

Oliver, Miquel; López-Fernandez, Daniel

Conference Paper

Bridging the Hype and Reality of Generative AI: Insights from Real-World Enterprise Implementations

ITS 33rd European Conference 2025: "Digital innovation and transformation in uncertain times", Edinburgh, UK, 29th June – 1st July 2025

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Oliver, Miquel; López-Fernandez, Daniel (2025) : Bridging the Hype and Reality of Generative AI: Insights from Real-World Enterprise Implementations, ITS 33rd European Conference 2025: "Digital innovation and transformation in uncertain times", Edinburgh, UK, 29th June – 1st July 2025, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/331294>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

ITS 33rd European Conference 2025

Edinburgh, Scotland, 29th June – 1st July 2025

Bridging the Hype and Reality of Generative AI: Insights from Real-World Enterprise Implementations

Authors: **Miquel Oliver**, **Daniel López-Fernandez**, Universitat Pompeu Fabra, Spain

Abstract

Generative AI (GenAI) has rapidly gained traction as a transformative technology, promising to revolutionize business processes across various sectors. However, the gap between theoretical potential and real-world implementation remains substantial, with many enterprises struggling to transition from experimental pilots to scalable applications. This study critically examines 31 GenAI projects from different industries evaluating the strategic, technical and operational factors that influence their success or stagnation.

To ensure a methodologically rigorous and comparable analysis, we adopt a mixed-methods approach that integrates primary data—collected through direct project involvement and stakeholder feedback—with secondary data from project documentation, industry reports, and academic research. Our case selection is based on practical exposure to enterprise GenAI implementations, focusing on projects where organizations have actively engaged in experimentation and deployment efforts. Rather than applying rigid selection criteria, we leverage first-hand access to real-world implementations to extract patterns, challenges, and success factors across different organizational contexts. This approach allows us to identify both industry-specific challenges and broader cross-sectoral trends in GenAI adoption.

Our findings indicate that most GenAI projects remain stalled at the PoC phase, with organizations facing significant obstacles in scalability, predictability, and enterprise-wide integration. The primary barriers include: (1) Unpredictability of AI outputs, limiting reliability in mission-critical applications; (2) Regulatory, ethical, and legal concerns, particularly in relation to data privacy, intellectual property, and compliance with evolving AI governance frameworks. (3) Fragmented AI strategies within enterprises, leading to siloed innovation efforts and a lack of alignment between technical development and business objectives.

Despite these challenges, early indicators of business value are emerging, particularly in process automation, knowledge augmentation, and customer interaction enhancement. However, without robust governance structures, clear strategic alignment, and adaptive technical strategies, enterprises risk failing to realize the long-term benefits of GenAI.

This research contributes to both academic and industry discourse by providing a nuanced, multi-sector examination of GenAI implementation. By identifying critical success factors and common failure points, we offer actionable insights for enterprises aiming to transition from experimentation to sustainable deployment. The study ultimately highlights the conditions under which GenAI can move beyond hype to deliver tangible enterprise value, emphasizing the importance of human-AI collaboration, regulatory foresight, and business-driven AI adoption strategies.

1. Introduction

Generative AI (GenAI) has emerged as a groundbreaking innovation with the potential to reshape enterprise functions through automated content creation, decision support, and operational efficiency (Yin et al., 2024). The technology, particularly driven by large language models (LLMs) like GPT-4, offers significant promise in natural language processing, multimodal reasoning, and human-like interaction. Companies have responded with enthusiasm, channeling large investments into pilots and internal tools aimed at improving productivity, customer interaction, and decision-making (Jen Langusch et al., 2024).

However, alongside this enthusiasm lies a complex reality: many GenAI initiatives stall after the initial phase, struggling to meet business expectations due to lack of control, scalability issues, and absence of a robust governance framework for this non deterministic technology (Andreas Kremer et al., 2023). This paper aims to explore this gap between hype and reality through a grounded study of enterprise experiences across sectors.

2. Methodology

This study is based on a mixed-methods approach combining direct involvement in 31 GenAI enterprise projects with a structured review of internal documentation, deliverables, and stakeholder interviews. The projects were conducted between mid-2023 and mid-2024, encompassing both public (26%) and private (74%) sectors.

