Delegation Interaction: A Paradigm for Human Oversight of Agentic AI

Thomas Herrmann

Ruhr-University of Bochum, 44780 Bochum, Germany, thomas.herrmann@rub.de

Albrecht Schmidt

LMU Munich 80337 München, Germany, albrecht.schmidt@lmu.de

This paper proposes and discusses the concept of delegation user interfaces. Designing delegation interaction represents a paradigm with the focus on task delegation in the context of human-AI interaction. Delegation interaction is an approach to employ agentic AI combined with pulling the human in the loop and to allow for effective human oversight. Delegation includes six core components. Relevant are: (1) definition and description of the task, (2) embedding into the context of the task, (3) testing to drive task refinement, (4) pulling the human in the loop for oversight and autonomy, (5) continuous quality evaluation, and (6) self-adaptation of AI.

CCS CONCEPTS Human-centered computing, Human computer interaction (HCI), Interaction paradigms

KEYWORDS Human-centered AI, delegation, human-in-the-loop, socio-technical design

1 INTRODUCTION AND BACKGROUND

Collaboration with AI tools is currently changing practices in many domains, including writing, programming, and design. Image generation based on a prompt, vibe coding an app, and creating articles from bullet points are implicitly delegating tasks to an AI. What we want in these contexts from the AI system is generally greatly underspecified. This is like delegation between people and every team leader has experienced that delegating is far from trivial. Working with agentic AI, delegation of tasks becomes the central strategy.

We understand agentic AI as an approach that takes high-level goals as input and the AI systems autonomously decompose them into subtasks, plan multi-step courses of action, and iteratively act in digital or physical environments. This requires orchestrating tools, services, or specialized agents. In contrast to traditional, single-step AI tools, agentic systems have persistent internal state and memory across interactions [1], [2], [3], [4], [5].
They also support dynamic adaptation of plans and can coordinate several agents. Consequently, part of the decision-making and coordination is shifted from humans to the AI system. However, we typically assume that the user is responsible and the AI systems are still operating within explicit governance and oversight frameworks.

To enable this, systems must have the option to get the human back into the loop and into the decision making. The pulling in of humans and human oversight takes place on a meta-level. It is reasonable to align Agentic AI with human centered AI (HCAI) [6] and AI-Governance as frameworks that help to mitigate risks, unfair bias or misuse [7] and in this way supports trust calibration [8], i.e. avoiding over- and under-trust by understanding the potential and limits of AI. Without understanding the limits of AI, people's motivation to contribute to AI governance will fade. A key element here is to keep the human in the loop within the socio-technical workflows in which AI-supported decision making takes place. Humans need to be encouraged and empowered to exercise control and oversight [9]. A prerequisite are interaction modes that allow users to intervene into AI driven procedures [10], veto on AI decisions [11], refine results [12], etc. However, these kinds of interaction modes need to be complemented by modes that are appropriate for agentic AI. Thus, we propose the concept of delegation interaction where AI can actively decide when and how it executes tasks, but humans are still in the mode of being a responsible supervisor as is the case with managers who delegate tasks to employees.

Delegation interaction represents a a paradigm shift that decouples task definition and goal description on the one hand from task execution on the other hand, both in terms of time and granularity of interaction. Such delegation potentially eases collaborative work and reduces the effort for the user. It also enables phases of interaction-free usage [13] where the system works towards the task, while the users are not active. Similar to how management principles like "management by delegation" and "management by objectives" [14] replaced the step-by-step task prescription characteristic of Taylorism [15], we expect a paradigm shift in the use of systems that incorporate generative and agentic AI. Within the delegation paradigm, the system can pull the user back into the loop, or the users can decide by themselves to intervene [10].

2 BASIC CHARACTERISTICS OF DELEGATION INTERACTION

We introduce “Delegation Interaction” as an interaction paradigm that focuses on creating interactions with systems that have a degree of autonomy.

Definition: Delegation Interaction represents a paradigm that enables users to specify what a system should achieve by defining tasks that are delegated to the system. The interface also enables the user to articulate goals of this delegation and why they are relevant. It additionally provides means for monitoring progress and results and offers mechanisms for meaningful interventions.

With delegation interaction we move beyond the how of the task execution as implemented when following a direct manipulation interaction paradigm (e.g. GUIs) or in dialog based interfaces. This interaction paradigm allows the user to describe the what and why. It gives opportunities to interactively refine and narrow down the delegators’ requirements and the constraints to be considered.

