Legitimate Delegation and the Preservation of Human Agency in Trustworthy Agentic AI

Alexandru Mateescu[1,*]

[1]Université Paris 1 Panthéon-Sorbonne, Paris, France

Abstract

Recent advances in AI make it increasingly useful to distinguish between assistive systems, which support human decision-making while leaving action under direct human control, and agentic systems, which may initiate, coordinate, or execute activities on behalf of users across extended chains of activity. As actions progressively shift from direct human execution toward delegated agentic workflows, questions arise concerning the extent to which delegation remains compatible with meaningful human agency.

Existing trustworthy AI and AI risk management frameworks primarily evaluate system properties such as reliability, transparency, oversight, and compliance with ethical and legal requirements. However, as agentic systems increasingly act on behalf of users, trustworthy AI must also address the legitimacy of the delegation relations established between users and AI systems.

Drawing on research on automation, organizational delegation, trust calibration, and intelligent AI delegation, we argue that legitimacy of delegation should be treated as a distinct criterion for assessing trustworthiness in agentic AI. We propose five conditions for legitimate delegation: authorization, bounded scope, intelligible oversight, contestability, and preservation of practical human capacities.

To facilitate practical evaluation, we operationalize these conditions through assessment questions, indicators, and representative failure modes. We then illustrate the framework through examples of emerging background delegation in contemporary agentic systems. Rather than evaluating such systems solely through efficiency or performance outcomes, we argue that they should also be assessed according to whether they preserve meaningful conditions for supervision, revision, contestability, and sustained practical engagement.

The paper contributes to trustworthy and human-centered AI by providing a framework for evaluating when delegated activity remains compatible with meaningful human agency.

Keywords

agentic AI, trustworthy AI, delegation, human agency, automation, AI governance

1. Introduction

Recent advances in AI make it increasingly useful to distinguish between assistive systems and agentic systems. Assistive systems primarily support human decision-making while leaving the execution of actions under direct human control. By contrast, agentic systems are increasingly capable of acting on behalf of users across extended chains of activity involving planning, coordination, monitoring, tool use, and adaptive execution. In this respect, agentic AI increasingly adopts a logic of delegation traditionally associated with situations in which one actor is authorized to act on behalf of another. As activities are progressively carried out through partially autonomous delegation chains, errors, capability mismatches, and failures of oversight may propagate across multiple stages of execution. Trustworthy AI can therefore no longer be assessed solely at the level of system outputs, but must also account for how delegated activity is organized, supervised, and governed across agentic workflows.

Delegation is not originally a technical concept, but an organizational one traditionally associated with transfers of authority performed under continuing conditions of supervision, accountability, and informational asymmetry. Classical theories of delegation describe situations in which principals transfer discretionary authority to agents in order to benefit from specialization, expertise, or informational advantages unavailable to the principal directly. At the same time, delegation introduces a persistent tension between the advantages of specialized execution and the risks associated with reduced direct

TRUST-AI 2026: The Workshop on Trustworthy Artificial Intelligence, co-located with ECAI 2026

alexandru.mateescu@etu.univ-paris1.fr (A. Mateescu)

0009-0007-7084-8284 (A. Mateescu)

control. Modern theories of delegation repeatedly characterize this relation as a trade-off between expertise and control, especially under conditions of asymmetric information. Agentic AI increasingly extends these organizational structures into computational environments in which systems may initiate, coordinate, or execute activities on behalf of users across partially autonomous chains of action.

This transformation raises questions that differ from those traditionally associated with assistive AI systems. In assistive systems, users generally remain directly involved in the execution and revision of the relevant activities themselves. By contrast, agentic systems may progressively reduce the practical involvement through which capacities for supervision, revision, and judgment are exercised and maintained. The issue is therefore not only whether agentic systems perform effectively, but whether the forms of delegation they enable remain compatible with the practical capacities on which human agency depends over time.

