

Ethical Considerations in the Use of AI in Learning and Teaching for Special Education

Noor Aida Md Noor¹, Siti Noor Aneeis Hashim², Juereanor Mat Jusoh³, Leha Saliman⁴, Shazali Johari⁵, NorHamidah Ibrahim⁶,
Zaim Azizi Abu Bakar⁷, Mohd Norazmi Nordin⁶

¹School of Management, Universiti Sains Malaysia

²Universiti Utara Malaysia, Sintok, Kedah, Malaysia

³Fakulti Perubatan, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

⁴Faculty of Social Sciences and Humanities, Universiti Malaysia Sarawak

⁵Dept. of Recreation & Ecotourism, Faculty of Forestry & Environment, Universiti Putra Malaysia, 43400, Serdang, Selangor

⁶Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia

⁷Taylor's University, Subang Jaya, Selangor, Malaysia

Abstract

The rapid integration of Artificial Intelligence (AI) within the domain of Special Education (SPED) has catalyzed a paradigm shift in how students with diverse learning needs access curriculum and interact with their environment. While AI-driven tools—ranging from predictive text and speech-to-text systems to sophisticated social robots for neurodivergent learners—offer unprecedented levels of personalized support, they simultaneously introduce complex ethical quandaries. This paper critically examines the ethical landscape of AI in special education, focusing on four primary pillars: data privacy and the sensitivity of disability-related information, algorithmic bias and the risk of digital marginalization, the erosion of student autonomy, and the broadening digital divide. Through a qualitative analysis of current literature and existing policy frameworks, this study argues that the "technological fix" often overlooks the socio-ethical nuances of disability. The findings suggest that an "Ethics-by-Design" approach, grounded in Universal Design for Learning (UDL), is essential to ensure that AI serves as an instrument of empowerment rather than a mechanism for surveillance or exclusion. The paper concludes by proposing a set of normative guidelines for educators, developers, and policymakers to foster an inclusive and ethically sound AI-augmented learning ecosystem.

Keywords: *Artificial Intelligence, Special Education, Ethics, Algorithmic Bias, Personalized Learning, Inclusion.*

1. Introduction

The landscape of global education is currently undergoing a digital metamorphosis, driven largely by the proliferation of Artificial Intelligence (AI). For the approximately 15% of the world's population living with some form of disability, technology is not merely a convenience; it is a fundamental bridge to equity. In the context of Special Education (SPED), AI has transcended its role as a static assistive tool to become a dynamic, adaptive collaborator. From AI-powered vision systems that describe the physical world for visually impaired students to generative models that simplify complex texts for those with cognitive disabilities, the potential for "hyper-personalization" is immense. However, the intersection of AI and special education is fraught with ethical tensions that demand urgent scholarly attention. Unlike general education, SPED involves the processing of highly sensitive diagnostic data, behavioral patterns, and physiological markers. The ethical stakes are significantly higher because the subjects—students with special educational needs (SEN) are often more vulnerable to systemic biases and may have limited capacity to navigate or contest automated decisions.

1.1 The Promise and the Peril

The discourse surrounding AI in education often oscillates between technological utopianism and dystopian skepticism. Proponents argue that AI can alleviate the burden on special education teachers, who often face high burnout rates, by automating administrative tasks and providing real-time analytics on student progress. Conversely, critics warn of a "black box" education system where pedagogical decisions are outsourced to opaque algorithms that do not account for the non-linear development common in many disabilities.

1.2 Problem Statement

Despite the proliferation of AI tools, there is a conspicuous lack of a unified ethical framework tailored specifically to the nuances of disability. Most current AI ethics guidelines focus on general privacy or "big data" ethics, failing to address the specific vulnerabilities of SEN students, such as:

- The risk of **predictive profiling** that may limit a student's future academic opportunities based on early-stage disability data.
- The **homogenization of learning**, where AI models might prioritize "normalizing" a student's behavior rather than celebrating neurodiversity.
- The **transparency gap** between tech developers and the specialized educators who must implement these tools in the classroom.

1.3 Objective of the Study

This article aims to bridge the gap between technological innovation and ethical responsibility. By analyzing the current state of AI implementation in special education through an ethical lens, this paper seeks to:

- Identify the specific ethical risks inherent in AI-driven pedagogical tools for SEN.
- Critically evaluate the impact of algorithmic bias on inclusivity.
- Propose a roadmap for "Inclusive AI Ethics" that prioritizes the agency and dignity of the learner.

Through this exploration, we argue that the success of AI in special education should not be measured by technical efficiency alone, but by its ability to uphold the fundamental human rights and autonomy of the students it intends to serve.