Key Human Roles and types of systems in the context of delegation interaction

Fig. 1 describes the context of delegation interaction. With the delegation interaction in the center of the system, domain experts can

a. provide and refine the description of the task by goals and the desired outcomes, and
b. supervise the task execution and possibly intervene into it.

These domain experts are the members of a delegator team who conduct the delegation interaction collaboratively. The result of the delegation interaction is a delegatee system that is not directly developed by the delegator team but by an AI-based system called delegation system that can be represented by agentic AI. This system eventually develops (and later on controls) the delegatee system based on the support of the delegation interaction.

The delegatee system comprises the AI-based functionality to execute the delegated tasks. During execution, usually interaction takes place with the clients of the company which delegates the task and employs the delegator team. This interaction with the clients is provided by the delegatee system to clarify the facts needed to complete the task etc. Under certain conditions —depending on the decision of the delegatee system or the delegator team— the clients’ interaction with the delegatee system can be temporally replaced by direct collaboration between the client and the delegator team. This collaboration is supported by the delegatee system, e.g. by providing the appropriate communication channel. To ensure oversight, Fig. 1 suggests considering agentic AI as a delegation system. It is prepared by delegators who have the domain-specific knowledge necessary for proper task delegation and monitoring of task execution to ensure quality.

To develop the delegatee system and enable delegation interaction, the delegation system is needed. It is developed by a developer team. The development of the delegation system can be initiated by the company that employs the delegator team, or by a software development house that wants to cover the needs of a certain domain. After its development, the delegator team can employ the delegation system to conduct the delegation interaction. The effects of the delegation

interaction can be partially controlled by the developer team. The extent of this control depends on the decisions of the delegators. They collaborate with the developer team to complete the delegation. This collaboration comprises the customization of the delegation system that needs to meet the conditions of the application domain. The collaboration between these teams is supported by the delegation system.

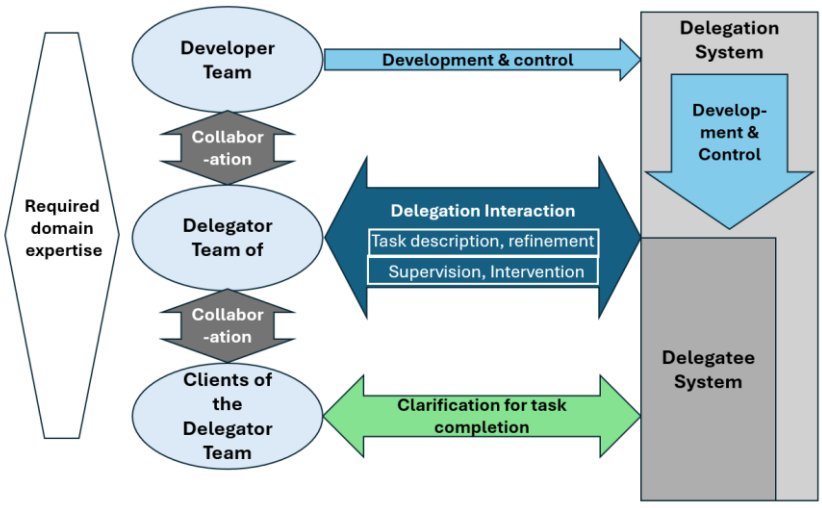

Figure 1: Roles and delegation interaction

The control of the delegatee system during its task execution is conducted by the delegator team —mainly by supervision and intervention— through the delegation system. As a design rationale, it is important that the control features are provided by the delegation system —and not by the delegatee syste—, since the features of the delegation system are developed by human decision making while the delegatee system is widely developed by agentic AI itself Thus, the delegation system is the central instance to support human oversight and autonomy during the whole process of delegating and running tasks.

The concept of delegation and the differentiation between the roles delegator and delegatee are taken from the literature on management strategies [14], [16], [17]. This literature emphasizes the idea that by providing objectives, requirements and criteria for task execution, the manager relies on the delegatees' capabilities instead of describing all the steps of task execution. This management strategy also includes the transfer of responsibility to the delegatee – here we suggest that the possibilities for transfer are limited, as the ultimate responsibility must remain with the human side.