Existing trustworthy AI frameworks [1] have made major contributions to the evaluation of AI systems by emphasizing reliability, transparency, human oversight, accountability, robustness, and adherence to ethical and legal requirements. However, these approaches primarily evaluate the properties of systems and decision procedures rather than the legitimacy of the delegation relations established between users and agentic systems. As agentic systems increasingly act on behalf of users, trustworthy AI must also address the conditions under which delegation remains compatible with meaningful human agency. This requires analyzing not only whether systems function correctly, but also how authority, supervision, contestability, and human competence are redistributed across human–AI relations.

Drawing on research on automation, organizational delegation, trust calibration, and intelligent AI delegation, this paper argues that legitimacy of delegation should be treated as a distinct criterion for assessing trustworthiness in systems acting on behalf of users. We propose five conditions for legitimate delegation in agentic AI systems: authorization, bounded scope, intelligible oversight, contestability, and non-erosion of users' practical capacities for supervision, revision, and judgment. We illustrate this framework through examples of emerging background delegation in agentic AI systems, including procedural and organizational environments in which systems increasingly coordinate, monitor, negotiate, or execute activities on behalf of users under conditions of intermittent supervision. Rather than evaluating such systems solely through efficiency or performance outcomes, we argue that they should also be assessed according to whether they preserve meaningful conditions for supervision, revision, contestability, and sustained practical engagement. The paper therefore contributes to human-centred trustworthy AI [2] by proposing legitimacy conditions for delegation in agentic systems. The paper first reconceptualizes agentic AI in terms of organizational delegation before developing legitimacy conditions for preserving human agency under delegated activity. It then examines emerging forms of background delegation in agentic AI systems in order to clarify how these legitimacy conditions become operationally significant in practice.

2. Delegation as an Organizational Relation

Delegation is not originally a technical concept, but an organizational one. It describes situations in which one actor authorizes another to act on their behalf while remaining concerned by the consequences of the delegated activity. Classical theories of organization analyze delegation as a response to limits in attention, expertise, and access to information. Principals delegate because they cannot directly manage every dimension of an activity themselves. Delegation therefore allows access to specialized competence while simultaneously creating new problems of supervision and control. In organizational settings, delegation does not suppress responsibility but redistributes it. The delegated actor receives discretionary authority within a bounded scope while remaining subject to supervision and possible revision. Delegation consequently introduces a persistent tension: the principal benefits from the competence or informational advantage of the agent while partially losing direct control over the delegated activity. Modern theories of delegation repeatedly describe this relation as a trade-off between expertise and control under conditions of informational asymmetry [3]. Delegation is therefore not simply a transfer of execution, but a structured relation organizing authority and supervision between

actors. This organizational perspective becomes increasingly relevant for agentic AI systems. Assistive systems primarily support deliberation while leaving action under direct human control. By contrast, agentic systems may increasingly act in users' name across partially autonomous chains of activity. In such contexts, AI systems no longer merely recommend or inform. They participate directly in the realization of delegated goals. Delegation thus becomes a constitutive feature of the human–AI relation rather than a secondary consequence of automation. In many emerging agentic systems, delegation increasingly operates in the background of practical activity. Users no longer interact with systems only through isolated commands or recommendations, but through persistent delegated processes operating continuously with limited direct attention. This form of "background delegation" becomes central to understanding how agentic systems may progressively transform the practical conditions under which human agency is exercised.

