



Ethical Criteria for AI Delegation: A Scoping Review from the Workers' Perspective

Abstract

Artificial Intelligence (AI) is fundamentally transforming decision-making processes across high-stakes domains; however, delegating complex tasks to automated systems raises profound ethical and psychological challenges. To address these criticalities, we conducted a scoping review to identify the factors determining whether a decision-making task should be delegated to an AI system by a human worker. Through a Reflexive Thematic Analysis (RTA) of contemporary literature (2015–present), we mapped six interrelated ethical criteria for AI delegation from the worker's perspective: (3.1) task moral delegability; (3.2) accountability and responsibility; (3.3) autonomy, agency, and professional identity; (3.4) transparency and explainability; (3.5) fairness, bias, and third-party harm; and (3.6) human-AI collaboration and governance. These dimensions form a multifaceted framework in which structural conditions such as transparency and governance determine whether human agency and professional integrity can be preserved. Applying this framework to the use case of medical radiology and deep learning triage, we demonstrate that total task delegation remains ethically unacceptable across all six criteria. We conclude that AI must be integrated strictly as an augmenting tool within a shared-control paradigm, ensuring that human professionals retain final accountability and moral agency over every consequential decision.

Keywords: AI Ethics, Algorithmic Delegation, Moral Distance, Accountability, Professional Identity, Shared Control.

1. Method

This study adopts a scoping review methodology to identify and map the main ethical themes related to AI task delegation in the workplace. The purpose of the review was not to evaluate the effectiveness of specific AI systems, but to explore the conceptual and ethical factors that influence whether a task can be legitimately delegated to artificial intelligence from the workers' perspective. The literature search was conducted across three academic databases: Google Scholar, ACM Digital Library, and Scopus. Search strings were designed to reflect the main themes of the review and were progressively refined through an iterative process to improve specificity and relevance. This approach allowed us to reduce the number of irrelevant results before the screening phase and to build a source pool already aligned with the conceptual framework of the study. The searches were performed between April 17th and April 21st, 2026. A publication date filter was applied to include only studies published from

2015 onward, ensuring the review focused on contemporary discussions related to modern AI systems. In Google Scholar, additional filters were used to prioritize academic and peer-reviewed sources, while on Scopus we limited results to scientific reviews, filtering by keywords such as 'scoping review', 'workers' and 'workplace'. The following search strings were used.

	Database	Search String
1	Google Scholar	intitle:"artificial intelligence" ("human competencies" OR "flow experience") (education OR classroom) AND ("workers" OR "work") AND ("practical" OR "practice") AND "ethics" AND "values" AND "principles" AND "employees"
2	Google Scholar	intitle:"artificial intelligence" ("human competencies" OR "flow experience") (education OR classroom) -physics -fluid -business -management
3	Google Scholar	("AI ethics" OR "responsible AI" OR "trustworthy AI") AND (guidelines OR standard OR policy OR framework) AND "perspective" AND "workers" AND "AI regulation" AND "case study" AND "practical" AND "human rights" AND "governance" AND ("augment" OR "substitute")
4	Google Scholar and ACM DL	("artificial intelligence" OR AI) AND "task delegation" AND (worker OR employee) AND ("human dignity" OR "worker autonomy" OR "moral responsibility" OR "meaningful work")

Database	Search String
5 ACM DL	("AI" OR "algorithm") AND ("task delegation" OR "decision making") AND "ethic" AND "trustworthiness" AND "transparency" AND "moral responsibility"
6 ACM DL	"artificial intelligence" AND ("working place*" OR "future of work" OR "occupational impact") AND ("explainable AI" OR "meaningful human control" OR "moral agency")
7 Scopus	artificial intelligence AND ("working place*" OR "future of work" OR "occupational impact") AND ("explainable AI")