3. CORE COMPONENTS FOR INTERACTING WITH THE DELEGATION SYSTEM

To describe the possibilities of delegation interaction and the features of the delegation system in more detail, we propose six components. Three of them mainly support the task delegation —i.e. the development of the delegatee system— before it is used. These components deal with task definition and description (1), contextual embedding (2), and testing and refinement (3). The other components are relevant for oversight and maintaining autonomy during the execution of the delegated tasks: pulling the human in the loop (4), continuous quality evaluation (5), and AI-based self-adaptation (6). The

component for quality evaluation also supports the testing of the delegatee system and triggers iterative cycles between re-design and usage of the delegatee system.

Human-Centered AI (HCAI) and AI-governance provide a series of principles to be taken into account when specifying the six components: These are mainly transparency and explainability of AI (XAI), human autonomy, safety and reliability, accountability, benefits and well-being, fairness, privacy, and variability. We expect that by following a delegation interaction for approach agentic AI, these requirements are easier to achieve than with traditional interaction paradigms.

1) Support for Task Definition and Task Description With GenAI, natural language dialogues and all kind of material can be used to specify goals [18], requirements [19] and constraints that describe tasks to be executed. Agentic AI helps to decompose high-level goals into appropriate subtasks [20]. The delegation system supports the delegators to detect and eliminate ambiguities of requirements, to specify criteria for measuring the progress of task execution, and to define guardrails for the delegatee system.

2) Embedding into the Context of the Task Domain The delegation system helps to specify information to be derived from the tasks' context. This can be incoming emails, updates of websites, webcams, sensor data, or more general factors such as holiday seasons, etc. Eventually, the delegatee system needs to be able to identify triggers that initiate specific behaviors as it is the case with implicit interaction [21]. The delegation system provides interfaces to connect the delegatee system with other applications. Agentic AI contributes to identifying the tools to be used [2].

3) Support for Testing to Drive Task Refinement A pilot testing is recommended for agentic AI solutions [4].With testing, the delegators can check how far a task description is complete and whether task execution provides intended results. The delegation system can provide a plan [22] of the process of task handling. Explainability [23] helps to understand how the plan is derived. Combined with testing, the plan helps to refine the way of task execution step-by-step. The plan also mirrors —as a socio-technical self-description— the collaboration between the involved roles (see Fig. 1).

4) Mechanisms for Pulling the Human in the Loop The delegators are not software but domain experts. Thus, the task delegation will not be perfect since some requirements or situational exceptions will not be anticipated from the very beginning. Therefore, humans need to stay in the loop to exercise oversight [24]. Research on Agentic AI discusses various approaches of maintaining AI governance and oversight [25].This includes supervision of task processing (based on awareness mechanisms) and —if necessary— intervention in task processing and thus its completion. Delegators can specify at which points they should be informed and triggered to check how the delegatee system performs. Agentic AI should have internal criteria to recognize situations when humans' attention should be involved. For the purpose of influencing task execution, the delegator team (see Fig. 1) should be able to intervene. Interventions have the following characteristics [10], [26]: They occur only exceptionally are not planned in advance; can be quickly initiated and their effect can be easily revised. They allow a time-limited phase of fine-grained interaction with or adaptation of the system.

5) Support for Continuous Quality Evaluation of Task Handling The delegatee system will be continuously evaluated by measuring the progress of task handling —e.g., the length of dialogues, number of requests for consultation, etc. Combining the plan of task execution with continuous assessment can support people's reflection on the system's quality. From a socio-technical perspective, organizational practices will have to be established that make this kind of reflection a standard routine. Agentic AI provides reasoning capabilities for self-reflection [20], together with humans.

6) AI-based Self-adaptation of AI-driven Task Handling AI can derive measures from the evaluation for customizing the system. Self-adaptation and learning of the involved AI-system is based on pre-specified criteria in relation to the measuring during evaluation, can take place. The delegatee system —in connection with the underlying delegation system– – must be able to learn from its faults during task handling or from missing the specified goals, or from violating constraints. Agentic AI has the history of its task execution available and pursues a continuous adaptation [25].

4. CONCLUSIVE REMARKS

We have defined delegation interaction as a new paradigm that supports users to have agentic AI execute their tasks by describing the goals, requirements, criteria and constraints that characterize the tasks. The users of delegation interaction act as delegators who use a delegation system to make a delegatee system run the tasks. This paradigm shift will be accompanied by future advancements such as:

- The takeover of tasks by IT is no longer a question of human programming – which could have been supported by GenAI – but programming is simply omitted.
- Input for task delegation can mainly be based on natural language and accompanying material, such as records of concrete task handling, other examples of similar tasks that are already delegated etc.

