HiLL PUBLISHING

Future Trends in AI Research, 2025, 2(1), 37-53
https://www.hillpublisher.com/Journals/ftair/
ISSN Online: 3066-3083

A Review of the Evolution and Application Scenarios of Generative Artificial Intelligence Technology

Matthew J. Collins

Mila-Quebec Artificial Intelligence Institute, Montréal H2S 3H1, Canada.

How to cite this paper: Matthew J. Collins. (2025) A Review of the Evolution and Application Scenarios of Generative Artificial Intelligence Technology. *Future Trends in AI Research*, 2(1), 37-53.
DOI: 10.26855/ftair.2025.12.007

Received: October 11, 2025
Accepted: November 6, 2025
Published: December 28, 2025

***Corresponding author:** Matthew J. Collins, Mila-Quebec Artificial Intelligence Institute, Montréal H2S 3H1, Canada.

Abstract

Generative Artificial Intelligence (GenAI), as one of the most transformative technological directions in the AI field, has recently undergone leapfrog development from statistical language models to large-scale pre-trained models, and then to multimodal unified generation systems. This evolution has profoundly reshaped the underlying logic of content production, knowledge acquisition, and industrial innovation. Cutting-edge models represented by the GPT series, Stable Diffusion, and Sora demonstrate powerful capabilities in text, image, and video generation, driving generative AI rapidly from laboratory research to diverse application scenarios such as education, healthcare, finance, security defense, and cultural creativity, sparking widespread attention from both academia and industry. However, the rapid technological iteration also brings complex challenges, including model hallucinations, algorithmic bias, copyright disputes, data privacy leaks, and the widening digital divide, necessitating systematic theoretical review and governance responses. Based on domestic and international frontier literature from 2023 to 2024, this paper adopts a systematic literature review method to comprehensively sort out and deeply analyze the current research status of generative AI from four dimensions: technological evolution trajectory, typical application scenarios, core challenges and countermeasures, and future development trends. Research shows that the technological evolution of generative AI presents stage characteristics of "rule-driven → statistical learning → deep generation → multimodal unification." Its application scenarios have expanded from general content generation to deep penetration into vertical domains, demonstrating significant value particularly in personalized education, intelligent medical diagnosis, financial advisory services, cybersecurity defense, and AIGC-enabled design. At the same time, technical reliability, ethical fairness, legal compliance, and system security constitute the key bottlenecks currently restricting its sustainable development. In the future, generative AI will evolve towards multi-technology integration, industry specialization, and systematic governance, achieving a transition from "tool empowerment" to "ecosystem reconstruction." This paper aims to provide a panoramic reference for the theoretical research and practical application of generative AI, supporting academic exploration and industrial decision-making in related fields.

Keywords

Generative Artificial Intelligence; Large Language Models; Multimodal Generation; Application Scenarios; Technological Evolution; Ethical Governance; AIGC

1. Introduction

1.1 Research Background and Significance

Artificial intelligence research has gradually shifted from predominantly symbolic approaches toward data-driven and connectionist paradigms, although symbolic methods continue to play an important role in specific domains, and from expert systems to deep learning. The rise of generative artificial intelligence (GenAI) marks a new historical stage in this technological evolution. Different from traditional discriminative models that classify and predict input data, generative models can autonomously create new content with high realism and semantic coherence by learning and simulating the underlying distribution of data, covering various forms such as text, images, audio, video, and cross-modal combinations. At the end of 2022, ChatGPT, released by OpenAI, became a global phenomenon with its fluent natural language interaction capabilities and broad knowledge coverage, reaching over 100 million active users within two months of launch, setting a new growth record for consumer internet products. This milestone event not only validated the technical feasibility of large-scale pre-trained models (Large Language Models, LLMs) but also revealed the penetration potential of generative AI as a general-purpose technology into various economic and social sectors.

The technological breakthroughs in generative AI stem from the synergistic evolution of several key elements: the proposal of the Transformer architecture solved the problem of long-distance dependency modeling, laying the structural foundation for subsequent models like GPT and BERT; the maturation of self-supervised pre-training and fine-tuning paradigms enabled models to learn general representations from massive unlabeled data; the optimization of generation mechanisms such as diffusion models and generative adversarial networks (GANs) significantly improved output quality and diversity; and advances in multimodal alignment techniques broke down semantic barriers between different modalities such as text, images, and video, giving rise to cross-modal generation systems like DALL-E 3, Midjourney, and Sora. These technological advances have collectively driven generative AI from a single-modal "content generation tool" to a multimodal "intelligent creative partner," with its application boundaries continuously expanding and its influence growing ever deeper.

Against this backdrop, systematically sorting out the technological evolution logic and application scenario distribution of generative AI, and deeply analyzing the challenges it faces and governance paths, has important theoretical value and practical significance for academia to grasp research directions, for industry to formulate development strategies, and for policymakers to improve regulatory frameworks. However, existing research mostly focuses on technical details or single application scenarios, lacking a comprehensive review that integrates technological evolution, application practices, and risk governance into a unified analytical framework. This paper attempts to fill this gap by integrating recent frontier literature to construct a systematic cognitive map of generative AI research.

1.2 Current Research Status at Home and Abroad

Internationally, research on generative AI presents a dual-track pattern of fundamental model innovation alongside security defense. At the foundational model level, researchers continue to explore model architecture optimization, training efficiency improvement, and multimodal expansion paths, driving simultaneous growth in model scale and capability. In the security defense domain, as generative AI is maliciously used in scenarios such as cyberattacks and disinformation generation, corresponding defense technologies are also developing rapidly. Babaey and Ravindran proposed the GenSQLi framework, which uses generative AI technology to automatically enhance Web application firewalls (WAFs) against SQL injection attacks, demonstrating the security governance approach of "using AI against AI" [1]. In the realm of higher education integration, Cordero et al. systematically summarized best practices for using generative AI in university teaching, emphasizing the importance of human-machine collaboration and critical thinking cultivation [2].

Domestic research places greater emphasis on educational application innovation and ethical governance exploration. Chen et al. designed and implemented an educational chatbot for project management training, integrating personalized adaptive learning functions, verifying the effectiveness of generative AI in vocational training [3]. Liu explored the application path of generative AI in environmental design practice teaching, proposing a teaching model that integrates "technology empowerment, design thinking, and practical ability" [4]. Ge et al. found through empirical research that generative AI, as a cognitive scaffold, can significantly promote collaborative knowledge construction among university students, enhancing learning outcomes and team interaction quality [5]. Additionally, domestic

scholars have actively discussed issues such as the digital divide, academic ethics, and copyright disputes caused by generative AI, forming an academic ecosystem where technological optimism coexists with risk prudence.

Overall, domestic and international research together outline a development picture of "rapid technological iteration, deep application expansion, and urgent governance follow-up" for generative AI, but cross-domain, cross-level systematic integration still needs strengthening.

1.3 Research Content and Methods

This paper focuses on frontier research achievements in the field of generative artificial intelligence from 2023 to 2024, adopting a systematic literature review method to analyze the following core questions: (1) What technological evolution process has generative AI undergone? What are the key characteristics and representative achievements of each stage? (2) In which areas has generative AI formed typical applications? What are its application models and value creation mechanisms? (3) What technical, ethical, legal, and security challenges arise from the widespread application of generative AI? What are the response strategies? (4) Looking ahead, how can the development trends and research directions of generative AI be defined?

Literature selection followed these criteria: research topics directly related to the technological evolution or application scenarios of generative AI; publication period 2023-2024 to ensure timeliness; research types covering theoretical analysis, empirical studies, technical frameworks, and case studies; literature sources including SCI/SSCI indexed journals, authoritative conference papers, and high-quality Chinese journals. The final 25 papers included in the analysis cover multiple dimensions such as security defense, educational applications, medical innovation, finance and business, cultural creativity, and ethical governance, demonstrating good representativeness.

1.4 Paper Structure

In addition to the introduction, the main body of this paper consists of four chapters: Chapter 2 outlines the technological evolution of generative AI from its early emergence to multimodal unification, analyzing key breakthroughs and internal logic at each stage; Chapter 3 systematically presents typical application scenarios and practical cases of generative AI in education, healthcare, finance, security, cultural creativity, and other fields; Chapter 4 focuses on four dimensions—technical reliability, ethical fairness, legal compliance, and system security—to analyze current core challenges and discuss response strategies; Chapter 5 forecasts future development trends of generative AI, including technological integration, industry penetration, and governance system improvement; Chapter 6 summarizes the core viewpoints of the paper, identifies research limitations, and suggests future research directions.

2. Literature Search Strategy

The literature search was conducted in Web of Science, Scopus, IEEE Xplore, SpringerLink, and CNKI databases. Search keywords included:("Generative AI" OR "Generative Artificial Intelligence" OR "Large Language Model" OR "AIGC") AND (application OR governance OR ethics). The publication period was limited to January 2023–December 2024.

3. Technological Evolution of Generative Artificial Intelligence

The technological evolution of generative AI is not a linear substitution process but a complex evolution with multiple parallel paths and interwoven paradigms. From early rule-based symbolic generation to current multimodal unified large models, technological iteration has always revolved around three core questions: "how to model data distribution more accurately, how to generate high-quality content more efficiently, and how to adapt flexibly to diverse scenarios." This chapter sorts out the evolution of generative AI from a technological history perspective, revealing its internal development logic and frontier trends.

3.1 From Rule-Driven to Statistical Learning: Early Technical Foundations

The seeds of generative thinking can be traced back to the very beginning of AI. From the 1950s to the 1970s, symbolic AI dominated research paradigms, and researchers attempted to achieve text generation through manually encoded grammatical rules and semantic templates. Typical systems included ELIZA (1966) and SHRDLU (1970); the former simulated psychological counseling dialogue based on pattern matching, while the latter generated natural language instructions within a limited blocks world. The fundamental limitation of such systems was insufficient rule

coverage and lack of scalability—once beyond the preset domain, generation quality deteriorated sharply, with "fragility" being a fatal flaw of rule-driven methods.

From the 1980s to the 1990s, the rise of statistical learning methods injected new vitality into generative technology. Researchers began learning probability distributions from large-scale corpora rather than relying on manual rules. The application of Hidden Markov Models (HMMs) in speech synthesis and text generation marked the initial establishment of the data-driven paradigm. N-gram language models predicted the next word by counting co-occurrence probabilities of word sequences, establishing the core idea of "probability-based sequence generation" despite their simplicity. However, N-grams faced severe data sparsity problems and weak long-distance dependency modeling capabilities, often generating text lacking global coherence.

The introduction of neural networks provided nonlinear modeling tools for statistical learning. The Long Short-Term Memory (LSTM) network proposed by Hochreiter and Schmidhuber in 1997 effectively alleviated the vanishing gradient problem, enabling Recurrent Neural Networks (RNNs) to capture longer-range contextual dependencies. Sequence-to-sequence (Seq2Seq) models based on LSTM achieved significant progress in tasks such as machine translation and text summarization, with the encoder-decoder architecture becoming the standard paradigm for early neural generation models. Nevertheless, the low training efficiency caused by serial computation characteristics and the bottleneck of long text modeling still constrained further improvement in generation quality.

3.2 Breakthroughs in the Deep Learning Era: GAN and VAE

In 2014, the Generative Adversarial Network (GAN) proposed by Goodfellow et al. opened a new era for deep generative models. GAN introduced a game-theoretic framework, enabling models to learn to generate highly realistic data such as images and audio through adversarial training between a generator and a discriminator. This "unsupervised generation" paradigm broke through the previous reliance on labeled data, demonstrating remarkable results in visual tasks such as image synthesis, style transfer, and super-resolution reconstruction. Subsequent variants like DCGAN, CycleGAN, and StyleGAN continued to optimize architecture stability and generation quality, with StyleGAN2 even capable of generating facial images indistinguishable from real ones.

Parallel to GAN's adversarial path, the Variational Autoencoder (VAE) proposed by Kingma and Welling offered a generative approach based on probabilistic graphical models. VAE maps data to a probability distribution in latent space through an encoder, and then reconstructs data from sampled points via a decoder, achieving "interpretable generation" and "smooth latent space interpolation." Improved versions such as β-VAE and VQ-VAE made progress in disentangled representation learning and discrete latent space modeling, laying technical foundations for subsequent large-scale image generation models.

The common contribution of GAN and VAE was proving that deep neural networks are capable of modeling complex data distributions, yet their inherent shortcomings have become increasingly apparent: GAN training is unstable, with mode collapse difficult to eradicate; VAE generates blurry samples of lower quality than GAN; both face challenges of low efficiency in modeling high-dimensional data. More importantly, these models were primarily designed for specific modalities (e.g., images), lacking the ability for unified cross-modal generation and making it difficult to support collaborative creation of multiple content types such as text, images, and video.

3.3 The Rise of Transformer Architecture and Large Language Models

In 2017, the Transformer architecture proposed by Vaswani et al. became a watershed moment in the history of generative AI. Its core innovation—the self-attention mechanism—fundamentally changed sequence modeling: by computing attention weights between any two positions in a sequence, the model can capture global dependencies in parallel, overcoming both the serial computation bottleneck of RNNs and the long-distance memory limitations of LSTMs. The design philosophy of "Attention Is All You Need" laid the architectural foundation for the subsequent explosive growth of large language models.

In 2018, OpenAI released GPT (Generative Pre-trained Transformer), establishing the two-stage paradigm of "pre-training + fine-tuning." The model first performs autoregressive language modeling on large-scale unlabeled text to learn general language representations, then fine-tunes for downstream tasks, significantly improving sample efficiency and generalization ability. GPT-2 in 2019 expanded the parameter scale to 1.5 billion, demonstrating surprising zero-shot generation capabilities; GPT-3 in 2020 pushed parameters to 175 billion, adapting to various tasks through in-context learning without needing gradient updates. The exponential growth in parameter scale and the

emergence of emergent abilities revealed the existence of "scaling laws"—model performance improves predictably with increases in computation, data, and parameters.

After 2022, large language models entered the stage of instruction tuning and Reinforcement Learning from Human Feedback (RLHF). ChatGPT, through instruction tuning and RLHF training on GPT-3.5, significantly improved instruction-following capabilities and output safety, achieving a qualitative transformation from "language model" to "conversational assistant." The open-source community became active, with models like LLaMA, Alpaca, and Vicuna promoting technological democratization; domestic models such as ERNIE Bot, Tongyi Qianwen, and iFLYTEK Spark also quickly followed, creating a highly diversified ecosystem of commercial and open-source foundation models.

The success of large language models lies not only in parameter scale but also in their versatility as "foundation models." Through prompt engineering and Retrieval-Augmented Generation (RAG), a single model can adapt to diverse tasks such as question answering, summarization, translation, and code generation, greatly lowering the development barrier for AI applications. However, the modal limitations of text-only models are increasingly evident—real-world information exists in multiple modalities such as vision, hearing, and touch—making cross-modal understanding and generation an inevitable direction for technological evolution.

3.4 Multimodal Generation and Frontier Trends

In 2021, DALL-E released by OpenAI first demonstrated the potential of the Transformer architecture for text-to-image generation, achieving unified modeling with text by discretizing images into visual tokens. DALL-E 2 in 2022 introduced CLIP (Contrastive Language-Image Pre-training) for text-image alignment, combined with diffusion models to improve generation quality and diversity. In the same year, Stable Diffusion was open-sourced, moving the computationally intensive diffusion process from pixel space to latent space, enabling high-quality image generation on consumer GPUs and greatly promoting the popularization of AIGC (AI Generated Content).

