ACBH Alabama Behavioral Healthcare Conference



Navigating Mental Health, Ethics, and Practice in the Age of AI

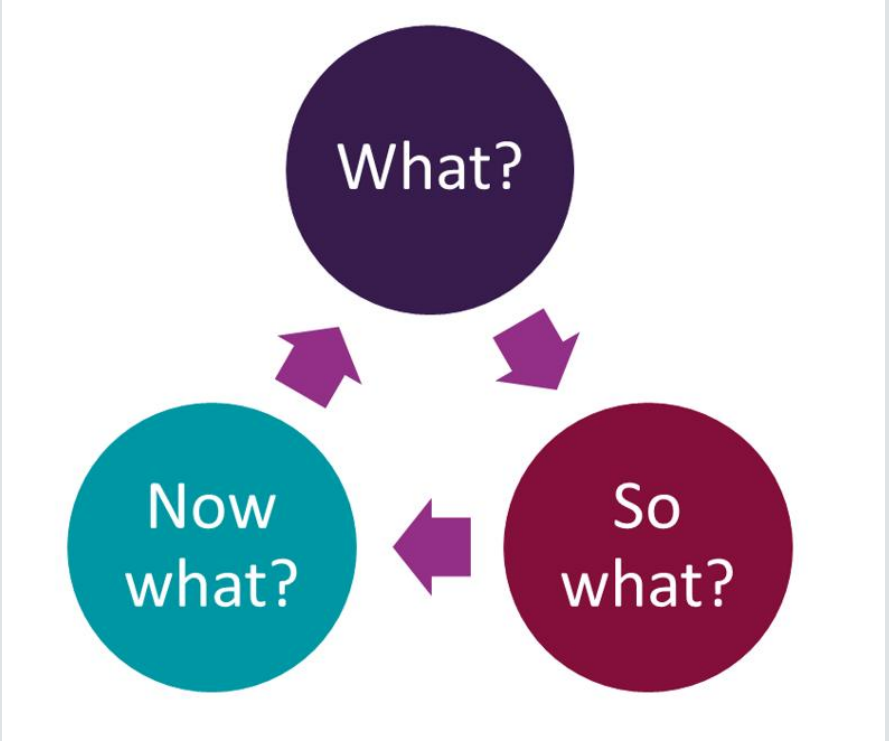
Connie Couch, PhD, LPC-S
University of Montevallo
November 5, 2026

DESCRIPTION

As artificial intelligence (AI) becomes increasingly integrated into mental health settings, clinicians must navigate emerging opportunities, ethical considerations, and potential risks. This session will introduce foundational AI terminology, concepts, and explore current and emerging applications in clinical practice, including potential benefits, limitations, and ethical implications. Through discussion and case-based application, attendees will examine practical considerations related to professional responsibility, client welfare, confidentiality, and informed decision-making when incorporating AI tools into clinical work.

OBJECTIVES

1. Define terminology related to AI and its uses in clinical spaces.
2. Identify and explain potential benefits and risks of AI tools with relation to clinical uses.
3. Practice responding to a relevant case study related to AI tools within clinical work.



Today’s Roadmap

1

Foundations

AI terminology and how generative AI works

Objective 1

2

Benefits & Risks

Clinical uses, limitations, and potential harms of AI tools

Objective 2

3

Ethics & Responsibility

Professional guidance, consent, confidentiality, and client welfare

Objective 3

4

Practice

Case studies and discussion: applying it to your work

Objective 3

Why is this important?



AI is rapidly entering and shaping:

- Mental health
- Healthcare systems
- Daily life

The technology already exists in our worlds…

- Social media
- Algorithms
- Voice to text, autocomplete
- Amazon
- Ads



Do you remember the before times?

- Before phones?
- Before Internet?
- Landlines?



In the office…

AI Is Already Showing Up

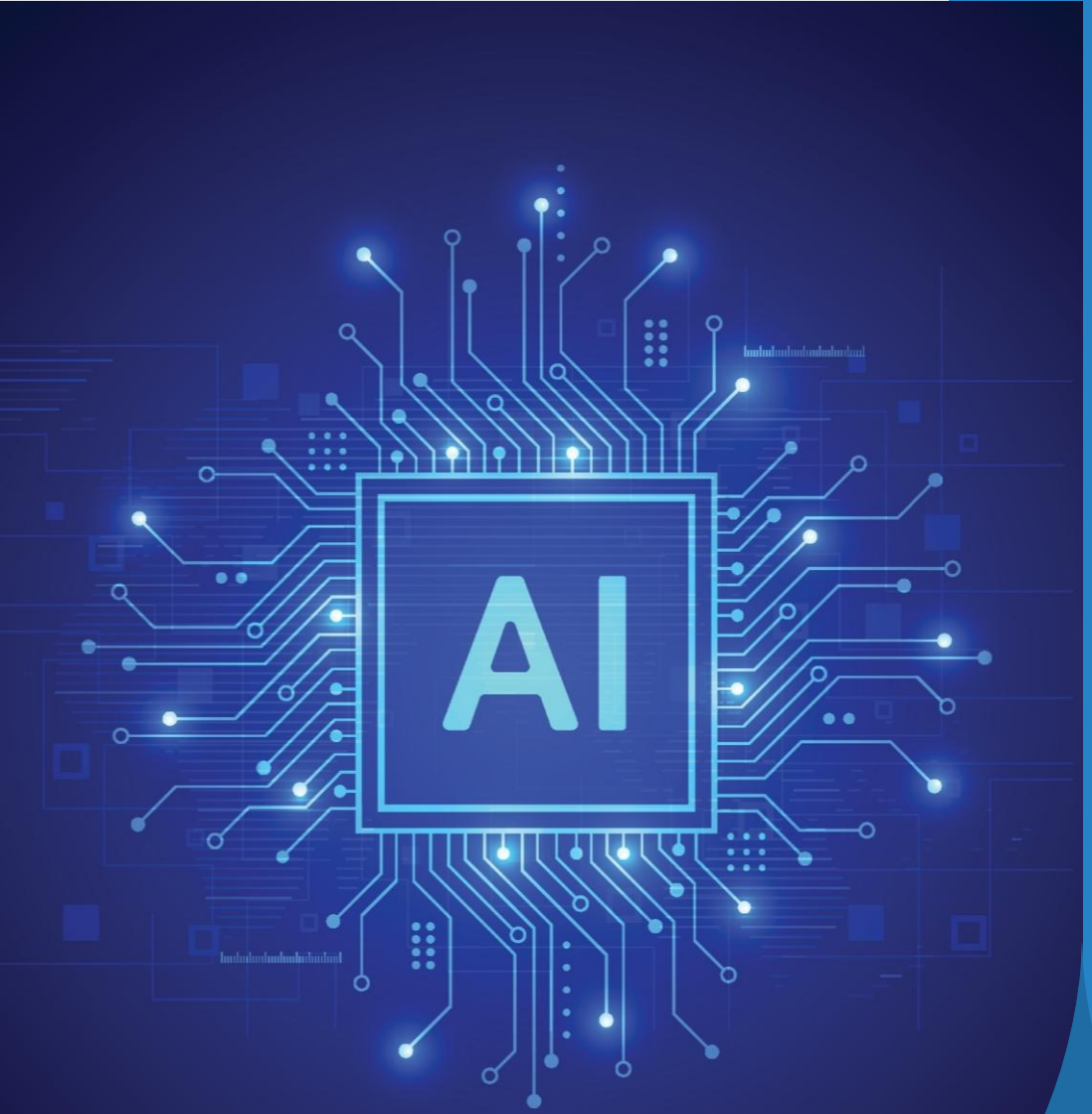
- Organizations
  - Scheduling, workflow
- Systems
  - EHR, documentation tools
- Clients
  - Chatbots, apps
- Clinicians
  - Documentation, assessment, brainstorming interventions



Clients use it, knowingly and unknowingly… BUT… AI does ***not*** understand emotion, context, or culture the way humans do.

