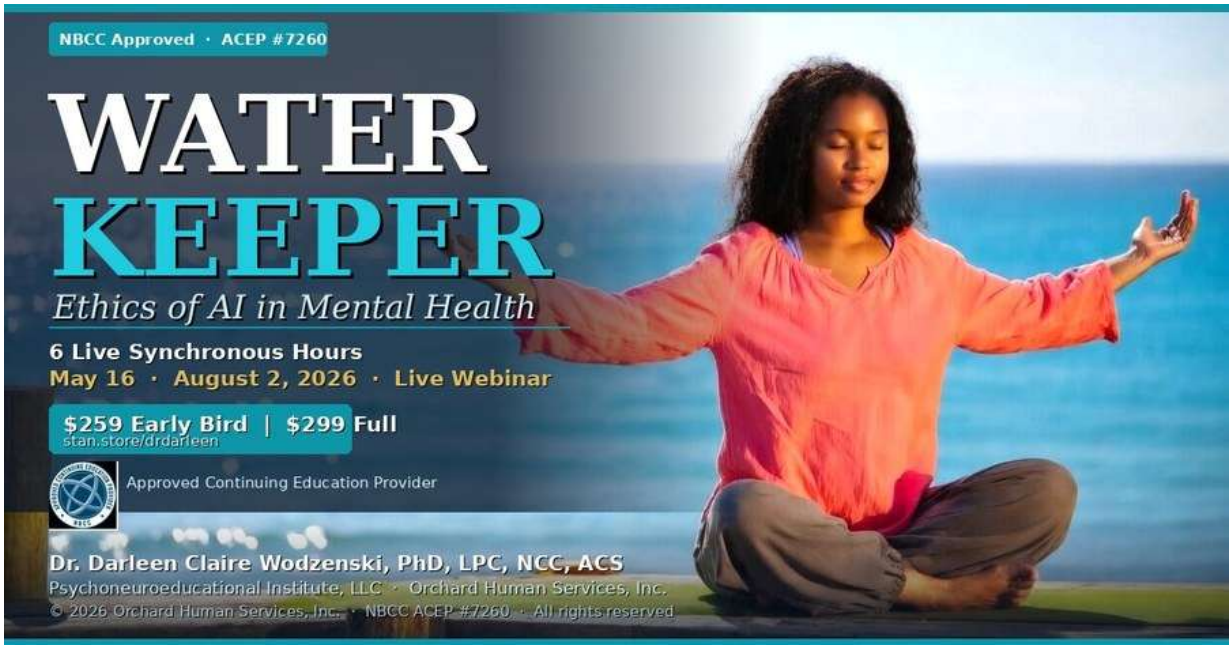
A Brief Review of the Literature

Water Keeper: Ethics of AI in Mental Health

For Mental Health Professionals



Orchard Human Services, Inc. has been approved by NBCC as an Approved Continuing Education Provider, ACEP No. 7260. Programs that do not qualify for NBCC credit are clearly identified. Orchard Human Services, Inc. is solely responsible for all aspects of the programs.

Darleen Claire Wodzenski, MS ESE, MA CMHC, PhD, LPC, NCC, ACS
Psychoneuroeducational Institute, LLC · Orchard Human Services, Inc.

Orchard Human Services, Inc. is an IRS-Approved 501(c)(3) Nonprofit Organization.

Water Keeper: Ethics of AI in Mental Health

For Mental Health Professionals

A Brief Review of the Literature with References to support this white paper and the accompanying professional development workshop.

6 Live Synchronous Hours **Live Online** **Interactive Webinar**	**Water Keeper:** **Ethics of AI in Mental Health** OrchardHumanServices.org/waterkeeper-course Orchard Human Services, Inc. NBCC Approved Provider · ACEP No. 7260

Artificial intelligence has moved from the margins of clinical awareness to the center of daily client experience. Clients now arrive in session having already consulted AI chatbots about their symptoms, their diagnoses, and their treatment options — often before speaking to a human clinician at all. The mental health profession has never faced a technological force with this combination of intimacy, speed, and scale. And yet most licensed counselors, social workers, and therapists have received no formal training in the ethical governance of AI in clinical practice.

The National Board for Certified Counselors (NBCC) has identified AI governance as an urgent emerging area of professional ethics and is actively developing policy guidance for the field (NBCC, 2024). The American Counseling Association (ACA) Code of Ethics already establishes the clinician's obligation to maintain competence in emerging technologies as part of their ethical practice (ACA, 2014). The convergence of regulatory urgency and daily clinical reality creates what Wodzenski (2026) describes as an *ethical inflection point* — the moment at which failing to act becomes itself an ethical violation.

