

FROM AI ADOPTION TO ORGANISATIONAL STABILISATION

Ryszard WIERTNIAK

Krakow University of Economics; cwiertr@uek.krakow.pl, ORCID: 0000-0001-9295-5549

Purpose: Artificial intelligence (AI) is increasingly framed as a transformative organisational capability whose benefits are assumed to follow from technological adoption. However, in probabilistic AI environments, adoption does not automatically lead to stable organisational integration. The purpose of this study is to find mechanisms that enable or hinder the organisational stabilisation of AI-supported work in startup environments.

Design/methodology/approach: This study adopts an explanatory mixed-method design to examine the organisational stabilisation of AI in startup environments. The quantitative part analyses relationships among environmental awareness, trust in AI, psychological sustainability and digital competence using Partial Least Squares Structural Equation Modelling (PLS-SEM) based on a survey of Generation Z employees in AI-enabled startups (N = 120). The qualitative part applies reflexive thematic analysis of open-ended responses and semi-structured interviews with founders and accelerator representatives responsible for AI implementation.

Findings: The results reveal a selective stabilisation structure in AI-supported work. Trust in AI appears as the only statistically significant predictor of psychological sustainability and acts as the key mechanism linking AI use with stable organisational routines. Environmental awareness and digital competence show no independent effect. The findings further show that AI systems may remain widely used while generating hidden validation effort and coordination costs, resulting in a condition described as *stabilisation drift*.

Research limitations/implications: This study has several limitations. First, the cross-sectional dataset from employees in AI-enabled startups limits the ability to capture long-term stabilisation dynamics. Second, the focus on Generation Z employees may influence feelings of AI use and trust formation. Future research should employ longitudinal and multi-source designs to examine how reliance on AI evolves over time and to explore stabilisation mechanisms across different organisational contexts and technological domains.

Practical implications: The findings suggest that organisations should treat trust in AI as a strategic infrastructure of digital transformation rather than a secondary outcome of implementation. Managers should evaluate AI initiatives not only through technical performance but also through reliance intensity and the validation effort generated in everyday workflows. AI systems requiring persistent checking may create hidden coordination costs; therefore, organisations should prioritise transparency, behavioural predictability, and clear communication of system limitations to support stable human–AI collaboration.

Social implications: The findings highlight trust as a key condition for sustainable human–AI collaboration. By focusing on stabilisation rather than simple adoption, the study contributes to debates on trustworthy and human-centred AI and may inform organisational governance and public policy on responsible AI deployment.

Originality/value: The study advances AI adoption research by distinguishing diffusion from organisational stabilisation and introducing the concept of *stabilisation drift*. It reconceptualises trust as an organisational coordination infrastructure in probabilistic AI environments, contributing to socio-technical perspectives on sustainable AI integration in startups and technology-driven organisations.

Keywords: Artificial intelligence; AI stabilisation; Trust in AI; Startup ecosystems; Diffusion of innovation.

Category of the paper: Research paper.

1. Introduction

Artificial intelligence (AI) has rapidly evolved from an experimental technology into a strategic capability across contemporary organisations. Firms increasingly invest in AI to enhance decision-making process, optimise operations, and accelerate innovation. Yet a continuing productivity paradox stays: many AI initiatives do not generate sustained organisational performance gains despite formal adoption. This problem highlights a conceptual limitation in management research, which often treats diffusion as a discrete technological event rather than a continuous process of socio-technical integration. As Brooks (2017, pp. 139-159) sees, the development of artificial intelligence has historically followed a nonlinear trajectory characterised by alternating periods of rapid progress and stagnation (Pilny et al., 2025).

The organisational challenge is intensified by the nature of present generative AI systems. Unlike deterministic legacy technologies, Generative AI (GenAI) works on probabilistic logic and may produce ambiguous or incorrect outputs, often referred to as hallucinations (Joshi, 2025, pp. 38-50). These characteristics introduce interpretive uncertainty and require ongoing validation of AI-generated results. Therefore, the managerial challenge increasingly shifts from technical optimisation toward setting up organisational conditions that enable dependable human–AI collaboration (Ćwiertniak, 2025, p. 80; Lipińska, 2024, pp. 369-383). From a socio-technical perspective, generative AI should therefore be understood not as a single technology but as a system running across interconnected levels of models, technical systems, and organisational applications (Feuerriegel et al., 2024, pp. 111-126).

To address this challenge now, the study conceptualises AI integration through the lens of organisational stabilisation. This balance is understood as the organisational condition in which AI-supported work becomes routinised without generating persistent validation effort, task duplication, or coordination costs that offset expected efficiency gains. Within this new perspective, trust functions not merely as an individual attitude but as a coordination mechanism enabling users to rely on probabilistic AI outputs without excessive monitoring. When trust is absent, organisations experience added coordination costs and technostress (De Cremer, Koopman, 2024; Kim, Lee, 2024).

These dynamics are particularly visible in startup environments running within rapidly evolving technological and regulatory ecosystems. In Europe, AI development is increasingly embedded in strategic initiatives such as the Strategic Technologies for Europe Platform (STEP) and the regulatory framework introduced by the EU Artificial Intelligence Act (European Commission, 2024). Poland's expanding startup ecosystem therefore provides a relevant empirical context for examining how organisational mechanisms enable or hinder the stabilisation of AI-supported work.

In this context, named challenges, the purpose of this study is to name and analyse the mechanisms that enable or hinder the organisational stabilisation of AI-supported work in complex environments. By distinguishing diffusion from stabilisation in the context of AI adoption, the paper contributes to management research in three ways. First, it extends the Diffusion of Innovation (DOI) perspective by introducing a process-oriented and cost-sensitive understanding of organisational AI integration. Second, it reconceptualises trust as a coordination mechanism necessary for navigating probabilistic AI environments. Third, it introduces the concept of *stabilisation drift* as a condition in which AI stays formally adopted yet does not become operationally stabilised within everyday organisational practices.