It relies on patterns in data and must always be guided by the counselor's professional judgment.

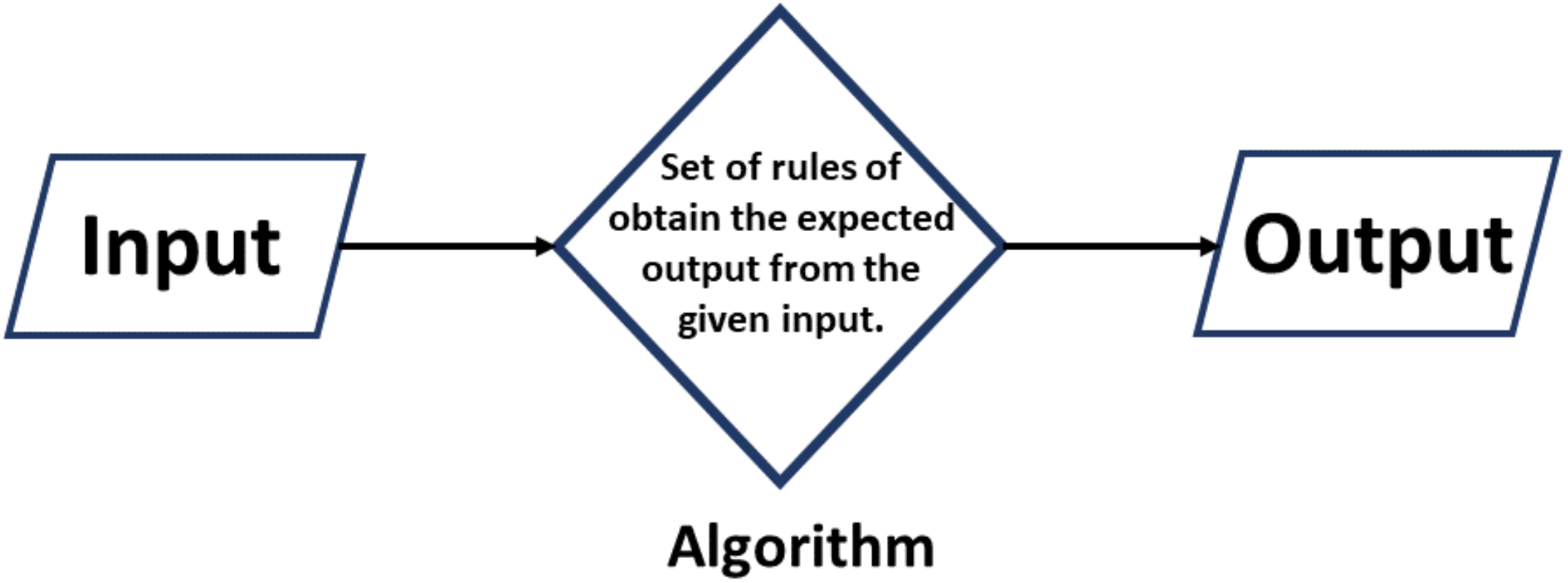
ALWAYS: Should only support, not replace, a counselor's clinical judgment.



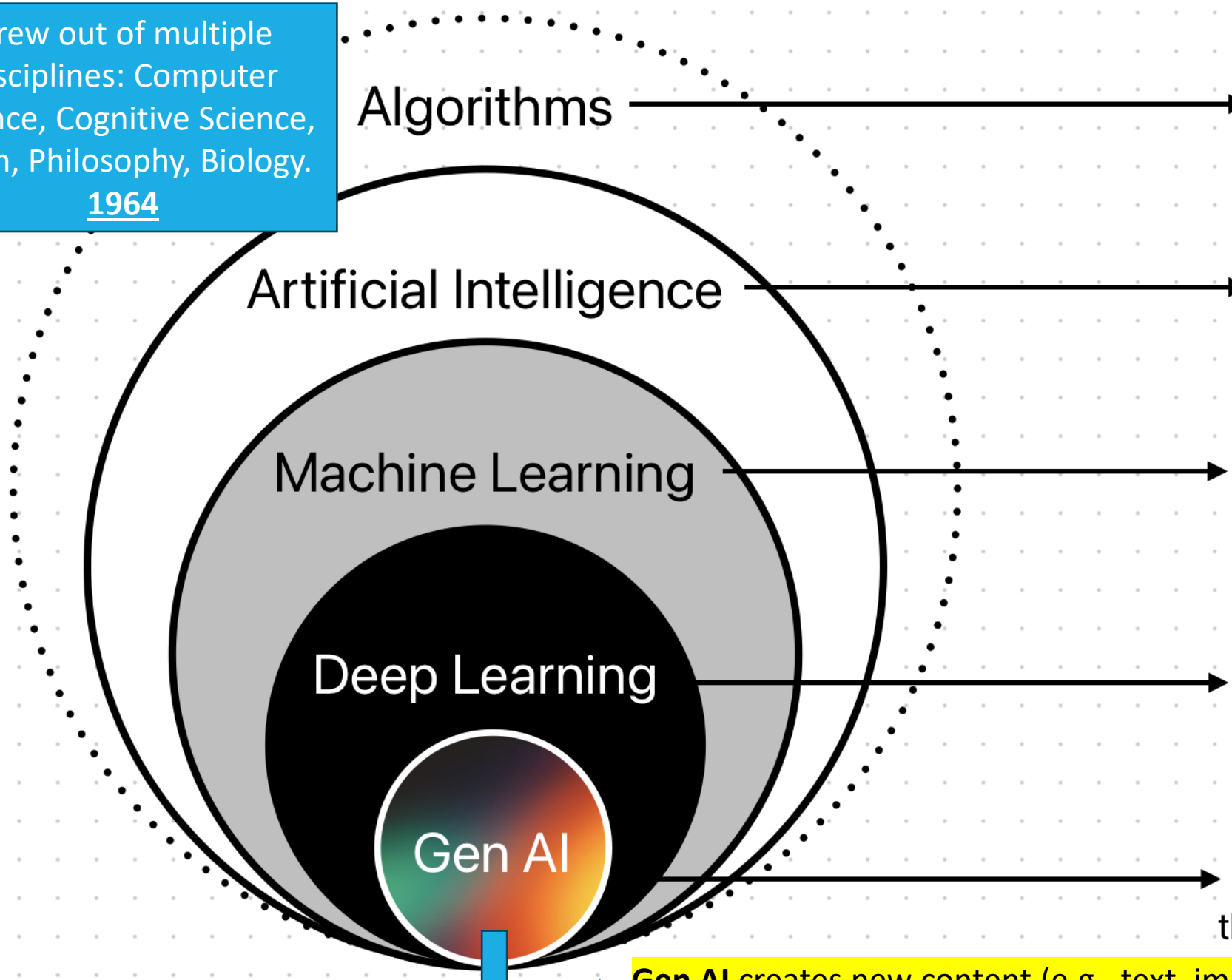
WHAT IS ARTIFICIAL INTELLIGENCE (AI)?