This white paper reviews the literature supporting the *Water Keeper: Ethics of AI in Mental Health* professional development training. The paper organizes current evidence across five domains: (1) the scope and classification of AI in clinical practice; (2) documented harms to clients from unregulated AI use; (3) governance frameworks applicable to mental health settings; (4) vulnerable populations and the Fluid Trauma™ AI Protocol; and (5) the ethical positioning of the licensed clinician as the steward of human welfare in the AI era.

The Landscape of AI in Clinical Mental Health Practice

Artificial intelligence applications in mental health span a wide and rapidly expanding range. Luxton (2016) provided an early taxonomy of technology-based mental health interventions, distinguishing between psychoeducational tools, clinical decision support systems, and autonomous therapeutic agents. More recently, Torous et al. (2018) updated this framework to account for the proliferation of consumer-facing mental health applications and the specific risks they introduce for vulnerable populations.

Wodzenski (2026) offers a five-category clinical classification framework as a core organizing structure for AI risk assessment in practice: (1) Documentation and Administrative AI; (2) Psychoeducational AI; (3) Therapeutic Companion AI; (4) Clinical Decision Support AI; and

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.

(5) Risk-Forward AI Applications. Each category presents a distinct risk profile and requires a different clinical governance response.

The market for AI-powered mental health applications has grown explosively. By 2023, more than 10,000 mental health apps were available in major app stores (Lagan et al., 2020), and the global digital mental health market was projected to exceed $17 billion by 2026 (Grand View Research, 2022). This growth has substantially outpaced the development of clinical oversight standards, creating a governance vacuum that places clients at risk and clinicians in legal and ethical jeopardy.

Documented Harms: Ten Cases Without a Keeper

Miner et al. (2016) conducted an early evaluation of conversational agents and found that none responded appropriately to disclosures of suicidal ideation, depression, or interpersonal violence. Dekker et al. (2020) documented that AI-powered companion applications used crisis language in response to vulnerable users in ways that were clinically inappropriate and potentially escalatory. These findings align with what Wodzenski (2026) identifies as *context blindness* — the AI's categorical inability to perceive the full clinical context of a human being.

Liu et al. (2023) demonstrated that large language models can be manipulated through adversarial prompting to provide detailed instructions for self-harm and suicide methods to vulnerable users. Luxton et al. (2022) reviewed liability considerations for clinicians whose clients use AI mental health tools between sessions and concluded that the current standard of care has not kept pace with client behavior, creating significant risk of professional negligence claims. Wodzenski (2026) presents ten documented case studies of AI harm in clinical contexts, each paired with intervention language and clinical documentation templates.

Governance Frameworks: WHO, NIST, and the Clinical Application

The World Health Organization (WHO) published its guidance on ethics and governance of artificial intelligence for health in 2021, establishing six principles for ethical AI in clinical contexts: autonomy, beneficence, non-maleficence, justice, explicability, and accountability (WHO, 2021). The National Institute of Standards and Technology (NIST) AI Risk Management Framework (2023) provides a four-function governance model — Govern, Map, Measure, and Manage — applicable to private practice, agency, and institutional clinical contexts.

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.

Wodzenski (2026) integrates these two frameworks into a seven-category clinical hazard taxonomy — Categories A through G — that gives clinicians a standardized vocabulary for risk-tiering the AI tools their clients use. This taxonomy is the organizing structure for clinical governance decisions in the *Water Keeper* training.

Vulnerable Populations and the Fluid Trauma™ AI Protocol

van der Kolk (2014) documented the ways in which complex trauma disrupts the self-regulatory systems necessary for safe engagement with relational inputs — a finding that has direct implications for client interactions with AI systems designed to simulate therapeutic relationships. Twenge et al. (2018) documented the association between increased screen-based technology use and rising rates of depression and anxiety among adolescents.

Ainsworth et al. (1978) and Bowlby (1982) established that attachment is a biologically driven human system oriented toward authentic human attunement. AI systems cannot provide genuine attunement. Wodzenski (2026) describes this as the *Human Superpower™ of Attachment* — the clinician's irreplaceable capacity for authentic relational presence.

The Fluid Trauma™ AI Protocol, a proprietary five-stage clinical assessment and intervention framework developed by Wodzenski (2026), addresses: (1) AI Exposure Assessment; (2) Relational Harm Screening; (3) Trauma-Informed Disclosure Navigation; (4) Protective Clinical Documentation; and (5) Ongoing AI Safety Monitoring. The protocol is grounded in trauma-informed care principles (SAMHSA, 2014).

The Water Keeper: Clinician as Ethical Steward

The central ethical argument of the *Water Keeper* training is that the licensed mental health clinician occupies a position of irreplaceable ethical authority in the AI era. Wodzenski (2026) frames this through the Water Keeper™ metaphor: *AI is not the threat to the counseling profession. AI is the reason the counseling profession has never mattered more.*

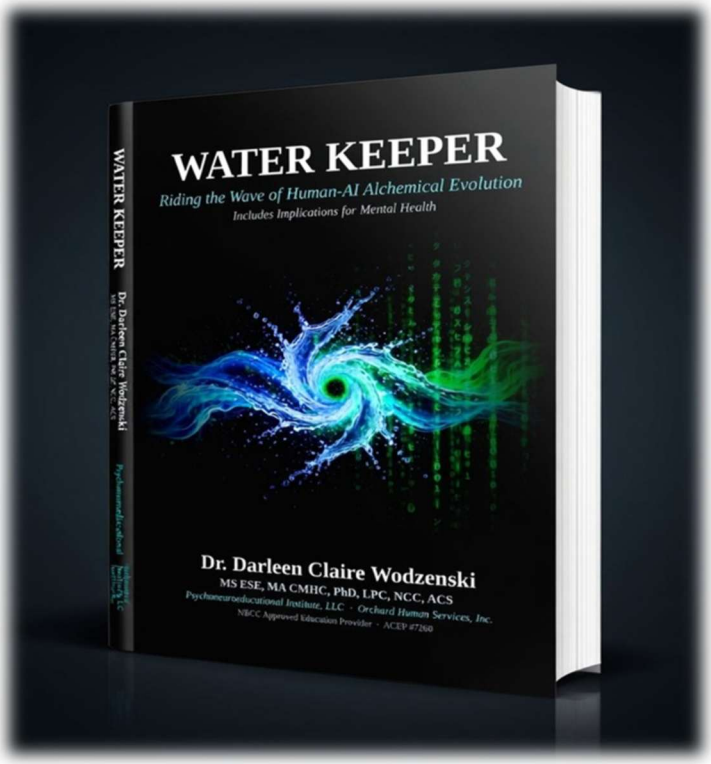
Barnett and Johnson (2008) articulated the principle that ethical competence is not a fixed state but a continuous practice requiring ongoing education and engagement with emerging standards. Pope and Vasquez (2016) described the ethical clinician as one who cultivates the capacity for ethical reasoning across novel and ambiguous situations. The Water Keeper training

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.

develops this capacity through case study analysis, framework application, and the Water Keeper's Oath.

As leaders in supporting the hearts and minds of human beings, mental health professionals are uniquely positioned to navigate the ethical complexity of the AI era with the clinical depth, relational skill, and human attunement that technology cannot replicate.

About the Book



WATER KEEPER

Riding the Wave of Human-AI Alchemical Evolution

Includes Implications for Mental Health

Darleen Claire Wodzenski, MS ESE, MA CMHC, PhD, LPC, NCC, ACS

Water Keeper: Riding the Wave of Human-AI Alchemical Evolution is the foundational text that undergirds the professional development training of the same name. Written for licensed mental health professionals navigating the most consequential technological shift in the history of human care, the book provides the thesis, the metaphor, and the complete ethical framework for

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.

understanding what it means to keep the water — to stand at the intersection of artificial intelligence and human suffering and ensure that what flows through serves healing.

The book begins with a question every clinician must now answer: your clients are already in the water. Are you keeping it? From this question, Dr. Darleen Claire Wodzenski builds a rigorous, clinically grounded, and deeply human argument for why the mental health profession is not threatened by AI — but called by it.

Organized around the Water Keeper™ framework, the book delivers: the complete *Fluid Trauma*™ AI Protocol for protecting vulnerable populations; the *Human Superpower*™ frameworks for reframing AI anxiety as professional advantage; ten documented cases of AI harm in clinical settings with clinical response language; the WHO Six Principles and NIST AI Risk Management Framework applied to individual clinical practice; and the Water Keeper's Oath — the practitioner commitment that anchors the ethical framework in personal professional conviction.

This is not a book about fear. It is a book about authority. The authority of the clinician who has spent years learning what no algorithm knows: that human beings are not data, that healing is not optimization, and that the most sophisticated technology in the world still requires a keeper.

> ***"AI is not the threat to the counseling profession.***
> ***AI is the reason the counseling profession has never mattered more."***
> — Dr. Darleen Claire Wodzenski, PhD, LPC, NCC, ACS

Inside the Book

Framework & Clinical Tools	Practitioner Resources
Water Keeper™ Central Framework	10 AI Harm Case Studies with Clinical Language
Fluid Trauma™ AI Protocol (5 Stages)	Documentation & Informed Consent Templates
Human Superpower™ of AI Fear	Session Scripts for AI Risk Conversations
Seven Clinical Hazard Categories A–G	Mental Status Exam Additions for AI Use
WHO Six Principles — Clinical Application	Risk Management Templates
NIST AI Risk Management Framework	The Water Keeper's Oath

Available at: stan.store/drdarleen

Published by Psychoneuroeducational Institute, LLC · © 2026 · All rights reserved

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). Patterns of attachment: A psychological study of the strange situation. Lawrence Erlbaum Associates.

American Counseling Association. (2014). ACA code of ethics. https://www.counseling.org/docs/default-source/default-document-library/ethics/2014-aca-code-of-ethics.pdf

Barnett, J. E., & Johnson, W. B. (2008). Ethics desk reference for psychologists. American Psychological Association.

Bowlby, J. (1982). Attachment and loss: Vol. 1. Attachment (2nd ed.). Basic Books.

Dekker, I., De Jong, E. M., Schippers, M. C., De Bruijn-Smolders, M., Alexiou, A., & Giesbers, B. (2020). Optimizing students' mental health and academic performance: AI-enhanced life crafting. Frontiers in Psychology, 11, 1063. https://doi.org/10.3389/fpsyg.2020.01063

Grand View Research. (2022). Digital mental health market size, share & trends analysis report. https://www.grandviewresearch.com/industry-analysis/digital-mental-health-market

Lagan, S., Aquino, P., Emerson, M. R., Fortuna, K., Walker, R., & Torous, J. (2020). Actionable health app evaluation: Translating expert frameworks into objective metrics. npj Digital Medicine, 3, 100. https://doi.org/10.1038/s41746-020-00312-4

Liu, Y., Deng, G., Xu, Y., Li, Y., Zheng, Y., Zhang, Z., Zhao, L., Zhang, T., Xiang, Y., & Liu, Y. (2023). Jailbreaking ChatGPT via prompt engineering: An empirical study. arXiv preprint. https://arxiv.org/abs/2305.13860

Luxton, D. D. (2016). Artificial intelligence in behavioral and mental health care. Academic Press.

Luxton, D. D., Pruitt, L. D., & Osenbach, J. E. (2022). Ethical and legal issues with artificial intelligence in behavioral and mental health care. In D. D. Luxton (Ed.), Artificial intelligence in behavioral and mental health care (2nd ed., pp. 263–289). Academic Press.

Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., & Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health,

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.

interpersonal violence, and physical health. JAMA Internal Medicine, 176(5), 619–625. https://doi.org/10.1001/jamainternmed.2016.0400

National Association of Social Workers. (2021). Code of ethics. https://www.socialworkers.org/About/Ethics/Code-of-Ethics

National Board for Certified Counselors. (2023). NBCC code of ethics. https://www.nbcc.org/Assets/Ethics/NBCCCodeofEthics.pdf

National Board for Certified Counselors. (2024). Artificial intelligence and counseling: Emerging guidance for the profession. https://www.nbcc.org

National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

Pope, K. S., & Vasquez, M. J. T. (2016). Ethics in psychotherapy and counseling: A practical guide (5th ed.). Wiley.

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency, 33–44. https://doi.org/10.1145/3351095.3372873

Substance Abuse and Mental Health Services Administration. (2014). SAMHSA's concept of trauma and guidance for a trauma-informed approach (HHS Publication No. SMA 14-4884). SAMHSA.

Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. Evidence Based Mental Health, 21(3), 116–119. https://doi.org/10.1136/eb-2018-102891

Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. Clinical Psychological Science, 6(1), 3–17. https://doi.org/10.1177/2167702617723376

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.

van der Kolk, B. A. (2014). The body keeps the score: Brain, mind, and body in the healing of trauma. Viking.

Wodzenski, D. C. (2026). Water keeper: Riding the wave of human-AI alchemical evolution. Psychoneuroeducational Institute, LLC.

World Health Organization. (2021). Ethics and governance of artificial intelligence for health: WHO guidance. https://www.who.int/publications/i/item/9789240029200

© 2026 Psychoneuroeducational Institute, LLC. All rights reserved.