

# Distinguishing AI as Autonomous Agent vs. Cognitive Partner in Scientific Collaboration

## 1. Introduction

The academic literature draws a clear distinction between AI as an **autonomous agent**—which executes delegated scientific tasks with minimal human intervention—and AI as a **cognitive partner**, which interactively scaffolds human creativity and decision-making in a symbiotic relationship. Autonomous agents are characterized by their self-sufficiency, adaptability, and ability to pursue complex goals independently, often excelling in structured, repetitive, or high-throughput scientific tasks (Acharya et al., 2025; Totschnig, 2020; Croon et al., 2022; R & Biradar, 2024; Panda, 2025). In contrast, cognitive partners are designed to augment human cognition, facilitate real-time collaboration, and enhance creative or complex problem-solving by leveraging complementary strengths of both human and machine (Chakraborti et al., 2017; Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Moura, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Kacprzyk, 2022; Schelble et al., 2022; Gao et al., 2024). Recent research highlights that while autonomous agents can outperform humans in certain domains, the most significant gains in scientific discovery and innovation often arise from interactive, human-AI teaming—where AI systems are transparent, explainable, and capable of adapting to human needs and workflows (Hauptman et al., 2022; Vaccaro et al., 2024; Endsley, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Gao et al., 2024). The literature also emphasizes the importance of trust, team structure, and shared mental models in maximizing the benefits of human-AI collaboration, and cautions against over-reliance on AI or the erosion of human autonomy (Seeber et al., 2020; Zhai et al., 2024; Prunkl, 2024; Fügener et al., 2021; Liu et al., 2021; Steyvers & Kumar, 2023; Atchley et al., 2024; Köbis et al., 2021; Bergdahl et al., 2023). Overall, the distinction is not merely technical but deeply conceptual, with implications for the design, deployment, and ethical governance of AI in scientific research.

## 2. Methods

A comprehensive search was conducted across over 170 million research papers in Consensus, encompassing databases such as Semantic Scholar, PubMed, and others. The search strategy involved 20 targeted queries grouped into 8 thematic clusters, focusing on conceptual, operational, and interdisciplinary distinctions between AI as autonomous agents and as cognitive partners. In total, 1,020 papers were identified, 609 were screened, 502 were deemed eligible, and the top 50 most relevant papers were included in this review.

## Search Strategy



FIGURE 1 Flow diagram of the literature search and selection process.

Eight unique search groups were executed, systematically exploring both technical and conceptual framings of AI autonomy and partnership in scientific collaboration.

## 3. Results

### 3.1 Conceptual Distinctions: Autonomy vs. Partnership

The literature consistently distinguishes between **autonomous agents**—AI systems that operate independently to achieve delegated goals—and **cognitive partners**—AI systems that interactively support, augment, or co-create with humans (Acharya et al., 2025; Chakraborti et al., 2017; Totschnig, 2020; Jarrahi, 2018; Sandini et al., 2024; Prunkl, 2024; Moura, 2023; Lemaignan et al., 2017; Kacprzyk, 2022; Gao et al., 2024). Autonomous agents are often described as tools or delegates, while cognitive partners are framed as collaborators or teammates, requiring shared mental models, transparency, and mutual adaptation (Hauptman et al., 2022; Chakraborti et al., 2017; Endsley, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Musick et al., 2021).

### 3.2 Operational and Design Implications

Autonomous agents excel in well-defined, repetitive, or high-throughput scientific tasks, such as data analysis, simulation, or automated experimentation (Acharya et al., 2025; Totschnig, 2020; Croon et al., 2022; R & Biradar, 2024; Abràmoff et al., 2020; Panda, 2025). In contrast, cognitive partners are most effective in creative, ambiguous, or complex problem-solving contexts, where real-time interaction, explainability, and adaptability are crucial (Chakraborti et al., 2017; Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Moura, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Kacprzyk, 2022; Schelble et al., 2022; Gao et al., 2024). The design of AI systems for cognitive partnership emphasizes transparency, explainability, and the ability to build trust and shared understanding with human users (Hauptman et al., 2022; Endsley, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Gao et al., 2024).