This transformation is especially significant because delegation in agentic AI is often continuous rather than isolated. Agentic systems may operate across adaptive workflows involving financial management, scheduling, contractual monitoring, procedural coordination, or interaction with multiple digital services. As these delegation chains become more extended, users may progressively rely on systems not only for assistance but for the operational management of parts of everyday activity. Recent systems such as Microsoft Copilot "Actions" already allow users to delegate concrete activities to AI systems operating in the background. Users may ask the system to book restaurant reservations, organize travel, purchase tickets, or manage recurring online procedures through natural-language instructions [4]. Microsoft explicitly presents these systems as capable of "taking action on your behalf" and completing tasks behind the scenes while users focus on other activities [5]. In such contexts, the central issue is no longer only whether systems function correctly, but whether the delegation relations they establish preserve the practical conditions required for sustained human agency. The issue raised by this transformation is not limited to efficiency or operational reliability. Delegation also modifies the conditions under which human agency is exercised. In assistive systems, users remain practically involved in the relevant activity even when supported by AI. By contrast, delegation within agentic systems may progressively reduce the involvement through which capacities for supervision and judgment are maintained. The central problem is therefore whether these delegation structures remain compatible with meaningful human agency over time. This issue cannot be reduced to familiar concerns about automation replacing human labor. What distinguishes agentic delegation is that systems increasingly operate under conditions in which they are authorized to act on behalf of users in environments characterized by informational asymmetry and limited direct supervision. The legitimacy of delegation therefore becomes itself a central normative issue. Trustworthy AI must consequently evaluate not only whether agentic systems function correctly, but also whether the delegation relations they establish preserve the practical conditions required for sustained human agency. Recognizing agentic AI as a form of organizational delegation transforms the problem addressed by trustworthy AI. The central issue is no longer limited to whether systems operate efficiently, reliably, or transparently, but whether the delegation relations they establish remain compatible with human agency over time. Once delegation is understood as a structured redistribution of authority and practical involvement, the question of legitimacy becomes unavoidable. The following section therefore proposes a framework for evaluating when delegation in agentic AI systems can be considered legitimate.

3. Legitimate Delegation in Agentic AI

The following conditions are not intended as an exhaustive set of criteria, but as a synthesis of normative concerns recurring across the literatures on delegation, automation, trust calibration, and human oversight. Once agentic AI systems act on behalf of users, trustworthy AI must evaluate not only system performance, but also the legitimacy of the delegation relations established between users and agentic systems. Otherwise, systems may remain technically reliable while progressively weakening the practical conditions under which meaningful human agency can be exercised. Delegation remains legitimate, that is, compatible with human agency, only if users preserve meaningful conditions for

supervision, revision, and judgment under delegated activity. This section proposes five conditions for legitimate delegation in agentic AI systems: authorization, bounded scope, intelligible oversight, contestability, and preservation of practical human capacities.

3.1. Authorization

Delegation presupposes authorization because acting on behalf of another requires a transfer of discretionary authority grounded in the intentions of the delegating actor. In agentic AI systems, delegated actions may involve financial transactions, procedural interactions, contractual commitments, or access to sensitive information. Delegation therefore cannot legitimately emerge from passive system use or broad acceptance of generic platform conditions. Users must be able to determine which activities are delegated and under what conditions systems are authorized to act. Authorization must also remain revisable. Delegation in organizational contexts does not permanently eliminate the authority of the principal. Similarly, users must remain able to suspend, modify, or revoke delegated authority when circumstances change. Without such reversibility, delegation risks becoming detached from meaningful user intention.

3.2. Bounded Scope

Delegation also requires bounded scope because discretionary authority becomes difficult to supervise once delegated systems progressively extend their operational domain beyond the conditions initially understood or intended by users. This risk becomes especially significant in adaptive systems capable of modifying goals, interacting with external services, or coordinating additional agents dynamically. Legitimate delegation therefore requires clear limits concerning the range of actions systems may perform autonomously. These limits may concern financial thresholds, categories of decisions, procedural domains, or contexts requiring renewed authorization. The purpose of bounded scope is not to eliminate autonomy entirely, but to preserve proportionality between delegated authority and user understanding.