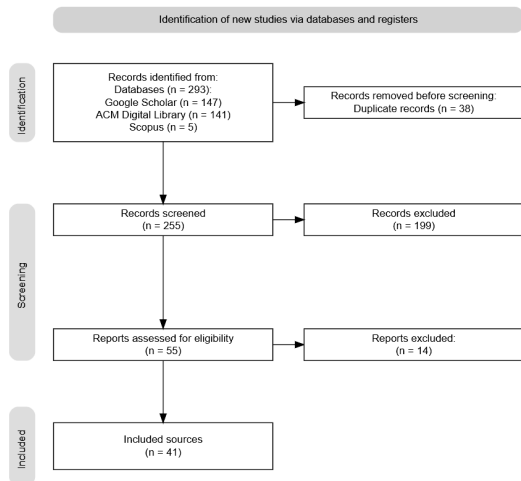
	Inclusion Criteria	Exclusion Criteria
Topic	Addresses factors influencing AI task delegation judgments; covers autonomy, responsibility, fairness, or human dignity in the context of AI	Purely technical papers on AI performance or engineering without normative/ethical dimension
Publication type	Peer-reviewed articles, conference papers (ACM, CHI, NeurIPS), institutional reports by major bodies (WHO, WEF)	Blog posts, non-peer-reviewed preprints without affiliation, gray literature of unclear origin
Language	English or Italian	Other languages
Date	2015–2026 (inclusive); pre-2015 foundational works retained if seminal (see justification below)	Pre-2015 works with no clear foundational status
Scope	General or sector-specific studies that allow cross-domain transferability of findings	Studies exclusively focused on a single narrow technical tool without generalizable conclusions

The initial search produced a total of 293 records. After removing duplicates across databases and overlapping search strings, only 255 records remained. A preliminary screening was then conducted, resulting in the exclusion of the majority of sources and leaving 55 articles for eligibility assessment. This screening was conducted based on title and abstract and excluded papers according to a few criteria. For example, purely technical papers that covered little to none ethical topics were removed, allowing us to include papers we identified as relevant. These articles were then examined through full-text reading, leading to the exclusion of 14 additional studies, as shown in the PRISMA diagram.

**Table with the inclusion/exclusion criteria*

2. Results

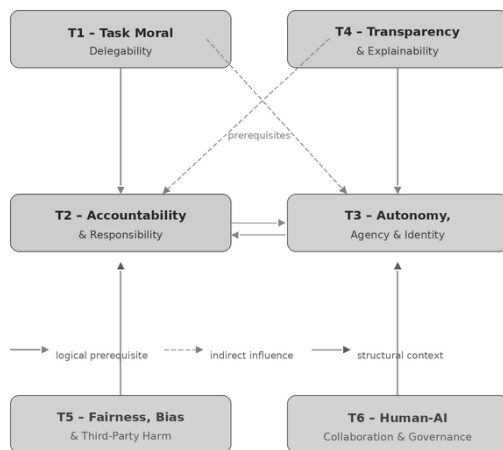
The review included 41 studies, which examine the ethical criteria for delegating tasks to Artificial Intelligence (AI) from the workers' perspective. The selected sources addressed the topic covering a wide variety of contexts, focusing both on generic or cross-domain work dynamics. Regarding the types of integration and the object of delegation, the studies range from automated systems based on deep learning to contemporary algorithmic assistants and agents. From the analysis of the studies included in the corpus, we developed a mapping of six interconnected ethical criteria that determine the delegability of a task to AI, framing them based on their cognitive and moral impact on the worker. This analysis was conducted to evaluate which factors establish when it is morally appropriate to delegate a task to an automated system and whether such delegation compromises professional integrity. In the following sections, we first described the six main criteria identified in the ethical framework: (3.1) task moral delegability; (3.2) accountability and responsibility; (3.3) autonomy, agency, and professional identity; (3.4) transparency and explainability; (3.5) fairness, bias, and third-party harm; and (3.6) human-AI



**PRISMA diagram*

collaboration and governance. Subsequently, we applied these concepts to the practical use case of medical radiology to examine how a shared-control paradigm allows AI to act as an augmenting tool rather than a replacement.

3. Framework



**Framework diagram: T1 and T4 as epistemic prerequisites; T2 and T3 as core concepts; T5 and T6 as structural context*

Core Question	
T1	What characteristics of a task determine its ethical delegability?
T2	Who is answerable when AI acts on behalf of a worker?
T3	How does delegation reshape worker freedom, skills, and meaning?
T4	Can the AI's reasoning be understood and verified?
T5	Does delegation risk producing unjust outcomes for others?
T6	What oversight structures make delegation legitimate?

3.1 Task Moral Delegability (T1)

Before questions of accountability, transparency, and governance can be meaningfully addressed, the worker must assess whether the nature of the task itself makes delegation to AI morally permissible. The literature converges on a multi-dimensional answer structured around the intrinsic properties of the task, its relational and emotional stakes, and the irreplaceability of human judgment in specific decision contexts. The most consistently identified criterion is the distinction between routine,