The rise of diffusion models marks a key turning point in the multimodal generation field. Unlike GAN's adversarial training, diffusion models achieve more stable training and finer generation control by gradually adding noise to data and then learning the reverse process to denoise. Models such as Stable Diffusion, Midjourney, and Imagen have been widely used in artistic creation, advertising design, and game asset generation. In 2023, tools like Runway's Gen-2 and Pika Labs achieved text/image-to-video generation, and although duration and consistency remain limited, they have already shown potential to disrupt traditional video production workflows.

In early 2024, Sora released by OpenAI pushed video generation to new heights, capable of generating high-definition videos up to 60 seconds long with complex scene changes and physical interactions. Its technical core lies in treating video as a spatiotemporally joint sequence of tokens, learning to simulate the physical world through large-scale training. Some researchers argue that generative AI may evolve beyond content generation toward simulation-oriented capabilities. If this trend materializes, it will profoundly impact many fields including film production, game development, autonomous driving simulation, and scientific computing.

At the same time, Multimodal Large Language Models (MLLMs) have become a research hotspot. Models such as GPT-4V, Gemini, and Qwen-VL can simultaneously process text, image, and video inputs, achieving cross-modal understanding, reasoning, and generation. Unified architectures (e.g., AnyRes, Flamingo) connect modality encoders with large language models, enabling models to possess comprehensive capabilities such as "seeing and speaking about images," "video analysis," and "text-image creation." Recent advances in multimodal foundation models suggest a gradual convergence of perception and generation capabilities across different data modalities, potentially enabling more unified intelligent systems, aiming to build general intelligent systems capable of perceiving, understanding, and creating arbitrary forms of information.

During technological evolution, reflections on ethical and social dimensions have also deepened. Liu and Huang from the perspective of fairness principles examined generative AI, pointing out that uneven distribution of technological dividends may exacerbate the digital divide, and that algorithmic bias and data discrimination cause systematic exclusion of vulnerable groups, calling for embedding ethical review mechanisms in technology development [6]. Li, taking the AI virtual anchor "Hua Xiaoshi" at Central China Normal University as an example, demonstrated the innovative application of multimodal generation technology in educational communication, while also prompting practical challenges such as the authenticity of virtual images and audience acceptance [7]. REN et al. discussed the application of AIGC technology in the packaging design of Sichuan intangible cultural heritage agricultural products, proving that generative AI can effectively integrate traditional cultural elements with modern design concepts,

providing technological empowerment for rural revitalization [8]. Schrepel and Pentland analyzed the competitive dynamics of AI foundation models from an industrial economics perspective, pointing out the tension between technological monopoly of leading companies and the vitality of the open-source ecosystem, and proposing policy recommendations to promote competition and innovation [9]. Xinyu deeply analyzed the internal mechanism of generative AI empowering ideological and political education in universities, arguing that technology can enhance the relevance and effectiveness of such education through personalized content push notification, immersive scenario simulation, and interactive learning experiences [10].

In summary, the technological evolution of generative AI shows clear stage characteristics: from rule-driven symbolic generation, to statistical learning and neural network data-driven approaches, to the era of large language models led by the Transformer architecture, and finally towards multimodal unified generation and world simulation. Breakthroughs at each stage are accompanied by core architectural innovations, training paradigm revolutions, and expansion of application boundaries, while the interaction between technology and society, ethics, and law is also becoming increasingly close, jointly shaping the development trajectory of generative AI.

4. Typical Application Scenarios of Generative Artificial Intelligence

The technological breakthroughs in generative AI are rapidly translating into real productive forces, with its application scenarios extending from early content creation assistance to core socio-economic sectors such as education, healthcare, finance, security, and cultural creativity. Different from traditional AI systems optimized for specific tasks, generative AI, with its general representation capabilities and flexible adaptability, presents an application pattern characterized by a foundation-model-centered ecosystem characterized by broad applicability and deep domain integration. This chapter systematically sorts out five typical application scenarios, analyzing their application models, value mechanisms, and practical cases, revealing the internal logic of generative AI empowering industrial transformation.

4.1 Education Sector

Education is one of the most extensive and deeply applied fields for generative AI. Its core application models can be summarized into three categories: intelligent tutoring systems, personalized learning platforms, and immersive teaching environments.

In intelligent tutoring, generative AI provides instant question answering, homework grading, and learning feedback through natural language interaction. Different from traditional rule-based expert systems, large language models can understand diverse expressions of student questions and generate targeted explanations and guidance, simulating "Socratic" dialogue teaching. Raitskaya and Tikhonova conducted a scoping review of the application of generative AI language tools in academic writing, finding that tools like ChatGPT significantly improve writing efficiency in grammar correction, structural optimization, and literature review generation, but also raise controversies about academic integrity and originality [11]. This study suggests that applications of generative AI in educational settings require establishing clear usage norms, distinguishing the boundary between "assistive tool" and "substitute completion."

Personalized adaptive learning is an advanced form of generative AI educational application. The educational chatbot designed by Chen et al. has already demonstrated the feasibility of dynamically adjusting content based on learners' knowledge levels, learning styles, and progress. Furthermore, generative AI can construct knowledge graphs based on learners' interaction histories, diagnose weak points, and generate customized exercises and explanatory paths. This "thousand people, thousand faces" learning experience is expected to break through the contradiction between standardization and personalization in traditional class teaching systems. However, research by Samuel et al. reveals users' risk perceptions regarding the sustainable use of ChatGPT—although the tool is convenient, users still worry about cognitive decline due to over-reliance, insufficient information accuracy, and privacy leakage risks [12]. This finding has important implications for educational applications: generative AI should be positioned as a "cognitive scaffold" rather than a "thinking substitute," and instructional design must preserve students' space for active construction.

Immersive gamified learning platforms represent the frontier of educational applications. The "LearningverseVR" platform developed by Song et al. integrates generative AI with virtual reality technology, constructing a highly immersive learning environment through AI-driven NPC dialogue, dynamic storyline generation, and personalized

task design. This study shows that multi-technology integration can significantly enhance learning motivation and knowledge retention, providing a new paradigm for fields requiring situational simulation, such as vocational education and medical training. Empirical research by Ge et al. also confirms that generative AI, as a cognitive scaffold for collaborative learning, can facilitate knowledge sharing and collective construction among university students in team projects.

4.2 Healthcare Sector

Healthcare is a high-value, high-sensitivity field for generative AI applications, involving multiple dimensions such as medical image analysis, clinical decision support, patient communication, and health equity.

In medical image classification and diagnosis, generative models exhibit unique advantages. Prusty et al. proposed a medical image classification framework based on a Latent Denoising Diffusion Probabilistic Model (LDDPM) and Wiener filtering, enhancing the model's representation capability for medical image features through generative pre-training, achieving classification accuracy superior to traditional discriminative methods on multiple datasets. This study reveals the potential of generative models in data-scarce scenarios—augmenting training data by generating high-quality synthetic samples to alleviate the high cost of medical image annotation [13].

Conversational health services represent another important form of generative AI medical application. Nadarzynski et al. proposed a roadmap for the design and implementation of inclusive healthcare chatbots from a health equity perspective, emphasizing the need to consider language diversity, cultural sensitivity, digital literacy differences, and accessibility needs during development to ensure technological dividends benefit marginalized groups. The study points out that most current health chatbots use English as the primary interaction language, with training data exhibiting demographic biases, potentially leading to systematic exclusion of non-English speakers and those with low digital literacy. Achieving health equity requires interdisciplinary collaboration, integrating expertise from medicine, linguistics, ethics, and design throughout the development process [14].