3.3. Intelligible Oversight

Delegation requires intelligible oversight because users cannot meaningfully supervise activities performed on their behalf if the operational logic of delegated actions remains inaccessible. Existing trustworthy AI frameworks often emphasize transparency, but delegation introduces a more specific requirement: users must be capable of reconstructing why relevant actions were performed and according to which delegated objectives. This does not require users to understand the internal technical architecture of AI systems. Users must instead remain capable of identifying what actions were taken, under which authority, and according to which operational logic. Oversight becomes unintelligible when delegation chains become too complex for users to follow or revise practically.

3.4. Contestability

Delegation further requires contestability because delegated systems may produce actions that remain operationally coherent while nevertheless conflicting with users' intentions, values, or contextual judgments. This condition becomes increasingly important in environments characterized by informational asymmetry or rapid automated execution. Contestability is essential because delegation never eliminates the possibility of error, conflict, or contextual misinterpretation. Agentic systems may act consistently according to delegated objectives while nevertheless producing outcomes users consider inappropriate once broader circumstances are taken into account. Delegation therefore remains legitimate only if users preserve effective possibilities for interruption, revision, and refusal.

3.5. Preservation of Practical Human Capacities

Finally, legitimate delegation requires preserving practical human capacities because repeated delegation may progressively reduce the involvement through which supervision and judgment are exercised and maintained. This condition distinguishes agentic delegation from more familiar discussions of automation focused primarily on replacement of labor or operational risk. As users increasingly rely on systems operating continuously in the background, practical engagement in relevant activities may progressively diminish. Over time, this may weaken capacities for supervision, procedural understanding, or independent judgment. The problem is not that users delegate isolated actions, but that repeated delegation may gradually transform the practical conditions under which meaningful human agency is exercised. This does not imply that delegation is inherently illegitimate. Delegation is often necessary in environments characterized by complexity, scale, or informational overload. However, trustworthy delegation requires preserving opportunities for continued human understanding and practical engagement rather than progressively eliminating them entirely. Taken together, these five conditions shift trustworthy AI from a framework focused primarily on system properties toward a framework also concerned with the legitimacy of delegation relations. Agentic AI systems should therefore be evaluated not only according to whether they function correctly, but also according to whether they preserve the practical conditions required for sustained human agency under delegated activity.

The five conditions proposed above are normative concepts intended to characterize legitimate delegation in agentic AI systems. However, normative concepts alone do not directly provide guidance for system design, governance, auditing, or evaluation. Before applying the framework to concrete cases, it is therefore useful to consider how these conditions may be translated into practical assessment criteria capable of guiding judgments about delegation relations in agentic AI systems.

3.6. Relation to Existing Trustworthy AI Frameworks

The legitimacy conditions proposed here are not intended to replace existing trustworthy AI (TAI) frameworks but to complement them. Contemporary TAI frameworks have made major contributions by emphasizing system properties such as reliability, robustness, transparency, accountability, human oversight, privacy, and compliance with ethical and legal requirements. These properties remain essential for evaluating the technical and organizational trustworthiness of AI systems.

The present framework adopts a complementary perspective by examining the legitimacy of the delegation relations established between users and agentic AI systems. From this perspective, human oversight is understood not simply as a mechanism for supervising system behaviour, but as one condition within a broader normative structure governing legitimate delegation. Authorization specifies who may delegate authority and under what conditions; bounded scope limits the extent of delegated authority; intelligible oversight enables users to reconstruct delegated actions; contestability preserves effective opportunities for interruption and revision; and preservation of practical human capacities addresses the long-term effects of repeated delegation on users' ability to supervise, understand, and exercise judgment.

Existing TAI frameworks therefore remain necessary as foundations for evaluating agentic AI systems. However, the present paper argues that, once AI systems increasingly act on behalf of users, evaluating system properties alone does not fully capture the normative issues raised by delegation. The proposed framework therefore complements existing TAI approaches by extending the object of evaluation from AI systems alone to the legitimacy of the delegation relations established between users and agentic AI systems, and whether these relations remain compatible with meaningful human agency.

This comparison also clarifies the role of human oversight within the proposed framework. Existing TAI frameworks typically treat oversight as an important design or governance mechanism for supervising AI behaviour. The present framework retains this concern but situates it within the broader structure of legitimate delegation. Oversight alone cannot establish the legitimacy of delegated action if users were not able to authorize the delegation, define its scope, contest its outcomes, or preserve the

practical capacities required for meaningful supervision over time.

4. Operationalizing Legitimate Delegation for Evaluation and Governance

Following classical approaches to operationalization in the social sciences, abstract concepts can be translated into assessment criteria and observable indicators before they inform practical evaluation. Operationalization therefore provides a bridge between conceptual analysis and empirical assessment by connecting theoretical constructs to observable features of concrete situations [6].

The purpose of operationalization here is not to reduce legitimacy to a simple metric. Rather, it is to identify assessment questions and possible indicators that may help determine whether a given delegation relation satisfies the conditions proposed above. Different application domains may require different indicators, but the underlying evaluative logic remains the same. In the context of agentic AI, operationalization provides a way for designers, governance teams, auditors, and regulators to assess whether delegated activity remains compatible with human agency.

Legitimacy condition	Assessment question	Possible indicators	Representative failure mode
Authorization	Can users determine which activities are delegated and modify this authorization?	Granular permissions, explicit delegation settings, revocation mechanisms	Delegated authority emerges from default settings or broad consent without meaningful user control
Bounded scope	Are clear limits placed on delegated authority?	Spending limits, domain restrictions, task boundaries, escalation procedures	Delegated activity expands beyond the scope initially understood or authorized by the user
Intelligible oversight	Can users reconstruct what actions were performed and why?	Activity logs, explanations, audit trails, action summaries	Delegation chains become opaque and difficult to understand or supervise
Contestability	Can users interrupt, revise, challenge, or reverse delegated actions?	Override mechanisms, rollback procedures, appeal functions, human review options	Delegated actions become difficult or impossible to contest once executed
Preservation of practical human capacities	Does delegation preserve opportunities for continued supervision, understanding, and judgment?	Periodic review requirements, active confirmations, user engagement mechanisms	Persistent background delegation progressively reduces practical involvement and supervisory engagement

Table 1
Illustrative operationalization of legitimacy conditions for agentic AI delegation.

Table 1 illustrates how the five legitimacy conditions may be translated into assessment questions, possible indicators, and representative failure modes. The objective is not to reduce delegation legitimacy to a simple metric, but to provide a practical framework capable of informing design review, governance assessment, trustworthiness auditing, and regulatory evaluation.

5. Background Delegation in Contemporary Agentic Systems

The structure underlying contemporary agentic AI systems predates recent LLM-based architectures. Earlier systems already implemented forms of delegated practical agency in which users transferred

portions of practical activity to computational systems operating on their behalf. A particularly revealing example is Calendar.help, a scheduling system allowing users to delegate meeting coordination to an assistant operating through email interactions. Users interacted with the system "as if it were a human personal assistant," while the system independently handled scheduling negotiations, follow-up messages, coordination constraints, and meeting organization across partially autonomous workflows [7]. The significance of such systems lies less in their technical sophistication than in the delegation structure they instantiate: practical activities previously performed directly by users become executed through persistent delegated processes operating with limited direct involvement. Importantly, however, such early delegated systems did not yet fundamentally challenge sustained human agency. Although users delegated portions of practical activity, they generally remained closely connected to the delegated process itself through frequent interaction, direct visibility over execution, relatively limited delegation scope, and continuous possibilities for intervention and correction. Delegation reduced procedural burden without fully withdrawing the relevant activity from the user's field of practical engagement. Contemporary agentic systems increasingly extend delegation beyond this model by allowing delegated activity to continue under conditions of reduced visibility and intermittent supervision, thereby progressively shifting practical activity into the background of direct human attention.

Contemporary agentic AI systems increasingly generalize and intensify this delegation structure across broader domains of activity. Recent systems such as OpenAI's Operator [8], Google DeepMind's Gemini 2.0 presented as "an AI model for the agentic era" [9], Microsoft Copilot Actions [10], and Anthropic's computer-use agents [11] increasingly illustrate this transition toward delegated background activity. Rather than merely providing isolated recommendations, such systems are designed to pursue delegated objectives across sequences of coordinated actions involving navigation, monitoring, workflow execution, transaction management, tool use, and interaction with digital environments on behalf of users. Recent OECD analyses [12] similarly characterize agentic AI systems through their capacity to pursue objectives, coordinate actions, use tools, adapt to changing environments, and operate under conditions of limited direct human intervention. In such systems, delegated activity progressively shifts from visible assistance toward forms of continuous background delegation in which execution increasingly occurs outside the immediate field of practical attention. The central transformation therefore concerns not only increasing technical autonomy, but a progressive reduction of continuous human involvement within the practical activity itself. Users may consequently remain formally responsible for delegated outcomes while becoming less continuously engaged in the processes through which those outcomes are produced.

Microsoft Copilot Actions illustrates how the proposed framework may be applied in practice. Delegation is explicitly initiated by the user, thereby illustrating the requirement of authorization. The delegated activity is also bounded by the specific tasks and services for which the system has been authorized to act, although the appropriate limits of such delegation may depend on the application domain. Intelligible oversight requires that users remain able to reconstruct which actions were performed, under what authority, and with what outcomes through activity summaries or audit trails. Contestability further requires that delegated actions can be interrupted, revised, or revoked whenever users judge them to be inappropriate. Finally, the long-term effects of repeated reliance on such systems remain an open question. As increasingly routine activities are delegated to background processes, it becomes important to examine whether opportunities for supervision, understanding, and practical judgment are preserved over time. This final condition illustrates how the proposed framework extends the evaluation of agentic AI beyond immediate system performance to the longer-term preservation of meaningful human agency. More generally, this transformation nevertheless raises emerging problems concerning practical visibility and supervisory engagement. As delegated activity increasingly operates in the background of direct user attention, users may become less continuously involved in the processes through which delegated outcomes are produced. Recent discussions surrounding agentic AI governance and human oversight increasingly identify difficulties associated with supervising extended delegated workflows operating under conditions of intermittent human intervention [12]. It emerges more immediately from the possibility that users remain formally responsible for delegated outcomes while becoming less directly engaged in the practical activity itself.

Even highly efficient systems may consequently create conditions under which intelligible supervision and meaningful revision become progressively more difficult to sustain. Such concerns are not entirely new. Earlier Human-in-the-Loop and automation literature already identified difficulties associated with intermittent supervision, automation bias, and "out-of-the-loop" performance in partially automated systems [13, 14]. Contemporary agentic AI systems, however, increasingly extend these tensions into broader forms of persistent background delegation operating across open-ended digital environments [12].

These immediate difficulties may become amplified over time through repeated reliance on delegated systems. Existing empirical research on AI reliance and cognitive offloading suggests that sustained dependence on AI systems may reduce critical engagement with tasks increasingly performed through automation and delegation. Recent quantitative work by Gerlich [15], based on a survey of 666 participants, identified significant associations between frequent AI-tool usage, cognitive offloading practices, and reduced critical thinking performance. Although such research does not directly study agentic delegation itself, it nonetheless supports the possibility that persistent background delegation may contribute to reduced critical engagement with capacities exercised less regularly in practice.

The examples above therefore illustrate why trustworthy AI can no longer be evaluated solely through criteria such as reliability, efficiency, or technical performance once systems increasingly act on behalf of users. The central issue becomes whether forms of background delegation remain compatible with sustained conditions for practical supervision and judgment under conditions of reduced continuous human involvement. The legitimacy conditions proposed in Section 3 consequently provide a normative framework for evaluating when delegated activity preserves meaningful possibilities for supervision, contestability, revision, and sustained practical engagement.