The advantages seem to be apparent: routine tasks can easily and efficiently be delegated. Communication between domain experts and software developers is critical in the specification of traditional software systems. However, with delegation interaction we expect that this gap between domain experts and software engineers and the communication issues between these two sides will become increasingly irrelevant, in particular with the advancement of agentic AI.

However, the risks are also obvious: all the problems discussed about lacking reliability, biases, intransparency etc. are highly relevant if expertise for designing IT-support is not systematically included. Thus, it is important to understand that the delegation of a certain task will not achieve sufficient quality through a single act of delegation interaction. Just telling agentic AI what one wants to have and then letting it go is a risky simplification —a kind of delegation in the wild—, since there is no checking in advance nor during task execution whether the task description is appropriate. Thus, we suggest employing agentic AI through systematically designed delegation systems to provide the users (consequently the delegators) with guard rails and enable them to exercise oversight and autonomy through supervision and intervention.

Agentic AI can be designed to refine the plan —together with a human supervisor— of how it proceeds during task execution. An ongoing socio-technical process of continuous refinement, quality evaluation and customization of the delegatee system is needed. The delegators, developers of the delegation system and clients, and their collaboration need be supported by the delegation system. A basis for evaluation and refinement is to provide a plan of how the task is executed. This plan is fed back to the involved stakeholders and can be considered as a self-description for the socio-technical process of task handling and collaboration. Thus, the relevance of the six core components described in the previous section should be considered in every case of implementing agentic AI.

To understand the risks of delegation interaction, it is important to assess the level of knowledge people who delegate have. If people who delegate a task know exactly how to perform it by themselves and what output they expect and what potential pitfalls they need to consider, delegation is much safer and easier than for people who have no clue about how to carry out the task themselves. With delegation interaction we believe that domain experts can be empowered to hand over tasks to IT-systems without them having to be able to program. They have domain expertise and know exactly what an appropriate task completion looks like and how task fulfillment can be monitored – e.g. by using indicators to measure whether progress is sufficient. Delegating tasks by domain experts, who know how these tasks are executed by humans, is also an act of documenting the current practice of running tasks and demonstrating typical ways of their execution. This is especially possible if the delegation system feeds back the plan of how it will run the task. Applying agentic AI-driven task delegation for many cases implies that we will have a corpus of examples for the future demonstrating the practice of task execution where human domain experts have been involved.

The main efforts to mitigate the risks of delegation interaction and bring it into compliance with the ethical guidelines and principles of HCAI are to further research how intervention user interfaces effectively support the exercising of supervision and how organizational practices ensure that oversight is actually conducted.

REFERENCES

[1] D. B. Acharya, K. Kuppan, and B. Divya, "Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey," *IEEE Access*, vol. 13, pp. 18912–18936, 2025, doi: 10.1109/ACCESS.2025.3532853.

[2] X. Li *et al.*, "DeepAgent: A General Reasoning Agent with Scalable Toolsets," Feb. 05, 2026, *arXiv*: arXiv:2510.21618. doi: 10.48550/arXiv.2510.21618.

[3] R. Sapkota, K. I. Roumeliotis, and M. Karkee, "AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges," *Information Fusion*, vol. 126, p. 103599, Feb. 2026, doi: 10.1016/j.inffus.2025.103599.

[4] S. Sivakumar, "Agentic AI in Predictive AIOps: Enhancing IT Autonomy and Performance," *int.jour.sci.res.mana*, vol. 12, no. 11, pp. 1631–1638, Nov. 2024, doi: 10.18535/ijsrm/v12i11.ec01.

[5] J. Wu, J. Zhu, Y. Liu, M. Xu, and Y. Jin, "Agentic Reasoning: A Streamlined Framework for Enhancing LLM Reasoning with Agentic Tools," Jul. 14, 2025, *arXiv*: arXiv:2502.04644. doi: 10.48550/arXiv.2502.04644.

[6] B. Shneiderman, *Human-Centered AI*. Oxford University Press, 2022.

[7] O. O. Garibay *et al.*, "Six Human-Centered Artificial Intelligence Grand Challenges," *International Journal of Human–Computer Interaction*, vol. 39, no. 3, pp. 391–437, Feb. 2023, doi: 10.1080/10447318.2022.2153320.

[8] K. Okamura and S. Yamada, "Adaptive trust calibration for human-AI collaboration," *PLoS ONE*, vol. 15, no. 2, p. e0229132, Feb. 2020, doi: 10.1371/journal.pone.0229132.

[9] European Commission, C. and T. Directorate General for Communications Networks, and High-Level Expert Group on Artificial Intelligence, *Ethics guidelines for trustworthy AI*. 2019. Accessed: May 23, 2021. [Online]. Available: https://data.europa.eu/doi/10.2759/346720

[10] A. Schmidt and T. Herrmann, "Intervention user interfaces: a new interaction paradigm for automated systems," *interactions*, vol. 24, no. 5, pp. 40–45, 2017.

[11] B. Rakova, J. Yang, H. Cramer, and R. Chowdhury, "Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–23, 2021.

[12] C. J. Cai *et al.*, "Human-centered tools for coping with imperfect algorithms during medical decision-making," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–14.

[13] T. Herrmann, "Evolution of interaction-free usage in the wake of AI," *i-com*, vol. 23, no. 2, pp. 179–187, Jul. 2024, doi: 10.1515/icom-2024-0005.

[14] D. Piplica, I. Peronja, and T. Luković, "Controlling in the Function of Management by Objectives," in *DIEM: Dubrovnik International Economic Meeting*, Sveučilište u Dubrovniku, 2024, pp. 131–142.

[15] C. Wright, "Taylorism Reconsidered: The Impact of Scientific Management within the Australian Workplace," *Labour History*, no. 64, p. 34, 1993, doi: 10.2307/27509164.

[16] B. Mathebula and B. Barnard, "The factors of delegation success: accountability, compliance and work quality," *Expert Journal of Business and Management*, vol. 8, no. 1, pp. 76–97, 2020.

[17] J. Ugoani, "Effective Delegation and Its Impact on Employee Performance," *International Journal of Economics and Business Administration*, vol. 6, no. 3, pp. 78–87, 2020.

[18] J. A. Maciariello, "Management by Objectives and Self-control," in *The Palgrave Encyclopedia of Strategic Management*, M. Augier and D. J. Teece, Eds., London: Palgrave Macmillan UK, 2016, pp. 1–4. doi: 10.1057/978-1-349-94848-2_165-1.

[19] C. Rupp, "Requirements and psychology," *IEEE Software*, vol. 19, no. 3, pp. 16–18, 2002.

[20] U. Nisa, M. Shirazi, M. A. Saip, and M. S. M. Pozi, "Agentic AI: The age of reasoning—A review," *Journal of Automation and Intelligence*, p. S2949855425000516, Aug. 2025, doi: 10.1016/j.jai.2025.08.003.

[21] A. Schmidt, "Implicit human computer interaction through context," *Personal technologies*, vol. 4, no. 2–3, pp. 191–199, 2000.

[22] C. A. Miller and R. Parasuraman, "Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control," *Hum Factors*, vol. 49, no. 1, pp. 57–75, Feb. 2007, doi: 10.1518/001872007779598037.

[23] U. Ehsan and M. O. Riedl, "Human-Centered Explainable AI: Towards a Reflective Sociotechnical Approach," in *HCI International 2020 - Late Breaking Papers: Multimodality and Intelligence*, vol. 12424, C. Stephanidis, M. Kurosu, H. Degen, and L. Reinerman-Jones, Eds., in Lecture Notes in Computer Science, vol. 12424. , Cham: Springer International Publishing, 2020, pp. 449–466. doi: 10.1007/978-3-030-60117-1_33.

[24] I. Georgieva, C. Lazo, T. Timan, and A. F. Van Veenstra, "From AI ethics principles to data science practice: a reflection and a gap analysis based on recent frameworks and practical experience," *AI Ethics*, vol. 2, no. 4, pp. 697–711, Nov. 2022, doi: 10.1007/s43681-021-00127-3.

[25] M. A. Ali and F. Dornaika, "Agentic AI: A Comprehensive Survey of Architectures, Applications, and Future Directions," *Artif Intell Rev*, vol. 59, no. 1, p. 11, Nov. 2025, doi: 10.1007/s10462-025-11422-4.

[26] T. Herrmann, "Intervenability as a Design Requirement for Autonomy and Oversight within Human-Centered AI," in *The Design of Human-Centered Artificial Intelligence for the Workplace*, C. K. Coursaris, J. Beringer, P.-M. Léger, and B. Öz, Eds., in Progress in IS. , Cham: Springer Nature Switzerland, 2025, pp. 143–166. doi: 10.1007/978-3-031-83512-4_9.