This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

AI agents as users: Managing innovation in an agentic era

Kamila Zahradnickova*

Faculty of Business Administration, Prague University of Economics and Business, nám. W. Churchilla 4, 130 67 Praha 3, Czech Republic

E-mail: kamila.zahradnickova@vse.cz

* Corresponding author

Adam StJohn Lawrence

IE Business School, Calle de María de Molina, 31, Chamartín, 28006 Madrid, Spain

Ladislava Zbiejczuk Sucha

Faculty of Arts, Masaryk University, Arna Nováka 1, 602 00 Brno, Czech Republic

Ferhat Demir

Faculty of Business Administration, Prague University of Economics and Business, nám. W. Churchilla 4, 130 67 Praha 3, Czech Republic

Abstract: As AI agents gain autonomy, organisations face a new design and management challenge: how to redesign services when the user is no longer exclusively human. This paper presents a fourteen-week action research study at a deep-tech startup, applying the Agentic Experience (AX) framework to transform a human-first platform into a dual-user service in which both human clients and AI agents can initiate, execute, and deliver research autonomously. Four findings emerge for innovation management: dual-user redesign requires rethinking service logic, not just the interface layer; agent usability is a problem of probability management; agents must enter stakeholder maps from the outset; and a structural gap between engineers who know how to build and managers who know what to build requires dedicated coordination artefacts to bridge. The AX framework, organised around accessibility, usability, and cooperation, provided the shared vocabulary that made this possible.

Keywords: agentic experience; AX framework; AI agent; innovation management; dual-user design; human-agent interaction; service innovation; action research; digital services; LLM.

1 Introduction

AI agents are becoming users. Once confined to backend infrastructure, large language model (LLM)-based agents can now navigate interfaces, interpret goals expressed in natural language, and execute multi-step tasks across digital services with increasing autonomy (Lu et al., 2024; Anthropic, 2023). For most organisations, this transition is first visible as a technical opportunity: agents can automate workflows that previously required sustained human attention (Haefner et al., 2021; Mariani et al., 2022). But the deeper implication is organisational. When agents participate in services not as tools but as users, entities with their own operational requirements, failure modes, and coordination needs, the assumptions underlying service innovation begin to break down.

Human-centered design (Norman, 2013; Nielsen, 1993), the dominant paradigm for digital service development, assumes a human user whose perceptual, cognitive, and behavioural needs guide every design decision. Innovation management frameworks for digital services share this assumption: services are designed for humans, tested with humans, and improved based on human feedback (Lusch and Nambisan, 2015). When AI agents enter the picture as co-users, this assumption is no longer sufficient. Services must simultaneously serve human clients and accommodate the structured access requirements, non-deterministic behaviour, and cooperation needs of autonomous agents. This is what we call the dual-user challenge.

This challenge is already arriving in practice. Platforms originally built for human professionals are being extended to support agents that can act on behalf of those professionals, often with greater speed and breadth than the humans themselves (Gartner, 2025; BCG, 2025). Yet the management literature on AI in innovation focuses predominantly on AI as an instrument of ideation, prediction, or process acceleration (Haefner et al., 2021; Cillo and Rubera, 2024), rather than AI as a participant in service use. How innovation teams should actually manage the redesign of a human-first service for dual-user operation remains largely unexplored empirically.

This paper addresses that gap through an action research study conducted at a Czech market research startup that builds AI-simulated consumer panels for enterprise clients. During a fourteen-week redesign project, the company applied the Agentic Experience (AX) framework to transform the company's human-first platform into a dual-user service. The central research question is: How can an existing human-first digital service be redesigned to support dual-user interaction between human clients and AI agents, and what organisational and management challenges arise in doing so? To address this challenge, we propose and apply the AX framework: a dual-user design approach organised around three minimum requirements: accessibility, usability, and cooperation.

The paper is organised as follows: Section 2 establishes the theoretical foundations. Section 3 describes the action research design. Sections 4 and 5 report the five cycles and their cross-cutting findings. Section 6 discusses implications for innovation management theory and practice.