### 3.3 Human-AI Teaming and Synergy

Meta-analyses and empirical studies show that human-AI teams can outperform either humans or AI alone, particularly in creative or content-generation tasks (Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Gao et al., 2024). However, the effectiveness of such teaming depends on task type, team structure, and the degree of interactivity and mutual adaptation (Hauptman et al., 2022; Vaccaro et al., 2024; Endsley, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Musick et al., 2021). Over-reliance on AI or poor team design can undermine human autonomy and critical thinking (Zhai et al., 2024; Prunkl, 2024; Fügner et al., 2021; Liu et al., 2021; Steyvers & Kumar, 2023; Atchley et al., 2024; Köbis et al., 2021; Bergdahl et al., 2023).

### 3.4 Ethical, Social, and Cognitive Considerations

The literature highlights ethical challenges, such as the risk of eroding human autonomy, the need for responsible delegation, and the importance of maintaining human agency in scientific decision-making (Totschnig, 2020; Seeber et al., 2020; Zhai et al., 2024; Prunkl, 2024; Bengio et al., 2023; Fügner et al., 2021; Liu et al., 2021; Steyvers & Kumar, 2023; Atchley et al., 2024; Köbis et al., 2021; Bergdahl et al., 2023). There is also a growing recognition of the need for new paradigms and governance frameworks to address the unique risks and opportunities posed by both autonomous agents and cognitive partners (Moura, 2023; Bengio et al., 2023; Kacprzyk, 2022; Gao et al., 2024).

#### Key Papers

Paper	Distinction Focus	Methodology	Key Results	Application Domain
(Vaccaro et al., 2024)	Human-AI synergy	Systematic review & meta-analysis	Human-AI teams outperform alone in creative tasks; synergy depends on task type	Broad (decision-making, content creation)
(Chakraborti et al., 2017)	Cognitive teaming vs. autonomy	Analytical review	Human-robot teaming requires new cognitive models beyond autonomy	Robotics, HCI
(Hauptman et al., 2022)	Adaptive autonomy in teams	Mixed methods (survey, interviews)	Dynamic autonomy enhances team performance and cohesiveness	Incident response, HCI
(Jarrahi, 2018)	Human-AI symbiosis	Conceptual analysis	AI augments human cognition, especially under complexity/uncertainty	Organizational decision-making
(Gao et al., 2024)	Levels of AI autonomy	Conceptual framework	Proposes levels from tool to autonomous agent; highlights risks and opportunities	Biomedical research

**FIGURE 2** Comparison of key studies on distinctions between AI as autonomous agent and cognitive partner.

## Top Contributors

Type	Name	Papers
Author	Nathan J. Mcneese	(Hauptman et al., 2022; Schelble et al., 2022; Musick et al., 2021)
Author	Beau G. Schelble	(Hauptman et al., 2022; Schelble et al., 2022; Musick et al., 2021)
Author	Michelle Vaccaro	(Vaccaro et al., 2024)
Journal	<i>Comput. Hum. Behav.</i>	(Hauptman et al., 2022; Seeber et al., 2020; Candrian & Scherer, 2022; Endsley, 2023; Walliser et al., 2019; Frischknecht, 2021; Musick et al., 2021)
Journal	<i>Nature Human Behaviour</i>	(Vaccaro et al., 2024; Köbis et al., 2021)
Journal	<i>Cell</i>	(Gao et al., 2024)

