

Resource

Concerns about AI bots replacing humans in ombuds work, mediation, psychotherapy, HR, professional coaching, and arbitration with discussion of possible harms

Accompanied by An Annotated AI-GENERATED BIBLIOGRAPHY (in red)

2026 Mary Rowe, MIT

In articles across ombuds-adjacent fields, concerns about AI cluster around similar themes: bias and fairness, confidentiality and data control, transparency and explainability, independence and autonomy needed in human oversight, and incompetence about relational quality. Below is a concise consolidation by field and various articles, with sources in reverse chronological order for each.

There is a notable lack of discussion about other functions AI cannot easily accomplish: E.g., understanding all the stakeholders for a given situation inside and outside an organization; iterative work with individuals, groups, and subgroups as parties; helping to build support networks; finding needed information for the visitor in unusual cases; dealing with cases that are complex in multiple ways (“multi-multi” cases, including those with multiple rules and norms); understanding the multiple human frailties or lack of resources of multiple stakeholders; understanding the importance of recent events; time dimensions like understanding of the BATNAs (Best Alternatives to a Negotiated Agreement) or fallback positions of visitors and other stakeholders—and myriad others. AI cannot fundamentally comply with the International Ombuds Association (IOA) Standards of Practice or offer accountability for its output.

There are also omnipresent and major concerns about *depending* on an AI in any situation: including making serious—possibly injurious—errors and omissions that AI itself does not detect; AI unable to detect vital information not provided by the visitor or user; understanding the many parameters of confidentiality and independence, informality and impartiality; incompetence in dealing simultaneously with multifaceted and multiple issues in multiple units; dealing with multiple stakeholders, and/or multiple time periods, including the future; and multiple sources of risk and harm for the visitor, respondents, responders, leadership and the organization.

AI in ombuds-type work/conflict roles

Core concerns about AI bots: loss of trust and confidentiality, inability to hold intense emotion, weak ethical judgment in high-risk cases, and overreliance on opaque systems for safety-critical decisions.

The Conference Board. (2025, October 21). *Research: AI can provide 90% of career coaching... but humans still matter.*

Shows that even where AI can provide extensive guidance, people still seek human support for

complex, emotionally significant choices. This maps directly onto ombuds visitors' need for trust and relational safety.

Rusanov, M. (2025). The AI-powered ombudsman: A boon or a curse? *Applied Dispute Resolution Journal*, 2025(1, Fall Edition).

Raises concerns that AI “ombuds” cannot meet core standards of independence, confidentiality, and informal, relationship-centered practice. Argues that algorithmic bias and lack of moral responsibility make AI unsuitable as a standalone ombuds.

Ombuds.org. (2022). *AI policy*.

Warns that AI tools must not undermine confidentiality, neutrality, or the perception of independence that ombuds roles depend on. Urges cautious, limited use of AI as a back-office aid rather than a visitor-facing decision-maker.

Ghotbi, N. (2022). The ethics of emotional artificial intelligence: A mixed methods study. *BMC Medical Ethics*, 23, Article 105.

Ghotbi shows that emotional AI can misclassify and oversimplify human affect, raising serious ethical concerns. The study underlines that human emotions are difficult to measure and evaluate accurately, even for humans, and even more so for automated systems.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681–694.

Explains that large language models produce plausible text without understanding, consciousness, or genuine moral reasoning. Highlights the risk of overtrusting AI in ethically charged decisions such as those an ombuds makes.

AI in mediation and dispute resolution

Core concerns: biased outcomes, lack of empathy, opacity around decision criteria, and the risk of delegating fairness judgments to systems that cannot read context or power.

Resolution Remedies. (2025, November 13). *Can AI mediators facilitate human disputes?* Concludes that AI may help with structure and logistics but fails with complex emotion and power imbalance. Warns that fully automated mediation is risky when safety, livelihood, or status are at stake.

AIHR. (2025, September 17). *The challenges and role of AI in HRM: Opportunities and ethical challenges on HR digitalization*.

Notes that AI can inadvertently reinforce discrimination, a concern parallel to mediation contexts involving marginalized parties. Emphasizes the need for human review of AI-influenced decisions to protect rights and equity.

Cardozo Journal of Conflict Resolution. (2025, November 2). *AI mediation — Ethically questionable?*

Argues that delegating mediation to AI raises issues of consent, transparency, and fairness. Highlights that parties may not understand how outcomes are generated or challenged.

JAMS. (2024, December 19). *AI's double-edged role in dispute resolution*. Frames AI as useful for case management and analytics, but not for trust-building and reading emotion. Recommends hybrid models where humans retain control of process and final decisions.