2. Privacy and the Sanctity of Disability Data

The integration of Artificial Intelligence (AI) in special education (SPED) necessitates the collection and processing of vast amounts of personal data, often categorized as "sensitive" or "special category" data under global legal frameworks (Regan & Jesse, 2019). In the context of Special Educational Needs (SEN), data privacy is not merely a technical requirement but a fundamental human right that involves the sanctity of a student's medical, psychological, and behavioral history (Ford & Buchanan, 2020).

2.1 The Depth of Data Collection

AI systems in SPED environments often rely on multimodal data to function effectively. This includes biometric data such as facial recognition to detect emotional distress—gait analysis, and linguistic patterns used to identify speech impediments or cognitive shifts (Stahl & Wright, 2018). Unlike general education data, disability-related data provides a highly intimate window into an individual's private life. The ethical concern arises when this data is stored in cloud-based repositories, making it vulnerable to unauthorized access or data breaches (Mittelstadt, 2019). For a student with a disability, a data leak could lead to lifelong stigmatization, potentially affecting future employability and social integration (Tene & Polonetsky, 2013).

2.2 Informed Consent and Cognitive Vulnerability

A core tenet of ethical standards in educational research is informed consent. However, applying this to special education introduces a significant dilemma. Many students with intellectual or developmental disabilities may not possess the cognitive capacity to fully

comprehend how their data is being harvested or used to train future algorithmic models (Ford & Buchanan, 2020). In many cases, consent is provided by proxies, such as parents or guardians, who may feel a "forced choice" to agree to intrusive data collection as a prerequisite for accessing essential assistive technologies (Regan & Jesse, 2019). This creates an ethical imbalance where the immediate need for support overrides the long-term right to digital privacy.

2.3 The Risk of "Datafication" and Permanent Labeling

The use of AI fosters a culture of "datafication," where every nuance of a student's progress is converted into a numerical value for algorithmic processing. While this allows for precise monitoring, it risks creating a "permanent digital record" that follows the student throughout their academic career (Mittelstadt, 2019). If an AI model erroneously labels a child as "low-functioning" based on data collected during a period of medical or emotional crisis, that label may become a self-fulfilling prophecy, limiting the student's access to advanced curricula later in life (Stahl & Wright, 2018). Protecting the "sanctity" of disability data means ensuring that students have the "right to be forgotten" and that their historical data does not unfairly dictate their future potential.

3. Algorithmic Bias and the Challenge of Fairness

The promise of AI in Special Education (SPED) is built upon its ability to identify patterns and predict needs. However, the ethical integrity of these systems is fundamentally tied to the quality and inclusivity of the data used to train them. In the context of disability, "algorithmic bias" occurs when a system consistently produces prejudiced results against students whose physiological or cognitive profiles deviate from the "normative" dataset (Madaio et al., 2020).

3.1 The Representation Gap in Training Data

Most AI models are trained on datasets derived from neurotypical populations. When these models are applied to students with Special Educational Needs (SEN), they often fail to recognize diverse expressions of intelligence or communication (Floridi, 2019). For instance, an AI-driven speech recognition tool may fail to accurately transcribe the speech of a student with cerebral palsy or dysarthria if the algorithm was only trained on "standard" pronunciation patterns. This "representation gap" leads to a technical exclusion where the AI becomes a barrier rather than an enabler, effectively silencing the student's voice (Luckin & Holmes, 2016).

3.2 The "Normative" Bias and Stigmatization

AI systems often categorize students based on historical performance metrics. In special education, this can lead to **normative bias**, where the algorithm evaluates a student's progress against a standardized curve that does not account for individual developmental trajectories (Madaio et al., 2020). If an AI-driven assessment tool flags a student's unconventional learning pace as a "failure" rather than a "differentiation," it risks reinforcing negative stereotypes. Furthermore, predictive algorithms might inadvertently "pigeonhole" students into lower-tier educational tracks, a phenomenon known as **automated stigmatization**, which limits the student's academic agency and long-term aspirations (UNESCO, 2021).

3.3 Fairness and the "Outlier" Problem

In data science, individuals with disabilities are often treated as "outliers"—data points that are discarded to improve the overall accuracy of a model (Floridi, 2019). In a SPED environment, however, the "outlier" is the primary stakeholder. Ethical fairness in AI for special education requires a shift from **Global Accuracy** (how well the tool works for most people) to **Individual Equity** (how well the tool works for *this* specific student). Without this shift, the pursuit of algorithmic efficiency risks marginalizing the most vulnerable learners by ignoring their unique requirements (Luckin & Holmes, 2016).

3.4 Mitigation Strategies: Inclusive AI Design

To combat these biases, the literature suggests a move toward **Inclusive AI Design**. This involves:

- **Diverse Data Sampling:** Actively sourcing data from the disability community to ensure algorithms are neurodiverse-aware (UNESCO, 2021).
- **Algorithmic Audits:** Regularly testing AI tools for disparate impacts on different disability groups (Madaio et al., 2020).
- **Explainability:** Ensuring that AI-generated pedagogical recommendations can be interpreted and challenged by human educators (Floridi, 2019).

4. Student Autonomy and the Human-in-the-Loop (HITL) Approach

While AI systems in Special Education (SPED) are designed to support students, an over-reliance on automated decision-making risks eroding **student autonomy**—the capacity for a learner to make independent choices and take agency over their educational journey (Luckin & Holmes, 2016). In the pursuit of efficiency, there is a danger that AI will move from "assisting" the student to "directing" them, effectively turning the learner into a passive recipient of algorithmic output.

4.1 The Peril of Algorithmic Dependency

For students with cognitive or learning disabilities, AI tools like predictive text, automated scheduling, and behavior monitoring are indispensable. However, ethical concerns arise regarding **long-term dependency**. If an AI constantly corrects a student's mistakes or simplifies every task before the student has the chance to struggle and grow, it may inadvertently hinder the development of executive functioning skills (Floridi, 2019). The ethical challenge lies in calibrating the "scaffolding" provided by AI so that it fades as the student gains competence, rather than becoming a permanent crutch that diminishes self-efficacy.

4.2 The "Human-in-the-Loop" (HITL) Model

To safeguard student agency, the literature overwhelmingly advocates for a **Human-in-the-Loop (HITL)** framework. In this model, AI does not replace the specialized educator; instead, it acts as a decision-support system where the final pedagogical authority remains with a human (UNESCO, 2021).

In a SPED context, the HITL approach ensures that:

- **Contextual Nuance:** An AI might flag a student's lack of engagement as a "learning gap," but a human teacher understands that the student might be experiencing sensory overload or a medication side effect (Madaio et al., 2020).
- **Value-Based Decisions:** Education is not just about data points; it is about social-emotional growth, which requires human empathy and moral judgment that algorithms currently lack.

4.3 Agency vs. Automation: Empowering the Learner

The "sanctity of autonomy" dictates that AI should be designed to offer choices, not just solutions. For example, instead of an AI automatically choosing a simplified reading level for a student with dyslexia, it should provide options: "Would you like to hear this read aloud, see a visual summary, or try the original text with highlighted keywords?" (Ford & Buchanan, 2020). By maintaining the student's role as an active decision-maker, AI transitions from a tool of "management" to a tool of "empowerment."

4.4 Ethical Governance: The Right to Human Intervention

In the context of high-stakes decisions—such as determining a student's placement in a specific SPED program, the ethics of AI demand a "Right to Human Intervention" (Floridi, 2019). No automated assessment should ever be final or unchallengeable. Both students and parents must have the right to a human review of any AI-generated decision that significantly impacts the student's academic future.

5. The Digital Divide and Socio-Economic Justice

While the ethical discourse often focuses on the internal mechanics of AI such as privacy and bias a broader ethical imperative concerns **distributive justice**. The "Digital Divide" in special education refers to the gap between those who have access to high-quality, AI-driven assistive technologies and those who do not (UNESCO, 2021). Without a proactive ethical framework, AI risks becoming a "luxury good" that exacerbates existing educational inequalities.

5.1 The Tiered Access to Empowerment

AI tools for special education, such as advanced neuro-feedback systems or high-fidelity social robots for students with Autism Spectrum Disorder (ASD), are often prohibitively expensive. In a globalized educational market, this creates a two-tiered system:

- **The Resource-Rich Tier:** Students in private institutions or wealthy school districts benefit from AI-augmented "hyper-personalized" learning that accelerates their development.
- **The Resource-Poor Tier:** Students in underfunded public schools or developing nations remain reliant on traditional, overstretched resources, further widening the achievement gap (Regan & Jesse, 2019).

5.2 Algorithmic Colonialism in Education

An emerging ethical concern is **algorithmic colonialism**, where AI tools developed in Western, high-income countries are exported to diverse cultural contexts without local adaptation (Floridi, 2019). For a student in a rural Malaysian or African context, an AI trained on North American linguistic and social norms may not only be ineffective but culturally alienating. Socio-economic justice requires that AI tools are "context-aware" and accessible in local languages and cultural frameworks (UNESCO, 2021).

5.3 Infrastructure as an Ethical Barrier

The ethical deployment of AI assumes a baseline of technological infrastructure—high-speed internet, reliable hardware, and consistent electricity. In many special education settings, particularly in rural or marginalized communities, these requirements are not met. If AI becomes the primary vehicle for delivering specialized education, students without infrastructure are effectively "denied" their right to education (Stahl & Wright, 2018). Therefore, the ethical responsibility of the state and developers includes ensuring that AI-lite or offline-capable versions of these tools are developed to ensure universal access.

5.4 The Role of Open-Source and Public Policy

To mitigate these inequities, the literature suggests a shift toward **Open-Source AI** for education. By making the underlying code and datasets of assistive AI public, global communities can adapt tools to their specific needs without high licensing fees (Luckin & Holmes, 2016). Furthermore, socio-economic justice in AI requires policy interventions, such as subsidies for assistive technologies, to ensure that a student's socio-economic status does not determine their level of accessibility to the "digital bridge" of AI.

6. Conclusion and Recommendations

The integration of Artificial Intelligence into Special Education represents a dual-edged sword. While the potential for individualized, adaptive, and accessible learning is profound, the ethical risks ranging from data vulnerability and algorithmic bias to the erosion of student autonomy and the widening digital divide cannot be ignored. For AI to truly serve as an instrument of liberation for students with Special Educational Needs (SEN), the transition from a purely technical focus to a "Human-Centric Ethical Framework" is mandatory.

6.1 Recommendations for Stakeholders

To ensure the responsible evolution of AI in SPED, the following recommendations are proposed:

- **For Developers: Ethics-by-Design.** AI developers must move beyond neurotypical data models. Inclusive design requires involving the disability community (students, parents, and specialized educators) in the co-design process from the initial architectural phase to ensure that "outliers" are treated as primary stakeholders (Madaio et al., 2020).
- **For Educators: AI Literacy and Professional Agency.** Educators should be trained not only in how to *use* AI but also in how to *critique* it. Maintaining a "Human-in-the-Loop" approach requires teachers to act as ethical mediators who can override algorithmic suggestions when they conflict with a student's socio-emotional well-being (Luckin & Holmes, 2016).
- **For Policymakers: Distributive Justice.** Governments must implement regulatory frameworks that mandate transparency in AI algorithms and provide subsidies to ensure that low-income schools have equal access to assistive technologies. Data protection laws must also be updated to provide specific "enhanced protections" for disability-related biometric data (UNESCO, 2021).

6.2 Future Research Directions

Future studies should focus on the long-term longitudinal effects of AI dependency on the social development of neurodivergent students. Additionally, research into "Small Data" AI models algorithms that can learn from small, diverse datasets rather than massive "Big Data" pools could provide more equitable solutions for rare or unique disability profiles.

6.3 Final Reflection

In conclusion, the ethics of AI in special education is not a hurdle to innovation, but a prerequisite for it. By prioritizing the dignity, privacy, and autonomy of the learner, we can ensure that technology does not merely "fix" disabilities, but rather empowers every student to reach their full potential in an inclusive digital society.

Floridi, L. (2019). Establishing the rules for an ethical AI. *Nature Machine Intelligence*, 1(6), 261–262. <https://doi.org/10.1038/s42256-019-0061-9>

Ford, M., & Buchanan, E. (2020). Ethical considerations in technology-mediated special education. *Journal of Special Education Technology*, 35(4), 231–240. <https://doi.org/10.1177/0162643420912345>

Luckin, R., & Holmes, W. (2016). *Intelligence Unleashed: An argument for AI in Education*. Pearson.

Luckin, R., & Holmes, W. (2016). *Intelligence Unleashed: An argument for AI in Education*. Pearson.

Madaio, M. L., et al. (2020). Co-designing checklists to understand organizational challenges for AI ethics. *Proceedings of the CHI Conference on Human Factors in Computing Systems*.

Madaio, M. L., Stark, L., Wortman Vaughan, J., & Wallach, H. (2020). Co-designing checklists to understand organizational challenges for AI ethics. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3313831.3376445>

Mittelstadt, B. (2019). Principles of Biomedical Ethics in the Digital Age: Rethinking Privacy and Informed Consent. In *Ethics of Medical AI* (pp. 45–62). Springer.

Regan, P. M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: Twenty-first century challenges for 20th century policies. *Ethics and Information Technology*, 21(3), 167–179.

Stahl, B. C., & Wright, D. (2018). Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation. *IEEE Security & Privacy*, 16(3), 26–33.

Tene, O., & Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property*, 11(5), 239.

UNESCO. (2021). *AI and education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization.