**INNOVATIONS IN BUSINESS AND STRATEGIC MANAGEMENT**
Jan - Mar 2026, VOL. 4, ISSUE 1, pp. 36-40
https://doi.org/10.61577/ibsm.2026.100004

RESEAPRO JOURNALS

**ORIGINAL ARTICLE**

OPEN ACCESS

# Generative AI and large language models: An integrative review of organizational productivity, technical limits, and responsible deployment

**Abhina C**

**Assistant professor, College of applied science IHRD, Cheemeni, India**

**ABSTRACT**

Generative Artificial Intelligence (AI) and Large Language Models (LLMs) have been one of the most consequential technological developments of the current decade, fundamentally changing the manner in which information is produced, disseminated, and consumed by individuals, organizations and institutions. Since late 2022, generative systems have received a lot of attention in academia and industry, from education to healthcare, business and public administration. This paper offers a conceptual and integrative review on the evolution, architecture, uses, and consequences of generative AI and LLMs in the post-2021 peer-reviewed and working-paper literature. This review synthesizes architectural evidence for modern transformer models, evaluates empirical adoption metrics including worker adoption rates of ~28% and controlled experimental task-completion time reductions of up to 40%, and critically analyzes persistent risks including hallucination, algorithmic bias, data privacy, and intellectual property infringement. The paper argues that generative AI offers enormous gains in efficiency, accessibility, and digital intelligence systems, but responsible dissemination requires an operationalized attention to governance frameworks, workforce skill development, and domain-specific evaluation. The paper closes with actionable directions for future interdisciplinary research on trustworthy and sustainable deployment of generative AI.

**KEYWORDS**

Generative Artificial Intelligence; Large Language Models; ChatGPT; Digital Transformation; Transformer Architecture; AI Ethics

**ARTICLE HISTORY**

Received 20 January 2026;
Accepted 07 February 2026;
Published 15 February 2026

## Introduction

The arrival of ChatGPT in November 2022 was a watershed moment for the public visibility of generative artificial intelligence, transforming a niche area of natural language processing (NLP) research into a mainstream digital phenomenon [1]. Generative AI is a branch of machine learning that can generate new content such as text, images, code, audio, and video, rather than just classifying or predicting outcomes from existing data. Large Language Models (LLMs), the language side of generative AI, are trained on huge collections of text and can generate coherent, contextual, human-like language for a growing range of tasks [2].

The rapid spread of these tools has resulted in an unprecedented volume of scholarly output. Bibliometric evidence shows a rapid increase in the number of publications related to LLMs between 2017 and 2023, with a notable acceleration after the release of GPT-3 and subsequently ChatGPT [1]. This growth is fueled not only by technical curiosity but also by the widespread recognition that generative AI is reconfiguring digital intelligence, the ability of information systems to understand, reason about, and act on complex, unstructured data.

This paper reviews the state of the art in generative AI and LLMs with three main objectives. First, to describe the conceptual and architectural underpinnings that distinguish generative systems from previous discriminative AI. Second, to synthesise empirical evidence on their applications and economic impact in education, business, and other digital-intelligence settings. Third, to critically discuss the ethical, legal and societal challenges arising from their adoption. The paper is structured as follows. Section 2 describes the conceptual and technical underpinnings of generative AI. Section 3 reviews the growth and structure of research on LLMs. Section 4 reviews applications across sectors. 5 discusses economic and organisational impact. Section 6 discusses the ethical and societal challenges, and Section 7 concludes with directions for future research.

## Methodology/Review Approach

This review takes a conceptual and integrative approach rather than a systematic (PRISMA-style) approach, because the purpose is to synthesize a large, fast-moving body of literature, not to catalog it exhaustively.

Search strategy: The sources were found by systematic searching of academic databases and repositories such as Google Scholar, SSRN, arXiv, National Bureau of Economic Research (NBER), Scopus, and Web of Science. Our retrieval strategy included keyword strings such as "generative AI", "large language models", "LLM adoption", "AI productivity", and "AI governance and ethics". The search parameters mainly focused on literature published in 2021 and after to capture empirical findings and architectural developments that arose in the wake of the widespread acceleration of generative model deployment.