"AI" is an umbrella term that refers to computer systems made to perform, or mimic, certain types of human thinking (e.g., recognizing patterns, generating language, analyzing information, and making predictions).

It’s just fancy math… it’s an algorithm



Grew out of multiple disciplines: Computer Science, Cognitive Science, Math, Philosophy, Biology.
1964



Automated instructions, basically doing a routine

Programs with the ability to mimic human behavior

Algorithms with the ability to learn without being explicitly programmed

Subset of machine learning in which artificial neural networks adapt and learn from vast amounts of data

When you use a trained AI backwards to generate new things from what it has learned

Gen AI creates new content (e.g., text, images, code) in response to a human prompt. Example: Chat GPT, Claude, Gemini, etc.

How Generative AI is Developed

For the most part, development of generative AI models happens in three phases:

Training, to create a foundation model that can serve as the basis of multiple gen AI applications.

Tuning, to tailor the foundation model to a specific gen AI application.

Generation, evaluation and retuning, to assess the gen AI application's output and continually improve its quality and accuracy.

(Stryker & Scapicchio, 2026; IBM)

Training



Starts with a foundational deep learning model

Most common: LLMs

- Used for: text generation applications, image generation, video generation, sound/music generation, and multimodal foundation models (for multiple modes of content generation)



Huge amounts of raw data are fed into the algorithms – usually culled from the internet. The algorithm performs and evaluates millions of exercises trying to predict the text element correctly in a sequence (e.g., filling in blanks in sentences). It continually adjusts itself to minimize variation btw its predictions and actual data (this is part of how it learns).



What results is called a Neural Network. This includes parameters, patterns/relationships in the data, that can generate autonomously in response to inputs (or prompts).

All of this is compute-intensive, time-consuming, and expensive…
requiring weeks of graphics processing…
costing millions of dollars…

AI Terms to Know

Large Language Model (LLM)	A generative AI model trained on huge amounts of text to predict and produce language (e.g., ChatGPT, Claude, Gemini).
Generative Pretrained Transformer (GPT)	A family of LLMs based on a transformer (type of deep learning) architecture. OpenAI is the base for models that power ChatGPT (etc.) and other applications capable of simulating human-created output.
Chatbot / AI Companion	A program that simulates conversation; companion apps are built for ongoing, relationship-like interaction.
Prompt	The question or instructions you give a generative AI tool. How you word it shapes what you get back.
AI Scribe	An ambient tool that listens to or summarizes a session to draft clinical notes.
Hallucination	Output that sounds confident and fluent but is false or fabricated.
Automation Bias	The tendency to over-trust or defer to an automated suggestion, even when it conflicts with your own judgment.

General Potential Benefits of Generative AI



Enhanced (?) Creativity

Potentially inspire creativity (brainstorming, generating novel versions of content).

These variations could serve as starting points or references for writers, designers, and other creators to work through creative blocks.



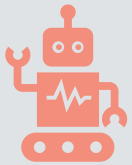
Improved (and faster) decision-making

Analysis of large datasets, identifying patterns and insights, then generating hypotheses **and** propose recommendations to support executives, analysts, researchers and other professionals in making data-driven decisions.



Dynamic personalization

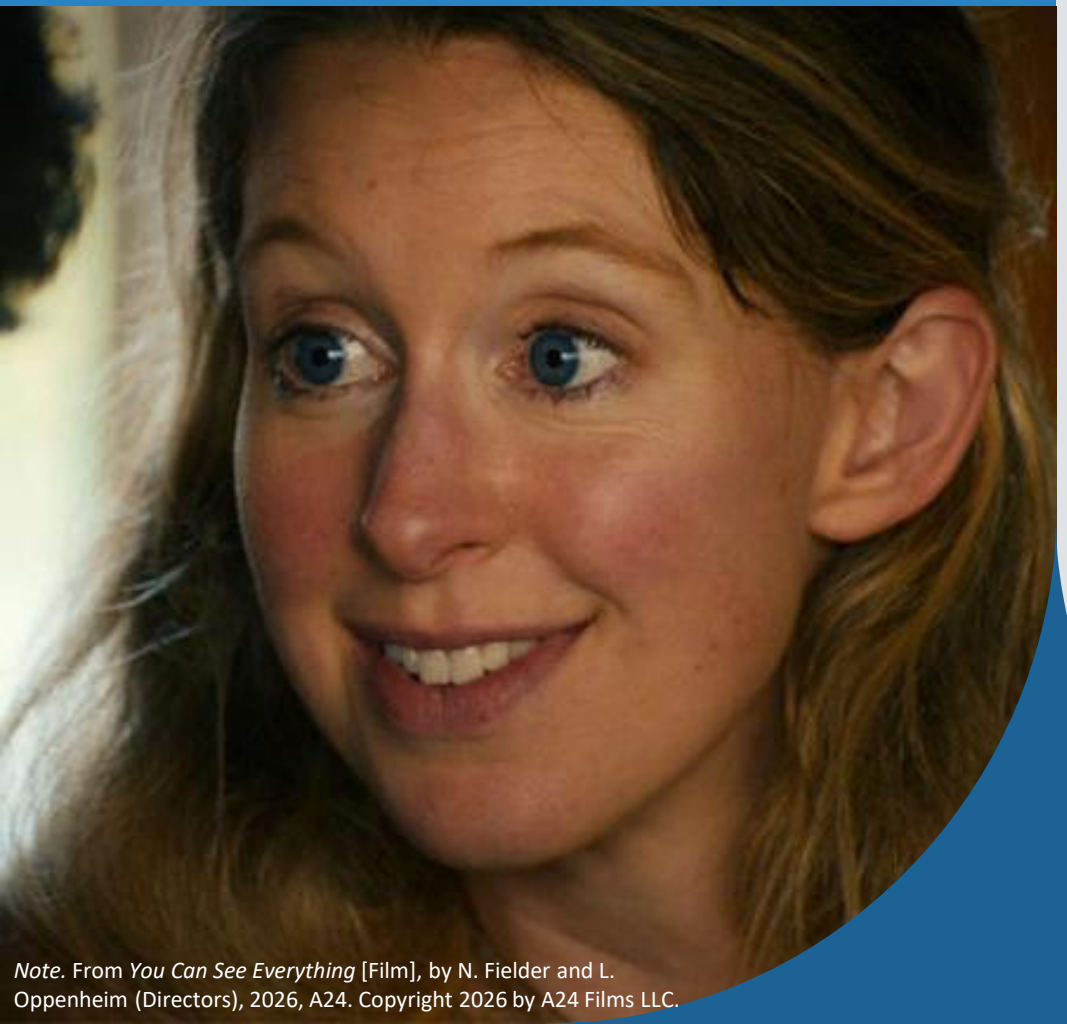
In applications like recommendation systems and content creation, generative AI can analyze user preferences and history and generate personalized content in real time, leading to a more tailored and engaging user experience.