This perspective also responds to broader critiques within AI research. Contemporary debates often focus on speculative scenarios of superintelligence (GAI), despite limited empirical evidence supporting such projections. Studies on sociotechnical risk suggest that too much attention to isolated technological futures may divert attention from immediate organisational, ethical, and governance challenges associated with current AI deployment (Stilgoe et al., 2013, pp. 1568-1580). By focusing on stabilisation, this study shifts analytical attention toward the organisational conditions under which AI systems become reliably embedded in routine work practices. An added contribution of the research is the development of an empirically grounded profile of next-generation AI users and the proposal of a Sustainable AI Adoption Canvas designed to support managerial decision-making in AI implementation processes.

2. Theoretical background and hypothesis development

2.1. Conceptual background

AI has been framed as a disruptive innovation (Christensen, 1997) and analysed through the lens of Diffusion of Innovation (Rogers, 2003). In this paradigm, adoption is treated as a function of perceived relative advantage, compatibility, and complexity. The implicit assumption is that once an innovation crosses the threshold of perceived utility, it naturally transitions into sustained use. However, contemporary probabilistic AI, specifically Large

Language Models (LLMs) and generative systems, breaks this deterministic connection. Unlike other IT tools, these systems are performative and inconsistent; they require ongoing validation and entail hidden coordination costs (Joshi, 2025, pp. 38-50). Consequently, formal adoption (diffusion) does not necessarily translate into routinised organisational integration. I therefore propose the concept of *stabilisation drift*. While diffusion is a discrete strategic event, stabilisation is a socio-technical process (Orlikowski, 1992, pp. 398-427) through which AI use becomes cognitively sustainable and institutionally normalised. Stabilisation is achieved only when the system ceases to require compensatory monitoring routines and defensive verification efforts that often offset the very efficiency gains AI was intended to provide (Ćwiertniak, 2024, p. 48).

Trust constitutes a central mechanism in AI diffusion. Multidimensional confidence models emphasise competence, integrity, and benevolence as key evaluative dimensions shaping whether AI is perceived as supportive or threatening (Dang, Liu, 2025; Marszałek-Kotzur, 2023, pp. 373-385). Although AI systems are often rated highly in terms of technical ability, scepticism toward fairness and goodwill persists, shifting responsibility toward the organisations that design and deploy them (Glikson, Woolley, 2020, pp. 627-660; Li et al., 2024). Within Rogers's (2003) framework, trust conditions perceived relative advantage and compatibility; when trust erodes, scaling AI becomes difficult regardless of technical performance. Trust therefore functions as a practical precondition for meaningful and sustained AI adoption (Bach et al., 2023).

Environmental considerations increasingly function as a compatibility filter in AI diffusion. The growing energy and water demand of AI infrastructures raise concerns about their legitimacy as sustainable innovation (Henderson et al., 2020, pp. 1-43; Verdecchia et al., 2023, pp. 1-26; Hosseini et al., 2024, pp. 1-3). In response, Green AI research focuses on energy-efficient architectures, carbon reporting, and environmentally responsible system design (Bolón-Canedo, Morán-Fernández, 2024; Tabbakh, 2024). From a strategic perspective, the effectiveness of AI-enabled sustainability initiatives depends on organisational green capabilities and credible ESG positioning (Rahman et al., 2025, pp. 2502-2525; Tetreiová et al., 2025, pp. 3311-3337). Importantly, environmental awareness does not translate directly into trust or positive psychological outcomes. Instead, it uses as a normative filter shaping how AI systems are evaluated. For environmentally conscious users, AI perceived as ecologically irresponsible may lack legitimacy despite strong technical performance, showing that environmental awareness functions as a contextual condition rather than a linear driver of trust.

In the European context, this tension is intensified by the technological sovereignty agenda. For startups in critical technology sectors, environmental awareness runs as an evaluative boundary condition. It sets the normative ceiling for AI initiatives. While sustainability may not directly drive adoption, it conditions the institutional justification of AI, influencing whether trust formation is viewed as socially responsible or ethically precarious. The Strategic Technologies for Europe Platform (STEP) shows AI as a critical sector for European

sovereignty (European Commission, 2024). However, as this study argues, institutional backing via STEP is insufficient without a robust AI Stabilisation Theory to manage the cognitive load of Gen Z employees.

I argue that AI stabilisation is as much a cognitive problem as an institutional one. I define psychological sustainability as the capacity of AI-supported work to support a practical effort-benefit ratio without increasing cognitive force. Existing research links AI opacity to technostress (De Cremer & Koopman, 2024). However, from a stabilisation perspective, the critical variable is not temporary stress but the accumulation of micro-costs (e.g., interpretive ambiguity, validation fatigue). If *the cognitive tax* of using AI persistently exceeds its performance utility, the stabilisation process collapses into *stabilisation drift*, leading to under-reliance or total rejection of the system (Ćwiertniak, 2025a, pp. 79-94; 2025b, pp. 57-72).

Generation Z plays a critical role in shaping AI diffusion as a cohort combining technological fluency with heightened ethical and environmental expectations (Varese et al., 2025). Their adoption of AI is conditional on perceived usefulness, trustworthiness, and alignment with ESG and fairness norms (Deloitte, 2024a, 2024b; Medina-Molina et al., 2025). At the same time, concerns about job displacement and technological unemployment influence their risk feelings (Nigar et al., 2025). Although AI adoption creates new career paths and entrepreneurial opportunities in the digital economy (Brynjolfsson, McAfee, 2017, pp. 66-77), digital competence within Generation Z is still heterogeneous. Beyond a minimum threshold that enables interaction with AI systems, higher competence does not necessarily translate into greater trust or improved psychological well-being. Digital competence therefore functions primarily as a baseline condition shaping how individuals interpret and interact with AI rather than as a reinforcing driver of positive adoption outcomes.

Organisations may experience *stabilisation drift*, a condition in which AI systems stay formally adopted and widely used yet do not achieve stable organisational integration. As users confront probabilistic outputs and uncertainty, they increasingly rely on other verification, monitoring, or parallel decision processes. These compensatory practices generate hidden validation effort and coordination costs, gradually eroding reliable confidence on AI despite continued system use.