well-structured work and tasks requiring subjective judgment, soft skills, or contextual interpretation [1, 2, 3, 4]. Hermann et al. [1] argue that delegating cognitively monotonous tasks is ethically appropriate, whereas delegating tasks constituting the core of professional value strips work of its ethical weight. Turunen [3] extends this boundary to the relational plane, while Bankins et al. [4] add that algorithm aversion intensifies when AI handles subjective or symbolically significant tasks. A second cluster of constraints concerns emotional depth and moral stakes. Singh [5] identifies empathy, moral judgment, and contextual wisdom as structural boundaries of delegation. Freisinger and Schneider [6] show empirically that in high-impact decisions, affected parties resist AI-driven outcomes because AI cannot assume vulnerability or bear moral responsibility. George et al. [8] document how over-automation produces quality decline and trust erosion. The degree of autonomy granted to the AI system is itself an ethical variable [9, 10]: full automation is morally appropriate only for narrowly defined, low-stakes tasks where meaningful human validation remains structurally possible [5, 9, 10]. Four dimensions determine task moral delegability: objective versus subjective nature; emotional and relational demands; severity and irreversibility of consequences; and structural possibility of human oversight. Where these signal that delegation is inappropriate, no accountability or governance mechanism can render it ethically legitimate.

3.2 Accountability and Responsibility (T2)

The literature converges on a foundational principle: delegating a task to AI does not transfer the worker's moral or legal accountability. Responsibility is consistently framed as inalienable — yet the structural conditions of AI-mediated work systematically undermine its exercise. A first cluster concerns the structural dilution of accountability. When opaque AI systems produce errors, clearly ascribing responsibility becomes analytically intractable [7], and workers may be held responsible for AI-driven outputs over which their effective control is limited [1]. Legal uncertainty further amplifies the problem: the absence of clear liability frameworks constitutes a systemic ethical obstacle to delegation across multiple professional domains [12]. When AI

systems lack transparent decision logs, organisational accountability is materially compromised [2], and the gap between ethical principles affirmed in governance frameworks and their empirical implementation weakens accountability structures in practice [17]. A second cluster addresses the instrumentalisation of AI as a mechanism for blame avoidance. Delegation is sometimes motivated not by efficiency considerations but by the intent to avoid personal responsibility for difficult decisions [13] — a "buck-passing" pattern identified as a morally impermissible delegation intent [6]. Qualitative studies further document moral disengagement mechanisms through which AI delegation functions as a vehicle for moral detachment rather than a legitimate operational tool [3]. A third cluster identifies non-delegable accountability floors. The EU AI Act mandates meaningful human oversight for high-risk applications including HR management, hiring, and dismissal [15]. Corporate governance frameworks frame fiduciary duties as legally requiring that final judgment remain a human prerogative [16], while in public administration human discretion over algorithmic recommendations is a condition of democratic legitimacy [14].

A fourth cluster addresses the practical oversight burden delegation places on workers [11, 18]. Ethical delegation is viable specifically when accountability structures are clearly defined *ex ante*, AI systems provide traceable reasoning, and oversight mechanisms are embedded in organisational workflows rather than left to individual discretion [7, 2].

3.3 Autonomy, Agency, and Professional Identity (T3)

The ethical delegability of tasks to AI is fundamentally centered on the preservation of the worker's autonomy, agency, and professional identity. Delegation is perceived as ethically appropriate when it reduces cognitive load and task monotony, allowing workers to focus on high-value roles such as creativity and critical thinking [1, 11]. Conversely, uncontrolled AI integration can erode human agency by progressively eliminating "choice and chance" in the workplace, reducing the

professional to a passive observer of automated processes [19, 20].

A critical risk is the "decoupling" between accountability and agency: workers are often held responsible for AI-generated outputs over which they have exercised significantly limited control [1, 12]. This ambiguity can trigger moral disengagement mechanisms, leading employees to shift responsibility onto the machine or use AI as a "scapegoat" to avoid blame [3, 13]. Furthermore, the systematic outsourcing of complex tasks threatens long-term professional growth through "deskilling," as workers progressively lose fundamental analytical and problem-solving skills [1, 23]. This erosion is further exacerbated by continuous algorithmic monitoring, which risks reducing professional work to merely measurable tasks [2].

Task delegability is also strictly conditioned by its nature: while AI excels in mechanical and analytical roles, its application in human-centric tasks — requiring empathy, intuition, and moral judgment — remains ethically problematic [5, 7]. Total automation is often resisted because work is not a mere economic transaction; it provides identity, purpose, and social connection — elements systematically stripped away when human labor is replaced [8].

To maintain ethical alignment, AI must function as a collaborative partner in a Human-in-the-loop model, where transparency and "causability" ensure that professionals can exercise informed judgment rather than acting as passive signatories [5, 24]. Ultimately, AI adoption must prioritise the augmentation of human capabilities, ensuring that technology supports the identity and meaning of work within a structured ethical framework [23, 22, 25].