**FIGURE 3** Authors & journals that appeared most frequently in the included papers.

## 4. Discussion

The distinction between AI as an autonomous agent and as a cognitive partner is well-established and increasingly nuanced in the literature. Autonomous agents are valuable for automating routine, well-structured scientific tasks, but their effectiveness is limited in domains requiring creativity, adaptability, and nuanced judgment (Acharya et al., 2025; Totschnig, 2020; Croon et al., 2022; R & Biradar, 2024; Abràmoff et al., 2020; Panda, 2025). Cognitive partners, by contrast, are designed to scaffold human creativity, support complex decision-making, and foster innovation through interactive, symbiotic relationships (Chakraborti et al., 2017; Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Moura, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Kacprzyk, 2022; Schelble et al., 2022; Gao et al., 2024). High-quality evidence, including systematic reviews and meta-analyses, supports the claim that human-AI teaming can yield superior outcomes in creative and collaborative contexts, provided that systems are designed for transparency, trust, and mutual adaptation (Hauptman et al., 2022; Vaccaro et al., 2024; Endsley, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Gao et al., 2024).

However, challenges remain. Over-reliance on AI can erode human autonomy and critical thinking, while poorly designed autonomous agents may introduce new risks or ethical concerns (Zhai et al., 2024; Prunkl, 2024; Fügner et al., 2021; Liu et al., 2021; Steyvers & Kumar, 2023; Atchley et al., 2024; Köbis et al., 2021; Bergdahl et al., 2023). The literature calls for new paradigms, governance frameworks, and empirical research to ensure that AI systems—whether autonomous or collaborative—are deployed responsibly and effectively (Moura, 2023; Bengio et al., 2023; Kacprzyk, 2022; Gao et al., 2024).

## Claims and Evidence Table

Claim	Evidence Strength	Reasoning	Papers
The literature clearly distinguishes between AI as autonomous agent and as cognitive partner	 Strong	Multiple high-quality reviews and empirical studies explicitly define and contrast these roles	(Hauptman et al., 2022; Acharya et al., 2025; Chakraborti et al., 2017; Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Prunkl, 2024; Moura, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Kacprzyk, 2022; Schelble et al., 2022; Gao et al., 2024)
Human-AI teaming outperforms either alone in creative/content-generation tasks	 Strong	Meta-analyses and systematic reviews show synergy in creative tasks, but not always in decision tasks	(Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Sowa et al., 2021; Lemaignan et al., 2017; Schelble et al., 2022; Gao et al., 2024)
Autonomous agents excel in routine, well-structured scientific tasks	 Strong	Empirical and conceptual studies show high performance in automation, data analysis, and repetitive tasks	(Acharya et al., 2025; Totschnig, 2020; Croon et al., 2022; R & Biradar, 2024; Abràmoff et al., 2020; Panda, 2025)
Cognitive partners are most effective in complex, creative, or ambiguous tasks	 Strong	Reviews and case studies highlight the value of interactivity, explainability, and adaptability	(Chakraborti et al., 2017; Vaccaro et al., 2024; Jarrahi, 2018; Ren et al., 2023; Moura, 2023; Sowa et al., 2021; Lemaignan et al., 2017; Kacprzyk, 2022; Schelble et al., 2022; Gao et al., 2024)
Over-reliance on AI can erode human autonomy and critical thinking	 Moderate	Systematic reviews and conceptual analyses warn of risks to autonomy and agency	(Zhai et al., 2024; Prunkl, 2024; Fügner et al., 2021; Liu et al., 2021; Steyvers & Kumar, 2023; Atchley et al., 2024; Köbis et al., 2021; Bergdahl et al., 2023)
The boundary between tool and agent is sometimes blurred in practice	 Moderate	Some systems exhibit both tool-like and agentic qualities, complicating clear categorization	(Prunkl, 2024; Albrecht & Stone, 2017; De Zúñiga et al., 2023; Bitterman et al., 2020; Mosqueira-Rey et al., 2022; Yıldız, 2025)

**FIGURE 4** Key claims and support evidence identified in these papers.

## 5. Conclusion

In summary, the academic literature robustly distinguishes between AI as an autonomous agent and as a cognitive partner, with each role offering unique strengths and challenges in scientific collaboration. The most promising advances in scientific discovery are likely to arise from thoughtfully designed human-AI partnerships that leverage the complementary capabilities of both.