2 Theoretical background

2.1 The organisational consequences of AI agents as users

Recent advances in LLM-based systems have produced agents that complete complex tasks by interpreting goals expressed in natural language, selecting strategies, and executing sequences of actions across digital environments (Achiam et al., 2023; Durante et al., 2024). Unlike classical automation, these agents are non-deterministic: the same input may produce different outputs across runs, and agents may generate strategies not explicitly anticipated by their developers (Atil et al., 2025). This characteristic, human-like variability in a machine, is what makes agentic systems both powerful and organisationally novel.

For innovation management, the shift from AI-as-tool to AI-as-user carries structural consequences. AI-as-tool frameworks assume that human innovators use AI to improve the quality or efficiency of their work (Haefner et al., 2021; Gao et al., 2025). When agents become users of a service, the organisational question changes: it is no longer how to integrate AI into the innovation team, but how to redesign the service itself so that two fundamentally different kinds of users, human and agent, can participate coherently. This demands reconfiguration of innovation capabilities, governance structures, and delegation logic (Bradshaw et al., 2017; Shneiderman, 2022).

A structural accessibility gap makes this challenge concrete. Digital services are built for human perception and motor control: visual layouts, click sequences, modal confirmations. These features are at best neutral and at worst obstructive for agents, which operate most effectively through structured APIs, machine-readable state signals, and explicit semantic cues (Lu et al., 2024; Chen et al., 2024). When agents attempt to use human-first interfaces, the result may be latency, high operational cost, and unreliable task completion, not because the agent is incapable, but because the service was never designed for it.

This challenge cannot be resolved by treating agent integration as a standard machine-to-machine engineering problem. Classical machine-to-machine communication was deterministic: fixed protocols exchanging predictable data streams between fully specified systems (Bradshaw et al., 2017). Agents break this assumption in two ways. First, they are non-deterministic, the same request produces variable outputs (Atil et al., 2025). Second, agents operate in environments designed around human intent, language, and workflow logic (Achiam et al., 2023; Durante et al., 2024), which means participation requires interpretation, not just transmission. A structured API solves the access problem but not the legibility problem (Lu et al., 2024; Chen et al., 2024). This is the gap that a dual-user design framework must address (Shneiderman, 2022; Borghoff et al., 2025).

2.2 The Agentic Experience (AX) framework

Existing models of human-agent interaction remain largely descriptive, offering taxonomies of roles and behaviours but little prescriptive guidance for practitioners redesigning real services (Lu et al., 2024; Battistoni et al., 2023; Chen et al., 2024). The proposed AX framework addresses the gap between descriptive models and the prescriptive guidance that innovation teams need. AX proposes a dual-user paradigm in which both humans and AI agents are treated as users with legitimate and distinct operational needs. Three pillars define the minimum requirements for any dual-user system:

- **Accessibility:** the degree to which both human and agent users can reach and interpret the functions necessary to complete a task. For agents, this requires structured API or MCP endpoints, consistent semantic labelling, and machine-readable state signals that do not depend on visual layout.
- **Usability:** the extent to which both user types can perform tasks effectively and efficiently. For agents, usability cannot be achieved through perceptual ergonomics;

This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

it depends on workflow clarity, action granularity, and explicit handling of edge cases, because agent task completion is probabilistic rather than deterministic.

- **Cooperation:** the quality of coordination between humans, agents, and other stakeholders to achieve shared goals. This encompasses human-in-the-loop design, transparency mechanisms enabling human oversight of agent behaviour, and eventually multi-agent coordination.

Each pillar is operationalised through a three-level scoring rubric (Poor / Moderate / Optimal) with observable evaluation indicators, enabling criterion-referenced assessment of dual-user system performance (Table 1).

Table 1 AX (Agentic Experience) scoring rubric: Evaluation criteria per pillar.

Pillar	Poor	Moderate	Optimal	Key Indicators
Accessibility	Agent has no access; cannot perform crucial actions	Agent accesses some features; limited scope hinders operation	Agent accesses and navigates all core features, tailored to its capabilities	Core actions exposed via API or MCP; consistent semantic structure; machine-readable interface states; no blocking modals
Usability	Agent frequently fails or generates errors	Agent completes tasks inconsistently or less efficiently than a human	Agent completes tasks with human-equivalent or superior efficiency	Linear goal-based workflows with explicit success/error states; structured inputs/outputs; redundant human-only steps bypassable; minimised latency and token usage
Cooperation	Agent cannot communicate with other agents or humans	Agent communicates with some stakeholders but is hindered	Agent communicates with all key stakeholders: human(s), internal and external agents	Shared activity logs; human-readable agent reasoning; human override/delegation mechanisms; context-aware collaboration; notification and feedback loops

The three pillars are not arbitrary. Each has roots in human-centred design, but their meaning shifts when applied to dual-user systems. Accessibility originates in inclusive design (Clarkson et al., 2013). For agents, the relevant impairment is structural: agents cannot perceive visual affordances, interpret implicit conventions, or recover from ambiguous states the way humans do. Usability, rooted in Nielsen's (1993) heuristics, shifts from an ergonomic to a probabilistic concern when applied to LLM-based agents (Atil et al., 2025; Xu et al., 2025). Cooperation extends the individual interaction frame to the systemic level (Hopf et al., 2024; Schmutz et al., 2024; Shneiderman, 2022).

The three pillars are also interdependent in a specific sequence. Accessibility is a precondition for usability: an agent that cannot reach a function cannot perform it. Usability is a precondition for cooperation: an agent that cannot complete tasks reliably cannot be trusted as a collaborative actor. Together they form a minimal sufficiency condition — a service can claim dual-user capability only when all three are achieved at some level.

3 Method

3.1 Research design

Action research integrates systematic inquiry with deliberate organisational intervention, generating knowledge through iterative cycles of diagnosing, planning, acting, observing, and reflecting (Lewin, 1946; Susman and Evered, 1978). It is particularly suited to contexts where the phenomenon of interest is emergent, the researcher has organisational authority over the intervention, and the aim is both practical improvement and generalisable knowledge (Coghlan and Brannick, 2014).

The study was conducted at a startup between July and October 2025. The startup is a Czech company with three years on the market and over 20 team members producing AI-generated research data for enterprise market research clients. As of the study period, the platform served more than 20 enterprise clients and processed over 20 million synthetic survey responses weekly. The platform offered two primary workflows accessible to clients through a graphical interface: quantitative surveys and qualitative interviews with AI personas. No agentic access existed at the project outset: all workflows required human navigation from start to finish.

The redesign project was led by the researcher in her dual role as designer and CEO, in collaboration with three cross-functional teams: frontend development (3 members), backend engineering (4 members), and AI/data science (3 members). The company operates as a remote-first organisation; most collaboration occurred through regular stand-ups and online co-creation sessions, with one three-day in-person retreat midway through the project.

3.2 Data collection

Data collection was embedded in the design process itself, following a hybrid inductive-deductive approach: deductively mapping observations to the predefined AX pillars, while inductively identifying emergent themes about organisational coordination, capability gaps, and delegation logic.

Data sources included:

- Observational field notes from workshops, co-creation sessions, and weekly team stand-ups across 14 weeks.
- Design artefacts: user stories, stakeholder maps, customer journey maps, high-fidelity Figma prototypes, technical briefs, and dual-user workflow diagrams.
- Structured testing sessions: an internal "Investigative Rehearsal" workshop and external user testing with enterprise clients, documented in a prioritised improvement spreadsheet.
- Team retrospectives at the close of each weekly development sprint, in which AX scores were assigned through consensus discussion against predefined pillar criteria.
- An "Unboxing Vlog" in which the researcher reviewed and narrated the first implementation draft, creating a documented record of design-implementation gaps.

This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

AX scores (Poor / Moderate / Optimal) were assigned through structured retrospective discussions in which team leads assessed current system behaviour against pre-defined, observable criteria for each pillar level (see Table 1). All participants provided informed consent; data collection adhered to client confidentiality policies.

3.3 Action research cycles

The study followed five sequential action research cycles, each corresponding to a phase of the redesign process. Table 2 provides an overview.

Table 2 Action research cycles in AX redesign.

Cycle	Phase	Weeks	Key Activities	Primary Outputs
1	Audit & Diagnose	27	Service mapping, intent/action decomposition, initial AX evaluation	User stories, stakeholder maps, AX baseline (all: Poor)
2	Idea Generation & Plan	28–29	HMW workshops, lo-fi prototyping, human-in-the-loop decisions, agentic stakeholder mapping	Idea portfolio, dual-user service blueprints
3	Technical Brief & Act	30	Hi-fi Figma prototypes, Loom explainer, feature list handover to engineering teams	Technical brief, annotated wireframes
4	Development & Observe	31–39	Agile sprints, Unboxing Vlog, Investigative Rehearsal, external client testing	Working product, fix timeline, client feedback spreadsheet
5	Evaluation & Reflect	40	AX retrospective evaluation, team reflection, identification of outstanding gaps	Final AX scores, cross-cycle findings, future roadmap

3.3.1 Audit

The first cycle established a precise diagnosis of the existing service. The platform offered clients two research workflows, both navigated entirely through a graphical interface. In the quantitative survey workflow, a client uploaded a research brief, selected a target segment, reviewed auto-generated questions, and awaited system-generated results. In the qualitative interview workflow, a client asked questions in natural language to AI personas and received structured responses with sentiment analysis. Both workflows ended at data delivery; the cognitive work of sense-making fell entirely to the human client.

Each workflow was deconstructed into intents (the underlying goal a user aimed to achieve) and actions (the discrete steps required to achieve it). Applying the AX evaluation to this baseline was unambiguous: the system received Poor scores on all three pillars. No API or MCP endpoints existed; no task could be executed autonomously; no cooperation mechanism between human and agent was possible. This established the scope of the redesign: not a UI upgrade, but a fundamental extension of who the service was designed for.

3.3.2 Ideate

The second cycle translated the diagnostic findings into a design direction. The central HMW challenge was deliberately agent-constrained: "How might agents help users conduct higher-quality research while reducing manual effort for clients and team?" Fixing "agents" as the actor kept all concepts anchored to the business direction while opening exploration of different forms and levels of agent involvement. The researcher first mapped agents onto the existing stakeholder diagrams, identifying where they could substitute for or augment human activity at each research intent.

One concept explored an "invisible UI": a human client submits a business goal in natural language, and the agent runs the entire research process backstage, from clarifying the brief through to report generation, without requiring any further human input. All five AX redesign steps (Table 3) were completed by the end of this cycle, yielding service blueprints that specified human-in-the-loop levels per intent, expanded stakeholder maps including agent actors, and dual-user journey maps for engineering handover.

Table 3 Five steps of AX redesign.

Step	Pillar	Outcome
Distil user intents	Accessibility & usability	User stories, technical brief
Distil actions per intent	Accessibility & usability	User journeys, wireframes
Decide human-in-the-loop level	Cooperation	Service blueprint
Suggest agentic stakeholders	Cooperation & accessibility	Stakeholder map
Define dual-user workflows	Usability & cooperation	Dual-user journey, wireframe

3.3.3 Technical brief

The third cycle operationalised the design direction into a formal engineering brief. The researcher handed over high-fidelity Figma prototypes, annotated with usage notes and accompanied by a Loom video explainer, plus a structured list of intents with their corresponding actions, required API or MCP endpoints, and specified cooperation mode for each workflow segment. A notable organisational dynamic emerged during this cycle: once engineering teams began translating dual-user workflows into state diagrams and API specifications, the designer's role inverted. Engineers became the structural authors; the designer moved to the position of domain expert and contextual clarifier. This role inversion would become a recurrent theme through the development cycle.

3.3.4 Development

The fourth and longest cycle covered twelve weeks of agile development, internal stress-testing, and external client validation. Three structured observation formats shaped this phase.

- Unboxing Vlog: The researcher reviewed and narrated the first working implementation of the agentic capabilities, providing timestamped commentary on gaps between design intent and built behaviour.
- Investigative Rehearsal: During the three-day in-person retreat, the team ran a structured role-play testing session in which each participant received a task and a user persona to embody. The session surfaced multiple bugs and usability gaps, organised into a prioritised fix timeline.
- External Client Testing: Enterprise clients performed key tasks under observation. Findings were documented in a prioritised wishlist spreadsheet. A key finding was that unformatted agent text in the chat interface created confusion about task completion, prompting a targeted redesign of the response display format.

3.3.5 Evaluation

The fifth cycle formally evaluated the redesigned system against the AX framework through a structured team retrospective in which each pillar was scored against pre-defined criteria. Table 4 summarises the before-and-after results.

Table 4 AX evaluation before and after redesign.

Pillar	Score Before	Score After	Change	Key Evidence
Accessibility	Poor	Moderate	+1	API and MCP endpoints added for core research intents; filters and external agent integrations remain inaccessible
Usability	Poor	Optimal	+2	Agent completes all supported research intents end-to-end with human-equivalent reliability once workflow decomposition reached sufficient granularity
Cooperation	Poor	Moderate	+1	Shared activity logs, human override mechanisms, and redesigned chat interface implemented; agent-to-agent coordination not yet achieved

The usability improvement was the most striking outcome: once workflow decomposition achieved sufficient granularity and explicit edge-case handling, the agent completed supported tasks with human-equivalent reliability across repeated runs. Accessibility improvements were real but partial, limited by time constraints on peripheral feature endpoints. Cooperation improvements were anchored in client-tested interaction changes but fell short of the optimal threshold because multi-agent coordination was not implemented within the project timeline.

4 Findings

Across the five cycles, four findings emerged with implications for how innovation teams should understand and manage dual-user service redesign.

4.1 Dual-user redesign changes service logic beyond the interface layer

The most consequential finding of the study is the scope of what changed when the service was redesigned for dual users. Before redesign, the platform supported a workflow in which human effort was required at both ends of the research process. A client prepared a research brief, selected a target sample, reviewed and adjusted questions, and submitted the task. The platform executed the middle portion automatically. But the client was required to return to interpret the data, draw conclusions, formulate recommendations, and decide what to research next. The platform delivered data; the client delivered insight.

After redesign, the agent can accept a business goal expressed in natural language (e.g., "I want to understand how Czech car owners think about electric vehicle adoption"), ask clarifying questions, translate the business goal into a research goal with appropriate survey questions, suggest and configure sampling criteria, conduct the research across all AI respondents, gather and analyse the responses, generate key findings, and synthesise them into a report with strategic guidance. A human client's operational input is reduced to submitting an intent; the agent manages everything between intent and insight.

This transformation reveals a key principle for innovation management: dual-user redesign cannot be accomplished by adding agent-accessible endpoints to an existing human interface. It requires rethinking which parts of the service logic should be exposed, at what level of granularity, and with what governance over agent autonomy. These are not engineering decisions; they are strategic service design decisions with direct implications for the organisation's value proposition, client relationships, and quality governance.

4.2 Usability is a problem of probability management

In human-centred design, usability is a threshold property: a well-designed interface either works reliably for its intended users or it does not. For LLM-based agents, usability is a continuous distribution. Every change to workflow structure, intent decomposition, action granularity, or edge-case specification shifts the probability that the agent completes a given task correctly.

In team retrospectives, engineers described the challenge in terms of recipe quality: the more precisely intents were decomposed into explicit actions, the more clearly actions were bounded by success and failure states, and the more thoroughly edge cases were pre-specified, the higher the agent's task success rate became. The AX framework, which structures workflow decomposition as a formal step before any engineering begins, proved directly functional here.

This finding has a direct implication for how dual-user service quality should be managed. Traditional UX metrics, task completion rate, error rate, user satisfaction — are insufficient for dual-user systems where agent performance varies stochastically. Innovation teams managing such systems need to add probabilistic performance metrics: task success distributions across repeated runs, error-type frequency by workflow step, and variance across different user inputs.

4.3 Agents belong to stakeholder maps

Introducing agents into the stakeholder map in the audit phase, before any engineering work began, changed the design conversation in ways that proved consequential for the entire project. Once agents appeared as named actors in the same artefacts used to communicate with clients and engineers, the team was forced to assign agents goals, limitations, and coordination responsibilities alongside human actors. The abstraction of "make the platform agent-accessible" became a set of specific design questions: What can this agent access? What decisions can it make autonomously? At what point must it be handed back to the human?