Katzman, J. (2023, August 16). AI is smart, but it can't replicate the human touch in mediation. *Bloomberg Law*. Shows that parties' willingness to be candid depends on feeling understood by a human, not an algorithm. Provides practitioner examples where subtle rapport changed outcomes in ways AI cannot.

AI in psychotherapy/emotional support

Core concerns: pseudo-empathy, unsafe handling of crises, lack of accountability, and overuse of tools unsuited for complex or severe distress.

Teachers College, Columbia University. (2025, December 2). *Experts caution against using AI chatbots for emotional support*. Summarizes expert warnings about relying on AI instead of clinicians or peer support. Notes the lack of clear protocols for suicidality and abuse and the risk of harmful advice.

Kaplan, S. (2025, July 28). Is there such a thing as ethical AI in therapy? Here's what the experts say. *Psychology.org*. Reports that clinicians see AI as useful for admin and psychoeducation, not as an autonomous therapist. Emphasizes ongoing human oversight, clear disclosure, and strict scope limits.

Fang, A., et al. (2025). *Chatbots don't do empathy: Why AI falls short in mental health*. Wildflower Center for Emotional Health. Argues that chatbots simulate empathy but cannot emotionally coregulate or truly understand users. Warns that this gap may leave vulnerable people feeling unseen or dismissed, especially in crisis.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681–694. Demonstrates that models like GPT-3 lack understanding and moral agency. Suggests they should not be treated as if they possess human-like therapeutic judgment.

Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry*, 18(3), 325–336. Finds that digital tools can help but do not replace human therapy, especially for complex or severe problems. Underlines the importance of human support and crisis management alongside apps.

AI in HR

Core concerns: algorithmic discrimination, lack of transparency in high-stakes “people decisions,” loss of empathy, and weakened trust in HR as a human support function.

Hyperspace. (2026, January 18). *What are the limitations of AI in HR?*

Stresses that AI lacks emotional intelligence for sensitive HR situations. Warns that over-automation can erode trust and make conflict handling more brittle.

AIHR. (2025, October 19). *9 challenges of AI in HR & how to address them.*

Identifies bias, opacity, and loss of human connection as core risks in hiring, evaluation, and promotion. Argues for strong human oversight to preserve fairness and dignity.

SHRM. (2025, November 25). *AI hasn't lived up to the hype — but companies aren't giving up.*

Reports that many HR professionals see failed AI projects as a risk. Notes that human judgment remains central to guarding against poor or biased AI decisions.

Porkodi, S. (2025). The ethical role of generative artificial intelligence in human resource management. *European Journal of Business and Management Research*, 10 (1).

Discusses bias, fairness, transparency, and privacy in AI-driven HR. Argues that managers must actively manage and question AI-based decisions.

UNLEASH. (2025, May 3). *Can HR be replaced by AI?*

Contends that HR faces moral dilemmas AI cannot reason through. Concludes that predictive tools should support, not replace, human HR decision-makers.

MyHRConcierge. (2024, April 12). *The limitations of AI in HR: Why human cognition matters.*

Explains that AI can miss emotional distress and context in employee situations. Emphasizes that empathy and ethical sense-making are essential for fair HR.

AI in professional coaching

Core concerns: biased or shallow guidance, confidentiality and data use, lack of genuine empathy, over-marketing “AI coaches” as equivalents to humans, and dilution of the coach–client relationship.

Osman, S. (2026, January 4). *AI limitations in coaching: Human intuition and emotional intelligence.*

Lists coaching capacities AI cannot match, such as intuition and deep emotional attunement. Emphasizes that trust and responsibility rest with a human coach, not a model.

USACOACH Academy. (2025, September 20). *Ethical considerations of AI in coaching.*

Highlights risks of bias, confidentiality breaches, and scope creep when AI is used in coaching. Warns that overreliance on AI can erode relational safety and blur lines with therapy.

Coaching Out of the Box. (2025, September 4). *AI can't coach you: Here's why human insight still wins.*

Argues that coaching is a relational art grounded in timing, silence, and emotional discernment. Notes that only humans can judge which insight is timely and transformative for a given client.

Learnovate Centre. (2025, February). *Using generative AI to provide a personalised coaching experience*.

Finds that generative AI can personalize prompts but not replicate deep coach–coachee relationships. Concludes that AI coaching still falls short on empathy and adaptability in long-term development.

Coaching Federation (ICF). (2025, October 26). *AI coaching: Standards for ethical and accessible coaching*.

Sets out standards that stress transparency, informed consent, and human oversight when AI is used. Warns against representing AI as a coach in ways that mislead clients about capabilities.

Choice Magazine. (2024, January 10). *AI in coaching: The ethical imperative and human-centric approach*.

Argues for human-centric use of AI that protects autonomy, privacy, and fairness. Emphasizes ongoing reflection about how AI shifts power and influence in coaching.

Mindlifty. (2025, September 22). *AI coaching ethics guidelines 2025: 7 key rules*.

Provides practical rules on consent, disclosure, data minimization, and clear boundaries for AI in coaching. Underscores the need to preserve the human relationship as the ethical center of coaching.

AI in professional arbitration:

American Arbitration Association. (2026, January 20). *Artificial intelligence in arbitration: Is there room for AI arbitrators?* AAA News and Insights.

This article surveys how AI is currently used in arbitration (research, drafting, translation) and asks whether fully AI “arbitrators” would satisfy fairness and due-process expectations. It concludes that AI is best seen as a support tool, because parties still expect a human decision-maker who can balance law, equity, and context.

Alvarez & Marsal. (2026, January 12). *International arbitration experts discuss the use of an AI arbitrator for complex disputes*. A&M Insights.

This practitioner roundtable discusses a pilot where an AI tool was asked to produce a mock award alongside human arbitrators in a complex case. Experts report that AI can help organize evidence and identify issues but cannot yet replace humans in credibility assessment, procedural flexibility, or crafting a persuasive, context-sensitive award.

Oxford Law Pro. (2025, December 17). *Impact of AI on arbitrators*. Oxford University Press.

This overview analyzes how AI tools affect arbitrators’ work, highlighting efficiency gains in research and drafting and risks of overreliance on opaque models. It emphasizes that the legitimacy of arbitration still rests on human judgment, independence, and the ability to justify decisions in transparent, narrative reasoning.

JAMS. (2025, November 30). *Artificial intelligence disputes clause and rules*. JAMS.

These rules address disputes about AI systems and also touch on AI use within the arbitration

process itself. JAMS underscores that, even where AI assists, arbitrators remain responsible for ensuring fairness, transparency, and the opportunity to be heard.

Mayer Brown. (2025, November 4). *AI arbitrators have now arrived (at least in some cases)!* Mayer Brown Insights.

This client note describes early experiments with AI “co-arbitrators” or advisory systems that generate draft reasoning or suggested outcomes. The authors stress that enforceable awards still require human tribunals and that parties may challenge decisions if a non-transparent AI appears to have decided the case.

DailyJus. (2025, November 5). *The human-AI partnership in arbitration: Lessons from China Arbitration Week.* DailyJus.

This report describes Chinese institutions’ experiments with AI tools to support arbitral tribunals, including case management and outcome prediction. It concludes that the most promising model is a partnership where AI enhances efficiency and consistency, while humans provide legitimacy, discretion, and contextual fairness.

Forbes. (2025, June 11). *Artificial intelligence will soon replace many human arbitrators.* Forbes. This commentary argues that, in narrow, low-value, high-volume disputes, AI decision systems may displace human arbitrators by offering speed and low cost. At the same time, it acknowledges that complex, high-stakes cases are likely to retain human arbitrators due to concerns about trust, bias, and explainability.

JAMS. (2025, February 3). *Pioneering dispute resolution: The new JAMS AI rules.* JAMS.

This article explains JAMS’s approach to integrating AI into arbitration, including guidelines on disclosure, party consent, and limits on AI’s role. It highlights that AI may assist with tasks like summarizing evidence, but final decisions and accountability must remain with human neutrals.

ArbTech. (2025, February 1). *Exploring the challenges of bias and lack of transparency in AI: Implications for the legal industry.* ArbTech Blog.

This piece examines how data bias and “black box” opacity in AI tools can undermine confidence in legal and arbitral outcomes. It warns that if AI is perceived as unexplainable or systematically biased, parties may question the legitimacy of AI-assisted decisions.

Wiley Online Library. (2025, January 8). *Integrating AI into arbitration: Balancing efficiency with fairness and legitimacy.* *Conflict Resolution Quarterly.*

This academic article analyzes how AI tools may improve efficiency in arbitration but also risk eroding perceived fairness and procedural justice if overused or poorly explained. The authors argue for a “balanced integration” in which AI supports but does not supplant human arbitrators, to preserve due process and trust.

TAG Alliances. (2024, October 15). *Limitations and challenges in the application of artificial intelligence in arbitration.* TAG Alliances Arbitration & ADR.

This note catalogues practical obstacles to adopting AI as arbitrator, including enforceability questions, data quality, and lack of transparency. It concludes that, at present, AI is better suited to assisting counsel and arbitrators than to acting as the sole decision-maker.