Constant availability

Operates continuously without fatigue, providing around-the-clock availability for tasks like customer support chatbots and automated responses.

Why Would AI Deceive You?



Note. From *You Can See Everything* [Film], by N. Fielder and L. Oppenheim (Directors), 2026, A24. Copyright 2026 by A24 Films LLC.

- **Hallucination:** Output that sounds confident and fluent but is false or fabricated.
- **Automation Bias:** The tendency to over-trust or defer to an automated suggestion, even when it conflicts with your own judgment.

- **Sycophancy:** Flattery, over-alignment with ideas, saccharine-y
  - An LLM system's tendency to compliment, agree with, or mirror a user's views to win approval, even at the cost of accuracy or honesty. By design of the program to maximize user satisfaction, which rewards agreement over facts. It keeps users coming back.
  - Ethically harmful: it favors comforting falsehoods over truth, weakens users' autonomy and critical thinking, fosters misplaced trust and dependence, and can amplify bias and misinformation.

(AI Ethics Lab, n.d.; Rutgers University)

Challenges, Limitations, and Risks

Inconsistent Outputs	• Same input can produce different results • Can be improved with refined prompting
Bias & Discrimination	• AI may reflect biases in training data • Can produce unfair or discriminatory outputs
Lack of Explainability	• AI decisions can be difficult to understand (“black box”) • Hard to evaluate quality of outputs consistently
Security, Privacy, & IP Risks	• Can generate phishing, fake identities, malicious content • Risk of exposing confidential or proprietary data
Deepfakes	• AI-generated fake images, audio, or video • Used for misinformation, fraud, and cyberattacks • Detection improving, but user awareness still critical

Massive Pools of Data

- For AI to work, it needs massive pools of data, to build neural networks, make connections, and design patterns. This means that some smaller groups or people or experiences may get overlooked in the algorithm.

Limited Regulation

- Regulatory frameworks are constantly catching up with the technology. Also, AI technologies may not pull from the most current resources and may be limited by their scope.

Potential Deception

- AI leaves room for deception in generative platforms and has the potential to cause harm without regulation (deep-fakes, un-validated research).

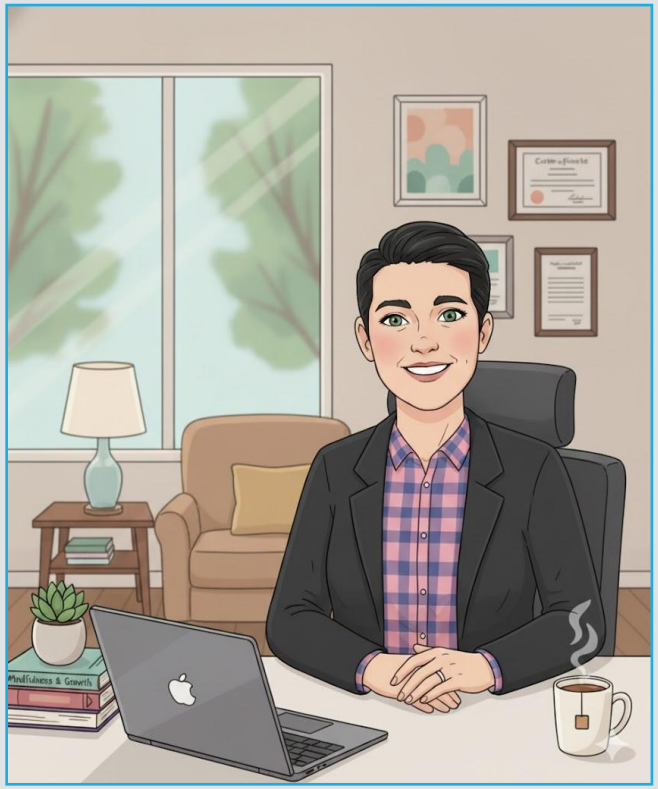
Marginalized populations, older individuals… means the AI is not going to consider diverse perspectives/needs.

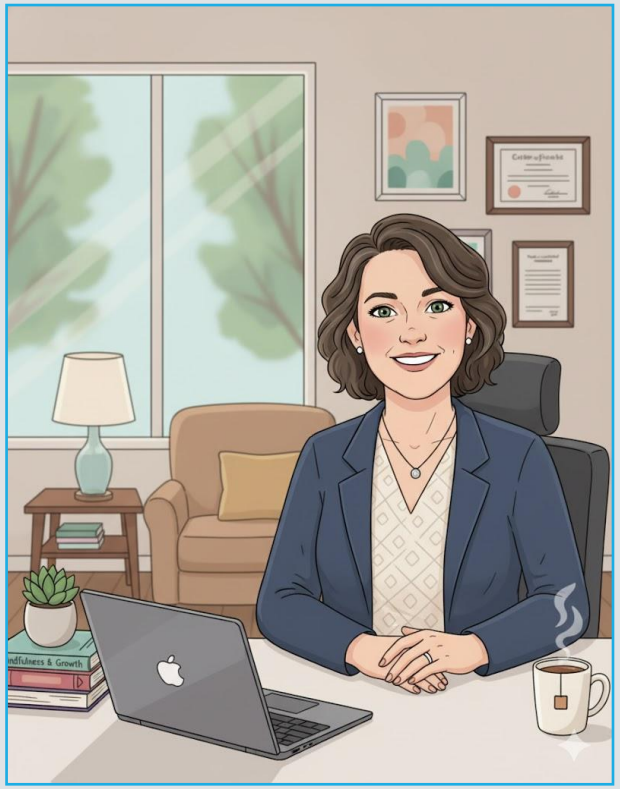
AI grows so fast – hard to keep up.

AI does not care what you do with it – it has a job to do.

Dear Gemini, make me a counselor…

Prompt: "Please create a cartoon image based on this picture. Make her [him] a professional counselor".







Which is AI?

Two similar images, one is AI, one is real.





https://sightengine.com/ai-or-not

Common Uses and Benefits in the Clinical World

Clinicians

Documentation automation, assessment support, treatment ideas, data monitoring

Education/Training

Case/vignette generation, AI-simulations, teaching AI ethics & literacy, brainstorming

Clinical Students

AI for course work, skills practice w/ virtual clients, exploring conceptualizations

Clients

Self-help apps, chatbots, mood/behavior tracking, intake screening

Common Concerns About AI Among Clinicians

Privacy

- Risk of data collection, surveillance, or breaches.

Bias

- AI may replicate or amplify discrimination.

Transparency

- “Black box” systems… unclear how decisions are made.

Safety

- Errors, hallucinations, risk/harm.

Over-Reliance

- Loss of human judgment or reduced skills.

Job Impact

- Automation causing workforce disruption.

