

# ETHICAL AND RESPONSIBLE USE OF ARTIFICIAL INTELLIGENCE IN SOCIAL WORK EDUCATION: A DESKTOP LITERATURE REVIEW PERSPECTIVE

GIFT SALADI<sup>1</sup> AND KUDZAI MWAPAU<sup>2</sup>

## Abstract

*The rapid advancement and integration of Artificial Intelligence (AI) into various professional domains present a transformative, yet complex, challenge for social work. While AI holds immense potential to enhance efficiency, streamline administrative tasks, and provide data-driven insights for social services, its adoption also introduces profound ethical and practical dilemmas. This desktop literature review explores the landscape of AI in social work education, synthesizing existing research, professional guidelines, and scholarly discourse, to provide a comprehensive overview. The article argues that the ethical and responsible use of AI is not merely a technical concern but a core pedagogical imperative that must be woven into the fabric of social work curricula. It examines key themes, including the potential benefits of AI in education and practice, the critical ethical challenges, such as algorithmic bias, privacy, and the de-emphasis of the human-centred relationship, and the necessary competencies social work students must acquire. The review concludes with recommendations for social work educators, highlighting the need for a revised curriculum that fosters AI literacy, critical thinking, and a*

---

<sup>1</sup> Department of Social Work and Applied Psychology, Zimbabwe Ezekiel Guti University, Bindura, Zimbabwe ORCID ID,<https://orcid.org/0009-0005-0335-4992>, [gsaladi@staff.zegu.ac.zw](mailto:gsaladi@staff.zegu.ac.zw)

<sup>2</sup> Department of Social Work and Applied Psychology, Zimbabwe Ezekiel Guti University, Bindura, Zimbabwe ORCID ID,<https://orcid.org/0000-0001-6621-8158> [kmwapaura@staff.zegu.ac.zw](mailto:kmwapaura@staff.zegu.ac.zw)

*robust ethical framework to prepare future professionals for a technologically mediated world.*

**Key words:** *pedagogy, digital competence, bias mitigation*

## INTRODUCTION

The 21st century is marked by an unprecedented digital revolution, with Artificial Intelligence (AI) emerging as a powerful and pervasive force. AI, encompassing technologies like machine learning, natural language processing, and predictive analytics, is no longer a futuristic concept but a present reality reshaping industries from healthcare and finance to education and social services (Wangdi, 2024; Zainuddin *et al.*, 2024). For the social work profession, which is fundamentally rooted in human relationships, social justice, and the well-being of individuals and communities, the rise of AI presents a unique duality of opportunity and risk (Rodríguez-Martínez *et al.*, 2024; Garkisch and Goldkind, 2025).

The potential for AI to aid social workers is significant. It can automate mundane tasks, analyse vast datasets to identify social trends, predict risks in areas like child welfare, and even provide initial support through chatbots in mental health. However, the integration of these technologies into a human-centred profession is fraught with ethical complexities (Garkisch and Goldkind, 2025). Questions of algorithmic bias, data privacy, informed consent, and the potential erosion of the therapeutic relationship loom large (Weber, 2020; Nurhasanah and Nugraha, 2024). As AI systems are increasingly deployed in real-world social service settings, social work education must respond proactively to prepare its students for this new landscape (Azman and Tümkeya, 2025; Musengimana, 2025).

This desktop literature review aims to provide a comprehensive analysis of the ethical and responsible use of AI in social work

education. Through synthesizing a range of scholarly articles, professional reports, and organisational statements, this article seeks to:

- identify the current and potential applications of AI in social work practice and education.
- examine the key ethical challenges and risks associated with AI adoption in the social services.
- discuss the existing and proposed frameworks for ethical AI in social work.
- highlight the pedagogical implications and necessary curricular changes for social work programs.
- propose recommendations for preparing future social workers to navigate the AI-integrated world with integrity and competence.

The article is structured as follows: an introduction; an overview of AI applications in social work and education; core ethical challenges; existing ethical frameworks and principles; pedagogical imperative and curriculum development; and a concluding summary and forward-looking recommendations.

## **THE LANDSCAPE OF AI IN SOCIAL WORK AND EDUCATION**

The use of technology in social work is not a new phenomenon, it has evolved from early computerised record-keeping to the widespread use of social media for outreach and communication (Rodríguez-Martínez *et al.*, 2024). AI represents the next frontier, with applications that are both administrative and clinical (Garkisch and Goldkind, 2025).

## **AI IN SOCIAL WORK PRACTICE**

In practice, AI applications are emerging across all levels of social work:

- **Micro-level (Individual Practice):** AI-powered tools are being developed to support direct client work. Chatbots and virtual assistants like Woebot and Wysa are designed to provide initial, low-acuity mental health support, offering cognitive-behavioural techniques and motivational interviewing. These tools can serve as a supplementary resource, assisting outside of traditional office hours. Predictive analytics are also used in risk assessments, for example, to help identify individuals at risk of suicide or homelessness, allowing for early intervention (*ibid.*).
- **Mezzo-level (Group and Community):** AI can be used to analyse large datasets to identify patterns and systemic issues within communities. For instance, it can help pinpoint areas with high rates of substance abuse or child welfare concerns, enabling social workers to allocate resources more efficiently and design targeted community-based interventions (*ibid.*).
- **Macro-level (Policy and Administration):** On an administrative level, AI can automate tasks such as case documentation, scheduling, and resource allocation. By sifting through a high volume of case files, AI can prioritise cases based on urgency, freeing up social workers to focus on direct client interaction. In policy, AI can analyse social trends and public sentiment to inform evidence-based policy development and advocacy efforts (*ibid.*).

## AI IN SOCIAL WORK EDUCATION

The integration of AI into social work education is still in its nascent stages, but holds significant promise (Nurhasanah and Nugraha, 2024).

- **Enhanced Learning and Assessment:** AI can personalise learning experiences by tailoring content to a student's pace and learning style (Wangdi, 2024). It can also automate the grading of certain assignments, such as quizzes and short-

answer questions, providing instant feedback (Chikumbirike and Chifamba, 2022).

- **Simulations and Virtual Realities:** AI-driven simulations can create realistic, immersive scenarios for students to practise clinical skills. Students could interact with virtual clients, family members, or other professionals in a safe, controlled environment, learning to handle difficult conversations or crises before entering the field (Contreras and Jaimes, 2024).
- **Data Literacy and Analytics:** Training future social workers to understand and utilise AI is becoming a core competency. AI tools can be used in the classroom to teach students how to analyse complex social data, identify trends, and evaluate the efficacy of interventions (Polat *et al.*, 2025). This prepares them for a data-driven world and empowers them to advocate for ethical data use.

## THE ETHICAL CHALLENGES AND RISKS OF AI IN SOCIAL WORK

While the potential benefits are clear, the literature highlights a number of critical ethical challenges which must be addressed to ensure AI's responsible use (Weber, 2020; Musengimana, 2025). These challenges directly intersect with the core values of the social work profession, including dignity and worth of the person, social justice, and integrity (Rodríguez-Martínez *et al.*, 2024).

### ALGORITHMIC BIAS AND INEQUITY

Perhaps the most significant ethical challenge is the risk of algorithmic bias. AI systems are trained on historical data, which often reflect existing societal biases, inequalities, and structural injustices. If an AI model is trained on data from a justice system that disproportionately penalises marginalised communities, the model will learn and replicate those biases (Barnes and Hutson, 2024). When this model is used to make decisions about parole, child custody, or resource allocation, it can perpetuate and even amplify systemic discrimination,

directly contradicting the social work commitment to social justice (Garkisch and Goldkind, 2025). The lack of diversity in the AI development workforce also contributes to this problem, as developers may not be aware of the biases inherent in the data they use.

#### **PRIVACY, CONFIDENTIALITY AND DATA SECURITY**

Social work is a profession built on trust and the promise of confidentiality. The use of AI, which often relies on vast amounts of personal and sensitive data, poses a significant threat to this principle (Leta and Vancea, 2023). Data collected from clients, including mental health information, family history, and personal struggles, could be used to train AI models without explicit, informed consent (Sywelem and Mahklouf, 2024). There are also risks of data breaches and misuse. The question of data ownership and how client information is stored, shared, and used by third-party AI developers is a critical and often opaque issue (Nurhasanah and Nugraha, 2024).

#### **LACK OF TRANSPARENCY AND EXPLAINABILITY**

Many advanced AI models, particularly "black box" algorithms, are opaque. Their decision-making processes are not easily understandable or explainable (Barnes and Hutson, 2024). For a social worker, this is deeply problematic. A professional must be able to explain the reasoning behind a particular decision to a client, a supervisor, or a court. If an AI system recommends a specific intervention or risk assessment score without providing a clear rationale, it undermines the social worker's professional judgment and the client's right to understand the basis of a decision that affects their life (Polat *et al.*, 2025). The lack of transparency erodes trust and accountability (Weber, 2020).

#### **DE-EMPHASIS ON THE HUMAN-CENTRED RELATIONSHIP**

The social work profession is fundamentally a relational one. Empathy, active listening, and the therapeutic relationship are at the heart of

effective practice (Garkisch and Goldkind, 2025). While AI can automate tasks, it cannot replicate human connection and judgment (Musengimana, 2025). Over-reliance on AI could lead to a de-emphasis on human-to-human interaction, reducing complex human experiences to data points and algorithms. This could lead to a "dehumanisation" of social services, where clients feel like they are being processed by a system rather than being seen and heard by a compassionate professional (Nurhasanah and Nugraha, 2024). A central concern is the potential for AI to replace the human element rather than augment it (Rodríguez-Martínez *et al.*, 2024).

### **INFORMED CONSENT AND AUTONOMY**

Securing true, informed consent for the use of AI in social services is a complex challenge (Rodríguez-Martínez *et al.*, 2024). Can a client truly understand the implications of their data being used to train a machine learning model, especially when the process itself is not transparent? The power dynamics inherent in the social worker-client relationship make this even more difficult. Clients, particularly those in vulnerable situations, may not feel they have the agency to decline the use of an AI tool, fearing it could negatively impact the services they receive (Garkisch and Goldkind, 2025; Leta and Vancea, 2023).

### **ETHICAL FRAMEWORKS FOR AI IN SOCIAL SERVICES**

Recognising these challenges, a number of frameworks and guidelines have been proposed by professional bodies and researchers to ensure the ethical use of AI (Mubofu and Kitali, 2024; Azman and Tümkeya, 2025). These frameworks often build upon existing principles of professional ethics but are adapted for the unique context of AI.

### **CORE PRINCIPLES OF RESPONSIBLE AI**

Many ethical AI frameworks share common principles that are directly relevant to social work (Chikumbirike and Chifamba, 2022; Zeer *et al.*, 2023):

- **Human Agency and Oversight:** The principle that humans must remain in control. AI should be a tool that augments, not replaces, human judgment. There must always be a "human-in-the-loop" who can override, question, or dismiss an AI recommendation (Barnes and Hutson, 2024).
- **Fairness and Non-discrimination:** AI systems must be designed and implemented in a way that avoids bias and promotes equitable outcomes. This requires a commitment to using diverse and representative data sets and regularly auditing algorithms for discriminatory outputs (Zainuddin *et al.*, 2024; Polat *et al.*, 2025).
- **Transparency and Explainability:** The processes and reasoning behind an AI system's decisions must be understandable and communicated in plain language (Weber, 2020; Zeer *et al.*, 2023).
- **Privacy and Data Security:** Strict data governance policies must be in place to protect client information. This includes securing informed consent, anonymising data where possible, and ensuring compliance with regulations like the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) (Leta and Vancea, 2023).
- **Accountability:** There must be a clear chain of responsibility for the development, deployment, and outcomes of AI systems. If a system causes harm, it must be possible to hold an individual or organisation accountable (Barnes and Hutson, 2024).
- **Beneficence and Non-maleficence:** The AI system should be designed to do well and "do no harm" (Chikumbirike and Chifamba, 2022; Contreras and Jaimes, 2024). This means a thorough risk assessment must be conducted before implementation to ensure the potential benefits outweigh the risks (Sywelem and Mahklouf, 2024).



## PROFESSIONAL STANDARDS AND GUIDELINES

Organisations like the National Association of Social Workers (NASW) and other professional bodies are beginning to release statements and standards for the use of technology in social work. These documents emphasize that all technological interventions, including AI, must align with the core values and ethical principles of the profession (Rodríguez-Martínez *et al.*, 2024). They call for social workers to be technologically literate, to critically evaluate new technologies, and to advocate for the rights of clients in a digital age.

## THE PEDAGOGICAL IMPERATIVE: REFORMING SOCIAL WORK EDUCATION

Given the rapid pace of technological change and the profound ethical implications of AI, social work education must adapt to prepare students for the future of practice (Mubofu and Kitali, 2024). This requires a shift in curriculum, pedagogy, and research (Polat *et al.*, 2025).

## INTEGRATING AI LITERACY AND DIGITAL COMPETENCE

Social work programmes must move beyond a basic introduction to technology and develop a robust curriculum focused on AI literacy (Chikumbirike and Chifamba, 2022; Wangdi, 2024). This involves more than just teaching students how to use specific software. It should include:

- **Understanding the Fundamentals:** Educating students on what AI is, how it works, and its various applications in social services. This includes concepts like machine learning, natural language processing, and predictive analytics (Contreras and Jaimes, 2024).
- **Ethical and Critical Thinking:** Fostering the ability to critically evaluate AI tools (Leta and Vancea, 2023). Students should learn to ask the right questions: Who developed this tool?

What data was it trained on? Are there potential biases? How does it align with the NASW Code of Ethics?

- **Hands-on Experience:** Providing students with opportunities to interact with AI tools in a controlled educational setting, such as through case simulations or data analysis projects. This demystifies the technology and allows them to develop practical skills (Azman and Tümkaya, 2025).

### EMPHASIZING THE HUMAN-CENTRED CORE

While integrating technology, social work education must re-emphasizes the unique and irreplaceable role of human-centred practice. The curriculum should:

- **Reaffirm Foundational Skills:** Strengthen training in core social work skills like empathy, relationship building, active listening, and critical thinking. These are the skills that AI cannot replicate and will become even more valuable in a technologically mediated world (Nurhasanah and Nugraha, 2024; Garkisch and Goldkind, 2025).
- **Teach Technology as a Tool, not a Solution:** Position AI as a supplementary tool to enhance a social worker's capabilities, not as a replacement for human judgment (Musengimana, 2025). The focus should be on how AI can free up time for more meaningful client interaction.
- **Promote Advocacy and Ethical Leadership:** Educate students to become advocates for ethical AI development and implementation. This includes understanding policy, participating in interdisciplinary dialogues with technologists, and advocating for the rights of vulnerable populations who may be impacted by AI-driven decisions (Zeer *et al.*, 2023; Barnes and Hutson, 2024).

### CONCLUSION AND RECOMMENDATIONS

The integration of Artificial Intelligence into social work education is not an option but an urgent necessity. AI holds the potential to significantly improve the efficiency and effectiveness of social services, but only if its adoption is guided by a strong ethical compass (Sywelem and Mahklouf, 2024; Polat *et al.*, 2025). This desktop review shows that the risks of algorithmic bias, privacy violations, and the de-emphasis of human relationships are real and must be actively addressed (Rodríguez-Martínez *et al.*, 2024; Garkisch and Goldkind, 2025).