2.2. Hypothesis development

While Diffusion of Innovation theory explains the first adoption of technological artifacts (Rogers, 2003), it often does not account for the post-adoption resistance inherent in probabilistic AI. Present AI systems, unlike deterministic software, impose hidden costs of validation and interpretive labour (Joshi, 2025, pp. 38-50). Stabilisation is threatened by compensatory monitoring, a redundant layer of human oversight that appears when trust is missing. To capture the micro-foundations of this process, I introduce psychological

sustainability as the core indicator of stabilisation, reflecting the user's ability to support high performance without reducing cognitive resources through constant system verification.

In the European startup ecosystem, AI deployment is increasingly analysed through the lens of institutional legitimacy. The high energy intensity of large-scale models creates a legitimacy gap that can trigger user scepticism (Bolón-Canedo, Morán-Fernández, 2024). I argue that environmental awareness is not merely an external ESG metric but an evaluative filter that conditions trust. When users perceive a misalignment between AI deployment and sustainability values, the resulting normative dissonance undermines the perceived integrity of the system. In opposition, alignment reinforces the moral dimension of trust, helping the transition from technical interaction to organisational confidence (Toderas et al., 2025). From an organisational legitimacy perspective, environmental awareness may function as a normative signal that shapes initial confidence in AI systems.

H1: Environmental awareness is positively associated with trust in AI systems

In probabilistic environments, trust serves as a cognitive heuristic that allows users to bypass exhaustive verification routines. When trust is high, users delegate agency to the AI, thereby reducing technostress and the cognitive load associated with monitoring (De Cremer, Koopman, 2024). From a stabilisation perspective, trust is the mechanism that prevents validation weakness. By lowering the perceived risk of incorrect outputs, trust enables a favourable benefit ratio, which is the hallmark of psychological sustainability in AI-supported work. Trust therefore acts as a coordination mechanism that reduces verification effort and supports psychologically sustainable human–AI collaboration.

H2: Trust in AI positively influences psychological sustainability

Impact of institutional values (environmental awareness) on operational outcomes (sustainability) is not direct but mediated by trust. Environmental evaluations shape the normative acceptance threshold, but they only influence day-to-day work patterns if they are internalised as trust in the system's integrity. Trust, therefore, translates macro-level legitimacy into micro-level cognitive relief. Without this mediation, even the greenest AI will not stabilise if users do not trust its operational reliability (De Cremer, Koopman, 2024; Huang, Rust, 2021, pp. 3-21). In contrast, ecological dissonance, recognising the environmental impact of AI while relying on it, may generate stress or moral tension. Trust can mitigate this dissonance by signalling that AI use stays compatible with sustainability values. Environmental awareness may provide a normative foundation for trusting AI systems, but its influence on psychologically sustainable human–AI collaboration is expected to run primarily through trust as a coordination mechanism.

H3: Trust in AI mediates the effect of environmental awareness on psychological sustainability

Digital competence is often mistaken as a direct driver of adoption. I re-conceptualise it as a moderating variable that governs the calibration of trust. Highly competent users have the interpretive schemas necessary to navigate AI's probabilistic nature. For these individuals, trust

is more effectively converted into psychological sustainability because they can distinguish between systemic uncertainty and operational failure. In contrast, for users with low competence, trust stays fragile and less effective in mitigating the strain of interpretive ambiguity. Digital competence may shape how effectively users translate trust in AI into sustainable collaboration, as more competent users are better able to interpret probabilistic outputs and integrate AI into their work routines (Feuerriegel et al., 2024, pp. 111-126).

H4: Digital competence positively moderates the relationship between trust in AI and psychological sustainability

The integrated model proposes a selective stabilisation structure that moves beyond traditional additive adoption frameworks. In this perspective, environmental awareness functions as an institutional precursor that provides the normative legitimacy necessary for the development of trust in AI systems. Trust is positioned as the central mechanism of stabilisation, being the primary pathway through which interaction with AI systems translates into psychological sustainability in everyday work practices. Rather than treating AI integration as a natural consequence of adoption, the framework conceptualises stabilisation as a fragile socio-technical process. Sustained collaboration with AI appears only when trust effectively offsets the hidden cognitive costs associated with probabilistic systems, including monitoring effort, validation work and coordination friction. The inclusion of digital competence as a moderating factor further suggests that stabilisation is not uniform across users. Instead, the strength of the relationship between trust and psychological sustainability depends on the user's ability to interpret and calibrate reliance on non-deterministic AI outputs.

Overall, the proposed framework shifts analytical attention from diffusion, understood as the spread of technology, to organisational stabilisation, known as the management of cognitive and organisational resources needed for sustainable human–AI collaboration in the high-velocity environments of the European startup ecosystem.

3. Methodology

This study employs a sequential explanatory mixed-methods design combining quantitative PLS-SEM analysis with qualitative thematic analysis. Such an approach reflects the socio-technical nature of AI integration: quantitative analysis shows relationships between constructs, while qualitative interviews provide insight into organisational routines and coordination mechanisms that shape AI-supported work.

The quantitative part examines Generation Z employees (N = 120) aged 20-27 working in technology-intensive startups, including fintech and AI-enabled services. This cohort was selected due to its important level of digital fluency and frequent interaction with AI systems. The sample size satisfies methodological requirements for Partial Least Squares Structural

Equation Modelling (PLS-SEM), exceeding the ten-times rule and ensuring adequate statistical power (Hair et al., 2021).

To complete the operational perspective, semi-structured interviews were conducted with 32 decision-makers from the Polish startup ecosystem, including founders, senior managers, and accelerator directors. Many of these organisations take part in European innovation programmes such as the Strategic Technologies for Europe Platform (STEP) (European Commission, 2024), providing insight into how AI deployment interacts with regulatory frameworks including the Artificial Intelligence Act (2024).

The survey instrument consisted of twenty-three Likert-scale items measuring four constructs: environmental awareness, trust in AI, psychological well-being in AI use, and digital competence. The conceptual definitions and operationalisation of these constructs are presented in Table 1. Prior to the main study, the questionnaire was pilot tested with AI-using startup employees to ensure clarity and contextual validity.

Table 1.
Overview of user-centred constructs shaping Sustainable AI Adoption

Construct	Description	Items	Scale
Environmental awareness	Sensitivity to AI's environmental footprint and sustainability expectations.	5	1-5
Trust in AI	Perceived reliability, transparency, fairness, controllability, and safety.	5	1-5
Psychological well-being in AI use	Cognitive load, technostress, autonomy, and emotional comfort when using AI.	6	1-5
Digital competences	Ability to interpret AI outputs, detect errors and understand basic AI functioning.	7	1-5

Note. Constructs capture environmental, psychological, and cognitive dimensions relevant to organisational AI adoption.

Source: own elaboration based on the study.

Quantitative data were screened for completeness, internal consistency, and response patterns. Reverse-coded items were adjusted, and composite construct scores were created. Descriptive statistics and correlation matrices were examined prior to structural analysis to assess distributional properties and potential anomalies. Qualitative responses were anonymised, cleaned, and prepared for coding.

Hypotheses were examined using PLS-SEM. This approach is right for exploratory research involving latent constructs, emerging theoretical relationships, and complex socio-technical phenomena where theory development and mechanism exploration are prioritised over predictive optimisation. In line with the study's systems-oriented aim, PLS-SEM was employed to find relational patterns and boundary conditions rather than to maximise model fit. Analytical procedures included assessment of collinearity (VIF), reliability and validity evaluation (indicator loadings, composite reliability, AVE, HTMT), estimation of direct and indirect effects, and bootstrapping with 5000 resamples. Model explanatory power was assessed using R^2 and SRMR, following established methodological guidelines.

The thirty-two managerial interviews and open-ended survey responses were analysed using reflexive thematic analysis. The process involved iterative coding, identification of recurrent themes, and development of higher-order interpretive categories. Coding focused on trust calibration, monitoring intensity, validation routines, sustainability framing, and perceived cognitive and organisational costs. Qualitative findings were used to contextualise and interpret quantitative patterns, particularly in relation to non-significant effects and cross-level tensions. All participants were informed about the purpose of the study and provided voluntary consent.

4. Results

4.1. Descriptive Insights

Descriptive statistics provide a first overview of AI use in early-stage startup environments. Respondents (M age = 23.1; SD = 2.4) held cross-functional roles spanning product development, analytics, operations, automation, and customer support. In line with the flat structures typical of startups, AI tools were embedded directly in task execution rather than peripheral experimentation. AI exposure was intensive: 81% of respondents reported daily interaction with AI systems. However, only 18% had received formal organisation-sponsored AI training. This imbalance shows that AI diffusion occurred faster than structured capability development. In acceleration-driven environments, such asymmetry may increase validation effort and reliance volatility. Trust in AI was moderate (M = 3.28), while environmental awareness was slightly higher (M = 3.41). Psychological sustainability scores revealed variability in perceived cognitive load and effort–benefit balance, suggesting heterogeneity in stabilisation experiences.

4.2. Measurement model assessment

Given the exploratory and theory-building nature of the study, construct retention was prioritised over strict threshold adherence. Indicator-level collinearity was not problematic, with VIF values ranging from 1.26 to 4.27. The psychological well-being construct showed strong internal consistency and convergent validity ($CR > 0.70$; $AVE > 0.50$). For the remaining constructs, composite reliability exceeded the recommended threshold of 0.60. Although AVE values for some constructs fell below the conventional 0.50 threshold, they were kept following the recommendations of Hair et al. (2021), who note that constructs with adequate composite reliability and strong theoretical grounding may be preserved in exploratory research. Discriminant validity was confirmed using the HTMT criterion, with all values below the 0.85 threshold. Overall, the measurement model was considered suitable for structural analysis.

4.3. Structural model results

The structural model was estimated using PLS-SEM with a bootstrapping procedure of 5000 resamples. Inner VIF values remained below 3.0, showing that multicollinearity did not bias parameter estimates. The model explained 10% of the variance in trust ($R^2 = 0.10$) and 30% of the variance in psychological well-being ($R^2 = 0.30$). In line with recommendations for exploratory PLS-SEM (Hair et al., 2021; Sarstedt et al., pp. 197-217), the explanatory power for psychological well-being can be considered moderate, given the complexity, subjectivity, and context-dependence of the phenomena under investigation. Of the four hypothesised relationships, one was empirically supported. Environmental awareness showed no significant association with trust (H1: $\beta = 0.130$, $p = .118$). Trust in AI showed a strong and statistically significant relationship with psychological well-being (H2: $\beta = 0.415$, $p < .001$), appearing as the sole stable predictor in the model. Neither the direct nor indirect effects of environmental awareness on psychological well-being were significant (H3: direct $\beta = -0.003$, $p = .970$; indirect $\beta = -0.044$, $p = .670$). The interaction effect between trust and digital competence was also non-significant (H4: $\beta = 0.128$, $p = .468$). Rather than saying weak explanatory power, this pattern reveals a selective structure of influence, in which trust in AI functions as the primary stabilising mechanism linking AI use to psychological well-being. Detailed coefficients and significance levels are reported in Table 2.

Table 2.
Structural path factors and hypothesis testing results

Hypothesis	Path	β	p-value	Supported
H1	Environmental Awareness \rightarrow Trust in AI	0.130	.118	No
H2	Trust in AI \rightarrow Well-being	0.415	<.001	Yes
H3	Environmental Awareness \rightarrow Well-being (direct)	-0.003	.970	No
	Environmental Awareness \rightarrow Trust \rightarrow Well-being (indirect)	-0.044	.670	No
H4	Trust \times Digital Competence \rightarrow Well-being	0.128	.468	No

Note. β = standardised path coefficients; bootstrapping based on 5000 resamples.

Source: own elaboration based on the study.

Overall, the findings reveal a selective pattern of relationships, where trust in AI serves as the primary stabilizing mechanism for user well-being. These results highlight the conditions under which trust-centered AI adoption may influence psychological outcomes in organizational environments. Moreover, the structural model explains 30% of the variance in psychological sustainability, suggesting that additional organisational conditions also contribute to the stabilisation of AI-supported work. Future research could therefore incorporate variables such as workload intensity, governance policies, or task complexity.

4.4. Qualitative analysis

The qualitative analysis was based on thirty-two semi-structured interviews conducted with startup founders and senior managers responsible for AI implementation and scaling decisions. The purpose of this analysis was to interpret the quantitative findings from the employee survey

and to examine how managerial governance practices align with or challenge the structural model (H1-H4). Four dominant themes appeared.

First, managers consistently described AI as efficiency enhancing but requiring structured oversight. Several interviewees indicated that AI-generated outputs are systematically reviewed before deployment, particularly in client-facing, regulatory, or data-sensitive contexts, e.g. *AI accelerates output, but final responsibility stays with us; We treat AI as support, not as decision authority; AI works like an intern in our company – it gets the job done quickly, but a senior must always check and approve it before it's sent to the client. Accountability simply cannot be transferred to an algorithm; Final responsibility always stays with the human lead. We use AI to accelerate the process, but the 'human-in-command' rule is non-negotiable for our high-stakes projects.* This confirms that monitoring routines are institutionally embedded. The centrality of trust in the quantitative model (H2) is therefore consistent with managerial emphasis on calibrated reliance rather than unconditional delegation.

Second, while managers acknowledged environmental concerns related to AI infrastructure, e.g. *Energy consumption is part of the ESG discussion; Ecology is a brand-building tool at the strategic level, but it has zero impact on our developers' daily choices. They need the most powerful tool available, regardless of its carbon footprint; Green legitimacy helps us secure funding from Brussels, but it does not translate into operational constraints. In the scaling phase, survival, and speed trump environmental concerns,* these considerations were primarily framed as legitimacy or investor-related issues. They were not described as determinants of day-to-day reliance decisions. This managerial perspective helps explain the non-significant association between environmental awareness and both trust (H1) and psychological sustainability (H3) seen in the structural model.

Third, trust was described as situational and task dependent. Managers showed that reliance levels vary according to risk exposure, task criticality, and reputational consequences, e.g. *We increase oversight when stakes are high; Trust depends on context; Our reliance on AI is dialled up or down based on risk. If a mistake could cost us a key client or a regulatory fine, we switch back to manual oversight at once"; Trust is not a permanent state; it is managed daily. We only rely on the system as far as we can predict and control its potential for failure.* Trust is therefore actively managed rather than assumed. This directly reinforces H2, trust functions as the primary mechanism linking AI use to sustained operational viability.

Fourth, managers emphasised the importance of employee training and AI literacy. However, competence was described as improving calibration and error detection rather than eliminating monitoring needs, e.g. *Training improves usage quality, not autonomy; Raining didn't make people trust the system more; On the contrary – a competent employee catches errors and AI hallucinations faster, so they check the system even more thoroughly; High AI literacy doesn't lead to blind faith; It leads to better calibration. Our most skilled team members are the ones who are the most sceptical of unverified outputs; We invested in digital skills to reduce technostress, but it has not changed our oversight. Even a pro knows that*

a probabilistic system always requires a safety net. This managerial view aligns with the non-significant moderation effect (H4): digital competence enhances interpretive capacity but does not structurally alter the trust–psychological sustainability relationship.

Table 3.

Managerial themes in relation to hypotheses H1-H4

Theme	Illustrative Manager Statements (n = 32)	Interpretation	Link to Hypotheses
Verification as governance practice	“AI accelerates output, but we always review”	Oversight is structurally embedded.	Supports H2
Environmental legitimacy framing	“ESG matters strategically, not operationally”	Sustainability shapes discourse, not reliance.	Explains non-significant H1 & H3
Conditional trust	“Trust depends on task criticality”	Trust is calibrated and managed.	Dedicated support for H2
Competence as calibration	“Training improves use, not autonomy”	Competence refines monitoring but does not cut it.	Clarifies non-significant H4

Note. Based on thirty-two semi-structured interviews with startup founders and senior managers.

Source: own elaboration based on the study.

Overall, the managerial interviews corroborate the selective structure found in the quantitative model. Trust appears as the key stabilising mechanism in AI-supported work, while environmental considerations and digital competence shape governance narratives and calibration practices rather than directly finding sustained reliance

4.5. Integrated interpretation

The integration of quantitative and qualitative evidence shows a coherent and selective structure of influence. The structural model (Table 2) finds trust in AI as the only statistically robust predictor of psychological well-being, while managerial interviews (Table 3) clarify the organisational mechanisms underlying this effect. Together, these findings delineate a trust-centred architecture of AI-supported work in startup environments.

Trust applies a direct influence on psychological sustainability (H2), confirming its role as the primary stabilising mechanism linking AI use with perceived control and manageable cognitive demand. Qualitative evidence shows that trust is actively calibrated through governance routines, oversight practices, and risk-sensitive deployment strategies. Managers do not cut monitoring, they regulate it. In this sense, trust runs as a control-adjustment mechanism rather than as unconditional reliance. While the interviews reveal the presence of monitoring routines and human-in-command governance structures, the present study does not quantify how frequently these verification practices occur across organisation.

In contrast, environmental awareness and digital competence do not produce stable linear effects in the structural model (H1, H3, H4). The qualitative findings clarify this pattern. Environmental considerations function mainly as legitimacy and positioning factors shaping evaluative discourse rather than operational confidence. Similarly, digital competence improves

users' interpretive capacity and calibration behaviour but does not fundamentally alter the trust–well-being relationship. Competence refines the quality of AI use but does not replace governance protections.

Taken together, the results reveal a differentiated pattern of influence in which trust functions as the central mechanism of psychological sustainability, while environmental awareness and digital competence run as conditional and context-sensitive factors shaping evaluation and calibration processes. Within this framework, Generation Z employees in startup environments can be characterised as adjusting AI users. Despite elevated levels of digital fluency, their reliance on AI is not driven by technological enthusiasm but by pragmatic, task-oriented evaluation of system reliability and cognitive efficiency. When trust infrastructures are present, AI becomes integrated into everyday work routines, when they are absent, users increase monitoring and validation efforts, which can undermine perceived efficiency.

Overall, the findings suggest that AI diffusion in startup environments is not decided by uniform adoption drivers but by calibrated reliance under conditions of organisational acceleration. Effective integration depends less on technical optimisation alone and more on the establishment of trust as a stabilising infrastructure within socio-technical systems.

5. Discussion and practical implications for AI-driven organisations

The confirmed relationship between trust and psychological sustainability (H2) reinforces the concept of trust as a coordination structure rather than a mere attitudinal variable. In the absence of bureaucratic buffering, a hallmark of early-stage startups, trust becomes the primary mechanism for delegating agency to AI. Qualitative evidence further clarifies that this trust is never blind, it is standardized. Managers do not cut monitoring; they regulate its intensity based on trust levels. Consequently, trust functions as a control-adjustment mechanism that reduces the cognitive tax on employees. When trust is set up, the transition from defensive monitoring to routine reliance enables psychological sustainability by preventing the depletion of cognitive resources.

The non-significance of environmental awareness (H1, H3) and digital competence (H4) in the structural model is a significant theoretical finding. The qualitative data suggest a structural decoupling between strategic discourse and operational reality:

- While sustainability is a critical legitimacy wrapper for startups seeking European VC funding or compliance with ESG standards, it stays an external-facing construct. It shapes the justification of AI adoption but does not yet perceive the micro-routines of daily work. This explains why it does not trigger measurable changes in trust or psychological well-being at the operational level.

- Similarly, digital competence tasks as a tool for interpretive mastery rather than a driver of autonomy. Managers view competence to better editing rather than less oversight. Therefore, while competence improves the quality of AI outputs, it does not structurally alter the psychological strain-reliance nexus, explaining the lack of a robust moderation effect.

Finally, the integrated evidence suggests that AI integration among Generation Z employees is driven by calibrated reliance under acceleration pressure. In startup ecosystems, where structural slack is minimal, AI stabilisation is not a product of uniform adoption determinants. Instead, it is a fragile balance. Effective AI diffusion depends less on technical optimisation or sustainability rhetoric and more on the establishment of trust as a stabilising infrastructure. Without this infrastructure, AI initiatives are prone to *stabilisation drift*, where the technology is formally adopted but is still functionally difficult due to the persistent need for compensatory human control.

AI integration is not an automatic byproduct of technical deployment but a strategically conditioned stabilisation process. For startup leaders, the primary challenge shifts from technical implementation to reliance management, as illustrated in Table 4 – the Sustainable AI Adoption Canvas (Agrawal et al., 2018). The canvas represents the conceptual result of this research, integrating the study’s empirical insights into a stabilisation framework that captures how trust calibration, monitoring practices, and cognitive sustainability structure durable human–AI collaboration:

- Efficiency in probabilistic AI environments – **Prediction** – is not driven by marginal gains in model accuracy, but by behavioural predictability. Managers should check validation intensity, the frequency and depth of human cross-checking. High validation intensity is a leading indicator of *stabilisation drift*, where the cognitive tax of using AI offsets its productivity gains.
- **Training** should move beyond prompt engineering toward trust calibration. Employees need to understand the risk-tiered nature of AI outputs, knowing exactly when to trust the system and when to intensify scrutiny. This interpretive mastery is what prevents burnout and sustains psychological sustainability.
- Based on the **Action and Outcome** dimensions, leaders must evaluate whether AI outputs are integrated into workflows without defensive duplication. True stabilisation is achieved when the system reduces the net coordination cost, allowing the lean structure of a startup to remain agile without bureaucratic oversight.
- Sustainability – **Feedback** – should be treated as an internal governance pillar, not just an external reporting requirement. While ESG awareness may seem decoupled from daily tasks, it functions as a legitimacy infrastructure. Transparency on AI’s environmental footprint reduces normative dissonance among Generation Z employees, reinforcing their long-term commitment to the system.

In resource-constrained startup settings, where hidden inefficiencies cannot be absorbed through bureaucratic buffering, stabilisation becomes linked to strategic existence.

Table 4.
Sustainable AI Adoption Canvas - example

Dimension	Guiding Question	Implication Based on Findings
A. Strategic Decision Layer		
Prediction	What uncertainty does AI meaningfully reduce for the user rather than merely increasing technical accuracy?	Predictive value is experienced in terms of reduced cognitive effort rather than technical accuracy alone; excessive prompting undermines perceived efficiency.
Judgement	How are efficiency gains weighed against cognitive load, technostress, and environmental ambiguity?	Trust mediates these trade-offs; when perceived effort outweighs benefit, AI use becomes frictional.
Action	Is AI deployed as automation, recommendation, or decision support, and with what degree of human oversight?	Decision-support configurations better sustain trust and psychological stability than full automation.
Outcome	How is success evaluated beyond short-term productivity metrics?	Sustained trust and adaptive psychological balance constitute the relevant indicators of successful AI adoption.
B. Interpretive and Learning Conditions		
Input	What data are needed without increasing perceptions of intrusiveness or cognitive overload?	Data sufficiency should be balanced against psychological and ecological concerns, not maximised.
Training	How is the system trained to prioritise consistency and transparency rather than aggressive optimisation?	Behavioural predictability and explainability support trust more than marginal performance gains.
Feedback	How are system behaviour, errors, and limitations communicated to users?	Feedback loops reinforce trust by supporting learning, perceived control, and technostress reduction.

Note. Author's adaptation of the AI Canvas framework informed by PLS-SEM results (H2 supported) and qualitative evidence highlighting trust as a psychological stabiliser.

Source: own elaboration based on Agrawal et al. (2018).

AI systems that drift toward effort-intensive use may not fail visibly, yet they erode execution speed and decision confidence. Continuous monitoring of reliance patterns and validation intensity is therefore essential for supporting competitive agility. However, the quantitative component of this study relies on cross-sectional survey data collected from employees of AI-oriented startups, which limits the ability to observe the long-term dynamics of organisational stabilisation.

6. Conclusion

This research addresses a central tension in the contemporary management of critical technologies, the disconnect between the rapid diffusion of AI and its long-term organisational stabilisation. Evidence from the Polish startup ecosystem suggests that AI integration is not a linear implementation milestone but a fragile socio-technical process. The findings reposition

trust as the core coordination infrastructure of human–AI collaboration and show it as the primary pathway to psychological sustainability in probabilistic AI environments.

In the high-stakes environment of critical technologies, where AI systems often govern dual-use infrastructures, sensitive data, or strategic decision-support, the failure to stabilise integration is not merely an operational setback but a strategic vulnerability. This empirical evidence reveals a selective stabilisation mechanism, while environmental awareness and digital competence are essential for institutional legitimacy, they do not directly drive operational reliance. Instead, the engine of permanent integration is the improvement of the cognitive tax on employees. When trust is effectively adjusted, it drops the need for the excessive, compensatory monitoring that often paralyzes early-stage startups under acceleration pressure. At the same time, the results say the emergence of *stabilisation drift*. Despite continued organisational use of AI systems, reliance on probabilistic outputs requires ongoing verification and compensatory coordination. The resulting validation effort generates hidden coordination costs and gradually weakens stable human–AI collaboration (Ćwiertniak, 2025a, pp. 79-94).

The implications of this research resonate with the European Union’s pursuit of technological sovereignty and the mandates of the EU AI Act. As Europe intensifies its focus on critical digital infrastructures, the robustness of the startup ecosystem becomes a proxy for systemic resilience. This research suggests that for Europe to keep its competitive edge, policy and management frameworks must move beyond diffusion metrics toward stabilisation depth.

Beyond the theoretical distinction between diffusion and stabilisation, the study also develops an empirically grounded profile of next-generation AI users working in startup environments characterised by high uncertainty and rapid technological iteration. On this basis, the research proposes the Sustainable AI Adoption Canvas as a managerial tool that supports organisations in diagnosing stabilisation barriers and designing implementation strategies that account for cognitive load, validation effort, and trust calibration.

Finally, the long-term success of AI-enabled capabilities in Europe depends not on the mere speed of deployment, but on the capacity of startups to foster environments where trust enables durable, transparent, and cognitively sustainable reliance. By distinguishing diffusion from stabilisation, this research provides a roadmap for ensuring that digital transformation in critical sectors leads to genuine organisational capability rather than superficial adoption. In the race for global technological leadership, the ability to stabilise AI within the human-centric framework of the European model may well be the ultimate competitive advantage.

Future research should therefore extend the present findings through longitudinal and multi-source research designs that would allow a more systematic observation of how trust, reliance, and monitoring practices evolve as AI systems become embedded in organisational routines and either stabilise or drift into compensatory verification cycle.

Acknowledgements

The publication was financed from the subsidy granted to the Krakow University of Economics – Project nr 055/ZJE/2025/POT.

References

1. Agrawal, A., Gans, J., Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.
2. Bach, J.S., Khan, N., Hallock, D., Beltrão, G., Sousa, M.J. (2023). *A systematic literature review of user trust in AI-enabled systems: An HCI perspective*. arXiv. <https://arxiv.org/abs/2304.08795>
3. Bolón-Canedo, V., Morán-Fernández, L. (2024). A review of green artificial intelligence: Towards a more sustainable future. *Neurocomputing*, 599, 128096, <https://doi.org/10.1016/j.neucom.2024.128096>
4. Borah, P.S., Dogbe, C.S.K., Marwa, N. (2024). Generation Z's green purchase behavior: Do green consumer knowledge, consumer social responsibility, green advertising, and green consumer trust matter for sustainable development? *Business Strategy and the Environment*, 33(5), pp. 4530-4546, <https://doi.org/10.1002/bse.3714>
5. Brooks, R.A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1-3), pp. 139-159, [https://doi.org/10.1016/0004-3702\(91\)90053-M](https://doi.org/10.1016/0004-3702(91)90053-M)
6. Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM*, 36(12), pp. 66-77, <https://doi.org/10.1145/163298.163309>
7. Christensen, C.M. (1997). *The innovator's dilemma: When new technologies cause great firms to fail*. Harvard Business School Press.
8. Ćwiertniak, R. (2024). *Sztuczna inteligencja w organizacji. Innowacje biznesowe w praktyce*, Gliwice: Helion.
9. Ćwiertniak, R. (2025a). Ludzie i boty sztucznej inteligencji w nowoczesnych projektach. *Studia i Prace Kolegium Zarządzania i Finansów SGH*, 199, pp. 79-94, <https://doi.org/10.33119/SIP.2024.199.4>
10. Ćwiertniak, R. (2025b). Zarządzanie innowacjami we wdrażaniu sztucznej inteligencji w organizacji. In: Doligalski, T., Kaszyński, D., (Eds.), *Sztuczna inteligencja w przedsiębiorstwach i gospodarce (AI Spring 2024)*, pp. 57-72. Warszawa: Oficyna Wydawnicza SGH. Retrieved from: <http://hdl.handle.net/20.500.12182/1384>, 03.03.2026.
11. Dang, J., Liu, L. (2025). Public trust in artificial intelligence users. *Current Opinion in Psychology*, 66, 102148, <https://doi.org/10.1016/j.copsyc.2025.102148>

12. De Cremer, D., Koopman, J. (2024). Research: Using AI at work makes us lonelier and less healthy. *Harvard Business Review*, June 2024. Retrieved from: <https://hbr.org/2024/06/research-using-ai-at-work-makes-us-lonelier-and-less-healthy>, 06.03.2026.
13. Deloitte (2024a). Generative AI trust actions correlate with better risk management and bigger rewards. *Deloitte Insights*. Retrieved from: <https://www.deloitte.com/us/en/insights/topics/risk-management/trust-in-ai.html>, 06.03.2026.
14. Deloitte (2024b). More US consumers think AI-generated health information should be left to the experts. *Deloitte Insights Magazine*, Iss. 33. Retrieved from: <https://www.deloitte.com/us/en/insights/deloitte-insights-magazine/issue-33/ai-generated-health-information-us-consumers.html>, 06.03.2026.
15. *EU Artificial Intelligence Act* (2024). Retrieved from: <https://artificialintelligenceact.eu/>, 06.03.2026.
16. European Commission (2024). Guidelines on certain provisions of Regulation (EU) 2024/795 establishing the Strategic Technologies for Europe Platform (STEP) (C/2024/3209). *Official Journal of the European Union*. Retrieved from: <http://data.europa.eu/eli/C/2024/3209/oj>, 06.03.2026.
17. Feuerriegel, S., Hartmann, J., Janiesch, C., Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), pp. 111-126, <https://doi.org/10.1007/s12599-023-00834-7>
18. Floridi, L., Holweg, M., Taddeo, M. (2023). *The ethics of artificial intelligence: Human-centric principles for a sustainable digital future*. Oxford University Press.
19. Glikson, E., Woolley, A.W. (2020). Human trust in artificial intelligence: A review of empirical research. *Academy of Management Annals*, 14(2), pp. 627-660, <https://doi.org/10.5465/annals.2018.0057>
20. Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE Publications.
21. Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. *Journal of Machine Learning Research*, 21(248), pp. 1-43. Retrieved from: <https://www.jmlr.org/papers/v21/20-312.html>, 06.03.2026.
22. Hosseini, M., Gao, P., Vivas-Valencia, C. (2024). A social-environmental impact perspective of generative artificial intelligence. *Environmental Science and Ecotechnology*, 23, 100520. <https://doi.org/10.1016/j.ese.2024.100520>
23. Huang, M.H., Rust, R.T. (2021). Artificial intelligence in service. *Journal of Service Research*, 24(1), 3-21. <https://doi.org/10.1177/1094670520902266>
24. Joshi, A. (2025). *A comprehensive review of AI hallucinations: Impacts and mitigation strategies for financial and business applications*. PhilArchive. Retrieved from: <https://philarchive.org/rec/JOSCRO-3>, 06.03.2026.

25. Kim, B.-J., Lee, J. (2024). The mental health implications of artificial intelligence adoption: The crucial role of self-efficacy. *Humanities & Social Sciences Communications*, 11, Article 1561, <https://doi.org/10.1057/s41599-024-04018-w>
26. Li, Y., Wu, B., Huang, Y., Luan, S. (2024). Developing trustworthy artificial intelligence: Insights from research on interpersonal, human-automation, and human-AI trust. *Frontiers in Psychology*, 15, 1382693, <https://doi.org/10.3389/fpsyg.2024.1382693>
27. Lipińska, J. (2024). The use and application of artificial intelligence in decision-making processes. *Scientific Papers of Silesian University of Technology. Organization and Management Series*, 211, pp. 369-383, <https://doi.org/10.29119/1641-3466.2024.211.22>
28. Marcinów, T. (2024). Ethics of artificial intelligence in the light of selected EPTA reports: Implications for CSR. *Scientific Papers of Silesian University of Technology. Organization and Management Series*, 199, pp. 401-410, <https://doi.org/10.29119/1641-3466.2024.199.30>
29. Marszałek-Kotzur, I. (2023). The ethics of artificial intelligence. Can we trust robots? *Scientific Papers of Silesian University of Technology. Organization and Management Series*, 183, pp. 373-385, <http://dx.doi.org/10.29119/1641-3466.2024.210.25>
30. Medina-Molina, C., Pérez-Macías, N., Rey-Tienda, S., López-Sanz, M.E. (2025). Do ethical issues influence the interest of young people in using artificial intelligence? *Sustainable Technology and Entrepreneurship*, 4(3), 100108. <https://doi.org/10.1016/j.stae.2025.100108>
31. Microsoft (2024). *Work Trend Index Research and data on the trends reshaping the world of work*. Retrieved from: <https://www.microsoft.com/en-us/worklab/work-trend-index>, 06.03.2026.
32. Nigar, M., Juli, J.F., Golder, U., Alam, M.J., Hossain, M.K. (2025). Artificial intelligence and technological unemployment: Understanding trends, technology's adverse roles, and current mitigation guidelines. *Journal of Open Innovation: Technology, Market, and Complexity*, 11(3), 100607. <https://doi.org/10.1016/j.joitmc.2025.100607>
33. OECD (2024). *OECD AI Principles and Policy Observatory: Advancing trustworthy and sustainable AI*. Retrieved from: <https://oecd.ai/en/>, 06.03.2026.
34. Orlikowski, W.J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science*, 3(3), pp. 398-427, <https://doi.org/10.1287/orsc.3.3.398>
35. Pilny, A., Endacott, C., Treem, J.W. (2025). *AI in the Workplace*. Wiley-Blackwell.
36. PwC (2024a). *Ready for artificial intelligence*. PwC International Limited. Retrieved from: <https://www.pwc.pl/en/publikacje/ready-for-artificial-intelligence.html>, 06.03.2026.
37. PwC (2024b). *Global NextGen Survey 2024: Success and succession in an AI world*. PwC International Limited. Retrieved from: <https://www.pwc.com/gx/en/issues/c-suite-insights/nextgen.html>, 06.03.2026.

38. Rahman, S., Adeel, S., Ali, M., Bajaba, S., Latan, H. (2025). Power of green capabilities and artificial intelligence (AI): Understanding how and when green innovation promotes sustainability. *Business Strategy and the Environment*. Advance online publication, <https://doi.org/10.1002/bse.70252>
39. Rogers, E.M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
40. Sarstedt, M., Ringle, C.M., Hair, J.F. (2019). Treating unobserved heterogeneity in PLS-SEM: A multi-method approach. In: H. Latan, R. Noonan (Eds.), *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications* (pp. 197-217). Springer.
41. Stilgoe, J., Owen, R., Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), pp. 1568-1580, <https://doi.org/10.1016/j.respol.2013.05.008>
42. Tabbakh, A., Al Amin, L., Islam, M., Mahmud, G.M.I., Chowdhury, I.K., Mukta, M.S.H. (2024). Towards sustainable AI: A comprehensive framework for Green AI. *Discover Sustainability*, 5, Article 408. <https://doi.org/10.1007/s43621-024-00641-4>
43. Tetreiová, L., Kotková Striteská, M., Kuba, O., Prakash, V., Prokop, V. (2025). When trust and distrust come into play: How green concern, scepticism and communication affect customers' behaviour? *Business Strategy and the Environment*, 34(3), pp. 3311-3337, <https://doi.org/10.1002/bse.4144>
44. Toderas, M., Baciuc, C., Ionescu, L. (2025). Artificial intelligence for sustainability: A systematic review and critical analysis of AI applications, challenges, and future directions. *Sustainability*, 17(17), 8049. <https://doi.org/10.3390/su17178049>
45. Verdecchia, R., Sallou, J., Cruz, L. (2023). A systematic review of Green AI. *WIREs Data Mining and Knowledge Discovery*, 13(4), e1507. <https://doi.org/10.1002/widm.1507>