3.4 Transparency and Explainability (T4)

Transparency and explainability define the epistemic threshold under which a task can be ethically delegated to AI. A task appears more delegable when the worker can understand the basis of the AI output, question it, and retain a meaningful capacity to accept, modify, or reject the recommendation. Perceived fairness depends not

only on statistical accuracy but on whether the system provides context-sensitive reasons that make decisions intelligible [26]. Trust must be calibrated: both overtrust and undertrust are problematic for responsible delegation [28]. The literature identifies interpretability as a condition for preserving worker agency. Workers are more willing to delegate when AI systems expose their logic transparently, enabling them to interrogate outputs rather than merely execute them [7]. Opaque systems create accountability gaps in high-stakes domains — including legal decision-making, finance, healthcare, and HR — where unexplained outputs may reinforce historical biases while making responsibility difficult to locate [27, 23]. The ethical issue is therefore not opacity in the abstract, but opacity in relation to tasks where the worker is expected to remain accountable. Several sources distinguish ordinary explainability from stronger causal understanding. In safety-critical settings, "causability" is necessary for workers to retrace why an AI made a recommendation rather than rubber-stamping its outputs [29]. Delegation is more appropriate when AI augments rather than replaces human performance [30], and when transparency is assessed alongside fairness, accuracy, and risks of dehumanization [31]. Delegation becomes ethically problematic when responsibility is ambiguous or when workers cannot determine who is answerable for outputs [32]. The regulatory recognition of a right to human explanation for significant algorithmic decisions supports transparency as a minimum condition for legitimate delegation [33], while controlled integration — in which workers retain active oversight and the capacity to challenge AI predictions — represents the clearest positive condition for ethical delegability [34].

3.5 Fairness, Bias, and Third-Party Harm (T5)

The moral delegability of a task to AI is deeply connected with the negative externalities such a choice may generate towards third parties. A primary obstacle is algorithmic bias: workers judge delegation as unethical when automation risks amplifying biases in decisions with high human impact [36, 2], particularly in automated screenings where delegation criteria may perpetuate pre-existing biases to the detriment of vulnerable groups [35], or in healthcare, where algorithmic

tools raise severe ethical concerns regarding reliability [23, 25].

The literature however reveals an interesting dichotomy: under specific conditions, algorithmic delegation is evaluated positively, as automated systems are believed to promote greater neutrality by mitigating the cognitive and social biases that typically affect human judgment [6]. Further complicating this framework is sensitive data, which creates an intrinsic privacy risk the worker must carefully assess before delegating [2, 36]. The literature thus positions the professional as an "ethical gatekeeper": delegation is morally appropriate only when the AI system can effectively balance bias reduction with rigorous mitigation of algorithmic discrimination and privacy risks.

3.6 Human-AI Collaboration and Governance (T6)

T6 maps the systemic and organisational conditions under which task delegation to AI can be considered ethically legitimate. The central finding is that moral adequacy cannot be evaluated at the level of individual tasks alone: it is co-determined by the structural oversight architecture within which workers and AI systems operate. Wang et al. [38] distinguish interaction from genuine human-AI collaboration — only the latter preserves the worker's capacity to oversee and remain accountable for AI-driven outputs. The literature converges on the human-in-the-loop model as the primary structural response: delegation is morally appropriate when the human remains in the loop, handling edge cases requiring contextual judgment [37], and within augmentation rather than substitution frameworks [5]. Empirical evidence shows that workflow misalignment and system opacity — rather than AI use per se — determine whether delegation is experienced as ethically problematic [39]. Organisational readiness and genuine stakeholder inclusion are further necessary governance conditions [40], while AI trustworthiness and cultural readiness constitute

