Article

Not peer-reviewed version

Advantages and Limitations of Open-Source Versus Commercial Large Language Models (LLMs): A Comparative Study of DeepSeek and OpenAI's ChatGPT

Wan Chong Choi and Chi In Chang *

Posted Date: 24 March 2025

doi: 10.20944/preprints202503.1081.v1

Keywords: large language models; LLMs; artificial intelligence generated content; AIGC; open-source AI; DeepSeek; ChatGPT; transparency and security; domain-specific adaptation; commercial AI platforms

Article

Advantages and Limitations of Open-Source Versus Commercial Large Language Models (LLMs): A Comparative Study of DeepSeek and OpenAI's ChatGPT

Wan Chong Choi [1] and Chi In Chang [2,*]

[1] Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616, USA; wchoi8@hawk.iit.edu

[2] Department of Psychology, Golden Gate University, San Francisco, CA 94105, USA

* Correspondence: cchang@my.ggu.edu

Abstract: Large Language Models (LLMs) emerged as powerful text-processing frameworks with diverse applications, including code generation, text summarization, and research support. Contemporary LLM development followed two approaches: open-source LLMs emphasizing transparency, collaborative refinement, customization, and commercial platforms leveraging proprietary methods within hosted service ecosystems. This paper systematically compared these paradigms through two representative cases—DeepSeek, an open-source LLM, and ChatGPT, a commercial offering. Following a structured analytical framework, we examined essential benefits, significant limitations, and implementation considerations for both systems, supported by recent academic literature. Our findings revealed that open-source LLMs offered superior transparency, customization flexibility, and data governance control, yet faced challenges related to infrastructure requirements, fragmented support ecosystems, and security vulnerabilities. Conversely, commercial LLMs provided robust baseline performance, streamlined deployment, and integrated safety mechanisms, but presented concerns regarding usage costs, architectural constraints, and vendor dependency. Domain-specific analyses demonstrated that DeepSeek excelled in specialized tasks after fine-tuning, particularly in computational domains, while ChatGPT offered higher out-of-the-box performance for general applications. Infrastructure considerations highlighted that while self-hosted models eliminated recurring fees, they demanded substantial technical expertise and computational resources. This comparative analysis aimed to provide valuable insights for organizations and researchers planning to implement LLM technologies in specialized or production contexts, offering a framework for strategic decision-making aligned with specific operational priorities, available expertise, and regulatory requirements.

Keywords: large language models; LLMs; artificial intelligence generated content; AIGC; open-source AI; DeepSeek; ChatGPT; transparency and security; domain-specific adaptation; commercial AI platforms

1. Introduction

Large Language Models (LLMs) have fundamentally transformed natural language processing by leveraging vast computational resources and extensive datasets. Built primarily on Transformer architectures [1], these models generate text with remarkable coherence and contextual awareness. Their applications span diverse domains, including healthcare, software development, and academic research, establishing them as central to AI-driven automation [2,3].

The proliferation of LLMs has given rise to two major types. The first approach embraces open-source architectures, like DeepSeek [4], which publicly disclose model parameters and training methodologies, thus enabling community-driven innovation and specialized adaptations [4–6]. The second approach comprises commercial platforms such as ChatGPT, which operate within

 doi:10.20944/preprints202503.1081.v1

proprietary cloud-based ecosystems. These solutions prioritize deployment efficiency and general-purpose performance but limit user modifications and provide minimal transparency regarding their internal mechanisms. While ChatGPT has gained recognition for its instruction-following capabilities and API flexibility, its training details remain largely undisclosed, restricting users' ability to modify its underlying architecture. Given these different approaches, this paper was guided by the following two research questions:

(1) What are the comparative advantages of open-source versus commercial LLMs?

(2) What limitations and challenges are inherent to open-source and commercial LLMs?