To ensure the ethical and responsible use of AI, social work educators and professional bodies must act on the following recommendations:

1. **Revise and Update Social Work Curricula:** Integrate AI literacy as a core competency. Develop specific courses or modules on "Technology and Social Work" that cover the fundamentals of AI, its ethical implications, and its practical applications (Mubofu and Kitali, 2024).
2. **Develop Interdisciplinary Partnerships:** Foster collaboration between social work schools and computer science, data science, and law departments (Contreras and Jaimes, 2024). This will allow for the co-creation of AI tools that are ethically sound and designed with the values of social work in mind.
3. **Promote Research on Ethical AI:** Fund and conduct more research on the impact of AI on social work practice, particularly on marginalised populations (Nurhasanah and Nugraha, 2024). This research should be led by social work scholars who can bring a social justice perspective to the technological conversation.
4. **Establish Clear Professional Standards and Guidelines:** Professional organisations must continue to develop and enforce clear, actionable standards for the ethical use of AI (Azman and Tümkaya, 2025). These guidelines should provide concrete direction on issues like informed consent, data

governance, and algorithmic transparency (Weber, 2020; Garkisch and Goldkind, 2025).

5. **Advocate for Responsible AI Policy:** Social work professionals and educators must engage in the broader public and political discourse on AI policy (Mubofu and Kitali, 2024). They must advocate for regulations that protect vulnerable populations and ensure that all AI systems used in social services are transparent, fair, and accountable (Barnes and Hutson, 2024).

Through taking these steps, social work education can fulfil its mission of preparing future professionals who are not only competent practitioners, but also ethical leaders capable of navigating a technologically complex world while upholding the core humanistic values of the profession (Zainuddin *et al.*, 2024; Musengimana, 2025). The future of social work is intertwined with AI, and it is the responsibility of educators to ensure that this future is one of justice, equity, and human dignity.

## REFERENCES

- Azman, Ö. and Tümkaya, S. (2025). Navigating the Ethical Landscape of AI Integration in Education: Balancing Innovation and Responsibility. *F1000Research*, 14, 299.
- Barnes, E. and Hutson, J. (2024). Navigating the Ethical Terrain of AI in Higher Education: Strategies for Mitigating Bias and Promoting Fairness. *Forum for Education Studies*, 2(2).
- Chikumbirike, E. P. K. C. and Chifamba, C. (2022). Ethics and Artificial Intelligence in Academia. *Journal of Computer Assisted Learning*, 38(6), 1543-1557.
- Garkisch, M. and Goldkind, L. (2025). Considering a Unified Model of Artificial Intelligence Enhanced Social Work: A Systematic Review. *Journal of Human Rights and Social Work*, 10(1), 23-42.

- Leta, F. M. and Vancea, D. P. C. (2023). Ethics in Education: Exploring the Ethical Implications of Artificial Intelligence Implementation. *Ovidius University Annals, Economic Sciences Series*, 23(1), 413-421.
- Mubofu, C. and Kitali, L. (2024). Artificial Intelligence in Education: Ethics and Responsible Implementation. *Journal of Interdisciplinary Studies in Education*, 13(2).
- Nurhasanah, S. and Nugraha, M. S. (2024). The Future of Learning: Ethical and Philosophical Implications of Artificial Intelligence (AI) Integration in Education. *Indonesian Journal of Multidisciplinary Research*, 4(2), 341-352.
- Polat, M., Karataş, İ. H. and Varol, N. (2025). Ethical Artificial Intelligence (AI) in Educational Leadership: Literature Review and Bibliometric Analysis. *Leadership and Policy in Schools*, 24(1), 46-76.
- Wangdi, P. (2024). Integrating Artificial Intelligence in Education: Trends and Opportunities. *International Journal of Research in STEM Education*, 6(2), 50-60.