Generative AI also shows application prospects in drug discovery, medical record generation, and medical literature review. However, the stringent requirements for safety, interpretability, and accountability in the medical field constitute major obstacles to large-scale generative AI application. Model hallucinations have particularly severe consequences in medical scenarios—incorrect diagnostic advice or drug recommendations could directly endanger patients' lives. Therefore, generative AI in medical applications mostly adopts a "human-in-the-loop" design, treating AI outputs as auxiliary references rather than final decisions, with medical professionals assuming ultimate responsibility.

4.3 Finance and Business Sectors

The finance industry's high demands for information processing efficiency and decision quality make it an important scenario for the commercialization of generative AI. Current applications are mainly concentrated in four directions: intelligent advisory, market analysis, risk management, and customer service.

Intelligent robo-advisors are a typical representative of generative AI financial applications. Jiang et al. systematically analyzed the application scenarios of generative AI-based financial robo-advisors as investment advisors, pointing out that they can understand clients' risk preferences, investment goals, and constraints through natural language interaction, generating personalized asset allocation plans and explanations of investment strategies. Compared with traditional robo-advisory systems based on rule-based questionnaires, generative AI supports dynamic, in-depth conversational demand discovery, capturing subtle changes and complex constraints in client preferences. However, the study also highlights regulatory challenges such as insufficient algorithm transparency, ambiguous accountability, and market manipulation risks, calling for the establishment of a financial regulatory framework adapted to AI characteristics.

In business decision support, generative AI analyzes massive amounts of unstructured data (financial reports, news, social media, research reports) to generate market insights and investment recommendations. Vomberg et al. discussed the application of digital knowledge engineering in strategy development, arguing that generative AI can assist enterprises in extracting knowledge from dispersed information sources, identifying patterns, and generating strategic options, enhancing the scientific rigor and agility of decision-making. Jacques' survey of marketing executives found that ChatGPT has significant value in content idea generation, consumer insight analysis, and marketing copy optimization, but executives generally worry about brand reputation risks, data privacy issues, and strategic homogenization due to over-reliance.

The library and knowledge service sector has also witnessed the penetration of generative AI. Gupta studied factors influencing librarians' adoption of ChatGPT technology to provide entrepreneurial support, finding that perceived usefulness, ease of use, and organizational support are key drivers of adoption intention, while ethical concerns and technological anxiety constitute major barriers [15]. This study reveals the application logic of generative AI in professional service fields: technological value depends not only on functional performance but is also embedded in the social construction of organizational culture, professional norms, and user trust.

4.4 Security Defense Sector

The double-edged sword effect of generative AI is particularly evident in the security domain: on one hand, attackers use generative AI to create more sophisticated cyberattacks, deepfakes, and disinformation; on the other hand, defenders also leverage generative AI technologies to enhance threat detection and response capabilities.

Web application security is an important scenario for generative AI defense applications. The GenSQLi framework proposed by Babaey and Ravindran pioneered the defense paradigm of "using AI against AI." This framework uses generative AI to automatically generate a large number of diverse SQL injection attack samples for training and testing Web application firewall (WAF) detection rules. Traditional WAFs rely on manually written regular expression rules, which struggle to cover the combinatorial explosion of attack variants. GenSQLi learns the latent distribution of attack patterns through generative models, capable of generating semantically valid attack samples not seen in the training set, significantly improving WAF's generalization defense capability. This study demonstrates the unique value of generative AI in the cybersecurity field—not only as an attack tool but also as a defense enhancement means, achieving a spiral escalation of "offensive and defensive intelligence."

In threat intelligence analysis, generative AI can automatically parse security bulletins, vulnerability reports, and attack logs, generating structured threat intelligence summaries and response recommendations, alleviating the cognitive load of Security Operations Center (SOC) analysts. However, the security field has extremely high requirements for the accuracy and auditability of generated content; model hallucinations may lead to incorrect threat assessments and response decisions. Therefore, generative AI in security applications mostly adopts the RAG architecture, anchoring the generation process to authoritative knowledge bases to ensure traceability and verifiability of outputs.

4.5 Cultural Creativity and Media Sector

The cultural creativity industry is one of the most commercially mature fields for generative AI, covering content production, visual design, virtual personas, new media art, and more.

In content production, generative AI is reshaping the creative workflows of news, advertising, film, and games. Yang deeply explored the collaborative mechanism between generative AI and content production, arguing that technology can undertake basic tasks such as "first draft generation," "material collection," and "style transfer," allowing creators to focus on creative conception and value judgment, achieving a new "human-machine co-creation" production relationship. However, the study also warns of risks such as content homogenization, creative inertia, and copyright disputes, emphasizing that technology should serve rather than replace human creativity.

Visual communication in new media art is a typical scenario of generative AI empowering cultural innovation. Zhao studied the synergistic effects of AI technology in the evolution of new media art visual communication, pointing out that generative AI acts not only as a production tool but also as a "collaborator" in the artistic creation process, giving rise to emerging genres such as algorithmic art and generative art. AI-generated visual content has unique "machine aesthetic" characteristics—surreal details, unconventional compositions, and infinite style migration possibilities—expanding the boundaries of artistic expression.

The application of AIGC in traditional culture preservation and rural revitalization is also gaining increasing attention. REN et al., using the packaging of Sichuan intangible cultural heritage agricultural products as a case study, demonstrated how AIGC technology can empower the integrated innovation of traditional handicraft elements and modern design concepts. By training generative models with regional cultural characteristics, designers can quickly generate packaging solutions incorporating elements of intangible cultural heritage such as Sichuan embroidery, paper cutting, and New Year paintings, preserving cultural authenticity while improving design efficiency and market appeal. This application model provides a technological pathway for the revitalization of cultural heritage in the digital age.

Virtual anchors and digital humans are cutting-edge applications of generative AI in the media field. Li, using the AI virtual anchor "Hua Xiaoshi" at Central China Normal University as an example, analyzed the innovative value of virtual anchors in educational communication. This virtual anchor, based on generative AI, achieves real-time generation of text-to-speech, expressions, and movements, capable of 24/7 news broadcasting, course explanation, and interactive Q&A, breaking through the time and space limitations of human anchors. However, issues such as the emotional authenticity of virtual images, building audience trust, and ethical boundaries still require deeper exploration.

In summary, the application scenarios of generative AI exhibit characteristics of "breadth expansion" and "depth penetration" occurring in parallel. From educational personalization to healthcare equity, from financial intelligence to security offense-defense, from cultural innovation to rural revitalization, generative AI is becoming a key enabling technology in the digital economy era. However, application practices across various fields also collectively reveal the complexity of technology implementation—multiple factors such as performance optimization, ethical norms, legal adaptation, and social acceptance intertwine, determining the transformation path of generative AI from "technologically possible" to "socially feasible."

5. Challenges and Countermeasures in the Application of Generative Artificial Intelligence

The widespread application of generative AI, while creating immense value, also triggers a series of complex challenges at the technical, ethical, legal, and security levels. These challenges intertwine and reinforce each other, constituting key bottlenecks restricting the sustainable development of the technology. This chapter systematically analyzes the core challenges currently faced from four dimensions—technical reliability, ethical fairness, legal compliance, and system security—and discusses corresponding response strategies and governance paths.

5.1 Technical Challenges

Technical reliability is the primary concern for generative AI applications. Although large language models perform well in most scenarios, their inherent "hallucination" problem—generating content that seems plausible but is actually false or erroneous—has not been fundamentally solved. In high-risk fields such as medical diagnosis, financial decision-making, and legal consulting, hallucinations can lead to serious consequences. Although Prusty et al.'s research demonstrated the effectiveness of diffusion models in medical image classification, it also pointed out the high sensitivity of generative models to training data distribution: when input data deviates from the training distribution, model performance significantly degrades, and reliable confidence estimates are difficult to provide [16]. This finding reveals the vulnerability of generative AI in open-world applications—the data distribution in real-world scenarios is often dynamically changing, while the static training assumptions of models struggle to adapt.

Computational resource consumption and environmental impact constitute another technical challenge. The training and inference of large language models require immense computational power; GPT-3's training is estimated to consume approximately 355 years of single-GPU time, with carbon emissions equivalent to the lifetime emissions of multiple cars. As model scales continue to expand, energy consumption and hardware costs become obstacles to technological democratization, exacerbating the "computing divide"—only a few tech giants and wealthy countries can afford training cutting-edge models, while most institutions and individuals are forced to rely on API calls or open-source small models, forming technological dependence and capability differentiation.

Insufficient interpretability limits the application of generative AI in high-risk fields. The current "black box" nature of models makes it difficult for users to understand the basis and logic of generated content, trace error sources, or establish effective trust mechanisms. Although techniques such as attention visualization and probing classifiers provide some interpretability, for the deep decision-making mechanisms of hundred-billion-parameter models, humans still lack a systematic framework for understanding.

Strategies for addressing technical challenges include: developing Retrieval-Augmented Generation (RAG) and tool learning architectures to anchor model generation to external knowledge bases and verifiable tools, reducing hallucination risks; advancing model compression, quantization, and distillation techniques to lower deployment costs; establishing dynamic evaluation and continuous learning mechanisms to enable models to adapt to distribution shifts; investing in interpretability research to develop transparent interaction interfaces for end users.

5.2 Ethical and Social Challenges

Ethical challenges are a core issue concerning the social impact of generative AI. Algorithmic bias and discrimination

are particularly prominent: historical biases, social stereotypes, and cultural master narratives embedded in training data may be learned and amplified by models, systematically excluding marginalized groups in generated content. Nadarzynski et al.'s research on inclusive healthcare chatbots clearly points out that current health AI systems have significant language and demographic biases, with non-English speakers, people with low education levels, and people with disabilities facing higher access barriers and poorer service quality [17]. If not addressed, this "digital health divide" will exacerbate rather than alleviate existing health inequalities.

UNESCO emphasizes that AI systems should promote fairness, inclusiveness, and non-discrimination while avoiding the amplification of existing social inequalities, pointing out that the uneven distribution of technological dividends is creating new social stratification: groups with access to technological tools and prompt engineering skills gain efficiency advantages, while technologically disadvantaged groups face dual pressures of job displacement and skill devaluation. This paradox of simultaneous "technological empowerment" and "technological exclusion" requires policymakers to establish digital literacy education and skills retraining systems alongside promoting technology diffusion to ensure inclusive technological transformation.

Academic integrity crisis is a special ethical challenge in the education sector. Raitskaya et al.'s scoping review found that the misuse of generative AI language tools in academic writing has raised widespread concern—students may directly submit AI-generated papers, researchers may use AI to write literature reviews without disclosure, and the authenticity foundation of academic evaluation is shaken. Response strategies include: developing AI detection tools (though their reliability is questionable), revising academic integrity policies to clarify the boundaries of AI use, and reforming evaluation methods from "outcome-oriented" to "process-oriented," emphasizing the cultivation of critical thinking and original argumentation.

Occhipinti et al., from the perspective of "mental wealth," proposed an analytical framework for social transformation in the intelligent era, arguing that the rapid adoption of generative AI may trigger large-scale job anxiety, identity crises, and declining social cohesion, requiring the construction of resilient societies from three dimensions: mental health, social safety nets, and lifelong learning systems [18]. This study suggests that technology ethics concerns not only algorithm fairness but also social psychological adaptation and institutional transformation capacity.

5.3 Legal and Copyright Risks

The lag in legal systems is a structural dilemma facing generative AI governance. Existing intellectual property frameworks, based on a "human author-centrism" presumption, struggle to adapt to the needs of defining ownership of AI-generated content.

WIPO AI and Intellectual Property Reports copyright infringement risks of works "created" by generative AI, pointing out three current legal dilemmas: copyright compliance of training data—does the use of massive copyrighted works for model training constitute fair use or infringement?; copyrightability of generated content—does AI output meet "originality" requirements and qualify for copyright protection?; allocation of rights between users and platforms—who among prompt designers, model developers, and platform operators holds rights to generated content? This study compared legal practices in the United States, the European Union, and China, finding significant divergence in the determination standards for "fair use" and the copyrightability of AI-generated content, highlighting the urgent need for international coordination mechanisms.

In practice, several landmark lawsuits have emerged: Getty Images sued Stability AI for unauthorized use of its images to train models; The New York Times sued OpenAI and Microsoft for copyright infringement; a group of artists sued Midjourney and other image generation platforms. The outcomes of these cases will profoundly affect the business models and legal risk structures of the generative AI industry.

Strategies for addressing legal challenges include: establishing "opt-out" mechanisms allowing copyright holders to declare their works prohibited for AI training; developing data licensing markets to achieve large-scale authorization through collective management organizations; exploring special legal categories for "AI-generated content" to clarify its scope of protection and rights ownership; promoting international coordination of intellectual property rules to avoid regulatory arbitrage and legal conflicts.

5.4 Security and Privacy Risks

The security dimension is the most direct manifestation of generative AI's double-edged sword effect. Attackers use generative AI to create more sophisticated threats: deepfake technology for political disinformation, financial fraud, and identity theft; large language models assisting in writing malicious code, phishing emails, and social engineering scripts; automatically generated disinformation spreading on an industrial scale, eroding the public information

ecosystem.

Although the GenSQLi framework by Babaey and Ravindran demonstrates the application of generative AI on the defensive side, it also inversely reveals the possibility of attackers using similar techniques to generate variant attack payloads—the technical gap between offense and defense is narrowing, and the security game is becoming increasingly complex. Goetsch's report on the workplace safety revolution noted that the application of generative AI in industrial safety can analyze sensor data in real time, predict equipment failures, and generate safety operating procedures, but system vulnerabilities and false alarms may also introduce new safety risks.

Privacy risks are equally significant. Large language models may "memorize" sensitive personal information during training, leaking it under specific prompts; user interaction data with AI systems is collected for model improvement, leading to conflicts between informed consent and data minimization principles; model inversion attacks may infer characteristics of training data from outputs, threatening data subject privacy.

Strategies for addressing security and privacy challenges include: developing adversarial training and red teaming mechanisms to improve model robustness; establishing model security auditing and certification systems, mandating security assessments for high-risk applications; implementing privacy-enhancing technologies such as differential privacy training, federated learning, and data anonymization; improving security incident response mechanisms and establishing cross-industry, cross-border intelligence sharing and collaborative response networks.

Vomberg et al.'s research on digital knowledge engineering provides an organizational perspective on security governance: enterprises need to establish a full lifecycle security system covering data governance, model management, and application monitoring, embedding AI security into strategic decision-making and operational processes, rather than treating it as a purely technical issue [19].

In summary, the challenges facing generative AI are multi-dimensional and systemic, making single technical or policy measures insufficient for effective response. Technical reliability requires collaborative improvement of algorithmic innovation, engineering optimization, and evaluation standards; ethical fairness requires the joint guarantee of value-sensitive design, inclusive development, and multi-stakeholder participation; legal compliance requires synchronized advancement of rule innovation, market mechanisms, and international coordination; system security requires comprehensive implementation of offense-defense collaboration, organizational governance, and ecosystem building. Building a trustworthy, controllable, and sustainable generative AI development ecosystem is a major issue facing both technological evolution and social progress [20].

6. Future Development Trends and Prospects of Generative Artificial Intelligence

Based on the systematic analysis of technological evolution, application scenarios, and core challenges above, this chapter forecasts future development trends of generative AI from three dimensions: technological integration, industry penetration, and governance system, providing forward-looking references for academic research, industrial practice, and policy formulation.

6.1 Technological Integration Trends: From Single Modality to Multi-Technology Synergy

The future development of generative AI will break through the limitations of single technology, moving towards deep integration with technologies such as Virtual Reality (VR), Augmented Reality (AR), blockchain, and the Internet of Things (IoT). The "LearningverseVR" platform developed by Song et al. has already demonstrated the initial effectiveness of integrating generative AI with VR technology: the combination of AI-driven dynamic content generation and immersive virtual environments creates contextualized, interactive learning spaces that traditional online learning can hardly match [21]. This integration trend will extend to more fields—in medical training, AI-generated virtual cases combined with surgical simulation provide risk-free skills training environments; in industrial design, AI-generated real-time 3D models combined with AR overlays accelerate prototyping iteration and client confirmation; in smart city construction, generative AI interfaces with IoT data streams to automatically generate urban operation reports and optimization plans.

Multimodal unified models have emerged as a prominent direction in current AI research. Current models such as GPT-4V and Gemini have achieved joint understanding and generation of text, images, and video, but there is still room for improvement in deep integration and collaborative reasoning across modalities. Future models are expected to possess more refined cross-modal alignment capabilities, able to generate output in any modality based on input from any modality, achieving truly "full modality general intelligence." Furthermore, exploration of world models—such as the physical world simulation capabilities demonstrated by Sora—may enable generative AI to evolve from

a "content creation tool" to a "world simulator," playing a fundamental role in scientific discovery, engineering simulation, social prediction, and other fields [22].

6.2 Deepening Industry Penetration: From General Tools to Vertical Specialization

Recent studies indicate a growing trend toward domain-specific adaptation of generative AI systems. Practices in the finance sector have already revealed this trend: the financial robo-advisors studied by Jiang et al. do not simply invoke general-purpose large models but perform domain adaptation based on financial knowledge bases, regulatory rules, and investment strategies, providing compliant, explainable professional services [23]. In the future, more "domain-specific LLMs" will emerge in professional fields such as law, healthcare, scientific research, and manufacturing. These models, building on general language capabilities, deeply integrate domain knowledge, industry standards, and practical experience to provide professional accuracy and reliability exceeding that of general models.

The deepening of industry penetration is also reflected in the extension from "front-end interaction" to "back-end processes." Currently, generative AI is mainly applied in customer service, content marketing, and other end-user-facing scenarios. In the future, it will be more embedded in core business processes such as product design, supply chain management, quality control, and risk assessment, realizing intelligent reconstruction of the entire value chain. Jacques' survey of marketing executives suggests that successful industry applications need to go beyond the "tool substitution" mindset, restructuring business processes and organizational architectures to truly integrate AI capabilities into enterprise core competitiveness [24].

6.3 Improvement of Governance Systems: From Lagging Rules to Ecosystem Co-construction

The contradiction between rapid technological development and lagging governance systems is a key bottleneck for promoting the sustainable evolution of generative AI. Future governance systems will feature "multi-level nesting and dynamic evolution." At the technical level, establish model safety standards, evaluation benchmarks, and certification mechanisms to ensure the reliability and controllability of foundation models. At the application level, develop industry-specific usage norms, accountability definitions, and information disclosure requirements to achieve risk-graded management. At the societal level, build governance platforms for public participation and multi-party consultation, balancing innovation incentives and risk prevention [25].

Malespina's research on the key role of school librarians in formulating AI use policies reveals the unique value of grassroots professional institutions in governance systems. Librarians, as information literacy educators and knowledge intermediaries, can connect technology developers, policymakers, and end users, promoting localized adaptation and democratic participation in AI governance. This finding suggests that effective AI governance should not rely solely on top-down institutional design but also requires activating the participation capacity of various social actors, forming a collaborative governance network of "government regulation, industry self-discipline, social organizations, and public participation."

The construction of international coordination mechanisms is particularly urgent. Yan and Wang's copyright research reveals the constraints of national legal differences on the development of the global AI industry; Schrepel and Pentland's competition analysis of foundation models points out that fragmentation of technical standards may lead to market segmentation and innovation efficiency losses. In the future, under multilateral frameworks such as UNESCO, WIPO, and ISO, it is necessary to promote international coordination of governance principles, technical standards, and ethical norms for generative AI, avoiding a "race to the bottom" and "regulatory arbitrage."

Jacques' research also highlights governance innovation at the enterprise level: marketing executives are exploring internal mechanisms such as "AI use committees," "ethics review processes," and "supplier due diligence" to embed AI governance into corporate strategic decision-making and daily operations. This "inside-out" governance consciousness, complementing the "outside-in" regulatory requirements, jointly shapes a responsible AI innovation ecosystem.

In summary, the future development of generative AI will present three major trends: technological integration, industry specialization, and governance systematization. Technological integration will expand capability boundaries and create new application possibilities; industry deepening will enhance value density, achieving a transition from "icing on the cake" to "timely help"; governance improvement will establish a foundation of trust, ensuring technological dividends benefit society as a whole. These three aspects support each other and evolve dynamically, jointly driving generative AI from a "phenomenal technology" to an "infrastructure-level capability," profoundly reshaping the production and lifestyle of human society.

6.4 Emerging Paradigms: Agentic AI and Autonomous Systems

Beyond the current wave of generative models, the emergence of Agentic AI has been described as a potential paradigm shift in recent literature that promises to fundamentally alter how humans interact with intelligent systems. Unlike passive content generators that respond to prompts, AI agents are designed to autonomously plan, execute multi-step tasks, and interact with external tools and environments to achieve specified goals. This transition from content generation to task execution represents an important development in contemporary AI systems.

The concept of agentic AI builds upon the foundation laid by large language models but extends their capabilities through several key architectural innovations. Key architectural components of agentic systems include tool-use mechanisms, memory architectures, and planning frameworks..

Agentic AI has potential applications across multiple domains, including scientific research and software engineering. In scientific research, autonomous agents can formulate hypotheses, design experiments, analyze results, and iterate on methodologies without constant human oversight. Systems such as AutoGPT, BabyAGI, and more recent enterprise-focused platforms demonstrate early manifestations of this capability, though current implementations still struggle with reliability and error accumulation in extended task chains. In software engineering, agentic coding assistants can not only generate code snippets but also debug errors, refactor architectures, manage dependencies, and deploy applications through continuous integration pipelines.

The transition to agentic systems also introduces novel challenges that extend beyond those faced by current generative models. The compounding of errors across multi-step autonomous processes creates significant reliability concerns—where a single hallucination or reasoning error in an early step can propagate and amplify throughout subsequent actions, potentially leading to cascading failures. The attribution of responsibility becomes increasingly complex when autonomous agents make consequential decisions: developers, deployers, users, and the agents themselves may all bear partial accountability, yet existing legal frameworks offer little clarity on how to distribute liability.

Furthermore, the alignment problem takes on renewed urgency in agentic contexts. Ensuring that autonomous systems pursue objectives genuinely aligned with human values—rather than optimizing for proxy metrics or exhibiting reward hacking behaviors—requires robust oversight mechanisms. The challenge of specifying comprehensive yet unambiguous objectives, known as the "King Midas problem," becomes particularly acute when agents possess broad action spaces and extended autonomy.

6.5 Hardware-Software Co-Evolution and Edge Deployment

The computational demands of generative AI have traditionally confined powerful models to centralized cloud infrastructure, creating dependencies on network connectivity and raising concerns about data sovereignty and latency. However, emerging trends in hardware-software co-evolution are gradually democratizing access to generative capabilities through edge deployment.

Specialized AI accelerators, including neural processing units (NPUs) in consumer devices, are rapidly increasing in capability while decreasing in cost. Apple's Neural Engine, Qualcomm's Hexagon DSP, and Google's Tensor Processing Units in mobile devices exemplify this trend. These hardware innovations enable on-device inference for increasingly capable models, reducing latency, enhancing privacy by keeping data local, and enabling AI functionality in connectivity-constrained environments.

Model optimization techniques have evolved in parallel with hardware improvements. Quantization methods that reduce numerical precision from 32-bit floating point to 8-bit or even 4-bit integers dramatically decrease memory requirements and computational load with minimal impact on output quality. Knowledge distillation enables the transfer of capabilities from large teacher models to compact student models suitable for resource-constrained environments. Mixture-of-Experts (MoE) architectures activate only relevant subsets of model parameters for specific inputs, improving inference efficiency without sacrificing overall model capacity.

The implications of edge deployment extend beyond technical performance to encompass broader socioeconomic dimensions. On-device processing addresses privacy concerns that have hindered adoption in sensitive domains such as healthcare and finance, where data residency requirements and patient confidentiality obligations often prohibit cloud transmission. Edge deployment also reduces the environmental footprint of AI inference by eliminating data center cooling overhead and network transmission energy costs.

However, the fragmentation of deployment environments—ranging from hyperscale data centers to personal smartphones to embedded IoT devices—creates challenges for model consistency and update management. Ensuring

that edge-deployed models maintain alignment with centralized versions, receive timely security patches, and operate within acceptable performance bounds requires sophisticated distributed systems infrastructure that remains underdeveloped.

6.6 Synthetic Data and Data Ecosystem Transformation

The insatiable appetite of generative models for training data has precipitated a fundamental transformation in how data is sourced, curated, and valued within the AI ecosystem. As readily available high-quality text data approaches exhaustion—estimates suggest that current large language models have already trained on a significant fraction of all publicly accessible text—the industry is increasingly turning to synthetic data generation as a complementary or alternative source.

Synthetic data, generated by AI systems themselves, offers several compelling advantages. It can be produced in virtually unlimited quantities, tailored to specific distributions and edge cases that may be rare in natural data, and generated with guaranteed privacy properties since it does not correspond to real individuals. In computer vision, synthetic training data has demonstrated particular value for scenarios involving rare events, dangerous conditions, or privacy-sensitive subjects. In natural language processing, model-generated text can augment training corpora with stylistic variations, multilingual content, and domain-specific terminology.

The self-improvement potential of generative models through synthetic data bootstrapping represents an intriguing frontier. Research on self-play, curriculum learning from synthetic tasks, and iterative refinement through model-generated feedback suggests pathways toward autonomous capability improvement that reduce dependence on human-annotated data. However, the risks of model collapse—where iterative training on synthetic data progressively degrades output quality and diversity—pose significant theoretical and practical challenges that remain incompletely understood.

The economics of data are also undergoing transformation. The recognition that high-quality training data constitutes a critical competitive moat has intensified commercial interest in proprietary datasets, licensing arrangements, and data marketplaces. Publishers, social media platforms, and content aggregators are increasingly asserting control over their data assets, either restricting access or demanding compensation for training use. This trend toward data commodification may reshape competitive dynamics in the generative AI industry, potentially advantaging incumbents with established data relationships while raising barriers for new entrants.

The emergence of data provenance standards and attribution technologies aims to address concerns about consent, compensation, and transparency in training data usage. Initiatives such as the Data Provenance Initiative and technical approaches including watermarking, fingerprinting, and blockchain-based attribution seek to create auditable records of data origins and usage, though widespread adoption remains nascent.

6.7 Human-AI Collaboration and Workforce Transformation

The integration of generative AI into professional workflows is catalyzing a fundamental reconfiguration of human labor, challenging traditional conceptualizations of expertise, creativity, and productivity. Rather than simple substitution of human workers by automated systems, the emerging paradigm emphasizes collaborative intelligence—synergistic partnerships that leverage complementary strengths of human and artificial cognition.

In knowledge work, generative AI functions increasingly as a cognitive amplifier, handling information retrieval, pattern identification, and preliminary synthesis while human professionals focus on judgment, contextual interpretation, and strategic decision-making. This division of labor promises to enhance productivity while preserving the uniquely human capacities for ethical reasoning, emotional intelligence, and creative insight that remain beyond current AI capabilities. However, realizing this collaborative potential requires significant investments in workforce reskilling, organizational redesign, and the development of effective human-AI interaction interfaces.

The transformation extends beyond individual task performance to encompass broader changes in organizational structures and labor markets. Traditional hierarchical models of expertise, where knowledge accumulation over years of experience conferred authority, are being disrupted by systems that can rapidly access and synthesize information across domains. Professional fields including law, medicine, journalism, and education are experiencing pressure to redefine credentialing, evaluation criteria, and career progression pathways in response to AI capabilities.

The distributional consequences of workforce transformation warrant careful attention. While aggregate productivity gains may benefit economies broadly, the transition costs—including job displacement, skill obsolescence, and wage polarization—may concentrate on specific demographic groups and geographic regions. Workers in routine

cognitive tasks face particularly high exposure to automation, while those in physical dexterity roles or complex interpersonal interactions remain relatively insulated. Proactive policy responses, including portable benefits, lifelong learning accounts, and targeted transition assistance, may be necessary to ensure equitable distribution of AI-generated prosperity.

The psychological and social dimensions of human-AI collaboration also merit consideration. Over-reliance on AI assistance may lead to skill atrophy, reduced critical thinking, and diminished capacity for independent judgment—a phenomenon analogous to the "automation complacency" observed in aviation and process control. Conversely, effective collaboration may enhance human capabilities through a process of "cognitive offloading" that frees mental resources for higher-order reasoning. Understanding the conditions that promote beneficial versus detrimental forms of human-AI interaction represents an important research frontier.

6.8 Open Science and Democratization of AI Research

The trajectory of generative AI development has been profoundly shaped by tensions between proprietary commercial development and open scientific research. The initial emergence of transformative models from well-resourced corporate laboratories—OpenAI, Google DeepMind, Anthropic—established a paradigm of closed development that limited external scrutiny and independent verification. However, countervailing forces toward openness have gained momentum, with significant implications for the future evolution of the field.

The open-source release of models such as Meta's LLaMA series, Mistral AI's offerings, and numerous community-developed derivatives has demonstrated that competitive capabilities can emerge from decentralized, collaborative development. These open models enable academic researchers to study system behaviors without restrictive commercial terms, allow developers to customize and fine-tune for specific applications, and reduce dependence on single vendors. The transparency afforded by open weights—though falling short of full training data and methodology disclosure—supports more robust safety research and democratic oversight.

The debate over openness versus safety encapsulates fundamental tensions in AI governance. Proponents of openness argue that concentrated control of powerful AI systems by a small number of corporations poses greater risks than widespread availability, and that democratic societies benefit from broad access to transformative technologies. Safety advocates counter that unrestricted dissemination of increasingly capable models may facilitate malicious use, accelerate competitive races that compromise safety precautions, and complicate regulatory oversight. Navigating this tension requires nuanced approaches that distinguish between different levels of capability risk and application contexts.

International dimensions of AI openness add further complexity. Geopolitical competition, particularly between the United States and China, has introduced strategic considerations that influence decisions about model sharing, export controls, and collaborative research. The emergence of national AI strategies that prioritize domestic capability development may fragment the global research community and impede collective progress on safety and governance challenges that transcend national boundaries.

The future trajectory of AI openness will likely involve continued experimentation with intermediate arrangements—selective disclosure, structured access programs, and tiered release strategies—that attempt to balance the benefits of broad availability against safety and security considerations. The development of effective mechanisms for responsible openness represents a critical governance challenge that will shape the competitive landscape and societal impacts of generative AI for years to come.

7. Conclusion

Based on domestic and international frontier literature from 2023 to 2024, this paper has systematically reviewed the technological evolution, application scenarios, core challenges, and future trends of generative artificial intelligence. The research shows that the technological development of generative AI has undergone stage-wise transitions from rule-driven to statistical learning, from deep learning to Transformer architecture, and from single-modal generation to multimodal unification, with breakthroughs at each stage accompanied by core architectural innovations and training paradigm revolutions. At the application level, generative AI has deeply penetrated core sectors such as education, healthcare, finance, security defense, and cultural creativity, demonstrating diverse values including personalized services, intelligent decision support, immersive experience creation, and offensive/defensive capability enhancement. However, a matrix of challenges consisting of technical reliability, ethical fairness, legal compliance, and

system security constrains the sustainable diffusion of the technology, urgently requiring integrated responses from technological innovation, institutional design, and governance coordination.

The core contribution of this paper is the construction of an integrated analytical framework of "technological evolution—application scenarios—challenges and countermeasures—future trends," bringing together research achievements dispersed across different disciplines and fields into a unified perspective, revealing the internal logic and external constraints of generative AI development. The research also points out that the value realization of generative AI depends not only on technical performance but is also embedded in complex interactions among social institutions, cultural norms, and interest patterns, making the analytical perspective of the sociology of technology indispensable for understanding its true impact.

This paper has the following limitations: First, the literature selection focuses on 2023-2024, with insufficient coverage of the historical evolution of early technological development, which may affect the completeness of the evolutionary analysis. Second, literature sources are mainly in English and Chinese, with limited attention to research from non-English speaking countries and non-academic literature (such as industry reports, policy documents), introducing geographical and type coverage biases. Third, as a review study, this paper primarily uses qualitative analysis, without incorporating quantitative analysis or case comparisons, leaving room for improvement in depth and precision.

Looking ahead, research on generative AI can be deepened in the following directions: first, the theoretical foundations and engineering implementations of multimodal unified models, especially the construction and validation of world models; second, systematic research on trustworthy AI, covering collaborative optimization of explainability, robustness, fairness, and privacy protection; third, cross-domain collaborative innovation mechanisms, exploring the potential integration of generative AI with cutting-edge technologies such as quantum computing, brain-computer interfaces, and synthetic biology; fourth, comparative studies and institutional design of global AI governance, providing academic support for building an inclusive and sustainable global AI development order. As generative AI continues to mature, its future impact will depend on the balance between technological innovation, governance mechanisms, and societal adaptation and it is hoped that collaborative efforts from academia, industry, and policy circles will promote this transformative technology to serve human well-being and sustainable development.

References

[1] Babaey V, Ravindran A. GenSQLi: A Generative Artificial Intelligence Framework for Automatically Securing Web Application Firewalls Against Structured Query Language Injection Attacks. Future Internet. 2024;17(1):8. DOI:10.3390/fi17010008.

[2] Cordero J, Zambrano TJ, Castillo CA. Integration of Generative Artificial Intelligence in Higher Education: Best Practices. Educ Sci. 2024;15(1):32. DOI:10.3390/educsci15010032.

[3] Chen LD, Aaltonen K, Lampela H, et al. The Design and Implementation of an Educational Chatbot with Personalized Adaptive Learning Features for Project Management Training. Tech Know Learn. 2024;30(2):1-26. DOI:10.1007/s10758-024-09807-5.

[4] Kasneci E, Sessler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ. 2023;103:102274. DOI:10.1016/j.lindif.2023.102274.

[5] Tlili A, Shehata B, Adarkwah MA, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learn Environ. 2023;10(1):15. DOI:10.1186/s40561-023-00237-x.

[6] UNESCO. Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO; 2021.

[7] Dwivedi YK, Kshetri N, Hughes L, et al. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int J Inf Manag. 2023;71:102642. DOI:10.1016/j.ijinfomgt.2023.102642.

[8] Ren H, Shen H, Ye X. Research on the Packaging of Agricultural Products in Sichuan Province Empowered by AIGC in the Context of Intangible Cultural Heritage. Costume Cult Stud. 2024;2(4):44-55. DOI:10.6914/CCS.020405.

[9] Schrepel T, Pentland A. Competition between AI foundation models: dynamics and policy recommendations. Ind Corp Change. 2024;34(5):1085-1103. DOI:10.1093/icc/dtae042.

[10] Xinyu L. The Internal Mechanisms and Practical Paths for Leveraging Generative Artificial Intelligence to Empower High-Quality Development in University Ideological and Political Education. Adult High Educ. 2024;6(8). DOI:10.23977/aduhe.2024.060812.

[11] Raitskaya L, Tikhonova E. Appliances of Generative AI-Powered Language Tools in Academic Writing: A Scoping Review. J Lang Educ. 2024;10(4):5-30. DOI:10.17323/jle.2024.24181.

[12] Samuel FW, Lopez STW, Hossana T. Why Should Users Take the Risk of Sustainable Use of Generative Artificial Intelligence Chatbots: An Exploration of ChatGPT's Use. J Glob Inf Manag. 2024;32(1):1-32. DOI:10.4018/JGIM.365600.

[13] Zhao Y. The synergistic effect of artificial intelligence technology in the evolution of visual communication of new media art. Heliyon. 2024;10(18):e38008. DOI:10.1016/j.heliyon.2024.e38008.

[14] Goetsch M. A workplace safety revolution. ISHN. 2024;58(6):33.

[15] Gupta V. Factors influencing librarian adoption of ChatGPT technology for entrepreneurial support: A study protocol. J Econ Technol. 2024;2:166-173. DOI:10.1016/j.ject.2024.04.006.

[16] Prusty RM, Sudharsan MR, Anand P. Enhancing medical image classification with generative AI using latent denoising diffusion probabilistic model and wiener filtering approach. Appl Soft Comput. 2024;161:111714. DOI:10.1016/j.asoc.2024.111714.

[17] Nadarzynski T, Knights N, Husbands D, et al. Achieving health equity through conversational AI: A roadmap for design and implementation of inclusive chatbots in healthcare. PLOS Digit Health. 2024;3(5):e0000492. DOI:10.1371/journal.pdig.0000492.

[18] Occhipinti AJ, Prodan A, Hynes W, et al. Navigating a stable transition to the age of intelligence: A mental wealth perspective. iScience. 2024;27(5):109645. DOI:10.1016/j.isci.2024.109645.

[19] Vomberg A, Haan DE, Fabian EN, et al. Digital knowledge engineering for strategy development. J Bus Res. 2024;177:114632. DOI:10.1016/j.jbusres.2024.114632.

[20] Bommasani R, Hudson DA, Adeli E, et al. On the Opportunities and Risks of Foundation Models. Stanford, CA: Stanford Center for Research on Foundation Models; 2021. arXiv:2108.07258.

[21] Song Y, Wu K, Ding J. Developing an immersive game-based learning platform with generative artificial intelligence and virtual reality technologies – "LearningverseVR". Comput Educ X Real. 2024;4:100069. DOI:10.1016/j.cexr.2024.100069.

[22] Malespina E. The vital role of school librarians in shaping artificial intelligence acceptable use policies. Knowl Quest. 2024;52(3):20-25.

[23] Jacques B. To ChatGPT or not to ChatGPT: A note to marketing executives. Appl Mark Anal. 2023;9(2):110-116.

[24] Yan M, Wang L. Assessing the copyright infringement risk of generative AI "created" works. In: Proceedings of the 2024 International Conference on Artificial Intelligence and Law. EWA Publishing; 2024. DOI:10.54254/2753-7048/53/20240010.

[25] Jiang W, Qian K, Fan C, et al. Applications of generative AI-based financial robot advisors as investment consultants. In: Proceedings of the 2024 International Conference on Finance and Economics. EWA Publishing; 2024. DOI:10.54254/2755-2721/77/2024MA0057.