4. Worked Use Case: The Healthcare Professional (Radiology)

To contextualize the six ethical criteria, consider the integration of a deep learning triage system — such as those deployed for chest X-ray or CT scan prioritization — in a hospital radiology department, where an algorithm flags cases as high-priority or benign before the radiologist reviews them, often processing hundreds of scans per shift. While preliminary scan sorting is sufficiently structured to permit partial delegation, radiological diagnosis requires contextual judgment that current systems cannot replicate: a nodule flagged as benign may be malignant in a patient with a specific smoking history or genetic predisposition that the algorithm never sees. Given the severity and irreversibility of a missed malignancy, full delegation remains impermissible. If the system produces a false negative, liability becomes analytically intractable: the radiologist who approved an opaque risk score becomes the "moral crumple zone," bearing professional and legal blame for a failure they could not audit. In many hospital settings, such systems are introduced by administrative mandate to reduce reporting backlogs, without clinical staff consultation — accelerating throughput at the cost of genuine oversight. The resulting time pressure transforms the radiologist's review into a rubber-stamp exercise, compounding deskilling over time as pattern-recognition skills atrophy from disuse. Systems trained predominantly on datasets from specific demographics or scanner models risk generating disproportionate false negatives for underrepresented patient groups, producing differential clinical harm that the radiologist — positioned as ethical gatekeeper — has no practical means to detect without demographic validation data. Finally, most deployments offer no co-designed override mechanism: the radiologist can technically reject the AI output, but workflow design, volume pressure, and institutional incentives structurally discourage doing so. Consequently, total delegation remains ethically unacceptable across all six criteria, permissible only within a strict shared-control paradigm where the radiologist retains sufficient time, structural authority, and interpretive access to exercise genuine clinical and moral oversight over every consequential decision.

5. Gaps and future work

This scoping review reveals a persistent gap between the theoretical articulation of ethical principles for AI delegation and their concrete operationalisation in professional environments. On task moral delegability, future research must

develop empirically validated thresholds applicable across sectors and roles. On accountability, interdisciplinary work must design liability models that redistribute responsibility along the actual causal chain rather than concentrating it on frontline workers. On autonomy and professional identity, longitudinal studies must measure the cumulative effects of AI adoption on deskilling over time. On transparency, safety-critical domains urgently require causability standards, as no minimum explainability threshold has yet been defined. On fairness, design solutions balancing algorithmic neutrality against structural bias reproduction remain absent, as does any framework addressing workers' right to refuse delegation when third-party data protection cannot be guaranteed. Finally, on governance, genuine Human-in-the-Loop architectures remain largely untested, and future research must verify whether co-designed workflows materially reduce the ethical friction documented in this review.

6. Conclusion

This scoping review examined the ethical conditions under which human workers can legitimately delegate tasks to AI systems, revealing that without structural safeguards the ethical burden of delegation is systematically absorbed by frontline workers, who risk becoming "moral crumple zones" for opaque algorithmic failures. The review makes three main contributions. First, it provides a six-criterion framework — covering task delegability, accountability, autonomy, transparency, fairness, and governance — to evaluate the moral permissibility of AI delegation from the worker's perspective. Second, it applies this framework to the radiology use case, demonstrating that total delegation fails across all dimensions simultaneously. Third, it maps the principal gaps in the current corpus — including the absence of longitudinal empirical studies and the inadequacy of liability frameworks — identifying these as the frontier for future interdisciplinary research.

We conclude that AI must be integrated strictly as an augmenting tool within a shared-control paradigm, ensuring that human-AI collaboration enriches professional identity rather than reducing workers to passive, stressful monitoring.

