope and sector-specific nature of the stakeholder pool involved. Because the co-design activities relied on a limited number of experts, predominantly from a single chosen sector, the generalizability of the checklist to diametrically opposed professions remains constrained. While the instrument proves adequate for similar professional domains, the dataset lacks the necessary heterogeneity to guarantee the checklist's robustness across radically different occupational fields. Furthermore, the sheer diversity of organizations precludes the formulation of a one-size-fits-all checklist. Numerous confounding variables come into play—such as sector, organizational size, and financial stability, to name but a few. Consequently, while a satisfactory degree of generalization can be achieved, such an instrument must ultimately be tailored to the specific requirements of individual enterprises. Another significant constraint stems from the restricted timeframe allocated for gathering user feedback. Evaluations were confined to immediate reactions and simulation exercises, thereby impeding the collection of robust data regarding long-term reliability. Ultimately, this time constraint precluded the opportunity to test the prolonged implementation of the tool within real-world contexts, making it currently impossible to empirically measure the actual practical impact of the checklist on daily workflows and organizational routines. To address these research gaps, future work will focus on two primary objectives:

- **Sample and Sector Expansion:** Implementing and validating the Checklist V5 across a broader range of occupational sectors and diverse socio-cultural work environments, in order to verify its cross-sector applicability.
- **Longitudinal Field Testing:** Conducting long-term case studies within adopting organizations to evaluate the instrument's performance when integrated into daily operational decision-making processes over extended periods.

For future work, researchers should prioritize data collection from real world practitioners across a wider range of job sectors, geographical locations, and broadest time window.The expansion is essential to address the diverse needs shaped by society and work culture. Additionally, developing a standardized metric is critical to evaluate the impact of AI task delegation across a variety of occupations. To gain a broader perspective, we also interviewed individuals from completely unrelated industries. Although these feedbacks highlighted a lack of focus on worker occupation and environmental interest within the research, these insights were excluded from our final checklist due to the limited number of interviews and their isolated nature. Nonetheless, these findings offer valuable topics of interest for future research.

7 Discussion

The main contribution of this study is the transition from abstract ethical criteria to an operational managerial framework. While literature widely defines principles like fairness and transparency, practitioners often face an implementation gap. This iterative co-design process proves that responsible AI integration cannot rely on top-down theories; it requires continuous empirical feedback from both managers and operational workers to align with real workplace dynamics. The checklist's evolution highlights a profound shift toward business pragmatism and evaluation nuance, driven by specific participatory phases:

- Peer Review & Storyboarding: Initial workshop feedback on A3 prototypes and role-playing simulations reduced structural confusion and redundancies. This led to the transition from V1 to V3, merging overlapping technical items (V2) and temporarily flattening the list to capture unanchored stakeholder inputs. Reviewers firmly pointed out that complex socio-technical impacts inhabit gray areas that simplified "Yes/No" prompts cannot capture, prompting the upgrade to a 5-point scale for qualitative items in V5.
- Strategic & Operational Interviews: interviews with stratified industry professionals grounded abstract ethics into tangible metrics. The Operational cluster warned against excessive complexity and overhead, ensuring that AI supervision does not double human labor effort or create data prep barriers. Simultaneously, the Strategic cluster requested strict legal accountabilities, shifting the focus toward rigid intellectual property protection and legally qualified human sign-offs.

Crucially, the final framework reframes the worker's role from a position vulnerable to displacement to a vital operational safety net. By mandating active manual fallback procedures, preserving core human expertise becomes a functional necessity to maintain resilience against AI hallucinations or systemic failures. Consequently, the tool evolves into a proactive blueprint, using design guidelines to embed worker augmentation directly into the AI development process.