Data Sources:

- **Direct Experience:** Article authors involved in presales, delivery, and closure of most projects.
- **Documentation Review:** Meeting minutes, kickoff decks, technical reports, and final presentations.
- **Stakeholder Feedback:** Talks with IT managers, business owners, and end-users.

The selection of cases prioritized diversity in use cases (e.g., software development, HR support, ticketing, document processing), sector (finance, government, telecom, etc.), and stage of maturity (PoC, pilot, production).

3. Key Findings

3.1 Implementation Status and Investment Patterns

Among the 31 GenAI initiatives analyzed, only 13% advanced to pilot or production stages. Most projects lasted less than six months and had budgets ranging from €10,000 to €50,000. These figures reflect the experimental nature of GenAI adoption and the widespread use of lightweight PoCs.

3.2 Classification of Use Cases

Projects were classified into the following categories:

- **Application Development:** Using GenAI copilots to accelerate coding, testing, and documentation (Pandey et al., 2024).

- **Documentation Automation:** Summarizing legal texts, contracts, RFPs, and converting unstructured data into structured reports (Wang et al., 2022).
- **Smart Assistants:** Building conversational agents for HR, IT support, and intranet navigation (Park et al., 2023).
- **Ticketing Optimization:** Automatically classifying and prioritizing incidents, detecting sentiment (Truss and Böhm, 2024).
- **Operational Workflow:** Natural language interfaces to manage inventory, production, and calendars (Decardi-Nelson et al., 2024).
- **Multimodal Processing:** Interpretation of handwritten documents and complex infographics (Combs et al., 2024).
- **Maintenance (AIOps):** Anomaly detection and root cause identification in IT operations (Ucar et al., 2024).
- **Data Anonymization:** Masking sensitive user data before exposure to LLMs (Yermilov et al., 2023).
- **Productivity Enhancement:** Embedding GenAI copilots in Office tools, email, and CRM systems (Acharya, 2023).

3.3 Analysis of GenAI Use Cases in Large Enterprises

A detailed cross-case analysis was conducted to synthesize the challenges and impacts associated with the various use cases observed. The table below summarizes the most representative GenAI application areas:

Use Case	Key Features	Challenges	Findings
Application Development	Code acceleration, legacy migration, documentation	Productivity highly context-dependent	Useful in controlled environments; scalability varies
Documentation Automation	RFP generation, summarizing legal/business texts	Limited scalability, integration complexity	Effective for entry-level GenAI use; requires customization
Smart Assistants	HR bots, intranet support, virtual educators	Knowledge grounding, interaction reliability	Widely adopted in internal use; external use limited by accuracy concerns
Ticketing Optimization	Categorization, prioritization, sentiment analysis	Complex workflows, limited natural language understanding	Boosts support team efficiency; accuracy still evolving
Operational Workflow	Natural language control of business tasks	ERP integration, command interpretation	High potential for democratizing process control; adoption still experimental
Multimodal Processing	Handwriting recognition, graphical interpretation	Ambiguity in data, need for contextual tuning	Valuable for public records and insurance; requires continuous learning
Maintenance (AIOps)	Pattern detection, predictive maintenance	Data preparation, real-time interpretation	Additive to traditional systems; not a standalone solution
Data Anonymization	Masking of sensitive fields on-the-fly	Awareness and integration gaps	Emerging use case with strong compliance value

Productivity Enhancement (Copilots)	Writing support, email drafting, spreadsheet automation	Trust issues, performance validation	High end-user satisfaction; limited rollout due to cost and control barriers
-------------------------------------	---	--------------------------------------	--

Table 1. Challenges and findings in application areas

This structured analysis reinforces that while GenAI demonstrates meaningful value in enhancing operational efficiency and support functions, the maturity of the tools and their ability to meet enterprise-grade standards vary significantly across use cases.

3.4 Strategic and Technical Barriers

a) Unpredictability and reliability

Even when GenAI models achieve high average accuracy, their occasional “hallucinations” can invalidate entire workflows in finance, healthcare, and public services. Interviewees reported that a single erroneous answer in loan-risk assessment or clinical documentation is enough to block a roll-out, regardless of PoC success. Enterprises therefore deploy GenAI behind human-review “guardrails,” reducing time-savings and making ROI less compelling.

b) Enterprise-wide integration gaps

Most projects were incubated within innovation or IT units and never connected to core enterprise systems (ERP, CRM, HRIS). Integration hurdles fall into three layers:

- **Technical** – secure API gateways, data-model alignment, latency constraints;
- **Process** – change-management for existing SOPs, retraining of frontline staff;
- **Data Governance** – lineage tracking and version control for dynamic GenAI outputs.

Projects that progressed to pilot typically solved the technical layer but underestimated the process and governance layers, leading to stalled adoption beyond a single department.

c) Fragmented AI strategy and ownership

Only 5 of the 31 organizations had a formal GenAI roadmap linking pilots to business KPIs. In the remainder, GenAI initiatives competed for resources, duplicated effort, and lacked executive sponsorship. This fragmentation was cited by business leaders as a bigger blocker than model accuracy itself.

d) Regulatory, ethical, and legal pressures

Across EU-based projects, pending AI Act requirements around transparency, bias testing, and risk classification already influenced design decisions—even before formal enforcement. Public-sector respondents paused two pilots that performed behavioral prediction after data-protection officers raised GDPR concerns. Private-sector firms emphasized IP risks when training foundation models on proprietary code or client documents.

e) Skills and cultural resistance

Managers with limited ML literacy struggled to specify acceptance criteria, while domain experts resisted automated decision support, fearing loss of authority. Successful teams created cross-functional “tiger squads” pairing data scientists with process owners and compliance leads, but this practice was the exception rather than the norm.

f) Cost–benefit ambiguity

GenAI's cloud-based consumption pricing makes long-term cost forecasting difficult. CFOs therefore require clearer TCO models that factor in prompt-engineering labour, human-review overhead, and potential liability reserves. Without such models, scale-up funding is rarely approved.

These six barriers reinforce one another: lack of strategy amplifies integration pain, which in turn heightens cost ambiguity and regulatory risk. Overcoming them demands a holistic governance framework—combining centralized AI steering committees, standardised MLOps pipelines, and clear escalation paths for ethical review—before GenAI can graduate from isolated PoCs to enterprise platforms.

4. Conclusion

This study has examined the practical realities of GenAI implementation across 31 projects in diverse enterprise contexts, revealing a clear gap between the transformative potential often associated with GenAI and the limited scale of its real-world deployment. While many projects successfully validated the technology's promise in proof-of-concept stages—especially in tasks like document summarization, support ticketing, and productivity assistance—only a small fraction progressed beyond pilot phases.

A key insight from this research is the existence of a structural “80-20 gap.” Out-of-the-box GenAI solutions can deliver up to 80% of the needed functionality quickly and affordably. However, the remaining 20%—the part that requires domain-specific adaptation, contextual understanding, and compliance with organizational and regulatory standards—demands significant additional investment. For most enterprises, especially in regulated or high-risk environments like finance, public services, or healthcare, this final mile is the most difficult to bridge.

One of the most consistent barriers is the unpredictability of AI-generated outputs, which limits trust and usability in mission-critical applications. This unpredictability is not just a technical issue—it raises legal, ethical, and operational concerns. Enterprises cannot afford to deploy systems that disclaim responsibility or require manual post-review of all outputs. This leads to a cautious preference for human-AI collaborative models, where GenAI acts as a recommender or assistant rather than a fully autonomous agent.

Moreover, the lack of robust governance structures has created fragmented innovation landscapes within many organizations. GenAI projects are often initiated by enthusiastic IT teams but fail to achieve strategic buy-in from business leaders due to the absence of standardized evaluation criteria, accountability frameworks, or ROI models. The resulting silos hinder integration with existing systems and reduce long-term sustainability.

Regulatory foresight is another critical dimension. As legislation evolves around AI ethics, privacy, and liability, enterprises are increasingly aware of the risks tied to compliance. Some projects included in this study were paused or cancelled entirely due to unresolved concerns over bias, privacy, or potential misuse—especially in cases involving behavioral prediction or customer profiling.

Despite these challenges, the study reveals emerging zones of early business value. Internal-facing use cases—particularly those supporting knowledge augmentation, customer interaction, or administrative automation—consistently delivered measurable benefits in efficiency, cost savings, and user satisfaction. These low-risk, high-impact scenarios offer an effective entry point for enterprises starting their GenAI journey.

For organizations aiming to transition from experimentation to sustainable deployment, this study offers several actionable insights:

- Establish cross-functional AI governance boards to align innovation with business needs and compliance requirements.
- Prioritize use cases where human-AI collaboration adds value without requiring full automation.
- Invest in data governance and contextual tuning to ensure GenAI outputs are accurate, explainable, and auditable.
- Develop scalable frameworks to move successful PoCs into production environments with structured validation and change management.

In conclusion, GenAI holds real promise—but its path to maturity requires a shift from opportunistic experimentation to responsible, business-driven adoption. For GenAI to move beyond the hype and deliver enterprise-scale value, organizations must treat it not just as a tool, but as a transformational capability that demands integration, oversight, and long-term strategic thinking.

5. References

- Acharya, V.V., 2023. The impact of Microsoft 365 Copilot on small and medium-sized businesses 10.
- Andreas Kremer, Angela Luget, Daniel Mikkelsen, Henning Soller, Malin Strandell-Jansson, Sheila Zingg, 2023. As gen AI advances, regulators—and risk functions—rush to keep pace. Risk & Resilience. URL <https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/as-gen-ai-advances-regulators-and-risk-functions-rush-to-keep-pace> (accessed 10.9.24).
- Combs, K., Bihl, T.J., Ganapathy, S., 2024. Utilization of generative AI for the characterization and identification of visual unknowns. Natural Language Processing Journal 7, 100064. <https://doi.org/10.1016/j.nlp.2024.100064>
- Decardi-Nelson, B., Alshehri, A.S., Ajagekar, A., You, F., 2024. Generative AI and process systems engineering: The next frontier. Computers & Chemical Engineering 187, 108723. <https://doi.org/10.1016/j.compchemeng.2024.108723>
- Jen Langusch, James Temple, Joshua Bellin, Gaurav Khetan, Konrad Suchecki, Sanjiv Kumar Sharma, James Velasquez, Guy Norwel, Helen Miller, Nadia Malik, Rachel Jerome, 2024. IA Generativa para el crecimiento del cliente. Accenture Insights. URL <https://www.accenture.com/es-es/insights/song/generative-ai-customer-growth> (accessed 10.9.24).
- Pandey, R., Singh, P., Wei, R., Shankar, S., 2024. Transforming Software Development: Evaluating the Efficiency and Challenges of GitHub Copilot in Real-World Projects.
- Park, J.S., O'Brien, J., Cai, C.J., Morris, M.R., Liang, P., Bernstein, M.S., 2023. Generative Agents: Interactive Simulacra of Human Behavior, in: Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology. Presented at the UIST '23: The 36th Annual ACM Symposium on User Interface Software and Technology, ACM, San Francisco CA USA, pp. 1–22. <https://doi.org/10.1145/3586183.3606763>
- Truss, M., Böhm, S., 2024. 8 AI-based Classification of Customer Support Tickets: State of the Art and Implementation with AutoML. <https://doi.org/10.48550/arXiv.2406.01789>
- Ucar, A., Karakose, M., Kırımça, N., 2024. Artificial Intelligence for Predictive Maintenance Applications: Key Components, Trustworthiness, and Future Trends. Applied Sciences 14, 898. <https://doi.org/10.3390/app14020898>
- Wang, A.Y., Wang, D., Drozdal, J., Muller, M., Park, S., Weisz, J.D., Liu, X., Wu, L., Dugan, C., 2022. Documentation Matters: Human-Centered AI System to Assist Data Science Code Documentation in Computational Notebooks. ACM Trans. Comput.-Hum. Interact. 29, 1–33. <https://doi.org/10.1145/3489465>
- Yermilov, O., Raheja, V., Chernodub, A., 2023. Privacy- and Utility-Preserving NLP with Anonymized data: A case study of Pseudonymization, in: Proceedings of the 3rd Workshop on Trustworthy Natural Language Processing (TrustNLP 2023). Presented at the Proceedings of the 3rd Workshop on

Trustworthy Natural Language Processing (TrustNLP 2023), Association for Computational Linguistics, Toronto, Canada, pp. 232–241. <https://doi.org/10.18653/v1/2023.trustnlp-1.20>

Yin, S., Fu, C., Zhao, S., Li, K., Sun, X., Xu, T., Chen, E., 2024. A Survey on Multimodal Large Language Models.