Received: 19.12.2025 | Revised: 30.01.2026 | Accepted: 02.02.2026

Intimate Machines, Disturbed Minds: Managing the Affective Cost of AI

Leona Chandra Kruse, Patrick Mikalef

Abstract: As artificial intelligence (AI) systems increasingly simulate empathy and engage relationally with users, they begin to influence the affective fabric of organizations. Drawing on insights from information systems research, organizational science, and management theory, we argue that existing theoretical and governance frameworks are ill-equipped to account for these developments. We distinguish external regulation from internal governance and propose a preventive–corrective framework for addressing affective costs across work contexts. Building on this framework, we outline a research agenda that challenges prevailing assumptions in management and strategy, calling for renewed attention to affect, relational dynamics, and emotional architectures in organizations shaped by affective AI.

Keywords: Affective AI, AI Governance, Human-AI interaction, Responsible AI, Organizational affect

Vertraute Maschinen, verstörte Geister: über die affektiven Kosten von KI und ihre Steuerung

Zusammenfassung: Da künstliche Intelligenzsysteme (KI) zunehmend Empathie simulieren und mit Benutzern in Beziehung treten, beginnen sie, die affektive Struktur von Organisationen zu beeinflussen. Auf der Grundlage von Erkenntnissen aus der Informationssystemforschung, der Organisationswissenschaft und der Managementtheorie argumentieren wir, dass bestehende theoretische und Governance-Rahmenwerke nicht ausreichend sind, um diesen Entwicklungen Rechnung zu tragen. Wir unterscheiden zwischen externer Regulierung und interner Governance und schlagen einen präventiv-korrektiven Rahmen vor, um affektive Kosten in verschiedenen Arbeitskontexten anzugehen. Auf der Grundlage dieses Rahmens skizzieren wir eine Forschungsagenda, die gängige Annahmen in Management und Strategie in Frage stellt und eine erneute Aufmerksamkeit für Affekte, Beziehungsdynamiken und emotionale Architekturen in Organisationen fordert, die von affektiver KI geprägt sind.

Schlüsselwörter: Affektive KI, KI-Governance, Mensch-KI-Interaktion, Verantwortungsvolle KI, Organisatorischer Affekt

The Shift: From Automation to Affective AI

For centuries, society has been concerned about the prospect of automation: intelligent machines might deskill workers or even render human labor obsolete. Today, however, the most consequential shift concerns intelligent machines entering affective domains once reserved for humans (Zao-Sanders, 2025). Contemporary artificial intelligence (AI) systems increasingly operate not as neutral tools but as affective conversational partners, offering reassurance, mentoring, or even companionship—they emerge as *affective AI*. Managers and employees increasingly rely on them for counsel, conflict navigation, and stress management at work (McKinsey, 2025).

Patterns previously observed in adolescents or clinical contexts (e.g., Chandra Kruse, et al., 2023) now echo emerging dynamics in organizational life, where affective AI is woven into everyday work (Langhof & Güldenberg, 2022). A large-scale study documents a marked rise in business use cases related to emotion, coaching, and companionship (Zao-Sanders, 2025). Yet it remains uncertain how these developments will reshape the business environment. As affective AI engages employees and stakeholders in more interactive and resonant ways, organizations face new dynamics of emotional contagion, trust, and collaboration.

If earlier technologies automated the hand and the mind, affective AI reaches for the heart. By simulating warmth and empathy, these systems blur boundaries between professional and personal, task and feeling, instrumentality and intimacy. They generate new forms of attachment and introduce dynamics that existing organizational theories and governance frameworks are poorly equipped to address. The questions are *(1) What happens when machines do not merely execute tasks but enter the affective domain of organizations? (2) How can we study and govern affective AI in organizations?*

The Asymmetry: Homo Sentiens vs. Affectus Simulatus

At the core of contemporary encounters between people and affective AI lies a fundamental asymmetry in what it means to feel. Humans are *homo sentiens*, embodied beings whose perceptions, emotions, and lived experiences shape interpretation, interaction, and action. Managers and employees do not merely process information; organizational life unfolds through affect as much as through rational calculation. Affective AI, by contrast, operate in the domain of *affectus simulatus*.

Drawing on Jean Baudrillard's Simulacra and Simulation (1981), AI-generated empathy is not authentic *affectus* but a simulation detached from any original emotional experience. What appears as warmth, patience, or reassurance is a performance—a sign of feeling without the feeling itself. Yet simulated affects can have real consequences. Once severed from their referents, signs may reorganize reality rather than merely represent it. Affective AI can reshape the emotional architecture in organizations.

Despite this, management and human resource management (HRM) theory offer limited conceptual grounding for these developments. Affective and relational processes are typically treated as outcomes rather than as mechanisms of organizing. Dominant theories—agency theory, transaction cost theory, and human capital theory—privilege incentives, contracts, and monitoring, while largely neglecting emotion, attachment, and the felt dimensions of coordination and collaboration. As a result, the emotional architecture of organizations remains undertheorized.