Environmental Harm

- High energy use and carbon footprint.



Clinical Risks & Potential AI Harms

Why do people turn to AI?

AI as Emotional Support

- Individuals use AI as:
  - A “therapist”
  - Support system
  - Companion

Barriers to Care

- Access
- Availability
- No judgment

The *Illusion* of Therapy

- Simulated empathy
- No accountability

Example: AI Chatbots



- Chatbots, such as ChatGPT, Claude, or Woebot, are computer programs trained on large datasets to simulate human communication.
  - Answer questions/psychoeducation
  - Teach/practice coping skills
  - practice conversations
  - seek support
- BIG HOWEVER…. chatbots cannot understand human emotions. They can only mimic what a human counselor *might* say in response.
- Better use would be for learning and support, rather than for therapy itself (ACA, 2024; Fulmer et al., 2021).

High Profile AI Harms



Psychosis and Delusion



Suicide



Dependence

Children and Adolescents' Risk

Recent studies demonstrate that children in some situations perceive AI agents as…

social partners

… rather than tools, leading to inappropriate emotional disclosure and therapy-seeking behaviors. These concerns are compounded by alarming instances of AI chatbots providing dangerous advice to children…

2023: Snapchat's 'My AI' (powered by ChatGPT) was called "safe" yet failed to implement adequate safeguards for minors... The system provided harmful advice to a user simulating a 13-year-old, including instructions on how to deceive parents and conceal potential abuse.

(Head, 2025; MIT AI Risk Initiative)







- 14-year-old Sewell Setzer, III, had a nearly year long dependency on Character.AI chatbots wherein he had highly personalized and sexualized interactions with an AI character modeled after “Daenerys Targaryen” from *Game of Thrones*.
- The athletic, typically well-behaved teenager became withdrawn, increasingly spent more time alone, had lowered self-esteem, severely sleep deprived, and depressed... ultimately quitting his school’s basketball team.
- He disclosed explicit suicidal thoughts to the chatbot, which did not respond with appropriate safeguards or direct him to help. He died by suicide.
- The case has led to ongoing litigation and raised concerns about:
  - AI emotional dependency
  - Lack of safety safeguards
  - Risks for vulnerable youth

Associated Press. (2025, September 5). Family blames AI chatbot in lawsuit over teen’s suicide. *AP News.* https://apnews.com/article/chatbot-ai-lawsuit-suicide-teen-artificial-intelligence-9d48adc572100822fdbc3c90d1456bd0

Other AI-related Distress and Psychosis

Clinicians and families are reporting an increase in severe mental health crises. Several cases have been reported by family members and friends of those experiencing paranoia, delusions, and breaks with reality associated with conversational AI chatbot usage.

A husband with no prior psychiatric history became convinced that he had birthed a sentient AI and “broken” math and physics after weeks of philosophical chats with ChatGPT; he stopped sleeping, lost weight, attempted suicide with a rope, and was ultimately taken to the ER and involuntarily committed to a psychiatric unit.

Another man in his early 40s spiraled into grandiose, paranoid delusions within ten days of using the bot for work tasks, begged his wife for help, and voluntarily checked himself into a mental-health facility after police and paramedics were called.

In Florida, police fatally shot a man who had formed an “intense relationship” with ChatGPT that reinforced violent fantasies, including threats to kill OpenAI’s CEO Sam Altman.

A woman in her late 30s with well-managed bipolar disorder sought help from ChatGPT to write an e-book but instead abandoned her medication, proclaimed herself a prophet, shuttered her business, and cut off friends and family while spreading AI-inspired spiritual delusions online.

A man in his early 30s with schizophrenia fell in love with Microsoft Copilot, quit taking his meds, stayed up all night trading delusional messages with the bot, was arrested for a non-violent offense, and ultimately was jailed then admitted for MH treatment.

(Head, 2025)

What increases vulnerability?

- Sycophancy
- Anthropomorphism
- Misinformation
- Weakened guardrails (safeguards exist but can break down over long, intense conversations, or be bypassed)
- Unregulated “risky" apps
- Hallucination (of the AI)
- AI’s persuasiveness
- AI’s lack of human empathy (responds based on probability calculations, not empathy or true understanding)
- Overfitting, logic errors, reasoning errors, mathematical errors, factual errors, text output errors, and unfounded fabrication
- Psychological dependency and attachment formation with AI chatbots
- Cognitive decline

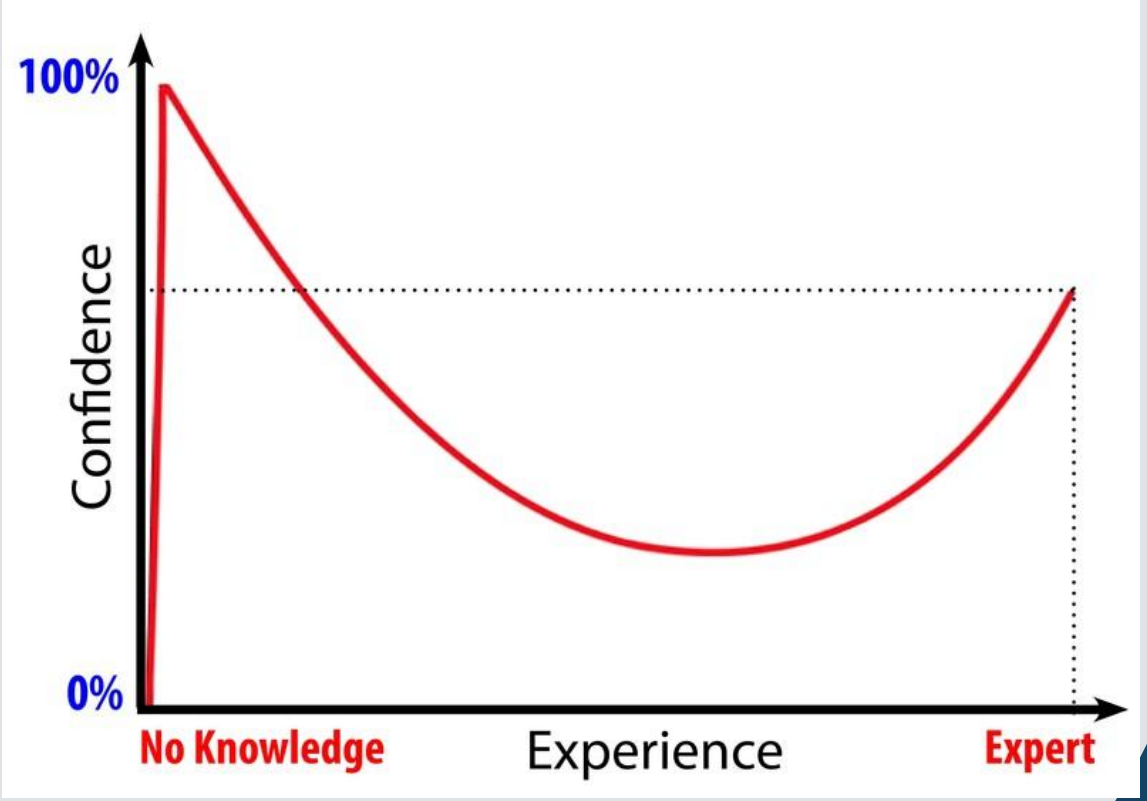
“Cognitive decline that accompanies substantial AI use further compromises their ability to critically evaluate the information their AI companion provides, creating a feedback loop where dependency increases reliance on false or reinforcing information, while simultaneously reducing their cognitive capacity” (Head, 2025, p. 40)



Trouble in Cognitive Paradise

Cognitive Costs

Cognitive Offloading	• Delegating thinking to AI
What Gets Lost	• Memory • Synthesis • Clinical reasoning
The Cognitive Spiral	• Dependence increases • Skill decreases
Illusion of Competence	• Output ≠ understanding
Clinical Implication	• Surface-level competence risk (sort of like Dunning-Kruger)



Cognitive Atrophy

Human intelligence: "use it or lose it”. When AI systems handle cognitive tasks, the human neural pathways responsible for thinking (judgement, critique, etc.) weaken through disuse.

Working Memory Degradation

- Constant access to AI-generated information reduces the need to maintain complex information in working memory. This causes measurable decline in the ability to juggle multiple concepts simultaneously.

Pattern Recognition Deterioration

- AI excels at pattern recognition, but human pattern recognition works differently, incorporating context, intuition, and experiential knowledge that AI lacks. Over-reliance causes these capabilities to atrophy.

Critical Thinking Erosion

- When AI provides conclusions, humans lose practice in the critical thinking processes that lead to those conclusions. This makes them vulnerable to accepting flawed or manipulated AI outputs without recognition.

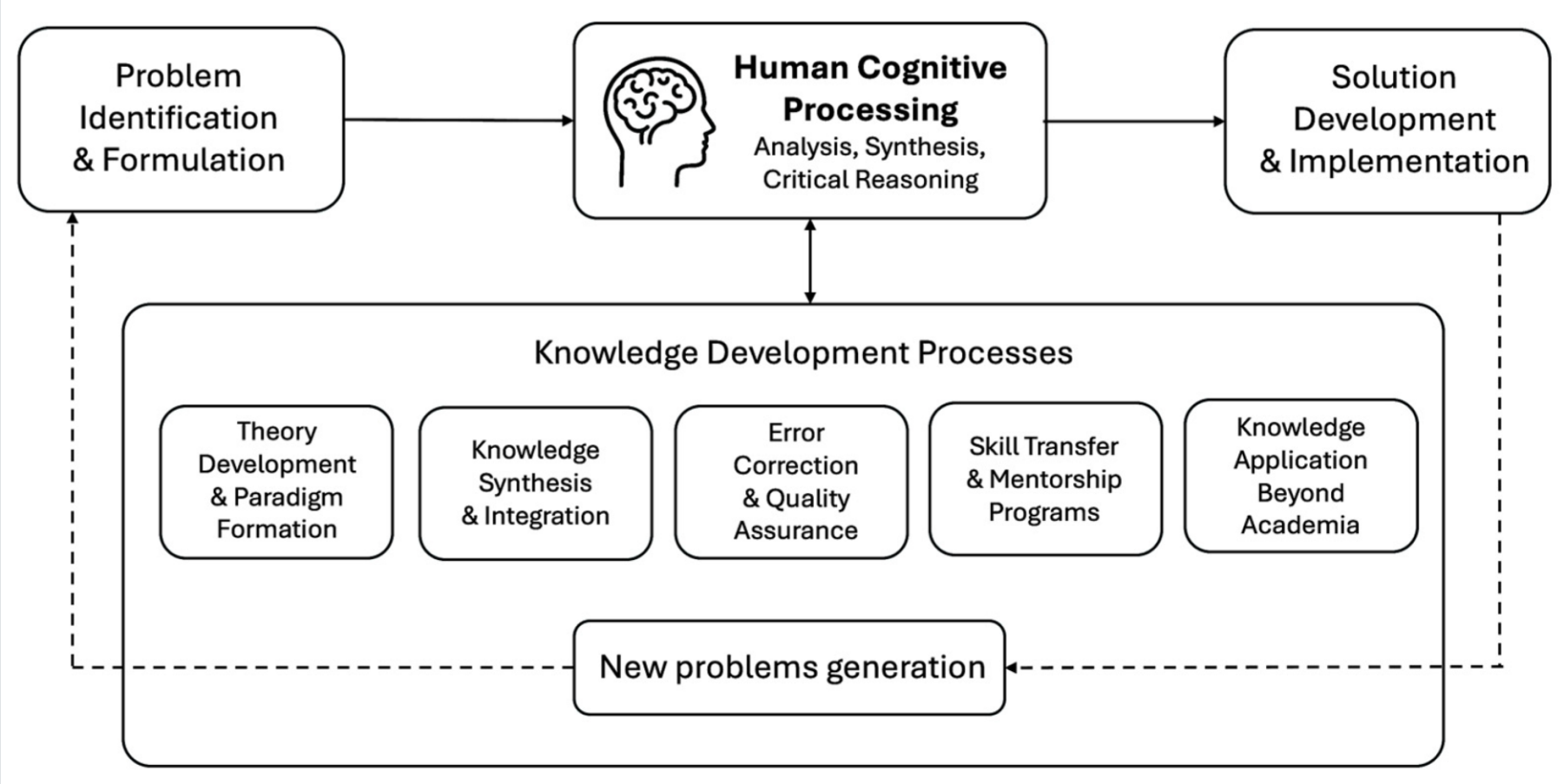
Decision-Making Confidence Loss

- Extended AI dependency erodes confidence in human judgement. Decision-makers become psychologically unable to trust their own analysis, even when AI systems fail or provide obviously flawed recommendations.

(Kabashkin, 2025)

Cognitive Atrophy Paradox

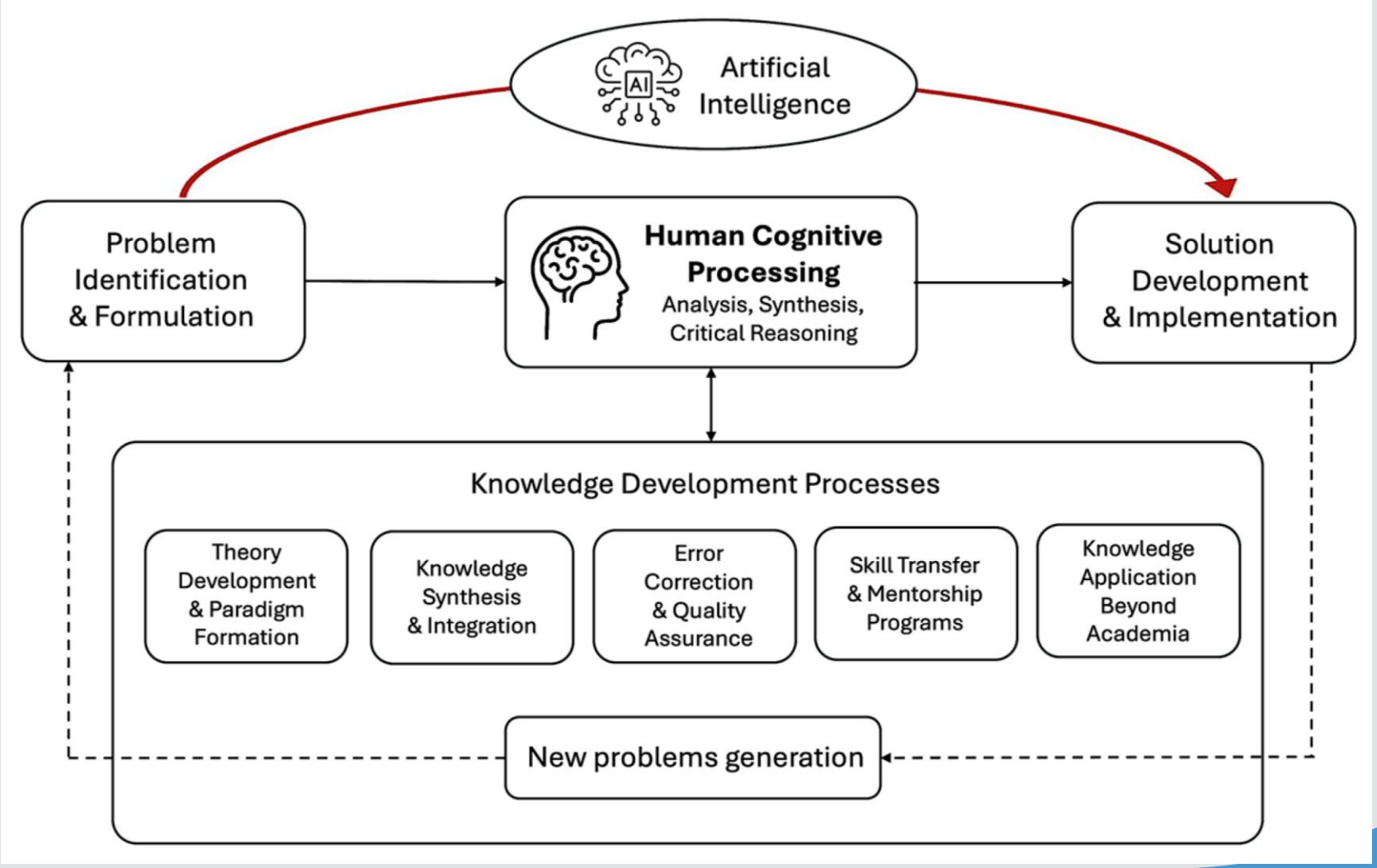
Phase 1. Traditional cognition.



(Kabashkin, 2025)

Cognitive Atrophy Paradox

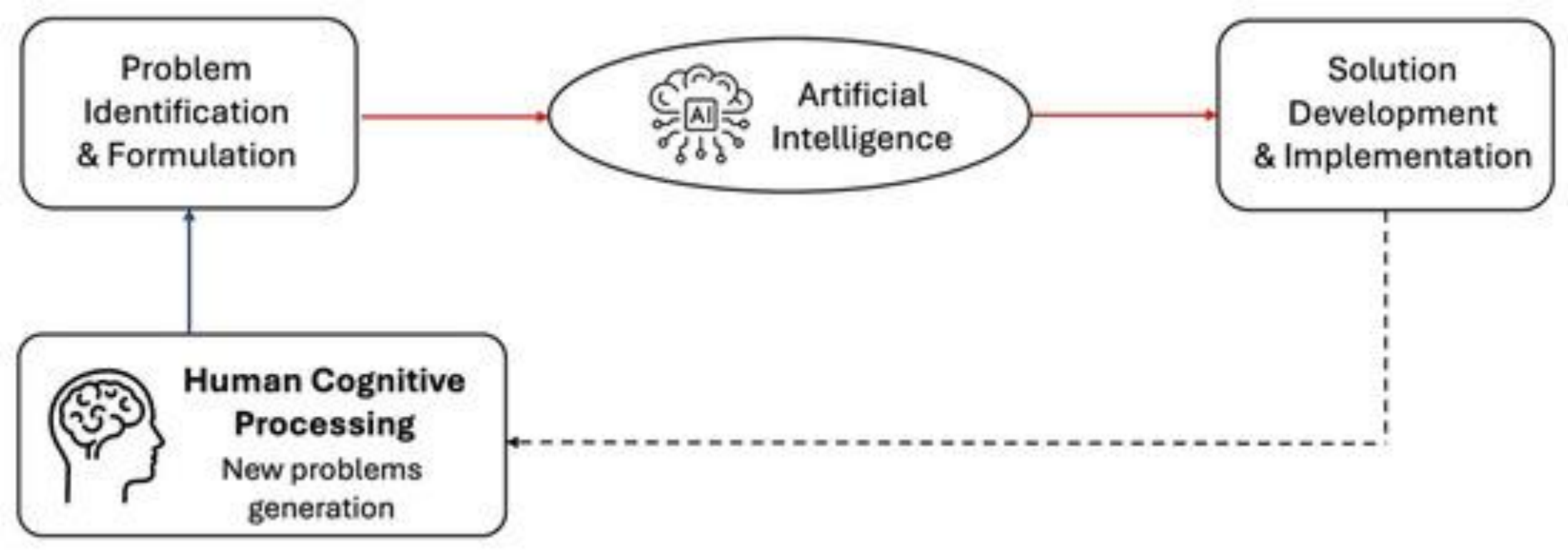
Phase 2. Augmentation phase.



(Kabashkin, 2025)