Inclusion/exclusion criteria: We focused on peer-reviewed articles, working papers from recognized institutions (e.g., NBER), and studies published in high-impact outlets (e.g., Science, Nature) or reputable preprint servers (arXiv, SSRN). Non-empirical sources, general web content, and material in non-English languages were excluded. Foundational or seminal works that pre-dated 2021 were retained for historical/technical context.

Synthesis approach: Instead of listing studies individually, sources were thematically grouped into the core sections of the paper (architecture, organizational adoption, productivity impact, and governance/ethics). Within each theme, empirical findings are presented, and the authors' interpretive synthesis links these to the overall argument of the paper. Explicit attribution is used to differentiate claims taken directly from the cited studies from the authors' own conceptual interpretations of those claims.

***Correspondence:** Dr. Abhina C, Assistant professor, College of applied science IHRD, Cheemeni, India. e-mail: abhinaculkl@gmail.com

© 2026 The Author(s). Published by Reseapro Journals. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Conceptual and Technical Foundations

Generative AI systems are fundamentally different from earlier discriminative machine learning models. Discriminative models are trained to classify or predict a label given an input, while generative models learn the underlying distribution of the training data in order to synthesise new, plausible outputs [3]. Most modern LLMs are based on the transformer architecture, which uses self-attention mechanisms to model the relationships between tokens over long sequences of text, allowing for parallelised training on enormous datasets, and leading to what researchers refer to as emergent capabilities as model scale increases [2].

Minaee et al. surveyed key LLM families such as the GPT, LLaMA and PaLM series and reported that model performance on general-purpose language understanding and generation tasks, in agreement with well-established scaling relationships, generally scales with increases in parameter count, training data volume and computational budget [2]. But the same review points out that scale alone does not overcome fundamental limitations such as factual inconsistency, brittleness in reasoning, and the tendency to generate plausible but wrong content, often called "hallucination". As discussed in Section 6, these limitations call for architectural improvements, which need to be accompanied by robust techniques for model reliability assessment [4].

The architecture's dependence on statistical sequence prediction produces a built-in tension between fluency and epistemic accuracy. Transformer models optimize for next-token probability distributions, not verifiable causal ground truth, so fluent syntax can hide factual errors or invented reasoning paths. As parameter scale increases, these models show emergent abilities on ambiguous linguistic tasks, but their failure modes become subtler and harder for human evaluators to detect at the same time. This technical limit matters for digital intelligence systems: high predictive fluency can engender misplaced user trust, as non-expert users treat the erroneous output as authoritative. Thus, the journey from foundational model capabilities to domain-specific reliability needs not only increased scale, but also architectural interventions such as constrained decoding, grounding against structured knowledge graphs, and verifiable output provenance.

## The Growth and Structure of LLM Research

The academic landscape of LLMs has evolved rapidly and unevenly across different fields. Fan et al. conducted a bibliometric and discourse analysis synthesising over 5,000 publications from 2017 to 2023 [1]. LLM research initially concentrated on core NLP tasks before spreading to applied domains such as medicine, engineering and the social sciences, while collaboration networks became more interdisciplinary and international. This pattern of diffusion is similar to earlier waves of diffusion of general-purpose technologies, where a base innovation is progressively embedded in different application domains as complementary expertise emerges.

Alongside this boom in applied research, there has emerged a separate body of methodological literature dedicated to the evaluation of LLMs themselves. Laskar et al. conducted a systematic survey of LLM evaluation practices and highlighted persistent challenges including inconsistent benchmarking protocols, contamination of training corpora with test data, and the lack of standardized metrics for factuality, and safety [4]. As the authors argue, without more robust and transparent evaluation frameworks, claims of model capability and reliability are difficult to compare across studies - a point with direct implications for how organisations choose and govern generative AI tools in practice.

## Applications Across Digital Intelligence Domains

### Business, Industry, and Knowledge Work

Generative AI has penetrated organisational contexts at what is being reported by industry surveys as an unprecedentedly rapid pace. McKinsey & Company's global survey work, conducted in 2023, found that generative AI had been deployed across nearly all business functions, with the highest concentration of applications in marketing, product development, and service operations [5]. A follow-up survey in early 2024 showed that about 65 percent of the respondent organisations were using generative AI regularly, roughly double the share found ten months earlier.

This is supported by field-experimental evidence with direct productivity measurement confirming these adoption trends. Noy and Zhang implement a randomised experiment on writing tasks, and find that access to generative AI reduced task completion time substantially, while improving output quality, with the largest gains accruing to lower-performing workers [6]. Similarly, Brynjolfsson et al. studied the rollout of a generative AI assistant among customer support agents and found measurable increases in issue resolution per hour, again concentrated among less experienced employees [7]. This suggests generative AI can narrow performance gaps within a workforce rather than uniformly amplifying existing advantages.

At the macro level, using nationally representative survey data, Bick et al. estimated that adoption of generative AI by U.S. workers had reached roughly 28 percent by late 2024 [8]. However, formal adoption at the firm level remained relatively low, suggesting that individual, informal experimentation is outpacing institutional integration at this point. Complementing this, Gulati et al. examined job postings between 2021 and 2023 and found that organisations recruiting explicitly for generative AI-related skills increasingly emphasised higher-order competencies such as critical evaluation and complex problem-solving, suggesting that the technology is reshaping not only task execution but also the skill composition demanded of digital-intelligence-related roles [9].

### Education and Academic Practice

Education has been among the domains most visibly affected by generative AI. Kasneci et al. argue that LLMs offer opportunities for personalised learning support, automated feedback, and accessibility enhancements for students with disabilities, while cautioning that these benefits are contingent on careful pedagogical integration and safeguards against over-reliance [10]. Chiu examined how ChatGPT and Midjourney influenced learning, teaching, assessment, and administration from the perspective of school teachers and leaders, concluding that generative AI is simultaneously perceived as disruptive and as an opportunity to reinforce more innovative educational practices [11].

In higher education specifically, Dempere et al. reviewed institutional and classroom-level responses to ChatGPT and found a wide range of reactions, from enthusiastic pedagogical


© 2026 Reseapro Scientific Services
https://doi.org/10.61577/ibsm.2026.100004

adoption to outright prohibition, reflecting unresolved tensions between the technology's capacity to support learning and legitimate concerns about academic integrity [12]. AL-Smadi similarly surveyed Scopus-indexed literature on generative AI in education published between November 2022 and July 2023 and reported a rapid expansion of empirical studies addressing student use, assessment design, and instructor training, alongside a comparative scarcity of longitudinal evidence on learning outcomes [13].

### Other Digital Intelligence Applications

Beyond business and education, LLMs are increasingly embedded within recommender systems, interactive media, and creative industries. Said reviewed the use of LLMs to generate natural-language explanations for algorithmic recommendations, finding that such explanations can improve user trust and transparency, though evaluation methodologies for explanation quality remain inconsistent across studies [14]. Sweetser conducted a scoping review of 76 papers on LLMs in video game research and development published between 2022 and early 2024, documenting applications in non-player character dialogue, procedural narrative generation, and game design tools, while noting that most work remains exploratory rather than deployed at scale [15]. Collectively, these applications illustrate that generative AI's transformation of digital intelligence extends well beyond text generation into interactive, multimodal, and decision-support contexts.

## Economic and Organisational Impact

Evidence on productivity presented in Section 4.1 suggests a broadly positive, but unevenly distributed economic impact of generative AI adoption. By combining observed usage intensity with the potential of generative AI, Bick et al. estimate that measurable aggregate labour productivity gains could be generated [8]. They caution that realised gains will depend on the extent to which time savings are redirected towards output versus absorbed in the form of informal leisure by workers whose employers do not have visibility into individual usage. Gulati et al. extend this economic perspective by demonstrating how the skill-based effects of generative AI adoption are role-specific, with some roles receiving skill enhancement and others experiencing task displacement [9]. This finding aligns with the broader theoretical perspectives that differentiate between the enhancing, displacing, and reinstating effects of new technologies on work.

The empirical divergence in organizational form implies an important distinction between pure technological substitution and task-level complementarity. Generative AI doesn't replace whole jobs. It automates specific subtasks like writing initial text, boilerplate code, and retrieving routine information. It also shifts the value of human work to checking tasks, synthesizing context, and providing strategic direction. These LLMs help lower-skilled workers surmount simple execution barriers, increasing the first floors of organizational capability. But sustaining those productivity gains across the enterprise units has a lot of operational friction. These unmonitored shadow AI practices that are threats to intellectual property and enterprise data integrity are often a byproduct of informal, bottom-up adoption by individual employees who are often ahead of formal enterprise governance. Second, the economic value of generative deployments depends on the extent to which institutions can organize workflows around augmented output. If productivity gains only result in unchecked slack time or the accumulation of redundant, unvalidated synthetic text, net organizational performance may stagnate despite high adoption metrics. Turning individual gains in efficiency into sustainable collective productivity requires purposeful managerial intervention. That means rewriting job descriptions to highlight the skills that will be evaluated, codifying procedures for cross-functional oversight and setting clear operational standards against which to measure the quality of the output and the speed of delivery.

Taken together, these findings point to the organisational implications of generative AI that are not necessarily captured by adoption rates but rather by the intensity, distribution and governance of use. Companies that treat generative AI as a productivity tool to be purchased and used without complementary investment in training, workflow redesign and quality control are likely to realise smaller and less durable benefits than those that embed it more deliberately into organisational processes.

A structured synthesis of the core empirical investigations evaluating generative AI adoption, task performance, and labour market impacts is summarized in Table 1.

**Table 1.** Summary of Empirical Evidence on Generative AI Adoption and Productivity

| Author(s) & Year | Domain / Focus Area | Sample & Methodology | Key Findings & Labor Implications |
|---|---|---|---|
| Noy and Zhang [6] | Professional writing and white-collar analytical tasks | Randomized controlled trial (RCT) with 453 college-educated professionals completing occupation-specific writing assignments | Substantially reduced task completion time while increasing average output quality; productivity gains were disproportionately concentrated among lower-performing workers, narrowing initial skill gaps. |
| Brynjolfsson et al. [7] | Customer service and enterprise software support | Staggered organizational rollout tracking 5,179 customer support agents using a generative AI conversational assistant | Resolution rates per hour increased significantly, with minimal impact on top-tier workers and substantial performance acceleration for novice and lower-skilled agents. |
| Bick et al. [8] | Economy-wide labor force adoption | Nationally representative survey of U.S. workers measuring direct utilization of generative tools across occupations | Found individual worker adoption reached ~28% by late 2024, yet formal enterprise adoption lagged; aggregate productivity gains depend on whether time saved is converted into productive output rather than unmonitored leisure. |
| Gulati et al. [9] | Labor demand, organizational skills, and job postings | Empirical analysis of firm job postings spanning 2021 to 2023 evaluating demand for generative AI capabilities | Identified role-specific polarization: routine execution tasks face substitution risks, while postings seeking AI skills increasingly require higher-order competencies such as critical evaluation and complex problem-solving. |

## Ethical, Legal, and Societal Challenges

### Hallucinations and Reliability


© 2026 Reseapro Scientific Services
https://doi.org/10.61577/ibsm.2026.100004

One of the biggest challenges of generative AI is its capacity to generate content that is fluent but factually incorrect or fabricated, also referred to as hallucination. Laskar et al. highlight hallucination as a significant problem in trustworthy LLM applications [4]. They note that existing evaluation benchmarks are frequently unable to identify minor factual errors, particularly in specialized or high-risk areas. This problem is exacerbated by the general-purpose evaluation problems mentioned above: inconsistent benchmarking makes it hard to compare the reliability of different systems across use cases.

### Fairness, Bias and Representation

Generative artificial intelligence systems are trained on large, imperfect training corpora that can encode and reproduce societal biases, learning statistical regularities. In their scoping review of the generative AI ethics literature, Hagendorff developed a taxonomy of 378 distinct normative issues, categorized into 19 topics, with fairness, safety, and harmful content being among the most prevalent [3]. However, the review notes that it is important not to interpret these findings as a complete picture of risk, as the underlying literature shows a negativity bias that may underestimate potential benefits relative to harms – an important methodological caveat for researchers synthesizing this body of work.