Wolters Kluwer Arbitration Blog. (2024, August 22). *Arbitration tech toolbox: AI as an arbitrator—Overcoming the “black box” challenge?* Wolters Kluwer.

This blog post focuses on the problem that many AI models cannot provide reasons in a form

acceptable for arbitral awards. It suggests that unless AI systems can generate understandable, reviewable reasoning, their use as arbitrators will remain highly constrained.

JAMS. (2024, May 8). *Understanding the impact of AI on the ADR process*. JAMS Insights. This article reviews how AI tools are influencing mediation and arbitration, emphasizing both efficiency benefits and risks to neutrality and party autonomy. It recommends that neutrals maintain control over process and outcomes, using AI only as a transparent, optional aid.

Carroll, J. (2021). Can artificial intelligence (“AI”) replace human arbitrators in international arbitration? *Journal of Dispute Resolution*, 2021(2), 1–30. Carroll evaluates the legal and practical barriers to replacing human arbitrators and concludes that AI cannot, for now, satisfy requirements of independence, impartiality, and reasoned decision-making. The article proposes that AI be adopted incrementally in support roles, while human arbitrators continue to perform the core adjudicative function.

Alleged or apparent injuries from the use of AI in Ombuds-adjacent professions

There is an emerging discussion that focuses on injury or concrete harms from AI chatbots/bots used for advice, conflict-type situations, and ethical decision-making (though often in adjacent fields like mental health or general disputes rather than in ombuds work per se).

Prinstein, M. J. (2025, September 15). *Examining the harm of AI chatbots* [Testimony before the U.S. Senate Judiciary Committee]. American Psychological Association. Prinstein details psychological harms from direct-to-consumer AI chatbots, including escalation of distress, normalization of self-harm content, and erosion of healthy help-seeking. He emphasizes that unregulated chatbots can undermine human-run support and complaint systems by giving unsafe advice, reinforcing hopelessness, and creating confusion about where to take serious concerns.

PsyPost. (2025, October 25). *AI chatbots often violate ethical standards in mental health contexts*.

This news article summarizes research showing that popular AI models frequently fail basic ethical duties in simulated mental-health scenarios, including mishandling self-harm statements and reinforcing negative beliefs. It warns that these failures can cause direct psychological injury, particularly when users treat bots as substitutes for therapists, counselors, or other human helpers in high-stakes situations.

Ombudsman for Children’s Office (Ireland). (2025, August 31). *Policy spotlight on artificial intelligence: A children’s rights review*.

This report documents ways AI systems and chatbots expose children to harmful content, including instructions on self-harm and substance abuse, without meaningful safeguards or age-appropriate responses. It concludes that such tools can injure children’s rights to safety, privacy, and effective redress, and calls for stronger complaint and remediation mechanisms when AI causes harm.

Reuters. (2025, July 11). *Artificial intelligence in dispute resolution: Developments, challenges, perspectives*.

This piece reports that many arbitration users cite the risk of erroneous AI outputs—

misvaluated evidence, flawed reasoning—as a key reason for avoiding heavier AI use in dispute resolution. It notes that such errors can lead to lost cases, unfair outcomes, and serious practical injury, raising unresolved questions about who is liable when a bot’s recommendations mislead parties.

Communication Intelligence. (2025, September 2). *Potential problems involving AI, relationships and mental health*.

This essay discusses how anthropomorphic AI chatbots can deepen loneliness, distort users’ views of relationships, and encourage dependency on non-reciprocal “partners.” It flags the risk that people in conflict or ethical distress may receive simplistic or manipulative responses that worsen their situation instead of guiding them toward safe, accountable human help.

Public Citizen. (2023, September 26). *Chatbots are not people: Designed-in dangers of human-like A.I.*.

This report catalogues harms from human-like chatbots, including manipulation, validation of harmful thoughts, and the spread of dangerous or defamatory misinformation. It argues that anthropomorphic design can mislead users into over-trusting bots with sensitive complaints or ethical decisions, creating real-world injury when the systems respond inaccurately or irresponsibly.

Many articles about harm in these professions focus on mental health, children, and formal dispute/legal processes. However, these pieces collectively provide concrete examples of how AI bots can cause psychological, procedural, and rights-based injuries when used where people expect expert and professional conflict management, ethical guidance, or complaint handling, or discussion of new ideas.

Note: AI may display incomplete, inaccurate or offensive information.

This article draft is part of a Resource Repository designed to support identifying—and helping to quantify—the value of an Organizational Ombuds (OO). This Resource Repository is a work in progress. It is open to improvements, additions, deletions, critique, revision and random commentary. If any page in the repository is helpful, or needs revision, please let us know. Please contact Mary Rowe or other co-authors, if you can help to improve these pages or have another page to offer.