

TECHNOLOGY

AI Initiatives Don't Fail - Organizations Do: Why Companies Need AI Experimentation Sandboxes and Pathways

by Tobias Straube, Olaf J. Groth, and Jöran Altenberg



Image Credit | Emiliia

Consider improving the organizational setup.

☑ INSIGHT | FRONTIER 22 May 2025

The Costs of Failure to Adopt AI

AI is projected to generate \$7 trillion in value through generative AI alone,¹ and is expected to boost U.S. labor productivity by 0.5 to 0.9 percentage points annually through 2030.² When combined with other automation technologies, generative AI could drive productivity growth to 3-4% per year. Experiments across 18 knowledge-based tasks show these gains are driven by increasing speed by 25% and quality by 40%.³ This has the potential to transform industries and accelerate sustainable development. However, in 2025, 30% of enterprise generative AI projects are expected to stall, due to poor data quality, inadequate risk controls, escalating costs or unclear business value⁴ - findings also echoed by Deloitte.⁵ RAND research highlights that over 80% of AI projects fail⁶ and Goldman Sachs questions whether the estimated \$1 trillion in AI capital expenditures over the coming years will ever deliver a meaningful return.⁷ AI isn't a magic bullet—without the right structures, companies will spend heavily, only to write those investments off when projects collapse.

RELATED CMR ARTICLES

Marcus Holgersson, Linus Dahlander, Henry Chesbrough, and Marcel Bogers, "**Open Innovation in the Age of AI**," California Management Review, 67/1 (2024): 5-20.

Rebecka C. Ångström, Michael Björn, Linus Dahlander, Magnus Mähring, and Martin W. Wallin, "**Getting AI Implementation Right: Insights from a Global Survey**," California Management Review, 66/1 (2023): 5-22.

The Real Bottleneck Isn't AI As a Technology -It's How Companies Set Up Experimentation Pathways

AI isn't waiting for legacy organizations to catch up. Technology advances exponentially, while legacy organizations remain stuck in outdated adoption cycles. So, timing is key. But the bigger challenge is often the missing capability of designing purposeful experiments with pathways to measurable impact. AI doesn't fail because the technology isn't ready but because enterprise structures aren't. Companies eager to embrace AI all too often launch pilot projects and encourage employees to acquire new applications, but they ignore that such early experimentation requires fertile ground, like a petri dish receiving cultures in a lab. Without the right substrate in terms of organizational design, the experiments can't grow. Instead, dysfunctional dynamics take hold and lead to proverbial tissue-rejection: Regulatory complexity, rigid IT architectures, worker council conflicts, and risk-averse governance slow decision-making to a crawl. Power structures meant to create stability instead become bottlenecks for renewal, resulting in AI adoption and impact demonstration happening too slowly or being disconnected from real business needs.

Meanwhile, employees eager to minimize frictions in their daily tasks or scared of falling behind on productivity are already using AI - 68% of professionals using ChatGPT at work don't inform their supervisors⁸, not allowing organizations to harness its full potential. Worse, unguided adoption exposes companies to compliance risks, data security challenges, and unmanaged AI-driven decision-making.

AI Sandboxes: The Missing Link Between Hype and Real Impact

The challenge isn't deciding whether to invest in AI—it's about building the internal capacity to continuously adapt to its rapid evolution within legacy systems. AI can't be treated like a traditional IT rollout; it must become an ongoing and constantly upgrading cognitive decision making capability, requiring an agile approach to experimenting,

refining, and integrating AI solutions at the pace of technological change. Yet, organizational structures are inherently resistant to disruptive ways of working—a welldocumented challenge in innovation management. The proven solution? Creating protected spaces where innovations can be nurtured, monitored and calibrated before being integrated into existing structures and processes. This concept of secure innovation spaces must now be applied to AI. AI Sandboxes serve this role, providing controlled environments to experiment with AI applications without disrupting core IT systems or business operations. This approach acknowledges a key reality: AI innovations outside companies evolve faster than legacy cultures and centralized IT architectures can adapt. By enabling ringfenced experimentation with greater permeability between external and internal expertise and technology, sandboxes not only upskill and motivate employees, create critical insights forAI procurement and governance decisions but they create a "ground truth" fact base that is a reality check for theoretical concerns and obstacles.

