Research Design Paper (Paper ID 197)

Submitted to the Innovation & Product Development Management Conference 2026

Human–AI Collaboration for Organisational Innovation in Times of Unrest

Rebecca Liu and Stefanos Mouzas

Lancaster University Management School, UK

1. Introduction and Motivation

Organisations today operate in a landscape of compounding disruptions—economic volatility, geopolitical instability, rapid technological change, and evolving social expectations. In this context, artificial intelligence (AI) is no longer a peripheral tool but a constitutive element of organisational decision-making, governance, and innovation. Yet the dominant lens through which AI integration is studied remains largely technical: researchers focus on efficiency gains, bias mitigation, and compliance with ethical frameworks (Rahwan et al., 2019; Beer, 2017; Batool et al., 2023). Far less attention is paid to how human-AI collaboration reshapes organisational behaviour, redistributes agency, and generates new forms of organisational innovation in volatile environments.

This paper presents the design of a research project that addresses this gap directly. It examines how human-AI collaboration fosters organisational innovation - understood as meaningful changes in structures, practices, governance, and collaboration patterns, and how it reshapes autonomy, responsibility, and legitimacy by redistributing agency between human and algorithmic actors. The project is situated in a moment of particular urgency: as organisations turn to AI to stabilise decision-making under pressure, the implications for authority, employee experience, and institutional legitimacy remain poorly understood.

The paper proceeds as follows. Section 2 maps the relevant literature and positions the study theoretically. Section 3 presents the research model and questions. Section 4 outlines the methodology. Section 5 discusses expected contributions and practical implications. Section 6

reflects on current developmental challenges on which conference feedback is particularly sought.

2. Literature Mapping and Theoretical Positioning

2.1 What We Know—and What Remains Unexamined

Existing scholarship has established AI's value in decision support, workflow optimisation, and algorithmic management (Elish & Boyd, 2018; Foster & Heeks, 2020; Rahman, 2023). Studies document how AI systems process data at scale, reduce cognitive load, and accelerate routine decisions. In innovation contexts, AI-enabled tools are shown to support ideation, market sensing, and iterative design processes.

However, empirical work on human–AI collaboration as a driver of organisational innovation remains fragmented. Three limitations are notable. First, employees are largely treated as passive recipients of AI-mediated decisions rather than active co-producers whose critical engagement shapes innovation trajectories. Second, much research is conducted in stable organisational settings, leaving open the question of how AI integration functions, and what it produces, when organisations are under stress. Third, the relational dynamics between AI systems and authority structures, particularly how AI redistributes legitimate power and shapes accountability, remain undertheorized.

2.2 Theoretical Framework: Entitlement Theory and the Capability Approach

To address these gaps, this study draws on two complementary normative frameworks. Amartya Sen's Entitlement Theory conceptualises access to resources and decision-making authority not as fixed endowments but as bundles of legitimate claims shaped by institutional arrangements (Sen, 1982). Applied to AI integration, this framework illuminates how algorithmic systems alter who has the legitimate standing to act, decide, and be held accountable, and how those shifts interact with existing organisational hierarchies.

Martha Nussbaum's Capability Approach shifts the analytical lens to individuals' real opportunities to achieve meaningful goals (Nussbaum, 2005). Where Sen's framework explains structural entitlements, Nussbaum's illuminates whether employees are genuinely able to exercise discretion, engage critically with AI outputs, and contribute to innovation in ways that reflect their capabilities. Together, these frameworks form an *Entitlement-Capability* lens that

is both normatively grounded and empirically tractable, capable of explaining how AI reshapes authority while attending to the lived experiences of those working alongside it.

This theoretical positioning bridges human-centred normative theory with organisational behaviour research. It moves beyond compliance-oriented ethics discourse to ask how AI integration substantively changes what people can do, what they are held responsible for, and how innovation emerges from these reconfigurations.

3. Research Model and Questions

The project operationalises the Entitlement–Capability framework across three analytical dimensions of AI integration in organisations:

1. **AI and authority in organisational innovation**: How do AI systems alter decision-making authority, accountability structures, and governance arrangements? Under what conditions do these changes enable versus constrain innovation?
2. **Employee perceptions and interactions**: How do employees perceive their agency when working alongside AI? How do role identity, prior experience, and organisational context shape critical engagement with AI-mediated processes?
3. **Critical engagement with AI in organisational innovation**: When and how do employees challenge, negotiate, or override AI recommendations? What are the organisational conditions that enable such engagement, and how does it contribute to innovative outcomes?

The core research question asks: *How does human–AI collaboration reshape organisational authority, responsibility, and legitimacy, and through which mechanisms does it drive adaptive organisational innovation in conditions of unrest?*

Six sub-questions operationalise this inquiry. They explore: (i) how AI systems alter behavioural norms and accountability expectations; (ii) how employees perceive their autonomy and discretion in algorithmically mediated work; (iii) how role identity and organisational context moderate engagement with AI recommendations; (iv) what forms of critical engagement emerge, and under what enabling conditions; (v) how these dynamics differ across periods of stability versus active organisational stress; and (vi) how governance structures adapt (or fail to adapt) in response to human–AI agency redistribution.

4. Methodology

4.1 Research Design

The project adopts a mixed-methods design, structured around the principle of methodological convergence: qualitative, quantitative, and behavioural data are collected in parallel and integrated at the analysis stage, enabling triangulation across levels of organisational experience. This design is appropriate given the complexity of the phenomena under investigation, which span structural, perceptual, and behavioural dimensions.

Fieldwork is conducted in UK-based organisations that have undergone significant AI integration in the past three years and have experienced periods of operational stress - whether from economic disruption, restructuring, or sector-level volatility. Site selection follows a purposive logic to ensure variation across industry type, organisational size, and AI maturity.

4.2 Data Collection Methods

Ethnographic observation provides the longitudinal backbone of the study. Researchers will spend extended periods embedded in participating organisations, attending meetings, following decision workflows, and documenting how human–AI interactions unfold in real time. This method is particularly well-suited for capturing the informal negotiation of authority and the situational dynamics of critical engagement.

Semi-structured interviews with employees at multiple organisational levels will elicit accounts of perceived agency, responsibility, and innovation-related experience. Think-aloud protocols, in which participants verbalise their reasoning while interacting with AI tools, will provide direct access to cognitive processes underlying decision acceptance or rejection. These qualitative streams will be thematically coded using an iterative approach, with codes derived from both the Entitlement-Capability framework and empirically emergent patterns.

A structured survey will be administered to a broader sample within participating organisations. The survey will assess perceived autonomy, trust in AI systems, innovation-related behaviours, and evaluative engagement, enabling generalisability testing of qualitatively derived propositions.

Finally, behavioural measures, including eye-tracking and clickstream analysis during AI-tool interactions, will provide non-self-report indicators of attention, deliberation, and critical

engagement. These measures address limitations of interview and survey data, which are susceptible to social desirability effects when respondents discuss AI use.

4.3 Analysis Strategy

Convergent mixed-methods analysis will integrate findings across data streams. Qualitative analysis will generate theoretical propositions about mechanisms linking AI integration, authority redistribution, and organisational innovation. Quantitative and behavioural data will test the prevalence and predictors of these mechanisms across the sample. Integration at the meta-inference stage will enable a holistic account of AI's influence across structural and experiential levels, attending both to what organisations do and what individuals experience.

5. Expected Contributions and Practical Implications

Conceptually, this study reframes AI as a constitutive organisational actor rather than an instrumental tool. By treating AI systems as entities that actively reshape authority, accountability, and agency, it advances a relational understanding of technology-in-practice. The Entitlement-Capability framework offers a replicable theoretical lens for future studies of AI integration in institutional settings.

Empirically, the project generates new insight into how employees co-produce decisions in algorithmically mediated environments, negotiate responsibility, and develop innovative practices under uncertainty. It provides one of the first systematic examinations of human–AI collaboration dynamics during periods of active organisational unrest.

Practically, the findings will inform managerial guidance on human-centred AI adoption: how to design AI integration processes that preserve meaningful discretion, foster critical engagement, and maintain institutional legitimacy. Policy implications include recommendations for governance frameworks that ensure accountability is not inadvertently dissolved as AI assumes greater decision-making roles.

6. Developmental Reflections and Questions for Discussion

As a research design paper, this contribution invites engagement with aspects of the project still in active development. Three areas are particularly open to discussion.

First, the operationalisation of 'organisational innovation' in this context requires further refinement. The study conceptualises innovation as changes in structures, practices,

governance, and collaboration. Yet, the boundaries of this definition, and how to distinguish AI-enabled innovation from AI-enforced standardisation, remain analytically challenging.

Second, the integration of behavioural measures (eye-tracking, clickstream) with ethnographic and interview data raises methodological questions about how to make sense of discrepancies across these registers. Feedback on convergence strategies and interpretive protocols would be welcome.

Third, the transferability of findings across organisational contexts is an open question. The UK-based fieldwork will generate contextually rich data, but the extent to which findings apply across different national institutional environments, regulatory contexts, and AI maturity levels warrants consideration.

The authors welcome comments on the theoretical framework, methodological integration, and the framing of 'unrest' as an analytical condition, as well as any suggestions for comparable empirical work that might inform the study design.

REFERENCES

Batool, B., et al. (2023). Ethical AI in organisations: Compliance, bias, and governance. Journal of Business Ethics, 182(3), 601–618.

Beer, D. (2017). The social power of algorithms. Information, Communication & Society, 20(7), 985–997.

Elish, M. C., & Boyd, D. (2018). Situating methods in the magic of big data and artificial intelligence. Communication Monographs, 85(1), 57–80.

Foster, C., & Heeks, R. (2020). Conceptualising inclusive innovation: Modifying systems of innovation frameworks to understand diffusion of new technology to low-income consumers. European Journal of Development Research, 25, 333–355.

Nussbaum, M. C. (2005). Women and human development: The capabilities approach. Cambridge University Press.

Rahwan, I., et al. (2019). Machine behaviour. Nature, 568(7753), 477–486.

Rahman, H. A. (2023). The invisible cage: Workers' reactivity to AI-generated recommendations. Administrative Science Quarterly, 68(1), 1–43.

Sen, A. (1982). Poverty and famines: An essay on entitlement and deprivation. Oxford University Press.