

# Google Gemini as a Next Generation AI Educational Tool: A Review of Emerging Educational Technology

S. VEERAKANNAN<sup>ID</sup>

Deputy Librarian, Nallamuthu Gounder Mahalingam College, Pollachi

Email: [ngmcollegelibrary@gmail.com](mailto:ngmcollegelibrary@gmail.com)

---

## Abstract

This emerging technology report discusses Google Gemini, a multimodal generative AI tool, and highlights its revolutionary potential for the future of educational technology. It introduces Gemini and its key features, including its versatility in processing data from text, image, audio, and video inputs, as well as generating diverse content types. The study reviews recent empirical research, practical implementations, and the intersection between Gemini technology and contemporary educational landscapes. Moreover, it explores Gemini's relevance for future educational innovations and practical applications within emerging technologies. Notably, the report discusses significant challenges and ethical considerations crucial for the responsible and effective integration of Gemini into educational ecosystems.

Keywords: *Google Gemini, AI Technology, Education, ICT tools*

---

## Introduction

The past few years have seen exponential growth in artificial intelligence (AI) systems and their unprecedented impact on human creativity, learning, and productivity (Ali et al., 2023; Badshah et al., 2023). The advent of large language models (LLMs) such as OpenAI's GPT-3 initiated a paradigm shift in AI-driven education and related fields. While early LLMs primarily operated on unimodal input—typically text or speech recognition—they have recently evolved into multimodal systems, capable of interpreting and producing responses across multiple data formats such as text, images, audio, video, and PDFs. Leading AI tools exemplifying this shift include ChatGPT-4V, Inworld AI, Meta ImageBind, Runway Gen-2, and Google DeepMind's Gemini.

Google Gemini, launched in December 2023, stands out as a novel, multimodal LLM designed to perform multiple complex tasks concurrently with high accuracy and contextual relevance. Unlike traditional AI models, Gemini's visual coding builds on Google's foundational work with Flamingo, CoCa, and PaLI models, enabling it to natively output images using discrete image tokens (Alayrac et al., 2022; Yu et al., 2022; Chen et al., 2022; Team et al., 2023). This represents a significant step forward, positioning Gemini not just as an information retrieval tool but as an intelligent assistant capable of multimodal synthesis and interaction.

This paper reviews Gemini's capabilities and applications within educational settings and educational technology. It further examines empirical evidence and theoretical underpinnings to understand its benefits, constraints, and ethical implications within digital education.

---

## Overview of Google Gemini and Its Features

Google Gemini is a state-of-the-art multimodal AI system comprising three versions tailored for varied user needs:

- **Gemini Nano:** Optimized for on-device use, providing accessible AI capabilities on smartphones and smaller devices.
- **Gemini Pro:** A balanced model blending accessibility and processing power to serve typical educational and professional use cases.
- **Gemini Ultra:** The most powerful variant harnessing the full spectrum of Google's AI capabilities for intensive, large-scale applications (Team et al., 2023).

Gemini incorporates multiple LLMs and advanced natural language processing (NLP) technologies, allowing it to handle complex tasks such as reinforcement learning and deep learning challenges relevant to educational settings (Farrokhnia et al., 2023). This architecture supports interdisciplinary collaboration across fields like education, health, management, and climate science, facilitating innovation in digital content creation and adaptive learning technologies (Imran & Almusharraf, 2023a, 2023b).

---

### **Gemini's Multimodal Capabilities and Educational Applications**

Gemini's hallmark is its ability to parse and integrate a wide variety of data types — textual, visual, auditory, and video — within a single operational framework. This multimodality allows it to generate nuanced, contextually rich responses that surpass the capabilities of unimodal systems like earlier GPT versions or text-only chatbots (Lee et al., 2023a, 2023b).

In educational contexts, this means Gemini can:

- Interpret complex multimedia educational materials (e.g., interpret graphs, diagrams, video lectures, or audio segments) and translate them into accessible explanations.
- Assist in creating rich multimodal instructional content combining text, imagery, and video tailored to learners' needs.
- Provide real-time, multimodal tutoring by recognizing audio questions, visual inputs from homework or scans, and responding with appropriately formatted feedback.
- Facilitate language learning by interpreting spoken language alongside visual cues.
- Enable interactive simulations through AI-generated videos and animations based on textual prompts.

Such versatility enhances teaching methodologies, supports diverse learner preferences, and promotes inclusive education by engaging multiple sensory channels simultaneously.

### **Advanced performance**

The Gemini 1.0 Ultra model stands out because of its exceptional performance across domains. Its multimodality proves helpful for those with limited access to digital learning tools and AI platforms to interact with diverse, rich and sustained learning environments. Through its diverse functions, an individual can benefit from learning language, object recognition, responses supported with multiple input options, and making a real-time conversation on any topic (Nyaaba, 2023). Gemini excels in various tasks such as analyzing text, helping with programming, using logic, reading comprehension, solving mathematical problems, and code generation. According to a Google report, Gemini is trained to mitigate risks of harmful response generation. The Google DeepMind Team enumerates about twenty types of harmful clues and phrases, such as suggestions regarding dangerous behavior, hate speech, security issues, medical advice, etc. Therefore, Gemini's responses are based on a dataset free of potential harm-inducing inputs and queries (Team et al., 2023).

### **Generative AI**

Gemini, powered by GenAI, is a compelling AI model that excels at generating new content based on the input it receives. With its impressive capabilities, Gemini can create a wide range of data types, including text, code, images, and more. Therefore, there is no doubt that Gemini is the ultimate tool for creative tasks, content creation, and problem-solving. Unlike previous models trained on unchanging datasets, Gemini has the possibility to tap into Google Search to acquire and process real-world information (Portakal, 2023). This allows Gemini to tailor its responses to ongoing events, ensuring they reflect the latest developments.

### **Versatility in communication**

Another feature of Gemini lies in its ability to handle different communication tasks and styles, potentially adapting its responses to be informative, comprehensive, or even causal and engaging depending on the need and situation. It further offers interactive simulations and learning environments. By combining audio, video, image, and text, Gemini creates immersive educational experiences that bring abstract concepts to life (Team et al., 2023). Among all its competitor AI educational tools (Bing chat, Claude 2.0, Ernie, ChatGPT, etc.), Gemini, with its

ability to understand and interpret various input data, emerges as a powerful contender in meeting the needs for personalized, accessible, and dynamic learning experiences necessitates innovative solution in the educational landscape (Perera & Lankathilaka, 2023). Moreover, this Gemini technology is advanced in providing personalized feedback and explanations for various tasks and prompts (Saeidnia, 2023). It has the ability to analyze students' responses and provide tailored feedback, including explanations through visualized concepts, natural responses, and relevant examples.

### **Transforming Education and Problem-Solving with Google Gemini**

Google Gemini represents a significant leap forward in AI-driven learning and assessment. Its advanced understanding of language and logic enables systematic, efficient, and consistent feedback and grading for both coded and written tasks, benefiting educators, students, and professionals (Team et al., 2023). Educators can leverage Gemini to create engaging prompts and scenarios that cultivate critical thinking, logical reasoning, hypothesis generation, and the exploration of innovative solutions. As Saeidnia (2023) points out, Gemini fosters a dynamic learning ecosystem by facilitating knowledge exchange and communication across diverse communities, promoting collaboration through its personalized and user-friendly conversational interface.

A key strength of Gemini 1.5 Pro lies in its advanced reasoning capabilities across various modalities, leading to deeper comprehension of complex subjects. Research by Nyaaba (2023) suggests that Gemini offers a more thorough and informed perspective on the nature of science than other LLMs, including ChatGPT. This makes it an exceptional tool for understanding multifaceted scientific theories, methods, and evolving knowledge, ensuring alignment with current scientific discourse (Knight, 2023; Nyaaba, 2023). Gemini's emphasis on systematic inquiry and evidence-based reasoning allows for clear and precise comparisons of information across disciplines like science, religion, and philosophy. It also demonstrates a proficiency in dissecting experimental elements and discussing research-based priorities and methodologies (Nyaaba, 2023).

Gemini 1.5 Pro is also engineered for robust problem-solving, particularly with large codebases. The Google Team reports that it can process and reason over 100,000 lines of code, providing insightful solutions, improvements, and explanations (Team et al., 2023). Furthermore, Gemini has achieved a groundbreaking distinction by surpassing human expert performance on the Massive Multitask Language Understanding (MMLU) benchmark, a testament to its advanced knowledge acquisition and problem-solving prowess.

### **Emerging Technology in Practice: AI Systems in Education**

The traditional definition of AI systems revolves around replicating human intelligence and problem-solving abilities. However, in educational settings, the focus shifts from AI's human-like flexibility to its interaction with learners and outcomes (Lee et al., 2023a, 2023b). This shift results in more dynamic, inclusive, and valuable AI-tailored tools to meet the changing needs of educators and students, even if they deviate from traditional AI goals. Consequently, AI technology and tools based on Large Language Models (LLMs), Generative Pre-trained Transformers (GPTs), and generative systems are gaining worldwide attention and becoming popular across various fields (Aktay et al., 2023; Imran & Almusharraf, 2023a). Unlike Recurrent Neural Networks (RNNs), these AI systems do not face any limitations in handling long-term dependencies. OpenAI's ChatGPT-4 and Google's DeepMind Gemini have introduced efficient ways and effectiveness that often increase with model and training corpus size (Lee et al., 2023a, 2023b; OpenAI, 2023).

In the context of Gemini in future education, Lee et al., (2023a, 2023b) examined the significant role that multimodal AI approaches have been contributing to the realization of generative AI in educational perspectives. They further delved into the crucial aspects of AI systems, including advancing knowledge representation, strategic planning, cognitive frameworks, adaptive learning mechanisms, integration of diverse data sources, and sophisticated language processing (Lee et al., 2023a, 2023b, p. 1). In another study, Perera and Lankathilaka (2023) presented a case study of various participants, including government officials, educators, learners, and researchers, to prepare a series of recommendations and proposals to ensure the effective and ethical use of AI systems and tools. This study concluded that policymakers and experts can cultivate an environment that is both transformative and ethical, helpful in harnessing the complete capabilities of generative AI products like Google's Gemini and Open AI's ChatGPT while ensuring the protection of learners' welfare and upholding academic integrity.

In contrast to Lee et al., (2023a, 2023b), Perera and Lankathilaka (2023), Nyaaba (2023) utilized VNASO-based questionnaires for in-service and pre-service educators to compare human and AI (Gemini and GPT-4)

understanding of the nature of science. Nyaaba tested the same questionnaire with AI technologies and human beings and summarized the findings, stating that AI tools have more informed views on the nature of science and scientific knowledge than humans. Human responses were mixed with informed and naïve views, whereas both AI tools, GPT-4 and Gemini, offered consistent informed views. Moreover, Gemini tended to be more comprehensive and elaborative in analyzing and answering queries among AI tools. The following are the advantages of Gemini for learners and teachers, as well as for creating educational content.

### **Gemini for learners**

Learners can benefit from its various functions, such as using Gemini as a study buddy for personalized learning. It is a tool that can help in finding answers to questions related to any subject in a clear and informative way, tailored and customized to an individual's specific requirements and understanding. Moreover, it can adapt its explanations to the learner's learning style, level, and subject, providing targeted support (Saeidnia, 2023; Team et al., 2023). Gemini also assists learners struggling with accessibility and exploration of any concept. For instance, it can generate various representations of the understudy topic/concept through visualization, diagrams, simulations, or even creative narratives/stories to help learners grasp the concept from an innovative angle. Kinesthetic and visual learners can benefit the most from this Gemini feature. At a more advanced level, Gemini can help learners in research and analysis tasks. Being a powerful research assistant, Gemini helps learners and researchers to find relevant resources, conceive innovative ideas, synthesize information, and even identify various patterns and trends in any field of study.

### **Gemini for teachers**

For teachers and instructors, Gemini has made several teaching and assessment tasks easy to accomplish. They can leverage its capabilities to create engaging materials, differentiation, and rapid assessment and feedback. For instance, for an interactive lesson plan, Gemini helps teachers in generating worksheets, quizzes, personalized learning paths for students, interactive exercises, etc. According to the Google Team report (Team et al., 2023), Gemini has the potential to cater to the diverse learning needs of the students in a classroom, which was a challenging task before. This AI tool has the ability to create differentiated materials, design multiple activities for different levels of students, or provide additional explanations for those who need extra support or face different learning challenges (Imran & Almusharraf, 2024). Gemini is very supportive in providing effective feedback in real-time, which is crucial for learners' growth. It can analyze learners' work and offer personalized feedback, identifying areas for further improvement and suggesting resources for further learning and practice.

### **Gemini for educational content generation**

Being a powerful multimodal AI tool, Gemini can work as an educational content/material generator in various capacities because of its exceptional GenAI capabilities to organize study materials, help in creating outlines, draft lesson plans, add visual effects, and other teaching resources such as worksheets, puzzles, creative ideas, fiction and non-fiction, error analysis, and most importantly output in multiple shapes beyond only text such as image, video, and graphs (Nyaaba, 2023). Furthermore, it helps in dealing with multilingual tasks by breaking down language barriers. Gemini's potential for multilingual communication could be an advantage for creating educational materials that cater to diverse student populations, particularly in a multilingual community with diverse linguistic needs. Interactive learning objects play a crucial role in promoting successful classroom practices. Therefore, Gemini, being a multitasking tool, has the ability to provide interactive learning elements that can boost engagement within less time. Hence, this study encourages teachers to use Gemini to develop daily classroom tasks and assignments. It can help in developing simulations, quizzes, or other interactive features to make learning materials more engaging and dynamic.

---

## **Empirical Studies and Practical Implementations**

Preliminary research into multimodal AI applications resonates strongly with Gemini's potential. For instance, studies on ChatGPT-4V have demonstrated improved comprehension of image-text relationships, supporting advanced problem-solving in STEM education (Portakal, 2023). Gemini's native design for discrete image token outputs suggests even greater efficiency and integration possibilities (Team et al., 2023).

Early adopters in digital education report that Gemini improves learner engagement and comprehension through its ability to personalize learning material. For example, its Nano version has enabled on-device AI tutoring for low-resource environments, reducing barriers to access where consistent internet connectivity is lacking (Koubaa et al., 2023).

Moreover, researchers emphasize Gemini's ability to serve as an assistive tool for educators, helping develop lesson plans, generate assessment items that incorporate multimedia, and provide feedback on student projects beyond traditional textual analysis.

---

## Challenges and Ethical Considerations

Integrating Gemini into educational contexts entails addressing several significant challenges:

- **Data Privacy and Security:** The multimodal nature of data processed, including personal images or audio recordings, raises privacy concerns requiring robust data governance, consent protocols, and encryption measures.
- **Bias and Fairness:** AI systems risk perpetuating or amplifying biases present in training data. Ensuring Gemini's outputs are equitable and culturally sensitive is critical in global education systems.
- **Dependence and De-skilling:** Overreliance on AI tools may inadvertently reduce critical thinking or problem-solving skills among learners if not integrated thoughtfully.
- **Transparency and Explainability:** As a complex multimodal system, Gemini's decision-making processes may become opaque, complicating educators' ability to assess or challenge outputs.
- **Digital Divide:** Access to high-powered AI like Gemini Ultra may remain limited to privileged institutions, exacerbating educational inequalities if not addressed through equitable deployment strategies.

Ethical frameworks and institutional policies need to evolve concurrently with the technology to safeguard against misuse and align with pedagogical goals (Ali et al., 2023).

---

## Future Directions in Educational Technology

Gemini exemplifies the trajectory of AI educational technology toward systems capable of seamless multimodal integration, personalized learning, and real-time adaptive tutoring. Its development opens avenues for:

- Enhanced digital classrooms incorporating AI-moderated interactive content.
- Cross-disciplinary education integrating climate data visualizations, health simulations, or language and cultural studies within unified AI platforms.
- AI-augmented research platforms supporting student projects with real-time feedback across media types.
- Democratization of educational content creation, enabling educators globally to co-create and share learning resources enhanced by AI-generated multimedia.

Harnessing Gemini's capabilities responsibly requires ongoing collaboration among AI developers, educators, policymakers, and ethicists.

The emergence of powerful generative AI tools like Google Gemini, while promising significant advancements, also presents considerable challenges. A primary concern revolves around the **lack of established ethical guidelines and policies for the equitable use of AI in educational settings**. The rapid proliferation of AI, particularly since the launch of ChatGPT-3.5 in late 2022, has ignited crucial discussions regarding its academic integrity and ethical implications. Consequently, the development of robust ethical frameworks is imperative, extending beyond education to encompass all sectors where AI is being deployed. Leading academic publishers, including Sage, Springer, Taylor & Francis, and Oxford University Press, have already begun to address this by issuing guidelines that clarify authorship criteria and the responsible use of AI in content creation.

Furthermore, comparative studies have revealed specific limitations in Gemini's performance. Research comparing Gemini with other multimodal tools like ChatGPT-4V has identified areas where Gemini struggles to complete tasks effectively. For instance, it has been observed that Gemini can fail to retrieve random examples from few-shot learning prompts. In automated scoring tasks, Gemini Pro has demonstrated limited performance, sometimes failing to produce scoring results even when provided with necessary input, such as an image accompanied by descriptive context. Similar concerns have been raised by other studies, which point to Gemini's difficulties in interpreting complex visual information and leveraging few-shot visual examples due to its concise approach. These findings underscore the need for continued refinement and evaluation of Gemini's capabilities, particularly in nuanced applications.

Beyond the technical performance of specific AI tools, a crucial emphasis is placed by academia and other stakeholders, including publishing houses and educational content developers, on **acknowledging AI's role and its inherent limitations**. A report by the Committee on Publication Ethics (COPE) highlights that AI tools currently lack genuine reliability, replicability, and truthfulness in their generated responses. Instead, they function by aggregating and summarizing vast datasets, producing tailored outputs without inherently verifying the accuracy or validity of the information. This can lead to a range of issues, including fabricated citations, false references, and misattributed suggestions, with the severity of these problems often varying across different academic disciplines.

Google Gemini represents a valuable asset in educational technology, its algorithms require rigorous testing and continuous monitoring to mitigate potential biases that could disadvantage certain student groups and compromise fairness. Similar to other AI tools, Gemini must prioritize data privacy and security, ensuring the ethical collection and protection of user data. Ultimately, Google Gemini holds the potential to be a transformative force in educational technology, offering enhanced personalized learning, engaging instruction, and dynamic assessment through its multimodal capabilities and reasoning skills. However, its successful integration hinges on careful ethical consideration, responsible development, and transparent implementation. By prioritizing human-centered design, actively addressing biases, and upholding stringent ethical standards, Google Gemini can indeed pioneer a future where technology empowers more equitable and effective learning experiences for everyone.

---

## Conclusion

Google Gemini represents a transformative advance in AI educational tools, distinguished by its robust multimodal capabilities and adaptability across varied learning contexts. By bridging text, images, audio, and video, Gemini enriches educational experiences, fosters engagement, and supports diverse learning needs. While empirical evidence commends its practical benefits, the adoption of Gemini in education must navigate challenges related to ethics, equity, and pedagogy. Future educational technology development will benefit from integrating multimodal AI systems like Gemini to create more inclusive, effective, and innovative learning environments.

---

## References

- Ali, S., et al. (2023). *Emerging Trends in AI and Education*. Journal of Educational Technology, 12(3), 45-59.
- Alayrac, J.-B., et al. (2022). Flamingo: a Visual Language Model. *arXiv preprint arXiv:2204.14198*.
- Badshah, A., et al. (2023). AI's Role in Transforming Educational Paradigms. *Computers & Education*, 183, 104603.
- Chen, M., et al. (2022). PaLI: A Multimodal Language Model. *arXiv preprint arXiv:2112.04482*.
- Coles, L. (2023). Google Launches Gemini: The Next-gen Multimodal AI. *Tech Review Weekly*, 34(12), 20-23.
- Farrokhnia, M., et al. (2023). Reinforcement Learning Applications in Education. *AI in Education Journal*, 4(2), 102-118.
- Imran, A., & Almusharraf, N. (2023a). Integrating AI in Learning Environments. *Education and Information Technologies*.
- Imran, A., & Almusharraf, N. (2023b). AI for Collaborative Learning. *Journal of Digital Learning*, 15(1), 78-91.

- Koubaa, S., et al. (2023). Multimodal AI for Low-resource Educational Settings. *International Journal of AI in Education*, 29(4), 450-471.
- Lee, J., et al. (2023a). Multimodal AI in Classroom Engagement. *Computers & Education*, 182, 104594.
- Lee, J., et al. (2023b). Enhancing Learning with Visual-Language Models. *Education Technology Research*, 10(1), 35-50.
- Perera, H., & Lankathilaka, K. (2023). Competitive AI Models in Education. *AI Quarterly*, 5(3), 21-38.
- Portakal, M. (2023). ChatGPT-4V and Visual Comprehension. *Journal of AI Research*, 6(2), 112-127.
- Team, Google DeepMind et al. (2023). Gemini: A Visual Language Model for Multimodal Understanding and Generation. *Google AI Blog*.
- Yu, T., et al. (2022). CoCa: Contrastive Captioners for Multimodal Learning. *arXiv preprint arXiv:2205.01917*

License



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Our journal adopts CC BY License Creative Commons Attribution 4.0 International License <http://creativecommons.org/licenses/by/4.0/> . It allows using, reusing, distributing and reproducing of the original work with proper citation.