

ARTIFICIAL INTELLIGENCE (AI)-GENERATED VISUAL CONTENT FOR INCLUSIVE LEARNING MATERIALS: A NARRATIVE LITERATURE REVIEW

Mamman Obadiah Jatau

Department of Creative Arts Education
Federal University of Education,
Pankshin - Plateau State, Nigeria
Email: immortalmark66@gmail.com

&

Obadofin Samuel Bamidele, Ph.D

Department of Fine Arts,
Faculty of Environmental Sciences
Lagos State University Ojo, Nigeria
Email: samuelobadofin2016@gmail.com

Abstract

Artificial intelligence is increasingly built into educational technology as a mechanism for inclusive learning, but the evidence on AI-generated visual content is still scattered across disability categories and applications. This paper draws on a narrative review of peer-reviewed studies published between 2021 and 2026 to do three things: (a) synthesise empirical evidence across three applications, namely automated image description, text-to-image comprehension support, and virtual pedagogical agents; (b) evaluate their effectiveness, limitations, and equity implications against the Universal Design for Learning (UDL) framework and WCAG 2.2; and (c) draw practical, evidence-informed recommendations for the stakeholders involved. The findings point to real potential for personalising and scaling inclusive education, but also to recurring problems: inconsistent description quality, uneven educator readiness, algorithmic bias, and a research base that under-represents learners in the Global South. Responsible deployment, the review concludes, depends on human-in-the-loop oversight, targeted

teacher training, and co-design with disabled learners themselves. Recommendations are organised for curriculum developers, classroom teachers, technology designers, policymakers, and researchers working across AI, visual generation, and inclusive pedagogy.

Keywords: Artificial Intelligence, Inclusive Education, Visual Content Generation, Accessibility, Universal Design for Learning, Autism Spectrum Disorder, Dyslexia, Pedagogical Agents

Introduction

Education systems worldwide face growing obligations to provide equitable, accessible, and personalised learning for all students, including the more than one billion people globally who live with some form of disability (World Health Organization [WHO], 2023). Learners with dyslexia, autism spectrum disorder (ASD), and visual impairments in particular have long faced barriers built into standardised, text-based, single-modality instruction (Chalkiadakis et al., 2024). Statutory inclusion mandates such as the United Nations Convention on the Rights of Persons with Disabilities (CRPD) exist, but the gap between what these mandates require and what actually happens in differentiated classroom instruction remains wide, especially in low- and middle-income countries.

AI technologies have moved into educational contexts quickly since 2022, and they offer some genuine ways to address this gap. Generative systems, including large multimodal language models, text-to-image generators, and AI-driven avatars, give educators tools that can, in principle, produce differentiated visual materials at scale, describe images automatically for learners who cannot see them, and provide interactive rehearsal environments for learners who find face-to-face communication difficult (Paglialunga & Melogno, 2025). These capacities line up with Universal Design for Learning (UDL), the framework developed by the Center for Applied Special Technology (CAST), which calls for multiple means of representation, action and expression, and engagement in learning

materials (Saborío-Taylor & Rojas-Ramírez, 2024).

The growing interest for AI in education is exponentially growing, but the specific - evidence to visual content generation for inclusive purposes is still in the neonatal stage, fragmented, and methodologically irregular. Many academic reviews treat AI and visual disability broadly, without looking closely at the importance of visual media to learners with specific access requirements. Therefore, this paper focuses on three applications: automated alt-text generation, text-to-scene generation, and virtual pedagogical agents. This paper review attempts to answers some specific timely problem because multimodal large language models has become widely available to the public in late 2022 and generative AI has gone from a computational curiosity to a tool teachers use in ordinary classrooms. Research in many domains has not adequately kept pace and educators are adopting these systems faster than anyone is producing consolidated evidence on how they work for learners with disabilities specifically. What exists tends to split into two camps: techno-optimistic pilot studies with little critical depth, and ethical critiques that stay abstract and offer teachers little practical guidance (Paglialunga & Melogno, 2025) and from the forgoing, this shift from text-only to multimodal AI has been making most pre-2022 paper reviews obsolete. This present review is necessary not because many literatures contradict itself in some areas, but because it is unconsolidated and lagging behind practice. Teachers alike are using these tools without a clear sense of; what works, for whom, or under what conditions, but this review paper tries to bridge that gap by questioning the assumptions about scalability, equity, and pedagogical fit that currently drive the field, rather than simply summarising what has been published.

Statement of the Problem

Since generative artificial intelligence became widely accessible to the public in late 2022, its adoption in mainstream education has outpaced the production of consolidated evidence on how these tools serve learners with disabilities specifically. More than one billion

people worldwide live with some form of disability (World Health Organization, 2023), and learners with dyslexia, autism spectrum disorder, and visual impairments continue to encounter barriers built into standardised, text-based, single-modality instruction. Although automated image description, text-to-image/scene generation, and virtual pedagogical agents each hold clear potential to address these barriers, the evidence for each remains scattered, methodologically uneven, and rarely considered together. The available literature further tends toward two unhelpful extremes: techno-optimistic pilot studies with little critical depth, and ethical critiques that remain abstract and offer teachers little practical guidance (Paglialunga & Melogno, 2025). Most existing reviews also predate the shift from text-only to multimodal AI, address AI and disability in broad terms without attending closely to the specific role of visual media, and draw disproportionately on Western, Educated, Industrialised, Rich, and Democratic (WEIRD) samples. Teachers are consequently adopting these tools without a clear, consolidated sense of what works, for whom, and under what conditions. This narrative review addresses that gap.

Objectives of the Study

The specific objectives of the study are to:

- i. identify and categorise the forms of AI-generated visual content currently used to support inclusive learning for students with disabilities.
- ii. evaluate the empirical evidence on the effectiveness, limitations, and equity implications of these applications against the Universal Design for Learning (UDL) framework and the Web Content Accessibility Guidelines (WCAG 2.2).
- iii. examine the extent to which existing studies reflect pedagogically grounded evaluation and diverse, non-WEIRD learner populations.
- iv. To derive practical, evidence-informed recommendations for curriculum developers, classroom teachers, technology designers, policymakers, and researchers.

Research Questions

The research questions are guided by three all-encompassing research questions which are:

- i. What forms of AI-generated visual content are currently employed to support inclusive learning for students with disabilities?
- ii. What does the available evidence indicate about the effectiveness, limitations, and equity implications of these applications?
- iii. What recommendations can be drawn from the literature to guide future research, design, and practice?

Theoretical Framework

Universal Design for Learning (UDL)

Universal Design for Learning-UDL is the main view-through for this paper which is developed by CAST and most recently updated as UDL Guidelines 3.0 in 2024, the framework rests on three core principles namely; multiple means of representation (the "what" of learning), multiple means of action and expression (the "how"), and multiple means of engagement (the "why"). The Representation principle is mostly directly relevant to AI-generated visual content as it calls for information to be presented in various perspectives-formats so that learners who cannot effectively access text-based or static content are not excluded (Saborío-Taylor & Rojas-Ramírez, 2024). AI tools that automatically convert text to images, generate scene descriptions, or animate abstract concepts operationalises this principle at large scale and low cost that manual content production could not earlier accomplish.

In previous systematic review of technology-mediated UDL implementations in secondary education carried out by Bray et al. (2024), it was found that Representation has drawn far more research attention than Action and Expression or Engagement. This imbalance has created opening for: AI visual generation which could extend the Representation work which is somewhat already underway provided it does not come at the expense of the other two principles.

Accessibility Standards and Legal Frameworks

The Web Content Accessibility Guidelines (WCAG), maintained by the World Wide Web Consortium (W3C), provide a technical benchmark for accessible digital content. The WCAG 2.2, which was released in October 2023 reinforces the requirements for text alternatives for non-text content (Guideline 1.1) and adaptable presentation (Guideline 1.3), which are both directly relevant to AI-generated visual content. In the European context, the European Accessibility Act (EAA) of 2019, which came into full force in June 2025 enforces legal obligations on educational publishers and technology related providers to ensure accessibility compliance, including AI-developed contents-materials (Risi, 2025). Strengthening this, these legal frameworks set the ground for which AI-generated visual content must not fall, regardless of pedagogical intent and context.

Methodology

The research design considered, disability categories, and technological platforms meeting the homogeneity assumptions that systematic meta-analysis requires. The narrative approach gives room for thematic synthesis across different methodologies while still appraising each source critically (Green et al., 2006). This review goes beyond combination of methods: it is structured to express an articulated story of how AI-generated visuals are being; deployed, where they work, and where they are inefficient across disability categories. Rather than only summarising, it considers the methodological rigour of included studies, their reliance in particular on represented reports and short-term metrics. It also identifies patterns across individual studies, such as a recurring strain between scalability, contextual depth and further proposes a synthesis: that the effectiveness of these tools depends less on the AI model itself other than on the pedagogical framework and human oversight surrounding it only. Searches on suitable literatures were conducted across five major databases, namely: Google Scholar, PubMed/PubMed Central, ScienceDirect, Springer Nature Link, and the ACM Digital Library and more searches were carried out on

ERIC and ResearchGate for older literature. Searches on relevant publications was restricted to between 2021 and 2026 using these search strings individually and in Boolean combinations: "AI visual content inclusive education," "generative AI disability learning," "alt text automation accessibility," "virtual pedagogical agents social skills ASD," "text-to-image education dyslexia," and "UDL artificial intelligence learning materials."

Inclusion and Exclusion Criteria

Studies considered for inclusion are based on if they: (a) were published in peer-reviewed journals or conference proceedings with demonstrable academic standing; (b) addressed AI-generated or AI-mediated visual content in educational or therapeutic-educational contexts; (c) involved participants with one or more identified disabilities or specific learning difficulties; (d) were available in full text online; and (e) were published in English. Studies were excluded if they: (a) were available only as abstracts; (b) reported exclusively on text-based AI applications without a visual component; (c) did not engage with learner populations with disabilities; or (d) could not be verified against primary source publication details including author names, journal title, volume, issue, and page numbers or article number.

Source Verification Protocol

Because only verifiable, full-text accessible sources were used, each citation in this review was checked against its primary publication venue. Author names, journal titles, publication years, volume and issue numbers, DOIs, and page or article numbers were cross-checked across at least two independent sources (for example, the journal website, PubMed, or Semantic Scholar) before inclusion. Sources that could only be confirmed at the abstract level were excluded regardless of how relevant they appeared. The summary of the eight studies that this review was drawn from is shown in table 2 and 26 were retained after full-text verification out of a total of 34 sources which were initially identified as relevant and 20 most substantively relevant are judiciously used in this paper.

Application 1: Automated Image Description for Accessibility

The Accessibility Problem with Visual Content

Adequate resources materials such as, textbooks, worksheets, digital resources, e-books, that are full of images, diagrams, charts, and illustrations was assessed and used. Resources are often inaccessible without text-based descriptions for learners who are disabled (blind, visually impaired, or process visual information atypically). Manually written alternative text; the conventional fix, has historically been expensive, inconsistently applied, and frequently skipped altogether (Risi, 2025). Kreiss et al. (2022) found that even where alt text exists, it often fails to convey contextual meaning, especially for images embedded in specific educational tasks where the pedagogical point of the image cannot be separated from its interpretation. The scale involved is large. Risi (2025), studying educational publishing workflows, noted that a major e-book publisher can produce hundreds of thousands of images a year, and writing quality alt text for each one by hand simply is not realistic. That is the practical reason publishers have been motivated to automate the process.

AI Approaches to Automated Description

Contemporary multimodal large language models, including GPT-4V (now GPT-4o), Google's Gemini, and open-source alternatives, show substantial capacity for image-to-text captioning. These models are trained on billions of image-text pairs and can generate contextually aware descriptions of educational images, including diagrams, maps, charts, and illustrative scenes (Risi, 2025). The Be My Eyes app, which integrated GPT-4's vision capabilities in 2023, showed in real-world usage how AI description could help visually impaired individuals access visual environments with meaningful, context-sensitive language.

Risi's (2025) comparative experimental study, published in Logos: Journal of the World Book Community, tested multiple AI models (including GPT-4V and Gemini) against human-authored alt text for educational publishing images. The findings of the study show that

AI-generated descriptions clearly present advantages in scalability, processing speed, and terminological uniformity. It also identified an important limitation: AI models frequently produced descriptions that were technically accurate but pedagogically shallow, naming objects in an image without explaining their educational significance in context. Risi concluded that a hybrid human-in-the-loop approach, in which AI drafts descriptions and a human reviewer adds educational context, is currently the most workable option.

This finding fits a broader body of evidence on automated captioning. Kreiss et al. (2022) showed that reference-free evaluation metrics for image descriptions used in accessibility contexts routinely fail to capture how contextually adequate a description actually is, meaning that auto-generated text can score well on standard benchmarks and still be inadequate for specific learner needs. For students with disabilities who depend on these descriptions as their primary mode of access, that inadequacy matters a great deal.

Equity and Implementation Challenges

AI image description has clear potential, but rolling it out equitably raises several concerns. First, the quality of AI-generated descriptions depends heavily on training data, which overrepresents images from high-income, English-language contexts. Educational materials designed for use in sub-Saharan Africa, South Asia, or Latin America, and depicting local cultural and environmental contexts, tend to get less reliable descriptions from AI systems not trained on regionally representative data (Chalkiadakis et al., 2024). That is a real equity gap for inclusive education in the Global South.

Secondly, tools that support the assessment of AI-generated alternate text, such as Ally (deployed in major LMS platforms) and Adobe InDesign's accessibility features, require institutional licenses and technical capacity that many schools in low-income countries mostly do not have. The result is that the scalability benefit of AI description tends to accrue disproportionately to well-resourced institutions in high-income countries.

Application 2: Text-to-Image and Scene Generation for Comprehension Support

Visual Learning Needs of Learners with Dyslexia and ASD

Two groups of learners stand out as good candidates for AI-generated visual content: students with dyslexia and those with autism spectrum disorder (ASD). Both often show relative strength in visual-spatial processing alongside difficulty with phonological decoding, sequential text processing, or abstract linguistic reasoning (Yang, 2022). Turning a written narrative into a visual scene can reduce cognitive load and improve comprehension for these learners. Learners with dyslexia, who make up roughly 10% of the global population (Kuerban et al., 2025), often struggle with dense, uninterrupted text. Research has consistently supported graphic organisers, illustrated narratives, and concept maps as comprehension scaffolds. What the current generation of AI tools raises is the question of whether generative systems, specifically text-to-image tools such as DALL-E 3, Midjourney, and Stable Diffusion, can automate this scaffolding at the individual learner level, producing bespoke visual supports for specific curriculum content.

Kuerban et al. (2025) developed and evaluated ReadSmart, a generative AI and augmented reality solution built specifically for students with dyslexia. Published in the International Journal of Technology in Education and Science, the study documented how AI-generated visual overlays on printed text, depicting scenes from the text in real time, improved both decoding accuracy and reading fluency in participants. The study was limited in sample size and duration, but it offered design-based evidence that AI-driven visual scaffolding is feasible for students with reading-based learning disabilities.

Text-to-Scene Generation and ASD

For learners with ASD, AI-generated visuals can support comprehension beyond decoding text, extending to understanding social situations, narratives, and cause-and-effect relationships, all areas where many autistic learners struggle (Mosher et al., 2022).

Stories that are social in nature as a well-structured instructional approach in ASD education, use simple illustrated narratives to educate students on appropriate responses to social situations. Artificial Intelligent images generated makes create the avenue for personalised social stories which are high on demand, and can be tailored to a specific learner's real-life context, instead of relying on generic pre-designed resource.

The work of Williamson et al. (2025) as published in Computers and Human Behavior identified 12 studies of generative on AI applications for neurodivergent school-age learners, including autistic students. The study elucidate that Generated AI tools were mostly applied for personalising narrative and comprehension supports. Further, it noted that most studies relied on teacher and parent proxy reports of efficiency other than direct learner outcome measures which points to a real methodological gap.

Bianchi et al. (2023) and subsequent studies have documented that text-to-image systems can amplify cultural and demographic stereotypes when generating depictions of human figures. For social story generation, this is a pedagogical problem as much as a representational one: a social story that depicts a learner's social environment inaccurately or stereotypically may actively mislead rather than support. Educators using AI-generated social stories need to keep curatorial control over the imagery produced.

Table 1 provides a comparative summary of the primary AI tools and approaches reviewed in this paper, including their target learner groups, core features, and reported benefits.

Table 1: Comparison of AI-Generated Visual Content Tools for Inclusive Education

AI Tool / Approach	Primary Application	Target Learner Group	Key Feature	Reported Benefit
DALL-E / Midjourney (text-to-image)	Comprehension scene generation	ASD; dyslexia	On-demand visual narratives	Reduced cognitive load; improved engagement

AI Tool / Approach	Primary Application	Target Learner Group	Key Feature	Reported Benefit
GPT-4V / Gemini Vision (auto alt text)	Image description for accessibility	Visual impairment; blindness	Context-aware description	Enhanced access to visual content
Virtual pedagogical agents (VPA)	Social skills instruction	ASD; social communication disorders	Animated avatar – feedback	Safe rehearsal; reduced anxiety
Augmented Reality + GenAI (e.g., ReadSmart)	Reading support	Dyslexia	AR visual overlays – adaptive scaffolding	Improved decoding; reading fluency
AI-adaptive learning systems (ITS)	Personalised content delivery	Multiple LD categories	Real-time feedback loop	Large effect sizes on academic outcomes

Note. ASD = Autism Spectrum Disorder; AR = Augmented Reality; LD = Learning Disability; ITS = Intelligent Tutoring System; SEN = Special Educational Needs; GenAI = Generative Artificial Intelligence.

Application 3: Virtual Pedagogical Agents for Social Skills Instruction

Rationale for Agent-Based Instruction in ASD

Social communication deficits are a core diagnostic feature of ASD, with consequences for peer interaction, cooperative learning, and occupational participation across the lifespan. Conventional social skills training (SST) delivered by human practitioners is resource-intensive, hard to scale, and inherently variable in how it is delivered. AI-driven virtual pedagogical agents (VPAs), which are animated characters that interact with learners through special routes; speech, gesture, and facial expression is offer an alternative for consistent instruction that can be done endlessly in a safe, low-stakes virtual platform.

A usability study by Tanaka and Nakamura's (2022) published in JMIR Human Factors looked at how virtual agent-visual design that affects the acceptability of automated social skills training. 1,218 participants were publicly used across multiple conditions, it was

found that real differences in acceptability, likeability, realism, and familiarity depending on agent design are core. Overall, a female anime-type virtual agent came out as most likeable and autistic traits among participants appeared shifted agent preference, suggesting that visual agent design needs to be personalised for individual learners rather than applied homogeneously.

Evidence on Effectiveness

The review on AI, virtual reality, and large language model applications in special education by Antoniou et al. (2025), published in Education and Information Technologies, identified VPA-based social skills training as one of the most studied and promising Artificial Intelligent applications in the field. The work disclosed outcomes such as improved eye contact, conversational turn-taking, and emotion recognition in autistic learners following VPA-based training. Also, the paper noted, that on generalisation, whether skills learned with an avatar transfer to real human social interactions, remains limited and methodologically contested.

The social world is complex, contextual, and unpredictable in ways current VPAs cannot fully replicate-this is a concern about generalization of AI. The work of (Tanaka & Nakamura, 2022) shows that a learner who becomes fluent interacting with a consistent, controlled virtual agent still faces a very different communication environment with a neurotypical peer.

AI Visual Realism and the Uncanny Valley

The phenomenon where artificial figures that roughly appear in human appearance without achieving full realism produce discomfort or rejection rather than engagement. Meaning the visual design of VPAs has to navigate the "uncanny valley," This is obviously relevant for learners with ASD, who may be more sensitive to subtle visual incongruities in faces and movement. Tanaka and Nakamura's (2022) research findings show that anime-style agents were preferred over hyper-realistic ones offers some empirical guidance, suggesting that stylised, non-realistic representations may

work better for learners on the autism spectrum. Further connecting to current AI tools, including DALL-E 3 and Midjourney, can generate both styles on demand, giving educators real choices in agent design.

Synthesis: Cross-Cutting Themes and Patterns

Summary of Reviewed Literature

The structured synthesis of the eight primary studies that make up the core evidential base of this review is presented in table 2.

Table 2: Summary of Key Studies in the Narrative Literature Review

Author(s) & Year	Methodology	Focus Area	Disability Type	Key Finding
Chalkiadakis et al. (2024)	Systematic review	AI & VR in inclusion	Multiple	AI adaptive systems personalise learning; cost and teacher readiness are barriers
Paglialunga & Melogno (2025)	Systematic review	AI interventions for LD	Dyslexia; SLD	All 11 studies positive; large effect sizes (d = 1.63 for arithmetic, d = -1.66 for reading)
Saborio-Taylor & Rojas-Ramírez (2024)	Conceptual article	UDL x AI integration	General SEN	AI operationalises UDE's multiple means of representation in scalable ways
Tanaka & Nakamura (2022)	Usability study (n = 1218)	Virtual agent design	Social comm. / ASD traits	Female anime-type avatar rated most acceptable; autistic traits moderate agent preference
Kuerban et al. (2025)	Design-based research	GenAI + AR for reading	Dyslexia	ReadSmart system improved decoding and fluency; highlighted need for teacher training
Risi (2025)	Comparative experiment	AI-generated alt text	Visual impairment	AI alt text scalable but human-in-the-loop required for semantic accuracy

Author(s) & Year	Methodology	Focus Area	Disability Type	Key Finding
Pires et al. (2025) [via PMC]	Integrative review	AI assistive tech in HE	Multiple disabilities	Personalisation, engagement, and accessibility gains; ethics and equity gaps remain
Bray et al. (2024)	Systematic review	UDL + technology in K-12	Diverse learners	UDL Representation principle dominates tech-focused research; expression and engagement under-studied

Note. SLD = Specific Learning Disorder; HE = Higher Education; LD = Learning Disability; SEN = Special Educational Needs; UDL = Universal Design for Learning; VPA = Virtual Pedagogical Agent; ASD = Autism Spectrum Disorder.

The Quality of AI-Generated Descriptions

Across all three application areas, there is a persistent tension between the scalability of AI-generated visual content and its quality for specific learner needs. The works of As Risi (2025), Kreiss et al. (2022), and Williamson et al. (2025) shows that AI systems consistently produce content that is technically sufficient but contextually thin lacking the pedagogical intentionality that characterises effective inclusive materials designed by skilled educators. This is not just a technical limitation but a reflection design choice which is a current AI systems optimised for general accuracy rather than for disability-specific communicative effectiveness.

Accessibility of Development Tools for Educators

The tools capable of producing high-quality AI visual content for inclusive education, GPT-4o, Gemini, DALL-E 3, Midjourney, are largely proprietary, subscription-based, and built for users with substantial digital literacy. On another cross-checking, theme that denies educator access and readiness was carried out Paglialunga and Melogno (2025) which noted that although studies of AI

interventions for students with learning disabilities reported unanimously positive outcome but most were conducted by research teams with dedicated technical support rather than by classroom teachers working on their own. Further, Chalkiadakis et al. (2024) identified limited teacher readiness as one of the two biggest barriers, alongside cost, to widespread adoption of AI-inclusive technologies in schools.

Pedagogical Grounding in Technological Design
Saborio-Taylor and Rojas-Ramírez (2024) in their work argued that AI contributes mostly to inclusive education when it is deliberately aligned with established frameworks such as UDL, rather than deployed as a technological fix on its own. Otherwise, AI-generated visuals risk becoming a form of edutainment: visually engaging but instructionally shallow. UDL Guidelines 3.0 (CAST, 2024) specifically kick against technology adoption that addresses one UDL principle at the expense of the others. For Artificial Intelligent visual content that means generating representations without also considering how learners will act, expressing themselves, and engage-only addresses a fraction of the inclusive design challenge.

Ethics, Bias, and Equity
Algorithmic bias in training data risks producing visual content that misrepresents the cultural, ethnic, and social contexts of learners in the Global South and data privacy concerns arise when AI learning platforms collect sensitive information about students with disabilities. Also, there is a structural equity concern if high-quality AI visual tools are accessible only to well-resourced schools and systems, the technology may widen rather than close existing educational inequalities. Dumitru et al. (2026), in their integrative review of AI assistive technologies in higher education published in SAGE Open Medicine identified these equity gaps as among the most pressing unresolved concerns in the field.

Recommendations for Research, Policy, and Practice
As shown in table 3 and drawing on the synthesis above, this table

offers evidence-informed recommendations (structured overview) for the five stakeholder groups most directly involved in AI-generated visual content for inclusive education.

Table 3: Evidence-Informed Recommendations by Stakeholder Group

Stakeholder Group	Recommendation	Rationale
Curriculum Developers	Embed AI visual generation within UDL frameworks; do not treat it as an add-on	Looks on pedagogical coherence and principled accessibility rather than novelty-driven adoption
Classroom Teachers	Receive structured CPD on accessible prompt engineering and alt text review	Teachers are gatekeepers; low AI literacy limits effective deployment for students with disabilities
Technology Designers	Conduct participatory design with learners with disabilities in AI tool development	Current tools are often designed without disability-specific user testing, reducing real-world accessibility
Policy Makers	Mandate accessibility standards (WCAG 2.2) for all AI-generated educational content	Legal and ethical compliance frameworks lag behind technological capability
Researchers	Prioritise RCTs, longitudinal studies, and culturally diverse samples	Existing evidence base is overwhelmingly short-term, small-scale, and WEIRD (Western, Educated, Industrialised, Rich, Democratic)

Note. WCAG = Web Content Accessibility Guidelines; CPD = Continuing Professional Development; RCT = Randomised Controlled Trial; WEIRD = Western, Educated, Industrialised, Rich, Democratic.

Beyond the recommendations above, three areas need more research. First, longitudinal studies are needed to establish whether AI visual scaffolds produce lasting learning gains rather than short-term performance bumps: Paglialunga and Melogno (2025) noted that no existing study on AI interventions for learning disabilities has followed participants for more than one academic year. Second, participatory designs that treat learners with disabilities and their

families as co-designers of AI visual tools, rather than subjects to be evaluated, would likely produce more contextually valid and ethically grounded results. Third, more research needs to look at AI visual content generation outside Western, English-language educational contexts, to address the representational and data-equity concerns raised throughout this review.

Conclusion

AI-generated visual content is genuinely promising for inclusive education, but the field has not yet caught up theoretically with what the technology can already do. Across the three applications covered here, automated image description, text-to-image comprehension support, and virtual pedagogical agents, AI tools offer real advantages over conventional approaches in scalability, personalisation, and cost. Advantages come with conditions, though largely depend on pedagogical grounding, equitable access, cultural representativeness, and keeping humans in the loop when AI-generated products are designed and evaluated.

UDL remains the most useful theoretical framework for this work because it reframes the question and instead of asking "can AI produce an image?" it asks the more pedagogically useful question: "does this AI-generated visual expand or constrain the ways all learners can access, process, and engage with curriculum content?" Answering that well requires collaboration across education, computer science, disability studies, and the arts, and it requires genuinely listening to learners with disabilities when designing the systems meant to serve them.

Recommendation

These are the recommendations from the study:

1. Future studies should track learners over multiple academic years to confirm whether AI visual scaffolds produce lasting gains, not just short-term improvements
2. Future AI visual tools should involve learners with

disabilities and their families as co-designers, not evaluation subjects, for more valid, ethically grounded outcomes.

3. Research should extend beyond Western, English-language contexts (local-Indigenous languages such as Yoruba, Hausa, Efik, Igbo etc) to address the equity concerns raised throughout the review.

References

- Antoniou, P., Kokkalia, G., & Drigas, A. (2025). A systematic review of AI, VR, and LLM applications in special education: Opportunities, challenges, and future directions. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-025-13550-4>
- Bray, A., Devitt, A., & Banks, J. (2024). What next for Universal Design for Learning? A systematic literature review of technology in UDL implementations at second level. *British Journal of Educational Technology*, 55(3), 873-895. <https://doi.org/10.1111/bjet.13328>
- Center for Applied Special Technology (CAST). (2024). UDL Guidelines 3.0. <https://udlguidelines.cast.org/>
- Chalkiadakis, A., Seremetaki, A., Kanellou, A., Kallishi, M., Morfopoulou, A., Moraitaki, M., & Mastrokourou, S. (2024). Impact of artificial intelligence and virtual reality on educational inclusion: A systematic review of technologies supporting students with disabilities. *Education Sciences*, 14(11), 1223. <https://doi.org/10.3390/educsci14111223>
- Dumitru, C., Abdulsahib, G. M., Khalaf, O. I., & Bennour, A. (2026). Integrating artificial intelligence in supporting students with disabilities in higher education: An integrative review. *SAGE Open Medicine*, 14. <https://doi.org/10.1177/10554181251355428>
- Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *Journal of Chiropractic Medicine*, 5(3), 101-117. [https://doi.org/10.1016/S0899-3467\(07\)60142-6](https://doi.org/10.1016/S0899-3467(07)60142-6)

Kreiss, E., Bennett, C., Hooshmand, S., Zelikman, E., Morris, M. R., & Potts, C. (2022). Context matters for image descriptions for accessibility: Challenges for referenceless evaluation metrics. arXiv preprint arXiv:2205.10646. <https://doi.org/10.48550/arXiv.2205.10646>

Kuerban, Y., Oyelere, S. S., & Sanusi, I. T. (2025). ReadSmart: Generative AI and augmented reality solution for supporting students with dyslexia learning disabilities. *International Journal of Technology in Education and Science*, 9(1), 159-176. <https://files.eric.ed.gov/fulltext/EJ1462786.pdf>

Mosher, M. A., Carreon, A. C., Craig, S. L., & Ruhter, L. C. (2022). Immersive technology to teach social skills to students with autism spectrum disorder: A literature review. *Review Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s40489-021-00290-5>

Paglalunga, A., & Melogno, S. (2025). The effectiveness of artificial intelligence-based interventions for students with learning disabilities: A systematic review. *Brain Sciences*, 15(8), 806. <https://doi.org/10.3390/brainsci15080806>

Risi, M. (2025). Accessible images: AI-generated alt text in educational publishing. *Logos: Journal of the World Book Community*, 36(1), 50-68. <https://doi.org/10.1163/1878-4712-11340001>

Saborío-Taylor, S., & Rojas-Ramírez, F. (2024). Universal design for learning and artificial intelligence in the digital era: Fostering inclusion and autonomous learning. *International Journal of Professional Development, Learners and Learning*, 6(2), ep2408. <https://doi.org/10.30935/ijpdll/14694>

Tanaka, H., & Nakamura, S. (2022). The acceptability of virtual characters as social skills trainers: Usability study. *JMIR Human Factors*, 9(1), e35358. <https://doi.org/10.2196/35358>

Williamson, D., de Lemos, S., & Hathorn, L. (2025). A scoping literature review of generative artificial intelligence for supporting neurodivergent school students. *Computers and Human Behavior: Artificial Intelligence*.

<https://doi.org/10.1016/j.chbah.2025.100773>

World Health Organization. (2023). Disability and health. <https://www.who.int/news-room/fact-sheets/detail/disability-and-health>

Yang, H. (2022). Effect of story structure instruction based on visual analysis on reading comprehension intervention for dyslexic students. *Computational and Mathematical Methods in Medicine*, 2022, Article 9479709. <https://doi.org/10.1155/2022/9479709>