The Affective Cost: Individuals and Teams in Organizations

Recent studies suggest that AI chatbots are reshaping how employees seek reassurance and feedback in organizational settings, introducing risks of affective dependency and blurred distinctions between relational support and simulated care (Richet, 2025). Another study reports that AI's tendency toward sycophancy distorts collaboration and undermines critical inquiry by privileging emotional validation over accuracy (Naddaf, 2025). Longitudinal studies further show that individuals experiencing anxiety or depression increasingly form affective attachments to AI systems (Huang et al., 2024).

These risks become critical for organizations as affective AI, being pseudo sentient, assumes roles historically tied to the *homo sentiens*. Their presence fosters affective dissonance, alters relational agency, and deteriorates psychological safety. Anthropomorphized interfaces of affective AI risk generating pseudo-relationships that replace human ones and undermining interpersonal trust. Recognizing this affective cost is essential for understanding how individuals and organizations can learn to self-regulate in environments where *affectus simulatus* becomes part of everyday work.

The cost of affective AI in organizations often emerges first at the individual level before diffusing across organizations. Affective costs are not understood here primarily in financial terms. Rather, they also encompass non-monetary dimensions such as individual well-being, perceived strain, productivity, and broader organizational effects. Workplace chatbots that flatter users or align with their assumptions can foster overconfidence and distort judgment. Over time, such altered dispositions propagate through interaction, reshaping how teams develop trust, assign responsibility, and manage vulnerability.

The Rethinking of Governance: Scope and Timing

As affective AI becomes integrated into everyday organizational life, the limits of existing regulation become clear. Frameworks, such as the EU AI Act primarily address technical and ethical risks through transparency, disclosure, and risk classification. Measures such as age assurance (European Parliament, 2025), user notification, and explanation of the artificial nature of AI companions (California Legislature, 2025) mark first attempts to protect users from false intimacy and affective dissonance. Yet these instruments regulate affective AI from the outside. They cannot, on their own, govern how simulated emotion is enacted, interpreted, and absorbed in daily work.

This gap makes governance essential. Governance refers to the internal practices, norms, and structures that shape how people engage with AI in daily work (Papagiannidis et al., 2025). Distinguishing regulation from governance clarifies that affective costs arise both *a priori* and *a posteriori* of human–AI interaction. Preventive approaches seek to limit problematic affective dynamics upstream, while corrective approaches respond to relational and emotional disruptions once AI becomes embedded in work practices. Importantly, many affective AI interactions unfold outside formally governed organizational systems, such as on personal devices or on general-purpose platforms. This fragmentation limits the reach of formal policies and renders governance an uneven context-dependent process rather than a uniformly enforceable one. Together, these approaches show that the affective cost of AI cannot be managed through compliance alone; it requires governance that actively attends to the emotional architecture in organizations. Table 1 summarizes this logic in a 2×2 framework.

		Scope	
		Regulation (External)	**Governance (Internal)**
Timing	**Preventive (Upstream)**	Rules that limit manipulative affective design, require disclosure of simulated emotional cues, and clarify system limitations.	Organizational design choices that promote emotionally responsible AI: moderating anthropomorphic design, calibrating affective responses, and signaling uncertainty.
	Corrective (Downstream)	Oversight, audits, and enforcement mechanisms addressing unintended emotional or relational harms.	Internal processes to monitor AI reliance, manage AI-mediated breakdowns, and reinforce human-to-human interaction where emotional labor is displaced.

Table 1 Preventive and corrective affective mechanisms across regulation and governance

The Research Agenda: Managing the Affective Cost of AI in Organizations

Affective AI does not arrive in a theoretical vacuum. Information systems research, management science, and human-computer interaction have produced rich accounts of how technologies shape work, coordination, and decision-making. Yet these traditions have largely treated technologies as cognitive or structural forces, not as affective participants. As a result, they offer limited guidance for governing systems that simulate empathy, engage relationally, and shape emotional experience. Advancing governance for affective AI therefore requires challenging and extending existing theoretical foundations across four interconnected levels (Table 2).

Layers	Prevailing Assumptions	Affective AI Shift	Governance Focus
AI Agent Layer	Technologies are neutral, technical tools.	AI behaviors acquire emotional meaning and are read as empathy, attention, or personality.	Design and regulate affective affordances to limit manipulative or misleading emotional signals.
Human Layer	Users are rational; deviations stem from cognitive bias.	Affective AI shapes emotion, attachment, and boundary perception.	Develop self-regulation, AI literacy, and affective awareness in human–AI interaction.
Work Layer	Automation improves efficiency without altering emotional labor.	Affective AI redistributes emotional work and reshapes roles and control.	Redesign work and oversight to manage emotional labor and preserve human judgment.
Relational Layer	Coordination is a human–human process.	AI mediates interaction, feedback, and trust in groups.	Build relational infrastructures that integrate AI while sustaining human connection.

Table 2 Our research agenda

At the ***AI agent layer***, information systems research on affordances, sociomateriality, and human–AI collaboration provides a strong starting point for understanding how

AI shapes organizational action. Governing affective AI requires a shift from functional to affective affordances—how behaviors such as mirroring, adaptation, or monitoring acquire emotional meaning and how they are interpreted. At the ***human layer***, work on human–algorithm interaction has largely focused on cognitive bias. Future research can build on organizational behavior work on emotional regulation and mindfulness to examine how employees learn to discern simulated empathy, manage attachment to AI agents, and maintain psychological boundaries in the face of constant affective feedback—and how to develop these skills.

At the ***work layer***, work design and digital transformation research show how technology reshapes roles, tasks, and control. Organizational change research can help design affective readiness for AI: how should roles be redesigned when traditionally human-centered roles are partially- or entirely- taken over by AI agents? What does "human in the loop" mean when the loop is affective as much as cognitive? Finally, at the ***relational layer***, theories of psychological safety, sensemaking, and coordination must be revisited to account for non-human actors that mediate feedback, trust, and interaction in teams and organizations. Research is needed on how governance structures, routines, and norms can cultivate healthy relational infrastructures that preserve human connection while productively integrating AI mediation.

Combined, these four layers frame a research agenda that addresses both the use of affective AI in organizational life and the design of affective AI systems. Advancing governance in this domain requires integrating design-oriented and use-centered research to ensure that *affectus simulatus* does not quietly reshape the emotional architecture in organizations.

The Epilogue: Intimate Machine, Disturbed Minds

The rise of intimate machines signals a reconfiguration of the emotional architecture in organizations. Interactions once grounded in relations between sentient beings are increasingly mediated by systems that simulate empathy without ever feeling it. Regulating the affective cost of AI is therefore not just a technical or legal challenge but a deeply human one, requiring that we preserve our capacity to feel, to discern, and to remain authentic in the presence of *affectus simulatus*.

References

Baudrillard, J. (1981). Simulacra and simulation. University of Michigan Press.

California Legislature. (2025). Senate Bill No. 243: Companion chatbots. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202520260SB243

Chandra Kruse, L., Bergener, K., Conboy, K., Lundström, J. E., Maedche, A., Sarker, S., Seeber, I., Stein, A., & Tømte, C. E. (2023). Understanding the Digital Companions of Our Future Generation. Communications of the Association for Information Systems, 52, 465-479.

European Parliament. (2025, October 16). New EU measures needed to make online services safer for minors. https://www.europarl.europa.eu/news/en/press-room/20251013IPR30892/new-eu-measures-needed-to-make-online-services-safer-for-minors

Huang, S., Lai, X., Ke, L., Li, Y., Wang, H., Zhao, X., ... & Wang, Y. (2024). AI technology panic—is AI dependence bad for mental health? A cross-lagged panel model and the mediating roles

of motivations for AI use among adolescents. Psychology Research and Behavior Management, 1087-1102.

Langhof, J. G., & Güldenberg, S. (2022). The rise of the robot servant-leaders? Next generation leadership. The International Journal of Servant-Leadership, 16(1), 381–424.

McKinsey (2025). Superagency in the workplace: Empowering people to unlock AI's full potential (AI in the workplace: A report for 2025). https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work

Naddaf, M. (2025). AI chatbots are sycophants – and it's harming science. Nature, 647, 13-14.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. The Journal of Strategic Information Systems, 34(2), 101885.

Richet, J. L. (2025). AI companionship or digital entrapment? Investigating the impact of anthropomorphic AI-based chatbots. Journal of Innovation & Knowledge, 10(6), 100835.

Zao-Sanders, M. (2025, April 9) How People Are Really Using Gen AI in 2025, Harvard Business Review. https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025

Leona Chandra Kruse, Prof. Dr., is a Professor of Information Systems at the University of Agder (UiA), Norway. Her research focuses on how people use information systems for work and leisure and how the technologies should be designed to provide people with support, pleasure, and security. Her research has been recognized with several awards, including the Young Researcher Award 2024 from Agder Academy of Sciences and Letters and the AIS Senior Scholar's best Information Systems publication in 2019 and 2025.

Address: University of Agder, Faculty of Social Sciences, Department of Information Systems, Gimlemoen 25, 4604 Kristiansand, Norway, Phone: +4738 14 18 27,
E-Mail: leona.chandra@uia.no
ORCID: https://orcid.org/0000-0002-9001-3870

Patrick Mikalef, Prof. Dr., is a Professor of Data Science and Information Systems at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway. His research examines digital transformation, artificial intelligence, data-driven decision-making, and responsible AI governance in organizations. He has published extensively in leading journals in information systems and management and is actively involved in research initiatives at the intersection of AI, strategy, and governance.

Address: Norwegian University of Science and Technology (NTNU), Department of Computer Science, Sem Sælands vei 7–9, 7034 Trondheim, Norway, Phone: +47 73558995,
E-Mail: patrick.mikalef@ntnu.no
ORCID: https://orcid.org/0000-0002-6788-2277