References

- [1] Hermann, A., Puntoni, S., and Morewedge, C.K. 2025. GenAI and the Psychology of Work. *Working Paper*.
- [2] Somé, L. 2023. Automated Decision-Making in Project Management. *Journal of AI and Management Studies*.
- [3] Turunen, M. 2025. Review of Employee-Focused Studies on Task Delegation to Service Robots. *Human-Computer Interaction*.
- [4] Bankins, S., Hu, J., and Yuan, Z. 2024. Artificial Intelligence, Workers, and Future of Work Skills. *Journal of Vocational Behavior*.
- [5] Singh, D. 2026. Advanced Human-AI Synergy for Effective Collaboration: Ethical Implications and Future Prospects. *Inteletica – Iberamia Journal*. <https://inteletica.iberamia.org/index.php/journal/article/view/61/33>
- [6] Freisinger, E. and Schneider, S. 2025. Decoding Decision Delegation to Artificial Intelligence. *Journal of Business Ethics*.
- [7] Al Masaeid, T., Al-Adwan, A.S., and Al-Debei, M.M. 2025. AI-Augmented Decision-Making: Ethical Dimensions and Accountability Gaps. *Decision Support Systems*.
- [8] George, A., Sagayarajan, S., and Baskar, T. 2026. The Fifth Wave of Automation: AI, Work Displacement, and Social Consequences. *Futures*.
- [9] Lubars, B. and Tan, C. 2019. Ask Not What AI Can Do, But What AI Should Do: Towards a Framework of Task Delegability. In *Proceedings of NeurIPS 2019 Workshop on Human-Centered Machine Learning*.
- [10] Candrian, C. and Scherer, A. 2022. Rise of the Machines: Delegating Decisions to Autonomous AI. *Computers in Human Behavior* 134 (2022), 107308. <https://doi.org/10.1016/j.chb.2022.107308>
- [11] Zhao, X. and Ma, L. 2026. How Do Interactions With Generative AI Impact Employees' Creative Performance? *Journal of Applied Psychology*.
- [12] Hassan, N., Kushniruk, A., and Borycki, E. 2024. Barriers and Facilitators to AI Adoption in Healthcare: A Systematic Review. *Journal of Medical Internet Research*.
- [13] Samland, J. 2025. When Do Decision-Makers Delegate Strategic Decisions to AI? *Organizational Behavior and Human Decision Processes*.
- [14] Dominek, A. 2025. The Interaction between Artificial Intelligence and Human Competencies in Public Sector Decision-Making. *Government Information Quarterly*.
- [15] European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.
- [16] Organisation for Economic Co-operation and Development (OECD). 2024. Corporate Governance and AI: Fiduciary Duties and Duty of Care in Algorithmic Decision-Making. *OECD Policy Note*.
- [17] Tang, C., Ross, K., Ghassemi, M., and Kaminski, M.E. 2023. Medical AI Ethics: A Systematic Review of Empirical Studies. *npj Digital Medicine* 6, 1 (2023). <https://doi.org/10.1038/s41746-023-00757-3>
- [18] Wang, F., Liu, X., and Chen, Y. 2023. Ethical Considerations of Using ChatGPT in Health Care: A Systematic Review. *Journal of Medical Internet Research* 25 (2023), e48009.
- [19] Schmidt, F.A. 2021. The End of Serendipity: Algorithmic Control in Digital Labour Platforms. *Work, Employment and Society* 35, 6 (2021), 1071–1087.

- [20] Shneiderman, B. 2020. Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems. *ACM Transactions on Interactive Intelligent Systems* 10, 4 (2020), 1–31. <https://doi.org/10.1145/3419764>
- [21] Villegas-Galaviz, C. and Martin, K.E. 2023. Moral Distance, AI, and the Ethics of Care. *AI & Society* 38 (2023), 1221–1232. <https://doi.org/10.1007/s00146-022-01412-x>
- [22] Veisi, S., Hajli, N., and Akhavan, P. 2025. AI Ethics and Social Norms: Exploring ChatGPT's Capabilities in Professional Work Contexts. *Technology in Society*.
- [23] Bankins, S. and Formosa, P. 2023. The Ethical Implications of Artificial Intelligence (AI) for Meaningful Work. *Journal of Business Ethics* 185, 4 (2023), 725–740. <https://doi.org/10.1007/s10551-023-05339-7>
- [24] Constantinides, M., Dowell, J., Johnson, D., and Papaioannou, T. 2025. The Future of Work is Blended, Not Hybrid: AI Co-workers and the Transformation of Human Roles. *ACM CHI 2025 Extended Abstracts*.
- [25] Marques, A., Almeida, M., and Pereira, R. 2024. The Medicine Revolution Through AI: Ethical Challenges and Governance Frameworks. *Frontiers in Medicine*.
- [26] Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., and Shadbolt, N. 2018. 'It's Reducing a Human Being to a Percentage': Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, Paper 377, 1–14. <https://doi.org/10.1145/3173574.3173951>
- [27] Asimiyu, A. 2025. The Ethics of Delegating Public Decision-Making to Autonomous Systems. *Philosophy & Technology* 38 (2025).
- [28] Rosenbacke, R., Sundström, N., Gyllenswärd, G., Alfredsson, J., and Oras, P. 2023. How Explainable AI Can Increase or Decrease Clinicians' Trust in AI Decision Support. *Frontiers in Digital Health* 5 (2023). <https://doi.org/10.3389/fdgth.2023.1170771>
- [29] Sorantin, E. 2022. The Augmented Radiologist: Artificial Intelligence in the Practice of Radiology. *Pediatric Radiology* 52 (2022), 2074–2086. <https://doi.org/10.1007/s00247-021-05130-6>
- [30] Tsiakas, K. and Murray-Rust, D. 2022. Using Human-in-the-Loop and Explainable AI to Envision Future Work Practices. In *Proceedings of the 14th International Conference on Automotive User Interfaces (AutoUI '22)*. ACM.
- [31] Park, J., Ahn, J., Hosanagar, K., and Lee, J. 2022. Designing Fair AI in Human Resource Management. *Journal of Management Information Systems* 39, 4 (2022), 926–960.
- [32] Koch, P., Mackay, W., and Schmidt, A. 2025. The Role of Design in the Future of Human-AI Interaction. *ACM Interactions* 32, 1 (2025).
- [33] European Union. 2024. Directive of the European Parliament and of the Council on Improving Working Conditions in Platform Work. *Official Journal of the European Union*.
- [34] Göndöcs, D. and Dörfler, V. 2024. AI in Medical Diagnosis: AI Prediction and Human Judgment. *AI & Society* 39 (2024), 123–139. <https://doi.org/10.1007/s00146-023-01659-2>
- [35] Raghavan, M., Barocas, S., Kleinberg, J., and Levy, K. 2020. Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. In *Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency (FAccT '20)*. ACM, New York, NY, 469–481. <https://doi.org/10.1145/3351095.3372828>
- [36] Ray, P.P. 2025. A Review on Agent-to-Agent Communication Protocols and Ethical Implications for Sensitive Data Delegation. *IEEE Access*.

- [37] Lazaros, E., Tzovla, E., and Kotsiantis, S. 2026. Human-in-the-Loop Artificial Intelligence: A Systematic Review. *Entropy* (MDPI). <https://www.mdpi.com/journal/entropy>
- [38] Wang, D., Churchill, E., Maes, P., Selman, B., Shen, Y., Shi, Y., and Terry, M. 2020. From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. ACM, New York, NY, 1–6. <https://doi.org/10.1145/3334480.3381069>
- [39] Wang, X., Shi, Y., Zhang, B., and Hua, Y. 2021. "Brilliant AI Doctor" in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, Article 43, 1–18. <https://doi.org/10.1145/3411764.3445432>
- [40] Sabat Sani, A., Fagbule, O.F., and Edemekong, P.F. 2025. Adoption of Artificial Intelligence in Primary Health Care: Systematic Synthesis of Stakeholder Perspectives. *BMC Primary Care* 26, 1 (2025). <https://link.springer.com/article/10.1186/s12875-025-03157-6>
- [41] Tangi, L., Rodriguez Müller, A.P., and van Noordt, C. 2024. Exploring the Challenges for Artificial Intelligence Adoption and Implementation in EU Public Administrations: Results from a Case Study Analysis. *ACM Journal on Responsible Computing* 1, 4 (2024). <https://dl.acm.org/doi/full/10.1145/3776547>

Appendix

A1. Team Members

Surname	Name	Course of study	University	Email
Accorroni	Lorenzo	Aerospace Engineering	Turin Polytechnic	s341682@studenti....
Casalegno	Luca	Management Engineering	Turin Polytechnic	s341804@studenti....
Leocata	Salvatore Daniel	Computer Engineering	Turin Polytechnic	s325651@studenti....
Liao	Livio	Management Engineering	Turin Polytechnic	s335883@studenti....
Lieggi	Thomas	Energy Engineering	Turin Polytechnic	s342299@studenti....
Mascherin	Jacopo	Management Engineering	Turin Polytechnic	s340620@studenti....
Muratori	Matteo	Aerospace Engineering	Turin Polytechnic	s338365@studenti....
Mustaj	Redi	Management Engineering	Turin Polytechnic	s341630@studenti....
Rinaldi	Pietro	Automotive Engineering	Turin Polytechnic	s341323@studenti....
Valori	Marco	Mathematical Engineering	Turin Polytechnic	s340261@studenti....
Ziarati Niasar	Davide	Computer Engineering	Turin Polytechnic	s321920@studenti....

A2. Databases Searched and Dates

The following databases were systematically searched to identify relevant literature.

Database	Query Focus	Date(s)	Results Found
Google Scholar	AI task delegation, ethics, autonomy, accountability, impact on work	~20 Apr 2026	147
ACM Digital Library	Human-AI collaboration, delegation framework, fairness, ethics	~20 Apr 2026	141
Scopus	AI impact on employment	~20 Apr 2026	5

A3. Search Strings

The following eight search strings were used across the identified databases. Strings were iteratively refined based on preliminary results to maximize coverage of ethical, normative, and empirical literature on AI delegation. In this table, you can see that the 4th string and the 7th are identical. We used the same string in both Google Scholar and ACM database as it returned a low number of results in both. To filter the resulting papers, we applied both a time filter (excluding any source that was published before 2015) and a “Only academic papers” filter. Additionally, using Scopus, we applied these filters: limited to review, keywords: scoping review, workers, workplace.

	Database	Search String
1	Google Scholar	intitle:"artificial intelligence" ("human competencies" OR "flow experience") (education OR classroom) AND ("workers" OR "work") AND ("practical" OR "practice") AND "ethics" AND "values" AND "principles" AND "employees"
2	Google Scholar	intitle:"artificial intelligence" ("human competencies" OR "flow experience") (education OR classroom) -physics -fluid -business -management
3	Google Scholar	("AI ethics" OR "responsible AI" OR "trustworthy AI") AND (guidelines OR standard OR policy OR framework) AND "perspective" AND "workers" AND "AI regulation" AND "case study" AND "practical" AND "human rights" AND "governance" AND ("augment" OR "substitute")
4	Google Scholar	("artificial intelligence" OR AI) AND "task delegation" AND (worker OR employee) AND ("human dignity" OR "worker autonomy" OR "moral responsibility" OR "meaningful work")
5	ACM DL	("AI" OR "algorithm") AND ("task delegation" OR "decision making") AND "ethic" AND "trustworthiness" AND "transparency" AND "moral responsibility"
6	ACM DL	"artificial intelligence" AND ("working place*" OR "future of work" OR "occupational impact") AND ("explainable AI" OR "meaningful human control" OR "moral agency")
7	ACM DL	("artificial intelligence" OR AI) AND "task delegation" AND (worker OR employee) AND ("human dignity" OR "worker autonomy" OR "moral responsibility" OR "meaningful work")
8	Scopus	artificial intelligence AND ("working place*" OR "future of work" OR "occupational impact") AND ("explainable AI")

A4. Inclusion and Exclusion Criteria

Sources were screened at title/abstract level against the following criteria.

	Inclusion Criteria	Exclusion
Topic	Addresses factors influencing AI task delegation judgments; covers autonomy, responsibility, fairness, or human dignity in the context of AI	Purely technical papers on AI performance or engineering without normative/ethical dimension
Publication type	Peer-reviewed articles, conference papers (ACM, CHI, NeurIPS), institutional reports by major bodies (WHO, WEF)	Blog posts, non-peer-reviewed preprints without affiliation, gray literature of unclear origin
Language	English or Italian	Other languages
Date	2015–2026 (inclusive); pre-2015 foundational works retained if seminal (see justification below)	Pre-2015 works with no clear foundational status
Scope	General or sector-specific studies that allow cross-domain transferability of findings	Studies exclusively focused on a single narrow technical tool without generalizable conclusions

A5. Screening Funnel

The figure below summarizes the record flow from initial identification to final corpus inclusion, following a PRISMA-inspired approach.

Stage	Count	Notes
Records identified across all databases	293	Google Scholar + ACM DL + Scopus
Records after duplicate removal	255	Approximate estimate of non-duplicated sources
Records screened (title + abstract)	55	Retained after first-pass relevance filter
Records excluded after title/abstract screening	14	Off-topic, purely technical, or inaccessible full text
Full-text sources included in final corpus	41	Used for framework synthesis and worked use case

Note: aggregate database counts are approximate due to overlap across platforms and manual de-duplication.

A6. Coding Scheme and Reader Assignments

Each source in the final corpus of 41 was assigned to a primary reader responsible for full-text extraction and structured annotation. The coding scheme recorded: identified factor, code/label, thematic category, evidence type (empirical/practical), sectoral context, methodological notes, and gaps.

The full coding table is provided in the ZIP archive.

A7. AI Use Disclosure

In accordance with academic integrity requirements, the following table documents all uses of AI tools in the production of this scoping review. All AI-assisted content was reviewed, verified, and validated by the human authors of this work, who bear full responsibility for its accuracy and conclusions.

AI Tool Used	How It Was Used	Verification Performed	Declaration
Claude (claude-sonnet-4-6)	Query formulation and identification of candidate sources across Google Scholar, ACM DL. Drafting of Appendix sections A1–A6 based on the group's search logs and screened corpus	Sources independently verified and accessed by the group; inclusion decisions made collectively	We explicitly declare that while AI assisted in identifying sources and drafting appendix sections, every source was independently verified. The human team members made all final decisions regarding source inclusion and retain full responsibility for the content.
Google Gemini Pro	For grammatical control and automatization on mechanical and repetitive tasks such as formatting the review or the tables.	After every AI intervention, a human team member checked if informations' accuracy was maintained.	We explicitly declare that only purely mechanical actions were delegated to AI, while every decision-making process has been conducted by a human team member.
Chat GPT Pro	Brainstorming and translating.	Discarded incorrect ideas and advice, and verified the translations using other AI tools.	We explicitly declare that all brainstormed concepts and translated sections were critically evaluated, and the human authors retain full responsibility for the paper's final content.