The framing of agents as entities with both specific limitations (sensitivity to ambiguity, bounded context windows, no embodiment) and superhuman capabilities (speed, memory, parallel processing) proved especially productive. For innovation managers, this suggests that stakeholder mapping can be productively extended to include agents with little additional methodological overhead, and that doing so early changes the quality of subsequent governance discussions.

4.4 The HOW/WHAT gap: design artefacts as coordination infrastructure

The most organisationally generalisable finding of the study concerns the knowledge gap that emerged at the intersection of design and engineering. Engineering team members understood how to build systems capable of agentic operation. What they lacked was guidance on what to build: which intents to expose, which cooperation patterns to support, where human oversight was essential and where full delegation was appropriate.

On the management and design side, the situation was a mirror image. The researcher and client-facing team members had a clear goal, make the platform accessible and useful for agents, but initially lacked the vocabulary and conceptual models to specify that goal precisely. What could an agent actually perceive, retain, and act on? What did context window constraints mean for workflow design? These questions were not answerable from a design or management perspective without technical input.

This HOW/WHAT gap — engineers possessing implementation capability without specification, managers possessing strategic intent without technical grounding — is likely to be a recurring feature of dual-user innovation across organisational settings. It is not resolved by adding one technically literate person to the design team, or by training managers in API design. It requires a structured shared vocabulary and a collaborative artefact — in this case, the dual-user workflow specification — that makes both sides' knowledge mutually legible.

5 Discussion

This study contributes to innovation management scholarship in three respects. First, it introduces dual-user service innovation as a distinctive organisational challenge. Existing frameworks treat AI as an instrument of innovation (Haefner et al., 2021), a driver of process efficiency (Gao et al., 2025) or a component of hybrid human-agent teams (Hopf et al., 2024; Schmutz et al., 2024). None of these framings captures the situation in which the AI agent is itself the user. Second, the study provides empirical evidence that a practical redesign framework can guide this transition in a real organisational setting. Third, the study characterises four specific organisational challenges that dual-user redesign introduces.

For practitioners, the study yields four concrete recommendations. First, begin with a service logic audit rather than a UI or API inventory: decompose the service into intents and actions before deciding what to expose to agents. Second, define cooperation governance at the level of each intent, specifying intent-level full delegation, action-level partial delegation, or human-only paths. Third, introduce probabilistic performance metrics from the outset. Fourth, invest in bridging the HOW/WHAT gap before implementation begins, through shared artefacts and structured workshops.

5.1 Limitations and future research

Three limitations merit acknowledgement. The single-case design limits generalisability; the AX framework requires validation in organisations with different design maturity, team structures, service domains, and levels of organisational authority. The researcher's dual role as CEO and designer created legitimate tensions between organisational decision-making velocity and reflective depth. Finally, the evaluation was conducted immediately after implementation; longitudinal effects remain unstudied.

Future research should address multi-agent configurations, in which the coordination challenges multiply. It should also examine organisations with less concentrated authority over the redesign process. The connection between dual-user redesign and broader innovation governance questions represents a productive agenda for management scholarship.

6 Conclusion

As AI agents become embedded across digital services, organisations face a genuinely new management challenge: how to design, govern, and improve services that serve two fundamentally different kinds of users simultaneously. This action research study demonstrates that the challenge is tractable: a human-first service can be systematically transformed for dual-user operation within a single project cycle, with measurable improvements in accessibility, usability, and cooperation. But it also demonstrates that the transformation is deeper than a technical upgrade. Dual-user redesign changes service logic, demands probabilistic quality management, requires agents to be treated as organisational stakeholders, and surfaces a distinctive capability gap between those who know how to build agentic systems and those who know what such systems should do. The AX (Agentic Experience) framework provides one scaffold for managing that organisational dimension.

References

Achiam, J. et al. (2023). *GPT-4 Technical Report*. arXiv preprint arXiv:2303.08774.

Anthropic (2023). *Claude model system card.* Anthropic.

Atil, B. et al. (2025). *Non-determinism of "deterministic" LLM settings*. arXiv:2408.04667v5.

Battistoni, P., Di Gregorio, M., Romano, M., Sebillo, M. and Vitiello, G. (2023). *Can AI-oriented requirements enhance human-centered design of intelligent interactive systems?* Multimodal Technologies and Interaction, 7(24).

Boston Consulting Group (2025). *Agentic commerce is redefining retail — here's* *how to respond.* BCG. Available at: https://www.bcg.com/publications/2025/agentic-commerce-redefining-retail-how-to-respond [Accessed April 2026].

Borghoff, U.M., Bottoni, P. and Pareschi, R. (2025). *Human-artificial interaction in the age of agentic AI.* Frontiers in Human Dynamics, 7.

Bradshaw, J.M., Feltovich, P.J. and Johnson, M. (2017). *Human–agent interaction*. In Boy, G. and Faucher, C. (eds.) The Handbook of Human–Machine Interaction. CRC Press, pp.283–300.

Chen, D., Wang, Y. and Xia, G. (2024). *Human-centered LLM-agent user interface: A position paper.* arXiv:2405.13050.

Cillo, P. and Rubera, G. (2025*). Generative AI in innovation and marketing processes: A roadmap of research opportunities*. Journal of the Academy of Marketing Science, 53(3), pp.684–701.

Clarkson, P.J., Coleman, R., Keates, S. and Lebbon, C. (eds.) (2013). *Inclusive Design:* Design for the Whole Population. Springer.

Coghlan, D. and Brannick, T. (2014). *Doing Action Research in Your Own Organisation.* 4th ed. SAGE.

Durante, Z. et al. (2024). *Agent AI: Surveying the horizons of multimodal interaction.* arXiv:2401.03568.

Gao, Y., Fan, J., Zeng, J. and Han, Z. (2025). *Artificial intelligence and innovation capability: A dynamic capability perspective.* Information & Management, 62(1), 103791.

Gartner (2025). *Gartner unveils top predictions for IT organizations and users in 2026 and beyond.* Gartner Newsroom, 21 October. Available at: https://www.gartner.com/en/newsroom/press-releases/2025-10-21-gartner-unveils-top-predictions [Accessed April 2026].

Haefner, N., Wincent, J., Parida, V. and Gassmann, O. (2021). *Artificial intelligence and innovation management: A review, framework, and research agenda.* Technological Forecasting and Social Change, 162, 120392.

Hopf, K., Nahr, N., Staake, T. and Lehner, F. (2024). *The group mind of hybrid teams with humans and intelligent agents in knowledge-intense work*. Journal of Information Technology, 40(1), pp.9–34.

Lewin, K. (1946). *Action research and minority problems*. Journal of Social Issues, 2(4), pp.34–46.

Lu, J. et al. (2024). *Turn every application into an agent: Towards efficient human-agent-computer interaction with API-first LLM-based agents*. arXiv:2409.17140.

Lusch, R.F. and Nambisan, S. (2015). *Service innovation: A service-dominant logic perspective.* MIS Quarterly, 39(1), pp.155–171.

Mariani, M.M., Machado, I., Magrelli, V. and Dwivedi, Y.K. (2022). *Artificial intelligence in innovation research: A systematic review, conceptual framework* *and future research directions.* Technovation, 102623.

Nielsen, J. (1993). *Usability Engineering.* Morgan Kaufmann.

Norman, D.A. (2013). *The Design of Everyday Things*. Basic Books.

Schmutz, J., Backmann, J. and Bittner, E.A.C. (2024). *AI-teaming: Redefining collaboration in the digital era.* Journal of Business Research, 172, 114162.

Schrills, T. and Franke, T. (2022). *Designing transparency for effective human-AI collaboration.* Information Systems Frontiers. https://doi.org/10.1007/s10796-022-10284-3.

Shneiderman, B. (2022). *Human-Centered AI.* Oxford University Press.

Susman, G.I. and Evered, R.D. (1978). *An assessment of the scientific merits of action research.* Administrative Science Quarterly, 23(4), pp.582–603.

Xu, F. et al. (2025). *TheAgentCompany: Benchmarking LLM agents on consequential real world tasks.* NeurIPS 2025 Datasets and Benchmarks Track. Available at: https://arxiv.org/abs/2412.14161 [Accessed April 2026].