### Confidentiality, Security, and Intellectual Property

Training and deployment of LLMs raise major issues of privacy and intellectual property. Mühlhoff argues that predictive artificial intelligence systems, including generative models, pose risks to the loss of "predictive privacy," where inferences drawn from aggregated data can reveal sensitive information about people who did not directly contribute data to the training process, and collective, not only individual, data protection strategies are required [16]. Wu et al. also mention security, privacy, and ethical issues related to ChatGPT, including the possibility of sensitive information being disclosed through user prompts and the use of generated material for purposes such as phishing or disinformation campaigns [17]. At the regulatory level, Novelli et al. (2 study how generative AI liability, privacy, intellectual property, and cybersecurity risks are addressed in European Union law and conclude that existing legal structures were mostly created before the advent of foundation models and thus need to be targeted to deal with issues such as the copyright status of training data and the allocation of liability for content produced by AI [18]. The findings demonstrate that generative AI governance is a moving target across jurisdictions, not an accomplished task.

### Towards Responsible Deployment

Persistent risks of epistemic unreliability, systemic bias, and data vulnerability show that safe integration cannot depend on isolated technical fixes but demands a multi-tiered, operationalized framework for responsible deployment. One of the key architectural pillars of this is the separation of pre-training alignment from ongoing operational oversight by humans. Reinforcement Learning from Human Feedback (RLHF) and direct preference optimization steer models toward a general policy compliance in the initial training, but do not guarantee accuracy in specialized environments. Therefore, workflows in the enterprise and public sectors need to integrate rigorous human-in-the-loop (HITL) verification when producing output. For high-stakes fields such as clinical healthcare, credit adjudication, and academic grading, human practitioners should be the authoritative gatekeepers who evaluate context, check citations, and validate decisions, rather than fully handing the end-to-end task over to autonomous agents. Human review and strong technical and auditing safeguards must be built directly into the software pipeline. Organizations are increasingly coupling foundation models with Retrieval-Augmented Generation (RAG) architectures to reduce hallucinations and domain-specific factual drift, grounding model generation in verified, real-time proprietary data repositories. Simultaneously, systematic adversarial evaluation, known as red-teaming, should be conducted before the model is released to discover latent vulnerabilities, jailbreaks, and triggers for harmful generation. Cryptographic watermarking schemes and synthetic text metadata are implemented at the output interface to enable provenance tracking. Mitigating risks related to deceptive automated communication and unauthorized academic plagiarism. Bias audits should also be standardized and continually maintained throughout the model's life-cycle to detect discriminatory representations as demographic distributions change. Finally, institutional governance mechanisms must translate high-level commitments to ethics into auditable compliance workflows that are aligned with evolving legal mandates. Novelli et al. point out that the EU AI Act and other transnational regulatory systems impose strict tiered obligations on general-purpose AI models, including technical documentation, algorithmic transparency, systemic risk assessments, and copyright compliance [18]. Organizations will need to establish clear data provenance policies, log model interactions, and establish clear accountability between model developers and downstream deployers to meet these statutory requirements. Systematically aligning technical controls, institutional protocols, and regulatory oversight will allow adopting entities to realize the productivity dividends of generative AI while safeguarding institutional integrity and public trust.

## Conclusions and Future Research Directions

Generative AI and Large Language Models have evolved, over a very short period, from a niche research frontier into a general-purpose technology transforming digital intelligence across education, enterprise, and beyond. The evidence synthesized in this review indicates consistent—yet unevenly distributed—productivity and learning gains, particularly in narrowing task-level performance gaps for novice workers, alongside persistent, non-trivial challenges around hallucination, algorithmic bias, data privacy, and intellectual property. Crucially, the literature demonstrates that these risks are not transitory flaws that will simply disappear with greater computational scale, but rather structural features of statistical sequence-prediction systems that require ongoing methodological, ethical, and regulatory attention.

Realizing the long-term dividends of generative AI without compromising institutional integrity requires pairing technical safeguards—such as retrieval-augmented generation, red-teaming, and human-in-the-loop verification—with deliberate organizational workflow redesign and clear legal accountability. Moving forward, three priorities should guide future research:

- **Longitudinal and Sector-Specific Impact Studies:** Research must move beyond initial adoption snapshots to measure sustained productivity, educational outcomes, skill polarization, and worker welfare over time.


© 2026 Reseapro Scientific Services
https://doi.org/10.61577/ibsm.2026.100004

- **Standardized Evaluation Frameworks:** The field requires transparent, domain-appropriate benchmarking protocols that allow claims regarding model reliability, factuality, and safety to be reliably compared across studies and jurisdictions.
- **Interdisciplinary Governance Translation:** Closer collaboration among computer scientists, ethicists, legal scholars, and domain practitioners is essential to translate normative ethics frameworks into auditable, actionable institutional governance practices.

Ultimately, the central research and policy challenge as generative AI continues to reshape digital intelligence is no longer whether to adopt these tools, but how to deploy them in ways that are reliable, equitable, and accountable.

## References

1. Fan L, Li L, Ma Z, Lee S, Yu H, Hemphill L. A Bibliometric Review of Large Language Models Research from 2017 to 2023. ACM Transactions on Intelligent Systems and Technology. 2024;15(5):1–25. https://doi.org/10.1145/3664930
2. Minaee S, Mikolov T, Nikzad N, Chenaghlu M, Socher R, Amatriain X, et al.. Large language models: A survey (arXiv:2402.06196). arXiv. . 2024.
3. Hagendorff T. Mapping the Ethics of Generative AI: A Comprehensive Scoping Review. Minds and Machines. 2024;34(4):39. https://doi.org/10.1007/s11023-024-09694-w
4. Laskar MTR, Alqahtani S, Bari MS, Rahman M, Khan MAM, Khan H, et al.. A Systematic Survey and Critical Review on Evaluating Large Language Models: Challenges, Limitations, and Recommendations. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024:13785–13816. https://doi.org/10.18653/v1/2024.emnlp-main.764
5. Michael C, Lareina Y, Bryce H, Alex S, Alexander S. The state of AI in 2023: Generative AI's breakout year. . 2023.
6. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science. 2023;381(6654):187–192. https://doi.org/10.1126/science.adh2586
7. Brynjolfsson E, Li D, Raymond LR. Generative AI at work (NBER Working Paper No. 31161). . 2023.
8. Bick A, Blandin A, Deming DJ. The rapid adoption of generative AI (NBER Working Paper No. 32966). . 2024.
9. Gulati P, Marchetti A, Puranam P, Sevcenko V. Generative AI adoption and higher order skills (INSEAD Working Paper No. 2025/26/STR). . 2025.
10. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al.. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences. 2023;103:102274. https://doi.org/10.1016/j.lindif.2023.102274
11. Chiu TKF. The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. Interactive Learning Environments. 2023;32(10):6187–6203. https://doi.org/10.1080/10494820.2023.2253861
12. Dempere J, Modugu K, Hesham A, Ramasamy LK. The impact of ChatGPT on higher education. Frontiers in Education. 2023;8:1206936. https://doi.org/10.3389/feduc.2023.1206936
13. M A. ChatGPT and beyond: The generative AI revolution in education (arXiv:2311.15198). arXiv. . 2023.
14. Said A. On explaining recommendations with Large Language Models: a review. Frontiers in Big Data. 2025;7:1505284. https://doi.org/10.3389/fdata.2024.1505284
15. Sweetser P. Large Language Models and Video Games: A Preliminary Scoping Review. ACM Conversational User Interfaces 2024. 2024:1–8. https://doi.org/10.1145/3640794.3665582
16. Mühlhoff R. Predictive privacy: Collective data protection in the context of artificial intelligence and big data. Big Data & Society. 2023;10(1):20539517231166886. https://doi.org/10.1177/20539517231166886
17. Wu X, Duan R, Ni J. Unveiling security, privacy, and ethical concerns of ChatGPT. Journal of Information and Intelligence. 2024;2(2):102–115. https://doi.org/10.1016/j.jiixd.2023.10.007
18. Novelli C, Casolari F, Hacker P, Spedicato G, Floridi L. Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity. . 2024;55:106066. https://doi.org/10.2139/ssrn.4694565


© 2026 Reseapro Scientific Services
https://doi.org/10.61577/ibsm.2026.100004