The remainder of this paper provides a structured analysis of DeepSeek and ChatGPT, highlighting their respective strengths, limitations, and implementation implications. Section 2 examines the distinct advantages of open-source and commercial approaches, supported by relevant performance metrics and empirical evidence. Section 3 explores the associated challenges, including computational demands, vendor dependency risks, adversarial vulnerabilities, and alignment with evolving data protection regulations. Section 4 synthesizes these insights to guide LLM selection strategies tailored to different organizational requirements.

2. Advantages of Open-Source versus Commercial LLMs

Before delving into the specific advantages of open-source LLMs, we conducted a comprehensive review of existing literature [4–21]. The detailed findings are shown below.

2.1. Advantages of Open-Source LLMs

2.1.1. Transparency and Community-Driven Innovation

Open-source LLMs, exemplified by DeepSeek, provide comprehensive disclosure of model weights and training methodologies, facilitating community-driven enhancement and scrutiny. This unprecedented level of transparency enabled external researchers to validate performance assertions, replicate experimental conditions, and conduct thorough analyses for potential biases or security vulnerabilities. Empirical investigations demonstrated that broader accessibility to weights and codebase architecture accelerated the identification and remediation of algorithmic inefficiencies, enhancing model stability over longitudinal implementation periods. Furthermore, this collaborative ecosystem positioned researchers to explore novel alignment techniques and quantization methodologies, which were subsequently integrated into the primary development branch for the collective advancement of the field.

2.1.2. Customization and Domain-Specific Adaptation

The accessibility of source code and weight parameters has rendered open-source models particularly amenable to domain-specific adaptations, including specialized applications such as biomedical text summarization and computational code generation. DeepSeek, in particular, introduced domain-targeted variants—DeepSeek Coder for programming tasks and DeepSeek Math for computational reasoning—each of which extended the foundational architecture through specialized fine-tuning protocols. Organizations successfully integrated proprietary datasets with direct parametric access, reconfiguring model behavior to align with narrowly defined tasks (e.g., legal document analysis). Such customization capabilities proved especially valuable in highly regulated environments where local hosting was mandated for compliance adherence or confidentiality preservation.

2.1.3. Lower Long-Term Costs

Despite the substantial computational requirements of training large-scale models, open-source LLMs demonstrated the potential for reducing aggregate expenditures by eliminating recurring licensing fees and per-token inference costs. Upon local deployment of models such as DeepSeek,

 doi:10.20944/preprints202503.1081.v1

high-volume usage no longer incurred incremental API charges, representing a significant economic advantage. For applications characterized by extensive or continuous workloads, self-hosting emerged as financially viable over extended operational periods, mainly when domain-specific inferencing constituted a predominant usage pattern. Recent economic analyses revealed that extensive utilization of open-source LLMs proved more cost-effective than commercial API rate structures once initial hardware investments were amortized over time.

2.1.4. Community Security Audits and Ethical Control

Beyond performance enhancements and feature extensions, open-source initiatives benefited substantially from crowd-sourced security evaluations. Collaborative efforts among researchers and practitioners facilitated the identification of potential vulnerabilities, including data poisoning vectors, adversarial training data, and exploitation techniques that could generate harmful outputs. The transparent nature of these models enabled third-party entities to develop security patches and specialized adversarial training regimens proactively. Additionally, local hosting configurations allowed organizations to implement context-specific moderation systems while maintaining complete autonomy over update cycles. In ethically sensitive domains (e.g., medical text generation), local control of both dataset composition and model parameters proved critical for adherence to stringent regulatory frameworks.

2.1.5. Flexible Hosting Options and Data Governance

Operating on self-managed infrastructure, open-source LLMs enabled organizations to process sensitive data within internal systems, providing a fundamental advantage for industries subject to stringent privacy regulations. Confidential information remained within organizational boundaries, mitigating risks associated with external API transmission. Models were deployed on specialized hardware configurations to optimize performance efficiency or implemented within air-gapped environments to maximize security protocols—deployment scenarios that proved infeasible with commercial-hosted solutions restricted to provider-controlled cloud environments.