### 5.1 Research Gaps

Despite significant progress, gaps remain in empirical research on the long-term impacts of human-AI teaming, the development of shared mental models, and the ethical governance of increasingly autonomous systems. There is also a need for more interdisciplinary studies and real-world evaluations of cognitive partnership models.

#### Research Gaps Matrix

Topic / Study Attribute	Empirical Studies	Conceptual Analyses	Meta-analyses	Case Studies	Ethical/Philosophical
Autonomous agents	12	8	2	4	3
Cognitive partners	9	10	3	5	4
Human-AI teaming	8	7	2	3	2
Trust & explainability	5	6	1	2	3
Ethics & governance	3	7	GAP	1	6

**FIGURE 5** Matrix of research topics and study attributes, highlighting areas with fewer studies.

### 5.2 Open Research Questions

Future research should focus on empirically validating cognitive partnership models, developing robust frameworks for trust and explainability, and exploring the ethical implications of increasingly autonomous AI in scientific collaboration.

Question	Why
How do different models of human-AI cognitive partnership impact long-term scientific creativity and innovation?	Understanding this will inform the design of AI systems that best augment human creativity and drive scientific breakthroughs.
What governance frameworks are needed to ensure responsible deployment of autonomous agents in science?	As AI autonomy increases, robust governance is essential to mitigate risks and ensure ethical, transparent scientific practice.
How can trust and shared mental models be effectively developed in human-AI teams?	Trust and shared understanding are critical for effective collaboration, but best practices for fostering them remain underexplored.

**FIGURE 6** Open research questions for future investigation in human-AI scientific collaboration.

In conclusion, the literature provides a strong conceptual and empirical foundation for distinguishing between AI as autonomous agent and as cognitive partner, but ongoing research is needed to optimize these roles for scientific advancement.

*These papers were sourced and synthesized using Consensus, an AI-powered search engine for research. Try it at <https://consensus.app>*

## References

- Hauptman, A., Schelble, B., Mcneese, N., & Madathil, K. (2022). Adapt and overcome: Perceptions of adaptive autonomous agents for human-AI teaming. *Comput. Hum. Behav.*, 138, 107451. <https://doi.org/10.1016/j.chb.2022.107451>
- Acharya, D., Kuppan, K., & Bhaskaracharya, D. (2025). Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey. *IEEE Access*, 13, 18912-18936. <https://doi.org/10.1109/ACCESS.2025.3532853>
- Chakraborti, T., Kambhampati, S., Scheutz, M., & Zhang, Y. (2017). AI Challenges in Human-Robot Cognitive Teaming. *ArXiv*, abs/1707.04775.
- Totschnig, W. (2020). Fully Autonomous AI. *Science and Engineering Ethics*, 26, 2473 - 2485. <https://doi.org/10.1007/s11948-020-00243-z>
- Seeber, I., Waizenegger, L., Seidel, S., Morana, S., Benbasat, I., & Lowry, P. (2020). Collaborating with technology-based autonomous agents. *Internet Res.*, 30, 1-18. <https://doi.org/10.1108/intr-12-2019-0503>
- Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 8, 2293 - 2303. <https://doi.org/10.1038/s41562-024-02024-1>
- Jarrahi, M. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*. <https://doi.org/10.1016/J.BUSHOR.2018.03.007>
- Sandini, G., Sciutti, A., & Morasso, P. (2024). Artificial cognition vs. artificial intelligence for next-generation autonomous robotic agents. *Frontiers in Computational Neuroscience*, 18. <https://doi.org/10.3389/fncom.2024.1349408>