The Road to Scalable AI Adoption: The Sandbox

To bridge the gap between AI experimentation and real-world execution, organizations need a structured approach that ensures pilots don't remain isolated tests but evolve into scalable, value-generating initiatives. The AI Capabilities Readiness Framework provides exactly that: a structured, metric-driven model to track AI projects from initial exploration to enterprise-wide adoption.



It maps AI initiatives along two key dimensions—technical feasibility and adoption readiness—and defines clear transition points where projects either progress towards fullscale deployment or risk stalling. By identifying three critical stages—(1) the current state of an AI project, (2) a realistic near-term target (3-6 months), and (3) the ideal state in 6-12 months—this framework enables companies to quantify the expected impact of AI investments and prioritize initiatives that deliver measurable ROI. The size of each stage's "bubble" represents the projected benefit, whether through faster time-to-market, strategic differentiation, or deeper technical capabilities. When integrated into AI Sandboxes, this readiness framework ensures that experimentation is not an end in itself but a structured pathway to execution, allowing organizations to systematically test, refine, and fast-track AI initiatives that demonstrate real business value.

From Experimentation to Execution: Building AI Strategy & Governance for Scale:

Much like an innovation funnel, an AI sandbox requires clear processes and structures to identify, test, and refine AI applications based on business needs and emerging ideas. Organizations must establish criteria for selecting use cases, define who has access to the sandbox, and determine how AI experiments are evaluated for potential integration into core systems. To ensure structured decision-making, organizations should define key performance indicators (KPIs) - such as speed improvements or quality gains - that determine whether an AI use case or application is considered successful and be considered for larger deployment. Experimentation should be supported, not left to chance, requiring AI coaches to provide guidance and expertise on demand. Finally, a lean,

agile decision-making committee with key stakeholders from IT, compliance, and business units should oversee the sandbox, ensuring that successful AI solutions are fast-tracked while maintaining alignment with strategic priorities and regulatory requirements. This governance model ensures AI experimentation is structured, accountable, and impactdriven—rather than an unfocused playground for AI enthusiasts.

Laying the Foundations for Scalable AI Integration: A well-designed AI sandbox must provide flexible access to a constantly evolving suite of AI tools and solutions, enabling employees to explore and experiment with the latest technologies. This includes end-user AI applications such as Canva, Invideo AI, DataRobot, or Midjourney, if these tools are not available within the organization's core IT infrastructure. Beyond ready-to-use applications, the sandbox should also offer API-level access to AI models, allowing teamswith the support of AI coaches—to integrate these capabilities into custom workflows. Whether deployed on-premise or in the cloud, the sandbox must be technically equipped to support secure AI experimentation, including model fine-tuning, customized RAG models, automation testing, and multi-tool interoperability. But sandboxes are also a proven mechanism for multi-stakeholder feedback and co-creation. For instance, depending on industry or use case, if corporate lawyers need to be involved for compliance monitoring or to adjust their own assessment of risk. or if workers councils are skeptical about impact on the future of their jobs, they can be invited to inputs or continue to innovate for improvements. Finally, AI coaches should play a crucial role in helping teams move beyond standalone AI applications, guiding them in embedding AI into end-to-end process automation. To enable this, the sandbox should feature scalable compute resources, secure API gateways, data processing pipelines, and integration frameworks that allow AI-driven solutions to seamlessly transition from testing to full-scale deployment.

Secure Data, Smarter AI - Laying the Groundwork for Safe AI Testing: Data is the backbone of AI development, especially for LLM-based solutions. However, one of the biggest challenges for organizations is ensuring secure handling of sensitive information. A major risk is that corporate data might unintentionally be absorbed into external AI models or leaked through AI-generated outputs. This concern is amplified by Retrieval-Augmented Generation (RAG), a technique that enriches AI responses with an organization's own data sources, making outputs less prone to hallucinations, but also

potentially increasing exposure to security vulnerabilities. To mitigate these risks, an AI sandbox should follow a phased approach. In the initial phase, the sandbox should operate exclusively with publicly available, scraped, or curated datasets, preventing any internal or confidential data from being exposed. In a second phase, organizations can introduce on-premise LLMs, allowing internal data to be leveraged for AI-driven solutions while maintaining full control over data security. Within this setup, a clear data classification framework is essential—distinguishing between non-sensitive data, sensitive but anonymized data, and restricted data that must never be processed, especially personal information governed by GDPR, CCPA, or industry-specific regulations. By structuring the AI sandbox in this way, companies can create a secure, compliant testing environment that fosters AI innovation without compromising data privacy or regulatory requirements. This approach not only enables responsible AI experimentation but also ensures that AI solutions can be developed and deployed with built-in security and governance from the start.