2.2. Advantages of Commercial LLMs

2.2.1. High Baseline Performance and Comprehensive Refinement

With ChatGPT as a notable example, commercial platforms leveraged extensive training corpora and substantial computational resources to achieve robust general-domain performance metrics. Their sophisticated development pipelines incorporated rigorous data cleaning procedures, safety-oriented fine-tuning, and multi-stage alignment methodologies (such as Reinforcement Learning from Human Feedback), yielding exceptional out-of-the-box functionality. Organizations lacking sufficient resources or specialized expertise to independently train or fine-tune large-scale models efficiently integrated ChatGPT into existing workflows, effectively outsourcing model optimization processes.

2.2.2. Managed Services and Technical Support

A distinguished characteristic of commercial LLMs was the provision of professional support infrastructure and managed technical environments. Providers such as OpenAI assumed responsibility for system updates, security patches, and scaling operations, alleviating end-user operational complexities, including GPU cluster management and capacity planning considerations. Comprehensive documentation resources, developer portals, and service-level agreements minimized potential downtime risks. For enterprises prioritizing rapid deployment timelines, API-based implementation's operational convenience frequently outweighed architectural transparency's benefits. In contexts where operational reliability was paramount, the streamlined nature of vendor-supported solutions offered substantial advantages.

2.2.3. Centralized Safety Features

Commercial LLMs consistently incorporated sophisticated moderation mechanisms to restrict potentially harmful or biased outputs. These protective measures, developed through extensive user interaction analysis and dedicated classifier models, enabled ChatGPT to reject or redirect problematic queries automatically. Organizations operating within regulated sectors recognized the significant value in these integrated guardrails that mitigated legal and reputational risks. Service providers continuously refined these safety frameworks to address emerging prompt exploitation techniques and malicious use cases—a process that demonstrated greater consistency when maintained by centralized technical teams.

2.2.4. Ongoing Updates and Incremental Upgrades

The hosted nature of commercial LLMs ensured that updates to foundational models or underlying infrastructure remained transparent primarily to end-users. Novel features, security patches, and model enhancements became immediately accessible, maintaining a consistently advanced operational environment. Improvements ranging from sophisticated, multilingual capabilities to refined instruction-following behaviors were implemented without requiring local modifications or retraining procedures. This continuous enhancement cycle appealed particularly to organizations that preferred to avoid establishing or maintaining internal machine learning infrastructures.

2.2.5. Streamlined Ecosystem Integration

Commercial providers typically offer comprehensive suites of complementary services within unified platform environments, including text-to-speech functionality and image recognition capabilities. ChatGPT, for instance, seamlessly integrated with additional OpenAI services, establishing a cohesive ecosystem that simplified application development processes. Efficient integration with established developer tools and project management platforms further accelerated market deployment timelines. This integrated environment represented a compelling advantage for teams seeking to minimize operational overhead when combining multiple artificial intelligence modules.

2.3. Advantages Comparison between DeepSeek and ChatGPT

Table 1 synthesizes the comparative advantages of DeepSeek (representing the open-source paradigm) and ChatGPT (exemplifying commercial approaches). Multiple scholarly investigations illustrated how these distinguishing factors manifested in practical implementation scenarios.

DeepSeek demonstrated exceptional capabilities in transparency, granular customization, and localized data governance, while ChatGPT exhibited distinctive strengths in deployment efficiency and sophisticated safety protocols. Although both approaches achieved comparable performance benchmarks in specific domains, their implementation pathways diverged significantly—open-source solutions necessitated substantial engineering investment, whereas commercial platforms offered expedited deployment trajectories.

Table 1. Advantages Comparison between DeepSeek and ChatGPT.

Advantage	DeepSeek (Open-Source)	ChatGPT (Commercial)
Transparency	Model visibility and modifiable architecture	Training methodology is largely proprietary; limited introspection
Cost Management	No per-token or monthly fees; self-hosted at scale	Pay-per-use or subscription-based model; costs can grow rapidly with heavy usage
Customization	In-depth fine-tuning and domain-level modifications	Fine-tuning available in limited or proprietary form; constrained model architecture
Security & Data Control	On-premise deployment; private data remains in-house	Cloud-based service by default; must trust vendor's data handling

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2025 doi:10.20944/preprints202503.1081.v1

Ease of Onboarding	Demands engineering expertise; the user must maintain infrastructure	Fast integration using managed API, vendor support, and documentation
Safety Mechanisms	Requires custom or community-based content filters	Pre-built moderation and refusal systems, regularly updated by the provider
Innovation Ecosystem	Community-driven patches and performance upgrades	Centralized R&D processes with less user insight; frequent iteration behind closed doors
Baseline Performance	Potentially robust in specialized tasks (e.g., code, math) after tuning	High out-of-the-box performance in general tasks; extensive alignment
Sustainability	Emphasis on more efficient architectures (Mixture-of-Experts, etc.)	Large-scale resources for training and hosting; energy usage not fully disclosed

3. Limitations of Open-Source versus Commercial LLMs

Through a detailed examination of existing literature [4–6,8,11–22], we identified several challenges associated with both open-source and commercial LLMs. The detailed findings are shown below.

3.1. Limitations of Open-Source LLMs

3.1.1. Infrastructure and Compute Overheads

While open-source LLMs circumvent recurring licensing expenses, they necessitate substantial computational infrastructure for both training and inference processes, particularly as model architectures expand to encompass tens or hundreds of billions of parameters. Deploying models like DeepSeek at scale required robust GPU clusters or specialized accelerator hardware. Research teams with limited resources often encountered difficulties managing these capital expenditures or acquiring the technical expertise necessary for implementing effective GPU parallelization strategies. When advanced architectural paradigms such as Mixture-of-Experts were incorporated, debugging or performance optimization complexity increased exponentially.

Furthermore, energy consumption emerged as a significant challenge in the open-source domain. Although initiatives like DeepSeek have explored architecturally efficient designs, large-scale training operations continued to generate substantial carbon emissions, exacerbating environmental concerns. A single training iteration for multi-billion parameter models potentially resulted in ton-level CO_2 emissions, mainly when conducted repeatedly or without optimization. This phenomenon underscored the importance of rigorous data center management protocols, including strategic green energy procurement and sophisticated scheduling algorithms.

3.1.2. Community-Led Quality Assurance

Open-source LLM development relied on volunteer contributions and peer review mechanisms. Despite community endeavors often demonstrating remarkable agility in identifying and rectifying issues, the outcomes exhibited inconsistent quality and methodological rigor. Specific code integrations and model checkpoints manifested variable reliability metrics. The dependence on volunteer-driven update cycles frequently resulted in heterogeneous documentation standards, complicating version rollback procedures, or systematic bug investigation. This non-uniform quality assurance framework presented substantial challenges for organizations requiring stable release cycles and stringent service-level agreements (SLAs).

3.1.3. Security Challenges and Adversarial Risks

Notwithstanding the advantages conferred by community-based security audits (Section 2.1.4), open-source LLMs inherently harbored elevated security vulnerabilities when malicious actors deliberately introduced adversarial elements—including contaminated training datasets or backdoor activation sequences—into public repositories. Given the transparent nature of the entire development pipeline, potential attackers possessed the capability to generate seemingly legitimate but compromised forks of models like DeepSeek. Organizations implementing these models necessitated rigorous validation protocols before adopting contributed code or pre-trained model weights.

 doi:10.20944/preprints202503.1081.v1

Moreover, local deployment architectures placed complete responsibility for content moderation, security policy implementation, and refusal mechanisms on the implementing organization. Many open-source projects released only foundational alignment checkpoints (e.g., DeepSeek Chat) without comprehensive safety guardrails. Teams adopting such models for specialized applications, such as educational interaction systems, were compelled to develop or refine their protective frameworks, requiring additional technical expertise and development cycles.

3.1.4. Fragmented Support and Documentation

Open-source LLM projects frequently operated with limited official documentation resources, dispersed knowledge bases, or decentralized user forums for technical assistance. When critical operational issues manifested—or when domain-specific fine-tuning requirements emerged—resolution processes potentially extended across multi-week or multi-month timeframes, contingent upon the availability and expertise of community volunteers. In contrast, commercial alternatives typically provided centralized support infrastructures, comprehensive knowledge repositories, and guaranteed response intervals. The unpredictability inherent in volunteer-coordinated support structures constituted a significant operational barrier for enterprises operating under stringent production deadlines.

3.1.5. License and Data Usage Uncertainties

Although most open-source models implemented permissive licensing frameworks (e.g., Apache 2.0), the composition of training corpora often lacked complete transparency. Questions regarding including copyrighted materials or sensitive personal information within training datasets persisted, raising substantial intellectual property and privacy concerns. From a jurisprudential perspective, datasets derived from open internet sources potentially contained copyrighted passages, personally identifiable information, or other restricted content. Organizations contemplating large-scale deployment of open-source LLMs required extensive legal consultation or implementation of rigorous data handling protocols, introducing additional operational overhead.

3.2. Limitations of Commercial LLMs

3.2.1. Potentially High Usage Costs

While commercial platforms facilitated expedited prototyping and seamless scaling, they frequently implemented consumption-based billing structures that became financially prohibitive under high-traffic or computationally intensive operational scenarios. For instance, ChatGPT employed per-token pricing models; as query volumes increased, costs accumulated rapidly. Such usage-dependent fee structures potentially created financial bottlenecks for emerging ventures or established enterprises conducting continuous inference at scale. Additionally, specific advanced capabilities—including higher-tier fine-tuning options or premium support services—incurred supplementary charges, potentially inflating total expenditures significantly.

3.2.2. Reduced Architectural Control

Proprietary models generally permitted superficial customization through partial fine-tuning or instruction-based prompting but precluded fundamental modifications to core architectural components. This limitation constrained organizations seeking deeper adaptations or domain-specific optimizations beyond basic instruction refinement. Researchers aiming to integrate specialized modules—such as proprietary numerical reasoning systems—encountered insurmountable barriers when attempting to modify internal layers of models like GPT. This dependence on closed, monolithic architectures potentially impeded advanced research initiatives or the development of highly specialized derivative models.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2025 doi:10.20944/preprints202503.1081.v1

3.2.3. Vendor Lock-In and Platform Dependency

Excessive reliance on individual commercial providers introduced several strategic vulnerabilities for organizations. Unexpected policy revisions, API deprecation cycles, or pricing restructuring potentially substantially disrupted dependent services. Migration away from proprietary platforms often presented complex technical challenges, as alternative solutions frequently failed to replicate equivalent performance characteristics or necessitated extensive architectural reconfiguration. Compliance with evolving regulatory frameworks became inextricably linked to vendor-specific protocols, introducing additional uncertainty for organizations operating in highly regulated sectors such as financial services or healthcare.

3.2.4. Opaque Training and Bias Discovery

Although commercial LLMs typically underwent extensive filtering and alignment processes, the methodological details and data sources utilized during training remained confidential. Users consequently possessed limited insight into potential biases or problematic content embedded within model parameters. For example, insufficient representation of specific demographic groups or languages within training corpora potentially resulted in performance disparities across different user populations. External evaluation procedures relying on "black-box" testing methodologies revealed only partial information, leaving significant unknowns regarding data provenance or preprocessing methodologies.

3.2.5. Data and Privacy Concerns

By architectural design, commercial offerings such as ChatGPT typically require transmitting user queries to provider-controlled cloud infrastructure for inference processing. This approach introduced potential exposure of confidential corporate information or personal data to external entities. Despite vendor commitments to robust security frameworks, the possibility of data breaches or unauthorized utilization could not be definitively eliminated. Furthermore, on-premise deployment options—while occasionally available for enterprise clients—lacked the straightforward implementation pathways characteristic of self-hosted open-source alternatives. Regulatory frameworks governing data residency, particularly stringent in sectors governed by HIPAA or GDPR, potentially limited the applicability of hosted commercial LLMs in specific contexts.

3.3. Limitations Comparison between DeepSeek and ChatGPT

Table 2 presents a comparative analysis of challenges associated with the open-source DeepSeek model and its commercial counterpart, ChatGPT. The cited research underscored that neither approach was devoid of significant limitations, with optimal selection typically contingent upon organizational priorities and constraints.

The comparative limitations enumerated in Table 2 revealed two predominant thematic patterns. First, the open-source implementation pathway necessitated substantial internal investment in computational resources, security frameworks, and community-based quality assurance mechanisms. Second, the commercial implementation approach alleviated numerous engineering challenges but potentially introduced opaque development cycles, escalating usage costs, and vendor dependency risks.

Table 2. Limitations Comparison between DeepSeek and ChatGPT.

Limitation	DeepSeek (Open-Source)	ChatGPT (Commercial)
Resource Demands	High GPU/TPU costs for large-scale training/inference	Pay-as-you-use pricing; cumulative fees can grow for large usage
Support Ecosystem	Volunteer-led forums; no formal SLAs	Dedicated support with SLAs and official documentation
Security and Moderation	Requires custom integration of filtering or adversarial training	Built-in refusal system and moderation layers, but user must accept vendor's approach
Data Governance and Compliance	Users must handle data privacy tools locally; uncertain training data sources	Data goes to the external cloud; vendor's compliance approach may not align perfectly with user's requirements
Architecture Control	Total access for specialized modifications; can be complex to maintain	Architecture locked; advanced domain adaptation is limited to partial fine-tuning or prompt engineering
Scalability & Sustainability	Greater flexibility in hardware but large-scale training is resource-intensive; possible heavy carbon footprint	Provider invests in HPC infrastructure; usage-based fees can hamper sustainable scaling if demand grows exponentially

4. Conclusions

Open-source LLMs, epitomized by DeepSeek, and commercial platforms, exemplified by ChatGPT, each presented a distinct equilibrium of advantages and constraints. Open-source initiatives championed transparency, enabling more granular control over security protocols and domain-specific optimization. This methodology proved valuable in sectors necessitating rigorous data governance or specialized NLP workflows. Nevertheless, self-hosting large-scale models incurred substantial infrastructure demands, fragmented support ecosystems, and vulnerability to unvetted community modifications that potentially compromised reliability or introduced security vulnerabilities.

Conversely, commercial solutions such as ChatGPT offered intuitive APIs, robust intrinsic safety mechanisms, and continuous vendor-orchestrated enhancements, rendering them attractive for organizations seeking expeditious deployment of advanced NLP capabilities. However, recurring utilization fees, non-transparent training methodologies, and reliance on external providers emerged as potential strategic limitations. Data privacy constituted another critical consideration: while enterprise clients might entrust sensitive information to a commercial cloud infrastructure, others found such arrangements excessively restrictive for regulated domains that mandated comprehensive on-premise control.

Ultimately, the selection between DeepSeek and ChatGPT necessitated aligning these trade-offs with organizational imperatives, available technical expertise, and regulatory parameters. Teams pursuing open collaboration, extensive customizability, or strictly localized data processing identified compelling value propositions in open-source LLMs, notwithstanding the associated overhead. Organizations prioritizing turnkey integration, consistent security updates, and comprehensive capabilities discovered more excellent utility in commercial ecosystems. As the discipline of large language modeling continues evolving rapidly, hybridizing open-source transparency with commercial-grade reliability may constitute an optimal solution for numerous adopters. This comparative analysis underscores how these two paradigms can be strategically integrated to address diverse real-world priorities, from fiscal constraints to compliance requirements.

References

1. A. Vaswani *et al.*, 'Attention is all you need', *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.
2. D.-M. Petroșanu, A. Pîrjan, and A. Tăbușcă, 'Tracing the influence of large language models across the most impactful scientific works', *Electronics*, vol. 12, no. 24, p. 4957, 2023.
3. Z. Guo *et al.*, 'Evaluating Large Language Models: A Comprehensive Survey', *ArXiv*, vol. abs/2310.19736, 2023, doi: 10.48550/arXiv.2310.19736.
4. F. Neha and D. Bhati, 'A Survey of DeepSeek Models', *Authorea Prepr.*, 2025.

5. G. Zhang *et al.*, 'Closing the gap between open source and commercial large language models for medical evidence summarization', *NPJ Digit. Med.*, vol. 7, no. 1, p. 239, 2024.
6. R. Sapkota, S. Raza, and M. Karkee, 'Comprehensive analysis of transparency and accessibility of chatgpt, deepseek, and other sota large language models', *ArXiv Prepr. ArXiv250218505*, 2025.
7. C. Davis *et al.*, 'Prompting open-source and commercial language models for grammatical error correction of English learner text', *ArXiv Prepr. ArXiv240107702*, 2024.
8. M. M. H. Manik, 'ChatGPT vs. DeepSeek: A Comparative Study on AI-Based Code Generation', *ArXiv Prepr. ArXiv250218467*, 2025.
9. D. Vake, B. Šinik, J. Vičič, and A. Tošić, 'Is Open Source the Future of AI? A Data-Driven Approach', *Appl. Sci.*, vol. 15, no. 5, p. 2790, 2025.
10. H. Wang *et al.*, 'Adapting Open-Source Large Language Models for Cost-Effective, Expert-Level Clinical Note Generation with On-Policy Reinforcement Learning', *ArXiv*, vol. abs/2405.00715, 2024, doi: 10.48550/arXiv.2405.00715.
11. X. Bi *et al.*, 'Deepseek llm: Scaling open-source language models with longtermism', *ArXiv Prepr. ArXiv240102954*, 2024.
12. C. I. Chang, W. C. Choi, and I. C. Choi, 'A Systematic Literature Review of the Opportunities and Advantages for AIGC (OpenAI ChatGPT, Copilot, Codex) in Programming Course', in *Proceedings of the 2024 7th International Conference on Big Data and Education*, 2024.
13. C. I. Chang, W. C. Choi, and I. C. Choi, 'Challenges and Limitations of Using Artificial Intelligence Generated Content (AIGC) with ChatGPT in Programming Curriculum: A Systematic Literature Review', in *Proceedings of the 2024 7th Artificial Intelligence and Cloud Computing Conference*, 2024.
14. W. C. Choi, I. C. Choi, and C. I. Chang, 'The Impact of Artificial Intelligence on Education: The Applications, Advantages, Challenges and Researchers' Perspective', *Preprints.org*, 2025, doi: 10.20944/preprints202501.1420.v1.
15. Y. Cao *et al.*, 'A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt', *ArXiv Prepr. ArXiv230304226*, 2023.
16. P. P. Ray, 'ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope', *Internet Things Cyber-Phys. Syst.*, vol. 3, pp. 121–154, 2023.
17. Y. Yao, J. Duan, K. Xu, Y. Cai, Z. Sun, and Y. Zhang, 'A survey on large language model (llm) security and privacy: The good, the bad, and the ugly', *High-Confid. Comput.*, p. 100211, 2024.
18. A. Liu *et al.*, 'Deepseek-v3 technical report', *ArXiv Prepr. ArXiv241219437*, 2024.
19. Y. Peng, Q. Chen, and G. Shih, 'DeepSeek is open-access and the next AI disrupter for radiology', *Radiology Advances*, vol. 2, no. 1. Oxford University Press, p. umaf009, 2025.
20. T. Gao, J. Jin, Z. T. Ke, and G. Moryoussef, 'A Comparison of DeepSeek and Other LLMs', *ArXiv Prepr. ArXiv250203688*, 2025.
21. Y. Peng *et al.*, 'From GPT to DeepSeek: Significant gaps remains in realizing AI in healthcare', *Journal of Biomedical Informatics*. Elsevier, p. 104791, 2025.
22. F. W. Liu and C. Hu, 'Exploring vulnerabilities and protections in large language models: A survey', *ArXiv Prepr. ArXiv240600240*, 2024.