- Candrian, C., & Scherer, A. (2022). Rise of the machines: Delegating decisions to autonomous AI. *Comput. Hum. Behav.*, 134, 107308. <https://doi.org/10.1016/j.chb.2022.107308>
- Ren, M., Chen, N., & Qiu, H. (2023). Human-machine Collaborative Decision-making: An Evolutionary Roadmap Based on Cognitive Intelligence. *International Journal of Social Robotics*, 15, 1101-1114. <https://doi.org/10.1007/s12369-023-01020-1>
- Zhai, C., Wibowo, S., & Li, L. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learn. Environ.*, 11, 28. <https://doi.org/10.1186/s40561-024-00316-7>
- Prunkl, C. (2024). Human Autonomy at Risk? An Analysis of the Challenges from AI. *Minds and Machines*, 34, 1-21. <https://doi.org/10.1007/s11023-024-09665-1>
- Endsley, M. (2023). Supporting Human-AI Teams: Transparency, explainability, and situation awareness. *Comput. Hum. Behav.*, 140, 107574. <https://doi.org/10.1016/j.chb.2022.107574>
- Moura, F. (2023). Artificial Intelligence, Creativity, and Intentionality: The Need for a Paradigm Shift. *The Journal of Creative Behavior*. <https://doi.org/10.1002/jocb.585>
- Sowa, K., Przegalinska, A., & Ciechanowski, L. (2021). Cobots in knowledge work. *Journal of Business Research*, 125, 135-142. <https://doi.org/10.1016/j.jbusres.2020.11.038>
- Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, Y., Zhang, Y., Xue, L., Shalev-Shwartz, S., Hadfield, G., Clune, J., Maharaj, T., Hutter, F., Baydin, A., McIlraith, S., Gao, Q., Acharya, A., Krueger, D., Dragan, A., Torr, P., Russell, S., Kahneman, D., Brauner, J., & Mindermann, S. (2023). Managing extreme AI risks amid rapid progress. *Science*, 384, 842 - 845. <https://doi.org/10.1126/science.adn0117>
- Walliser, J., De Visser, E., Wiese, E., & Shaw, T. (2019). Team Structure and Team Building Improve Human–Machine Teaming With Autonomous Agents. *Journal of Cognitive Engineering and Decision Making*, 13, 258 - 278. <https://doi.org/10.1177/1555343419867563>
- Croon, G., Dupeyroux, J., Fuller, S., & Marshall, J. (2022). Insect-inspired AI for autonomous robots. *Science Robotics*, 7. <https://doi.org/10.1126/scirobotics.abl6334>
- Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Cognitive Challenges in Human-Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Inf. Syst. Res.*, 33, 678-696. <https://doi.org/10.1287/isre.2021.1079>
- Liu, H., Lai, V., & Tan, C. (2021). Understanding the Effect of Out-of-distribution Examples and Interactive Explanations on Human-AI Decision Making. *Proceedings of the ACM on Human-Computer Interaction*, 5, 1 - 45. <https://doi.org/10.1145/3479552>
- Albrecht, S., & Stone, P. (2017). Autonomous agents modelling other agents: A comprehensive survey and open problems. *ArXiv*, abs/1709.08071. <https://doi.org/10.1016/j.artint.2018.01.002>
- De Zúñiga, H., Goyanes, M., & Durotoye, T. (2023). A Scholarly Definition of Artificial Intelligence (AI): Advancing AI as a Conceptual Framework in Communication Research. *Political Communication*, 41, 317 - 334. <https://doi.org/10.1080/10584609.2023.2290497>
- Steyvers, M., & Kumar, A. (2023). Three Challenges for AI-Assisted Decision-Making. *Perspectives on Psychological Science*, 19, 722 - 734. <https://doi.org/10.1177/17456916231181102>
- Bitterman, D., Aerts, H., & Mak, R. (2020). Approaching autonomy in medical artificial intelligence.. *The Lancet. Digital health*, 2 9, e447-e449. [https://doi.org/10.1016/s2589-7500\(20\)30187-4](https://doi.org/10.1016/s2589-7500(20)30187-4)



- Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2022). Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review*, 56, 3005-3054. <https://doi.org/10.1007/s10462-022-10246-w>
- Lemaignan, S., Warnier, M., Sisbot, E., Clodic, A., & Alami, R. (2017). Artificial cognition for social human-robot interaction: An implementation. *Artif. Intell.*, 247, 45-69. <https://doi.org/10.1016/J.ARTINT.2016.07.002>
- R, B., & Biradar, V. (2024). Empowering Autonomous Decision-Making Through Quantum Reinforcement Learning and Cognitive Neuromorphic Frameworks. *2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, 1-7. <https://doi.org/10.1109/ICMNWC63764.2024.10872223>
- Kacprzyk, J. (2022). AI-enabled Decision Aid and Decision Support for Symbiotic Autonomous Systems. *2022 IEEE 21st International Conference on Cognitive Informatics & Cognitive Computing (ICCICC)\**, 4-5. <https://doi.org/10.1109/iccicc57084.2022.10101649>
- Yildiz, T. (2025). The Minds We Make: A Philosophical Inquiry into Theory of Mind and Artificial Intelligence.. *Integrative psychological & behavioral science*, 59 1, 10. <https://doi.org/10.1007/s12124-024-09876-2>
- Schelble, B., Flathmann, C., Mcneese, N., Freeman, G., & Mallick, R. (2022). Let's Think Together! Assessing Shared Mental Models, Performance, and Trust in Human-Agent Teams. *Proceedings of the ACM on Human-Computer Interaction*, 6, 1 - 29. <https://doi.org/10.1145/3492832>
- Atchley, P., Pannell, H., Wofford, K., Hopkins, M., & Atchley, R. (2024). Human and AI collaboration in the higher education environment: opportunities and concerns. *Cognitive Research: Principles and Implications*, 9. <https://doi.org/10.1186/s41235-024-00547-9>
- Abràmoff, M., Tobey, D., & Char, D. (2020). Lessons learnt about Autonomous AI: Finding a safe, efficacious and ethical path through the development process.. *American journal of ophthalmology*. <https://doi.org/10.1016/j.ajo.2020.02.022>
- Köbis, N., Bonnefon, J., & Rahwan, I. (2021). Bad machines corrupt good morals. *Nature Human Behaviour*, 5, 679 - 685. <https://doi.org/10.1038/s41562-021-01128-2>
- Bergdahl, J., Latikka, R., Celuch, M., Savolainen, I., Mantere, E., Savela, N., & Oksanen, A. (2023). Self-determination and attitudes toward artificial intelligence: Cross-national and longitudinal perspectives. *Telematics Informatics*, 82, 102013. <https://doi.org/10.1016/j.tele.2023.102013>
- Frischknecht, R. (2021). A social cognition perspective on autonomous technology. *Comput. Hum. Behav.*, 122, 106815. <https://doi.org/10.1016/J.CHB.2021.106815>
- Gao, S., Fang, A., Huang, Y., Giunchiglia, V., Noori, A., Schwarz, J., Ektefaie, Y., Kondic, J., & Zitnik, M. (2024). Empowering biomedical discovery with AI agents. *Cell*, 187, 6125-6151. <https://doi.org/10.1016/j.cell.2024.09.022>
- Musick, G., O'Neill, T., Schelble, B., Mcneese, N., & Henke, J. (2021). What Happens When Humans Believe Their Teammate is an AI? An Investigation into Humans Teaming with Autonomy. *Comput. Hum. Behav.*, 122, 106852. <https://doi.org/10.1016/j.chb.2021.106852>
- Panda, M. (2025). Agentic AI: A Paradigm for Autonomous Decision-Making. *International Journal of Innovative Research in Science Engineering and Technology*. <https://doi.org/10.15680/ijirset.2025.1401077>