The AI Dilemma: Innovate or Regulate? Why You Need Both: While compliance defines which organizational practices are (or are not) legally allowed, AI ethics guides what we should (or should not) do to ensure responsible and fair AI use. In Europe, the AI Act, which entered into force in 2024 with full enforcement by 2026, establishes a risk-based regulatory framework for AI applications. Beyond legal obligations, many organizations have implemented AI ethics or trustworthiness frameworks to build trust, prevent reputational risks, and mitigate harms such as bias, misinformation, and privacy violations. This dual imperative-regulatory compliance and ethical responsibilityrequires that the AI sandbox evaluate models not only for performance and efficiency but also for explainability, transparency, and accountability. To achieve this, depending on the use case, organizations may have to integrate bias detection tools, model-cleaning, data logs, automated audits, stakeholder impact analysis, red-teaming, and explainability mechanisms into the sandbox infrastructure. A well-structured AI sandbox is more than just a practical testing ground—it acts as a governance systems builder which fosters a regulatory environment fit for purpose, ensuring AI solutions that get adopted also align with legal mandates and ethical principles. By embedding compliance and ethics into sandbox workflows, organizations can de-risk AI adoption while fostering trust and longterm viability. The EU AI Act not only sets a compliance framework but also encourages Regulatory Sandboxes as a way to foster innovation, reduce uncertainty, and support

SMEs, enabling AI development in a controlled, compliant environment (Art. 57). But while the Act focuses on national-level sandboxes, the slow pace of AI adoption suggests that organizations can't afford to wait for these to trickle down to individual organizations. Rather, these organizations should experiment with sandboxes bottom-up. In this way, data from sandboxes can also drive productive dialogue with regulators on how to improve regulation.



It's time to bring the idea of Sandboxes inside legacy organizations—not to replace centrally managed IT, but to complement it, creating protected spaces where AI can be tested, refined, and fast-tracked for real impact. The future of AI won't be built in central IT and HQ units alone—it will be shaped by those who give their teams the freedom to experiment in thoughtful and monitored ways

References

- 1. https://www.goldmansachs.com/insights/articles/generative-ai-could-raiseglobal-gdp-by-7-percent.html
- 2. https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-ofwork-in-america
- 3. https://www.hbs.edu/ris/Publication Files/24-013_d9b45b68-9e74-42d6-a1c6c72fb70c7282.pdf

- 4. https://www.gartner.com/en/newsroom/press-releases/2024-07-29-gartnerpredicts-30-percent-of-generative-ai-projects-will-be-abandoned-after-proof-ofconcept-by-end-of-2025
- 5. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/usstate-of-gen-ai-q3.pdf
- 6. https://www.rand.org/pubs/research_reports/RRA2680-1.html
- 7. https://www.goldmansachs.com/images/migrated/insights/pages/gsresearch/gen-ai-too-much-spend,-too-little-benefit-/TOM_AI 2.0_ForRedaction.pdf
- 8. https://www.fishbowlapp.com/insights/70-percent-of-workers-using-chatgpt-atwork-are-not-telling-their-boss/



Tobias Straube (Follow)

Tobias Straube is VP Analysis at Cambrian Futures, Assistant Lecturer at UC Berkeley Executive Education, as well as a lecturer at Steinbeis Business School. He is a co-founder of Scio-Network and previously served as an Innovation Manager for the German Society for International Collaboration (GIZ).



Olaf J. Groth (Follow)

Olaf Groth, PhD, is a Professor of Practice at UC Berkeley Haas School of Business, Adjunct Professor at UT Malaysia, and a member of the AI Governance Alliance of the World Economic Forum. He also serves as CEO of concept studio Cambrian Labs and of Cambrian Futures, a global network of analysts, strategists and futurists for transformations of enterprises and economies driven by AI, data and emerging technologies. He is the author of two books - The AI Generation and The Great Remobilization.



Jöran Altenberg (Follow)

Transformation Leader at GIZ Jordan, working at the intersection of change management, digital transformation, and AI adoption. He explores how the global digital divide shapes innovation, governance, and economic transformation. This article reflects his personal views and does not represent the positions of Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ).