

The Chat Trap

Rethinking the Interactional Default in Organizational AI

Martin H. Möhle · Frédéric Thiesse

University of Würzburg, Faculty of Management and Economics, Department of Management, Chair of Business Information Systems Engineering, Sanderring 2, 97070 Würzburg, Germany



MOTIVATION The Limits of AI Usage

Generative AI now reaches knowledge workers almost exclusively through one interaction form: the chat window. Across three rigorous study designs, capable models systematically reduced the cognitive engagement workers bring to their tasks.

PERFORMANCE -19 pp <i>out-of-frontier tasks · Dell'Acqua et al. 2023</i>	KNOWLEDGE RETENTION -17 % <i>controlled exp. · Shen & Tamkin 2026</i>	CRITICAL THINKING ↓ <i>self-reported · Lee et al., CHI 2025</i>
---	--	--

These results share a common cause, and it is **not capability**: the models pass the benchmarks. The cause is **genre**. Following Yates & Orlikowski (1992), communicative genres stabilize patterns of social action and leave their norms unexamined. Omizo & Hart-Davidson (2024) name the corollary for AI, **genre signaling**: surface conversational competence without epistemic accountability. The chat genre imports *conversational* norms (turn-taking, repair, charitable acceptance) into work that demands *interrogation*: verification, citation, audit trail.

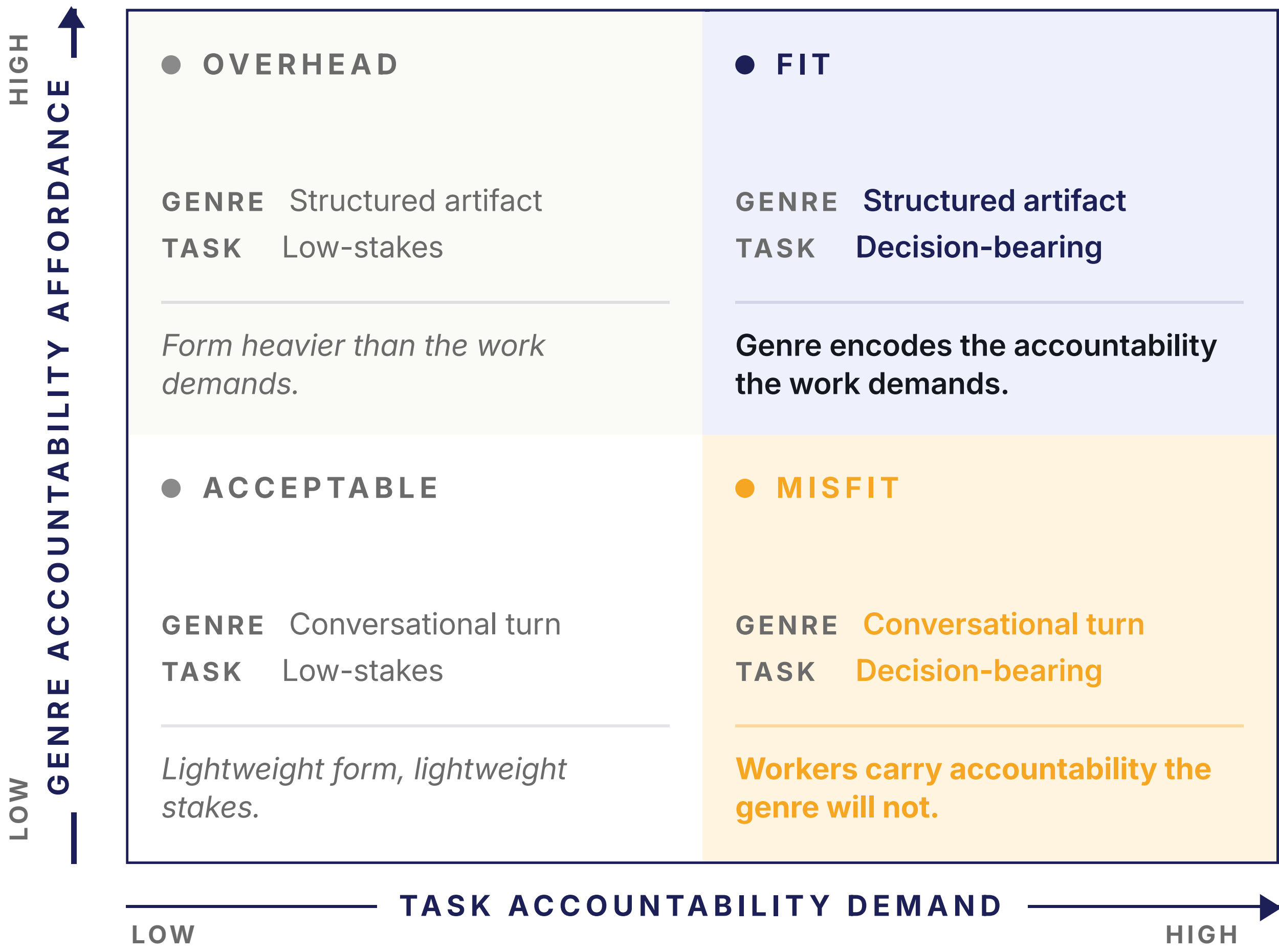
THESIS

Conversation is not a genre of interrogation. Routing complex, accountable work through a chat window invites workers to accept rather than verify. The bottleneck is the interface, not the model.

FRAMEWORK Genre-Technology Fit

We introduce **genre-technology fit**: the degree to which an interaction genre's structural properties match a task's coordination and accountability demands. Where **task-technology fit** asks whether a technology can *perform* a task, GTF asks whether the interaction form matches how work must be *interrogated* and *accounted for*.

Two observable dimensions. Norm alignment: do the genre's norms support critical scrutiny of outputs? Accountability alignment: does the artifact form capture the decision process?



EVIDENCE What the Data Show

Three independent study designs (field experiment, survey, controlled lab) converge on the same result. The mechanisms each team proposes describe conversational-genre import from three angles.

STUDY	DESIGN	EFFECT	MECHANISM PROPOSED
Dell'Acqua et al. 2023 · SSRN	758 BCG consultants; tasks on/off the GPT-4 capability frontier.	-19 pp	Conversational fluency masks unreliability off the frontier; workers cannot detect when the model is wrong.
Lee et al. 2025 · CHI	Knowledge-worker survey on AI-assisted cognitive effort.	↓ critical thinking	High confidence in AI displaces verification effort ; users delegate cognition instead of interrogating output.
Shen & Tamkin 2026 · arXiv	Controlled experiment; chat-assisted condition vs. unaided baseline.	-17 %	Passive consumption of finished chat answers weakens encoding and downstream recall.

Misfit signals in practice

Workers copying chat outputs into Word, maintaining parallel chat-log repositories, or annotating logs by hand are not idiosyncratic, they are diagnostic signals of accountability misfit. Embedding a chat panel in an existing tool (e.g. Copilot in Word) does not resolve the problem: the container changes, the genre does not.

IMPLICATIONS From Default to Repertoire

The fix is not to abandon conversational AI, which accelerates low-stakes work. It is to stop treating conversation as the genre-blind *default* and to cultivate an **AI genre repertoire** (extending Orlikowski & Yates, 1994): interaction forms matched to the work and governed by genre-technology fit.

Audit · Compliance	AI-integrated documents with traceable citations.
Strategic Planning	Spatial canvases over linear turns.
Routine Operations	Embedded task agents.
Exploratory Ideation	Chat is acceptable; low task demand.

Research challenges for IS

- Treat workarounds as **misfit signals**, not user idiosyncrasies.
- Connect genre choice to **governance & accountability** contexts.
- Map genres empirically to **classes of organizational work**.

Open questions for the TREO round table

- 1 **Operationalization**: survey, behavioral trace, or vignette experiment?
- 2 **Dimensionality**: two dimensions enough? Continuous or typological?
- 3 **Empirical first step**: experiment, field study, or pattern observation?
- 4 **Theoretical anchor**: genre theory, media synchronicity, or structuration?
- 5 **Boundary conditions**: when does chat fit (task, step, or workflow level)?
- 6 **Program scope**: one integrated paper or a three-paper program?