8 Conclusion

This work focused on creating a checklist to help decide when a workplace task can be supported, augmented or automated by an AI system.The development of the checklist evolved through a strict process, starting from the criteria of a previous scoping review and then being refined through peer reviews, a storyboarding activity, and feedback from stakeholders with both managerial and technical roles. The analysis of this data followed a coding approach, allowing themes to emerge based on their recurrence rather than preset schemas.To limit personal influence, the comments were examined independently by three researchers, whose interpretations were subsequently discussed as a group. Structural changes were applied following a majority criterion or through the guidance of the research group, ensuring that every decision was guided only by the tracked data. The results show that deciding whether to use AI is not just a technical question. It is not enough to ask whether the AI can do the task or whether the company has enough data: it is also important to consider how the technology changes the workflow, whether workers still need to supervise the task, whether people trust the system, and whether privacy, safety, legal, or ethical problems may appear. The checklist clearly shows that workers must remain central to the decision process: AI should not be introduced only to make work faster or cheaper; its effects on workers' skills, responsibilities, workload, and well-being must also be considered. In this way, the checklist helps avoid decisions where the technology works in theory but creates problems in real daily work. Feedback from different stakeholders confirmed that both managers and workers offer fundamental points of view: managers can evaluate the usefulness and feasibility of AI for the company, while workers can better explain how the tool would affect the actual task and everyday workflow. For this reason, the checklist can also work effectively as a communication tool between these groups. Overall, the final checklist offers a simple and structured way to reflect before introducing AI into a balanced and human-centred approach. For this future work, it is essential to focus on data collection and input from real-world practitioners across a wider range of job sectors and locations, alongside the need to develop a common metric to study the impact of AI task delegation in diverse occupational fields.

A Appendix

A.1 Checklist V1

A.1.1 AI Advantages and Task Characteristics.

1. Is there already a data set to train AI on the task?
2. Does the task retain value from human creativity?
3. Is the implementation economically sensible?
4. Does the task legally require human sign-off?
5. Does implementing AI upskill workers and improve efficiency?
6. Is the AI model available and capable of doing the task?

A.1.2 Organizational Changes and Labour Market Impact.

1. Is the task fully defined by a set of rules?
2. Does the task avoid creative and/or emotional components?
3. Is the task's workflow flexible in its structure?
4. Has AI implementation already been socially rejected in regard to the task, or not?
5. Does the task require human supervision?
6. Is the task physically augmentable or automatable?
7. Is the output of the task objectively measurable?
8. Does implementation create job displacement?
9. Has AI implementation been technically accepted?

A.1.3 Human-AI Relationship and Worker Skills/Training.

1. Does task delegation impact the worker's skills and experience in the long term?
2. Does task delegation impact the worker's mental health?
3. Does task delegation require specific knowledge from the worker?
4. Does the task delegation strategy preserve the necessary human expertise required to safely oversee the automated task?

A.1.4 Governance, Ethics and Trust.

1. Does the worker trust the AI model for the task?
2. Can privacy and data security be assured?
3. Does AI implementation lead to ethical concerns?
4. Does the government allow the AI model needed?
5. Is the AI model transparent for the task?

A.2 Storyboard Activity

Storyboard Question: An Industrial and Commercial designer need to design a new product. Certain specifications are given by the client. The designer wants to use an AI tool to develop this product. The designer uses the Checklist to evaluate the design process. How could the client react knowing AI was used?

Figure 2: Storyboard vignette drafted in the co-design activity.

These are the main feedback we received in this activity:

- **Reorganization of Categories:** The four original themes, such as "Ethics, Legality and Sustainability" grouped together, were too confusing. It was suggested to divide them into more specific categories.
- **Reduction of Redundancies:** Very similar points of V1, such as "physically automatable task" and "AI capable of dealing with it", had to be merged to streamline the list.
- **Nuances beyond Yes/No:** Colleagues pointed out that some ethical questions could not have a binary answer, suggesting the inclusion of open-ended conditions.

A.3 Checklist V2

This intermediate checklist was generated following the first round of peer review (Activity A). Based on the post-it feedback, redundant items regarding technical capabilities were merged. Furthermore, reviewers highlighted that complex ethical implications could not be answered with a strict Yes/No binary format. Consequently, the questions were transformed into reflective statements and a preamble was introduced, while retaining the original four-theme structure of V1.

Pre-amble. This checklist is meant to guide the decision process. Most items are phrased so that the eventual positive outcomes favour the implementation of the AI model. Some items have an if-condition that does not necessarily bring to a definitive yes/no situation.

A.2.1 AI Advantages and Task Characteristics.

- Verify if a quality dataset to train the AI on the task already exists or is easily obtainable.
- Assess whether the task retains value from human creativity and involvement.
- Ensure that the implementation is economically sensible and sustainable.
- Evaluate if the task legally requires a human signature or sign-off.
- **Technical Feasibility:** Confirm that the task is physically automatable and the AI model is fully capable of executing it. *(Merged item)*

A.2.2 Organizational Changes and Labour Market Impact.

- Verify that the task is fully defined by a set of rules that the AI model could easily follow.
- Check that the task avoids particularly deep emotional components.
- Verify that the model implementation harmonizes with the workflow structure without completely overhauling it.
- Check if the output of the task is objectively measurable.
- Verify that AI implementation has already been technically accepted in regard to the task.

A.2.3 Workforce Adaptation and Expertise.

- Verify that the possible mental health risks for the workers are minimal or none.
- Ensure that the implementation of the AI model does not cause a radical job displacement.
- Analyze the potential effect on the worker's skill level (will the implementation upskill or downgrade the worker?).
- Verify that the delegation strategy preserves the workers' expertise necessary to safely oversee the automated task.
- Check if the task delegation requires specific knowledge from the worker.

A.2.4 Governance, Ethics and Trust.

- Ensure that the workforce trusts the AI model that could be implemented.
- Verify that privacy and corporate data security can be assured and are not put at risk.
- Verify that no ethical concerns could arise from the implementation of the AI model.
- Ensure that the AI model is as transparent as possible in regard to its inner workings.

A.4 Checklist V1–V2 Changelog

Table 2: Checklist V1–V2 changelog

V1→V2			
Item Affected	**Change Type**	**Feedback Source**	**Result**
V1.1.6;V1.3.1	Merge	Workshop Feedback	Merged into a single "Technical Feasibility" statement within the first section.
V1.2.5	Merge	Workshop Feedback	It disappears as a separate point, only the focus on preserving skills for supervision remains
V1.2.7;V1.2.9; V1.3.3;V1.4.4	Removal	Workshop Feedback; Storyboard Activity	Were deemed superfluous, too difficult to evaluate or irrelevant
General Stament	Improvement	Workshop Feedback; Storyboard Activity	Refinement of vocabulary

A.5 Checklist V3

This checklist is meant to guide the decision process of someone who needs to delegate a worker's task to an AI tool, either by augmenting it or automating it.

The statements are organized in two ways: some can give positive or negative results, namely yes or no answers. Positive results favour the implementation of AI, whilst negative results do not. Other statements are meant to induce thought processes so that the checklist user can draw their own conclusions.

A.5.1 Checklist Statements.

(1) Verify that the task is fully defined by a set of rules that the AI model could easily follow.
(2) Check that the task does not involve particularly creative and/or emotional aspects.
(3) Verify that the model implementation does not completely overhaul the task's workflow.
(4) Verify that AI implementation has already been deemed socially acceptable in regard to the task.
(5) Assess whether the task inevitably requires human supervision.
(6) Ensure the task is physically augmentable or automatable.
(7) Verify that the output of the task can be objectively quantified and measured.
(8) Ensure that the implementation of the AI model does not cause radical job displacement.
(9) Verify that AI implementation in this task has been accepted on a technical level.
(10) Check whether there is already a quality data set to train the AI on the task, or whether it is easily acquirable.
(11) Validate whether the task retains value from human involvement and creativity.
(12) Make sure that the implementation of the AI model is economically viable.
(13) Evaluate whether the task legally requires human sign-off.
(14) Analyze the potential effect on worker skill level: will the implementation of the AI model upgrade the worker's skillset or downgrade it?
(15) Make sure that the AI model is fully capable of tackling the task.
(16) Verify that there are low to no possible mental health risks for the workers in case of AI model implementation.
(17) Assess whether the implementation of the AI model requires specific knowledge on the worker's behalf.
(18) Verify that the task delegation strategy preserves the worker's expertise necessary to safely oversee the task if needed.
(19) Ensure the workforce trusts the AI model that might be implemented.
(20) Verify that privacy and data security are not put at risk.
(21) Verify that no ethical concerns could arise from the implementation of the AI model.
(22) Ensure that the AI model is as transparent as possible in regard to its inner workings during the operation of the task.

A.5.2 Interview Questions.

For all stakeholders.

- Does this checklist represent the necessary criteria for the effective delegation of a task to an AI model?
- Should it be formatted in a different way?
- Are there any criteria that should be added?
- Are there any criteria that should be subtracted?

For worker roles.

- Do you think this is an appropriate decision process to delegate one of your tasks to AI?
- Imagine that you are a manager using this tool to decide whether you should delegate a task to AI: would this checklist serve its purpose effectively?

For managerial roles.

- Would you use this tool on a day-to-day basis to decide how to delegate a task to AI?
- Does the checklist work efficiently for its intended purpose?

A.6 Checklist V2–V3 Changelog

Table 3: Checklist V2–V3 changelog

V2→V3			
Item Affected	**Change Type**	**Feedback Source**	**Result**
Structure	Addition	Group Brainstorming	Addition of questions for specific stakeholders to gain more feedbacks
Structure	Removal	Group Brainstorming	Removal of subheadings among statements
V2.1.5	Split	Group Brainstorming	Splitted into separated items: "ensure the task is physically augmentable or automatable" and "Make sure that the AI model is fully capable of tackling the task"

A.7 Checklist V4

A.7.1 I. Intended Uses.

- **Algorithmic Rule Definition:** Verify the tak is governed by a stable, systematic rule set. *(Yes/No)*
- **Technical Feasibility:** Confirm the task is computationally automatable and the AI is capable of executing it. *(Yes/No)*
- **Objective Metrics:** Define measurable goals to quantify AI output and distinguish intentional results from errors.
- **Boundary Evaluation:** Confirm the task lacks deep creative or emotional nuances that require human involvement. *(Yes/No)*
- **Workflow Framing:** Verify the AI functions as a supportive tool rather than a manager or a total workflow replacement. *(Yes/No)*

A.7.2 II. Harms.

- **IP and Asset Protection:** Identify legal risks regarding copyright and ensure data is harvested transparently and legally.
- **Workplace Well-being:** Evaluate potential psychological or mental health impacts on workers.
- **Empirical Trust Building:** Demonstrate tool capabilities to build genuine social acceptance rather than demanding blind trust.

A.7.3 III. System and Data.

- **Dataset Availability:** Confirm a high-quality dataset is accessible or easily obtainable for training. *(Yes/No)*
- **Cost-Benefit of Data Prep:** Determine if data preparation costs exceed the economic gains of automation. *(Yes/No)*
- **Contextual Explainability:**
Establish technical transparency sufficient to justify the model's inner workings for its specific use.
- **Infrastructure Safeguards:** Verify the infrastructure ensures robust data security and corporate privacy. *(Yes/No)*

A.7.4 IV. Oversight.

- **Regulatory Compliance:** Determine if the task legally requires a qualified professional's final approval. *(Yes/No)*
- **Operational Overhead:** Confirm that oversight requirements do not make the AI workflow more labor-intensive than the human-driven version. *(Yes/No)*

A.7.5 V. Team.

- **Long-term Skill Impact:** Analyze whether AI deployment will professionally upskill or downgrade the workforce.
- **Workforce Adaptation:** Assess the need for task restructuring or specialized training programs.
- **Socio-Economic Stability:** Verify the implementation is managed to prevent unmitigated job displacement. *(Yes/No)*
- **Safe Oversight Maintenance:** Confirm the strategy preserves workers' core expertise so they remain capable of supervision. *(Yes/No)*
- **Resilience and Safeguards:** Maintain manual practices and workforce flexibility to ensure safe intervention during AI failures or hallucinations.

A.8 Checklist V3–V4 Changelog

Table 4: Checklist V3–V4 changelog

V3→V4			
Item Affected	**Change Type**	**Feedback Source**	**Result**
Structure	Addition	Stakeholders	Addition of titles and tag to group questions from the same category
Structure	Addition	Stakeholders	Addition of title before every statement to pointing out
Structure	Addition	Stakeholders	Addition of explicit "yes/no" answer after some statements
V4.2.1	Addition	Stakeholders (Electronic Designer)	The IP and Asset Protection point is added to assess legal risks on stolen or copyrighted data.
V4.4.2	Addition	Stakeholders (Product Manager)	AI must not make the workflow more laborious and expensive than manual work.
Added Item	Addition	Stakeholder (Company Consultant)	Added a specific criterion for "Intellectual Property and Asset Protection".
V4.5.5	Addition	Stakeholder	A new point has been added to the Team: manual procedures must be kept active to be able to intervene safely when the AI makes mistakes or stops working

A.9 Checklist V5

This checklist is made to underline the most important criteria regarding task delegation to an AI tool. Its ultimate goal is to promote the design of a more responsible and effective AI tool.

The checklist uses two scales: one binary system (yes or no), and a five-item Likert scale system (strongly agree, agree, neutral, disagree, strongly disagree), so as to align the checklist with a commonly used rating standard.

For each category, the items have been ordered by importance, going from top (most important) to bottom (least important).

A.9.1 I. Intended Uses.

- Verify that the task can be broken into simple steps. *(Yes/No)*
  The AI tool must be able to recognize the parts that make up its task.
- Confirm the AI is capable of executing the task. *(Yes/No)*
  The AI tool must have the technical capabilities to automate or augment the task, without becoming a burden to who implements it.
- Confirm the task does not involve deep creative or emotional aspects. *(Yes/No)*
  It has been observed that it is best not to delegate a task with such characteristics to an AI tool, even if it can realistically mimic human emotions and creativity.
- The task goals are measurable. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The AI tool must be able to give a quantifiable output, so as to rate its abilities.
- Verify the AI functions as a supportive tool. *(Yes/No)*
  The AI tool should be designed as an augment for the worker, especially if the task specifically gains value from being executed by a human.

A.9.2 II. Harms.

- Data security is guaranteed and is harvested transparently and legally. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The tool must guarantee the privacy of all stakeholders involved.
- There are potential risks regarding the mental health of workers. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The AI tool must not cause distress to the workers in any way.
- There are legal risks regarding potential copyrighted work. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The AI tool cannot rely on protected work for its output without proper authority.

A.9.3 III. System and Data.

- Confirm a high-quality dataset is accessible to train the tool. *(Yes/No)*
  The tool must be able to elaborate the available data to train for the task.
- The tool is transparent enough with its processes. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The tool must be able to clearly show the operations that were done to obtain an output, so as to be as transparent and trustworthy as possible.
- Determine if data and tool preparation costs exceed the economic gains. *(Yes/No)*
  The tool must be designed with an entry barrier that is as low as possible.

A.9.4 IV. Oversight.

- Determine if the tool's output legally requires a qualified professional's final approval. *(Yes/No)*
  The tool should be able to operate as independently as possible, without slowing down processes while waiting for outside approval.
- Confirm that oversight requirements do not make the AI workflow more labor-intensive. *(Yes/No)*
  The tool should be designed to need as little oversight as possible, so as not to complicate a process instead of simplifying it.

A.9.5 V. Team.

- Verify the implementation is managed to prevent job loss. *(Yes/No)*
  Either the tool gives the worker the possibility to requalify for a new job or it serves as a support and is not, most importantly, a cause for massive layoffs.
- Implementing the tool will professionally upskill or downgrade the workforce. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  Ideally, the tool should be designed with the intent of upgrading a worker's capability, not degrading it.
- The implementation of the tool brings significant changes to the workflow. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The AI tool should be built to insert itself seamlessly in an already established workflow.
- Confirm the tool implementation preserves workers' expertise so they remain capable of supervision. *(Yes/No)*
  Implementing an AI tool should not have negative consequences on the fundamental skills of a worker.
- Manual practices and workforce flexibility are maintained to ensure safe intervention during AI failures. *(strongly agree, agree, neutral, disagree, strongly disagree)*
  The tool should be designed so that workers can still retain fundamental skills, needed to intervene on the task if the AI malfunctions.

Decision Logic and Final Verdict
To translate the checklist responses into an actionable managerial decision, users must apply the following evaluation rubric:

(1) **Absolute Blockers (Yes/No Items):** The binary items act as fundamental operational and ethical gateways. If any critical prerequisite fails (e.g., the AI is incapable of executing the task, data preparation costs exceed gains, or workers' expertise is not preserved), the integration process must be halted. A single failure on a critical Yes/No constraint automatically triggers a "No-AI Delegation" verdict.
(2) **Weighted Scoring (Likert Items):** For the items evaluated on the 5-point Likert scale, assign points from 1 to 5 (where 5 represents the most favorable/safe condition for AI adoption).
(3) **Final Triage:** Based on the aggregation of the scores and the clearance of all blockers, the final verdict will fall into one of three categories:
   - **Full Automation / Augmentation Viable (High Score):** All blockers are cleared and the majority of Likert responses fall into the favorable spectrum (4-5). The AI integration can proceed as planned.
   - **Partial Automation Recommended (Medium Score):** All blockers are cleared, but Likert responses are clustered around "Neutral" (3) or reveal isolated vulnerabilities (e.g., high need for upskilling). The AI tool can be integrated, but only with strict human-in-the-loop oversight and targeted mitigation strategies.
   - **No-AI Delegation (Low Score or Blocker Triggered):** Negative/Risky responses dominate the Likert scale, or a Yes/No absolute blocker is triggered. The task must remain fully human-led.

A.10 Checklist V4–V5 Changelog

Table 5: Checklist V4–V5 changelog

V4→V5			
Item Affected	**Change Type**	**Feedback Source**	**Result**
All items	Lexicon Simplification	Peer Reviews/Lecture	Simplified lexicon for every item to improve clarity
Non-binary items	Response Format Update	Peer Reviews/Lecture	Introduced a 5-item Likert scale for all items that did not have a binary response system
All items	Content Addition (Design Guidelines)	Peer Reviews/Lecture	Added a design requirement under every item to explain how the development of an AI tool should tackle each criteria
All categories	Structural Reordering	Peer Reviews/Lecture	Ordered items in each category by importance.
Checklist preamble	Content Addition	Peer Reviews/Lecture	Added a brief description of the checklist and its structure.

A.11 Anonymized Notes

A.11.1 Interview 1: Strategic Cluster (Vice President & Commercial Director).

Structure & Usability. The participants noted that while the checklist works efficiently for its intended purpose, it felt "too idealistic on some topics". They requested to "enumerate the points" and "put the points in order of importance" to improve navigability.

Workforce Impact. They emphasized that expecting unconditional worker trust is an "unrealistic scenario", as lower-rank operational workers might not explore new technologies proactively. Furthermore, they explicitly stated that "for most companies job displacement is not a relevant topic".

Legal Accountability. They criticized generic human oversight items, requesting that the requirement for human sign-off be rewritten specifically as: "Evaluate if the task requires sign-off from a legally qualified subject".

Skills & Training. They suggested merging criteria regarding long-term skill impact, workforce adaptation, and the preservation of worker expertise into a single, comprehensive point. They stressed the need to include considerations for "eventual training required or the need of additional workforce".

A.11.2 Interview 2: Operational Cluster (Product Development Manager / Company Consultant).

Format & Usability. The participant advised to keep the tool brief and "don't overcomplicate the checklist, so as to have a clear output", suggesting to cut a few things if needed. They strongly recommended grouping criteria by specific decision areas, such as Tasks, Responsibility, and Risks/Compliance, rather than having a flat list.

Decision Process & Output. They noted that the checklist is a strong starting point and "will help in the delegation process, but must not be the only decision tool used". They also suggested explicitly stating the expected type of answer and inserting a final outcome evaluation, such as Automation possible, Partial automation, or No-AI, to make it a true managerial decision tool.

Customization. They emphasized that the checklist "must be personalized according to sector and task" to remain effective in real-world scenarios.

A.11.3 Interview 3: Operational Cluster (Electronic Designer / Product Manager).

Operational Effort & Cost-Benefit. The participant strongly criticized requirements for constant human supervision or human signatures, repeatedly noting that they "would require double the effort, making it (the automation) useless". They also warned that acquiring quality datasets or performing complex data preparation "risks costing more than the result".

Intellectual Property (IP). They raised significant concerns regarding copyright and intellectual property. They highlighted the risk of AI models relying on the "misappropriation of information", making it difficult to protect against "intellectual thefts".

AI as a Tool & Trust. They firmly stated that AI should be integrated strictly as "a tool and not a manager". Instead of demanding blind trust a priori, they suggested building acceptance through a "practical demonstration of potential". They also recommended focusing on possible job "remodulations and qualifications" rather than just job displacement.

Metrics & Outcomes. They recommended setting pre-established objectives to quantify outputs and clearly understand "what has been achieved voluntarily and what involuntarily" by the AI.

References

[1] Co-Design AI Team. 2019. AI Fairness Checklist and Workbooks. (2019).
[2] DrivenData. 2018. Deon: An ethics checklist for data scientists. (2018). http://deon.drivendata.org/
[3] European Union High-level Expert Group. 2019. Ethics Guidelines for Trustworthy AI: Building trust in human-centric AI. (2019). https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines
[4] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé III, Miro Dudik, and Hanna Wallach. 2019. Improving fairness in machine learning systems: What do industry practitioners need?. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, 1–18. DOI: http://dx.doi.org/10.1145/3290605.3300830
[5] Machine Intelligence. 2019. AI Ethics Framework. (2019). https://www.migarage.ai/ethics-framework/
[6] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. Nature Machine Intelligence (Sept. 2019), 1–11.
[7] Mike Loukides, Hilary Mason, and DJ Patil. 2018. Of Oaths and Checklists. (2018). https://www.oreilly.com/ideas/of-oaths-and-checklists
[8] Arvind Narayanan. 2018. Translation tutorial: 21 fairness definitions and their politics. In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*). New York.
[9] Google Research. 2019. What if Tool. (2019). https://pair-code.github.io/what-if-tool/
[10] Pedro Saleiro, Benedict Kuester, Loren Hinkson, Jesse London, Abby Stevens, Ari Anisfeld, Kit T. Rodolfa, and Rayid Ghani. 2018. Aequitas: A Bias and Fairness Audit Toolkit. (2018). http://arxiv.org/abs/1811.05577
[11] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. (2019), 59–68.
[12] Luke Stark and Anna Lauren Hoffmann. 2019. Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture. Journal of Cultural Analytics (May 2019).
[13] United Kingdom, Department of Digital, Culture, Media and Sport. 2019. Data Ethics Workbook. (2019). https://www.gov.uk/government/publications/data-ethics-workbook/data-ethics-workbook
[14] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In Proceedings of the 2018 CHI conference on human factors in computing systems. 1–14. DOI: http://dx.doi.org/10.1145/3173574.3174014

A.12 AI Disclosure

Verification. We confirm that all team members personally read, verified, and checked all cited sources. Also all AI suggestions were checked by team members.

AI Tool	Usage
Gemini	Gemini was used for spell checking and grammatical correction. It was also used to help draft graph structures and LaTeX formatting inside the paper.

A.13 Team Information

Team Name. In ChatGPT We Trust.

Team Members.

- Gianfranco Molino (s335970@studenti.polito.it)
- Giacomo Dettori (s325913@studenti.polito.it)
- Alessandro D'Ettorre (s340380@studenti.polito.it)
- Mario Zoda (s343450@studenti.polito.it)
- Luca Costa (s340699@studenti.polito.it)
- Darla Shahu (s323199@studenti.polito.it)

- Yalda Ahmadi Soumeh Saraei (s345681@studenti.polito.it)
- Nicolas Panetti (s340240@studenti.polito.it)
- Alessio Delli Santi (s337380@studenti.polito.it)
- Jia Tang (s284162@studenti.polito.it)