6. Conclusion

This paper argued that the emergence of agentic AI makes delegation a central concern for trustworthy AI. As systems increasingly operate on behalf of users through persistent forms of background delegation, the normative issue is no longer limited to whether such systems function efficiently, reliably, or transparently. It also concerns whether the delegation relations they establish remain compatible with meaningful human agency.

To address this problem, the paper proposed five legitimacy conditions for delegation in agentic AI systems: authorization, bounded scope, intelligible oversight, contestability, and preservation of practical human capacities. It further argued that these conditions can be operationalized through assessment questions, indicators, and representative failure modes capable of informing governance, auditing, and evaluation practices.

The broader implication is that trustworthy AI cannot be assessed solely at the level of system properties once systems increasingly act on behalf of users. Evaluating agentic AI also requires examining the legitimacy of the delegation relations through which authority, responsibility, supervision, and practical involvement are redistributed between humans and AI systems. As forms of background delegation continue to expand across digital environments, preserving meaningful human agency may therefore become a central normative requirement for trustworthy AI.

Declaration on Generative AI

During the preparation of this work, the author used ChatGPT to assist with language revision, stylistic improvement, structural discussion, and editorial feedback. All conceptual arguments, interpretations, and final editorial decisions were made by the author. After using this tool, the author reviewed and revised the manuscript and takes full responsibility for its content.

References

[1] High-Level Expert Group on Artificial Intelligence, Ethics guidelines for trustworthy ai, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, 2019. European Commission.

[2] B. Shneiderman, Human-centered artificial intelligence: Reliable, safe & trustworthy, International Journal of Human-Computer Interaction 36 (2020) 495–504.

[3] J. Bendor, A. Glazer, T. Hammond, Theories of delegation, Annual Review of Political Science 4 (2001) 235–269.

[4] Microsoft, Take back your time with copilot actions, https://www.microsoft.com/en-us/microsoft-copilot/for-individuals/do-more-with-ai/general-ai/take-back-your-time-with-copilot-actions, 2025. Accessed: 2026-05-13.

[5] Microsoft, Your ai companion, https://blogs.microsoft.com/blog/2025/04/04/your-ai-companion/, 2025. Accessed: 2026-05-13.

[6] E. R. Babbie, The Practice of Social Research, 15 ed., Cengage Learning, 2020.

[7] J. Cranshaw, et al., Calendar.help: Designing a workflow-based scheduling agent with humans in the loop, Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (2017) 2382–2393.

[8] OpenAI, Introducing operator, https://openai.com/index/introducing-operator/, 2025. Accessed: 2026-05-13.

[9] Google DeepMind, Introducing gemini 2.0: Our new ai model for the agentic era, https://blog.google/innovation-and-ai/models-and-research/google-deepmind/google-gemini-ai-update-december-2024/, 2024. Accessed: 2026-05-13.

[10] Microsoft, Copilot actions, https://copilot.microsoft.com/labs/experiments/copilot-actions, 2025. Accessed: 2026-05-13.

[11] Anthropic, Computer use and claude 3.5, https://www.anthropic.com/news/3-5-models-and-computer-use, 2024. Accessed: 2026-05-13.

[12] OECD, The agentic ai landscape and its conceptual foundations, https://www.oecd.org/en/publications/the-agentic-ai-landscape-and-its-conceptual-foundations_396cf758-en.html, 2026. Accessed: 2026-05-13.

[13] J. D. Lee, K. A. See, Trust in automation: Designing for appropriate reliance, Human Factors 46 (2004) 50–80.

[14] M. R. Endsley, E. O. Kiris, The out-of-the-loop performance problem and level of control in automation, Human Factors 37 (1995) 381–394.

[15] M. Gerlich, Ai tools in society: Impacts on cognitive offloading and the future of critical thinking, Societies 15 (2025) 6.