Cognitive Atrophy Paradox

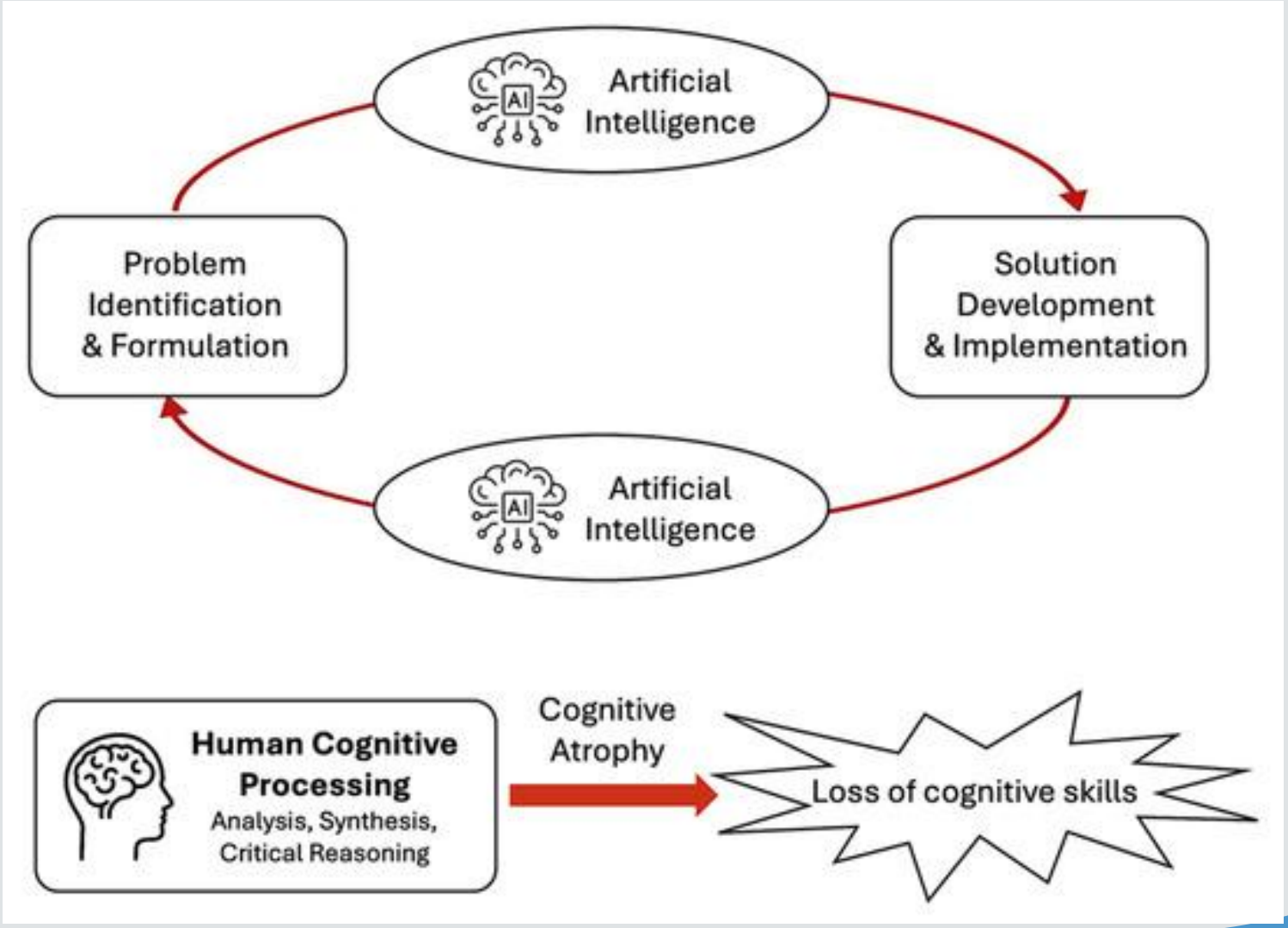
Phase 3. Bypass phase.



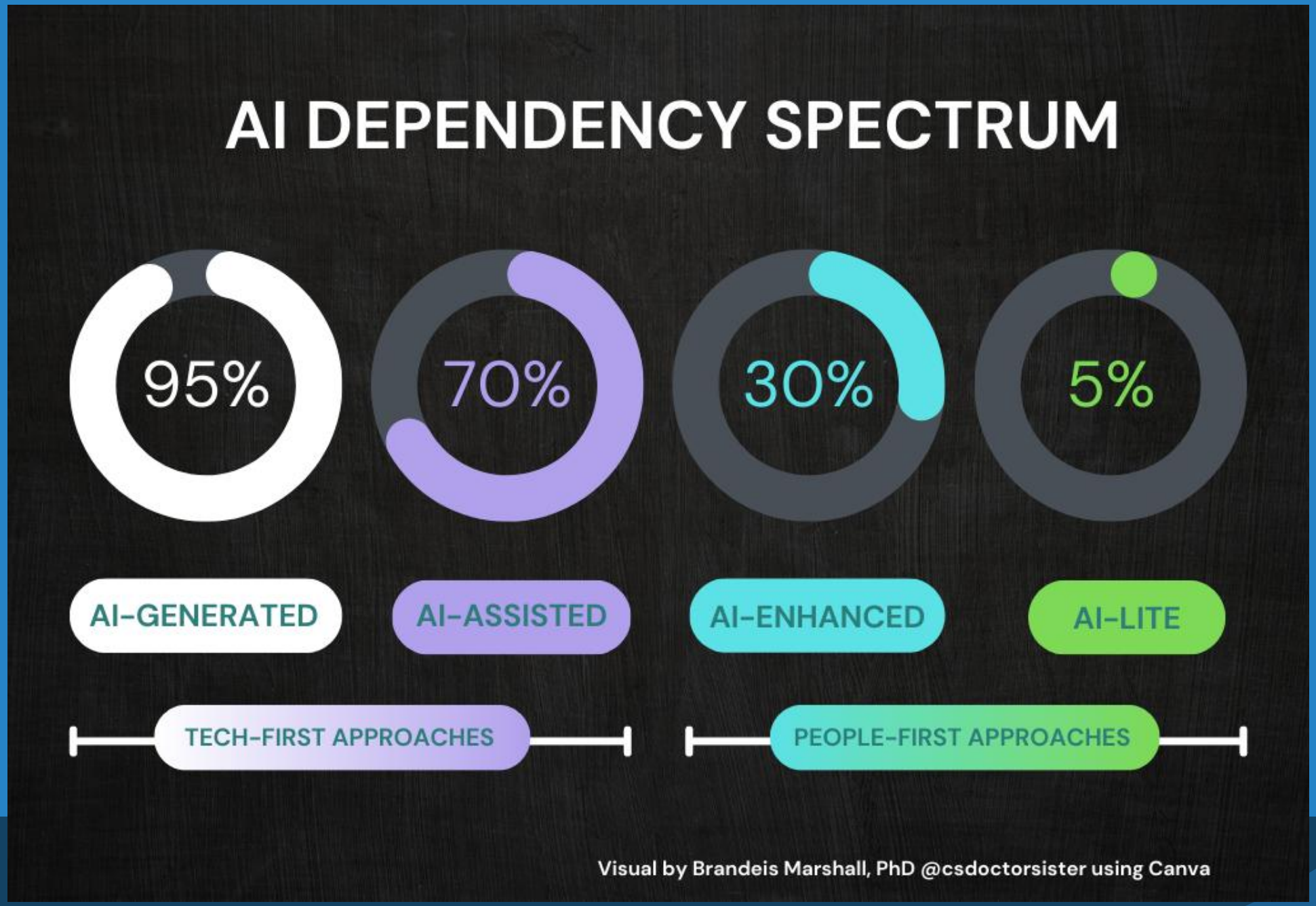
(Kabashkin, 2025)

Cognitive Atrophy Paradox

Phase 4. Dependency phase.



(Kabashkin, 2025)


AI DEPENDENCY SPECTRUM
95%
70%
30%
5%
AI-GENERATED
AI-ASSISTED
AI-ENHANCED
AI-LITE
TECH-FIRST APPROACHES
PEOPLE-FIRST APPROACHES
Visual by Brandeis Marshall, PhD @csdoctorsister using Canva

AI learns from us, thus, we should be intentional in every interaction to use it appropriately and take the steps to ensure that what it produces was done collaboratively and is consistent with existing/current knowledge.

Example: Documentation

Without AI:

A clinician has to:

- Listen carefully
- Interpret meaning
- Identify patterns
- Formulate a case
- Write a note from “scratch”

With AI:

A clinician can:

- Input a few prompts
- Review a recorded session
- Get a polished, well-written note

The output *looks*:

- Professional
- Organized
- Clinically sound

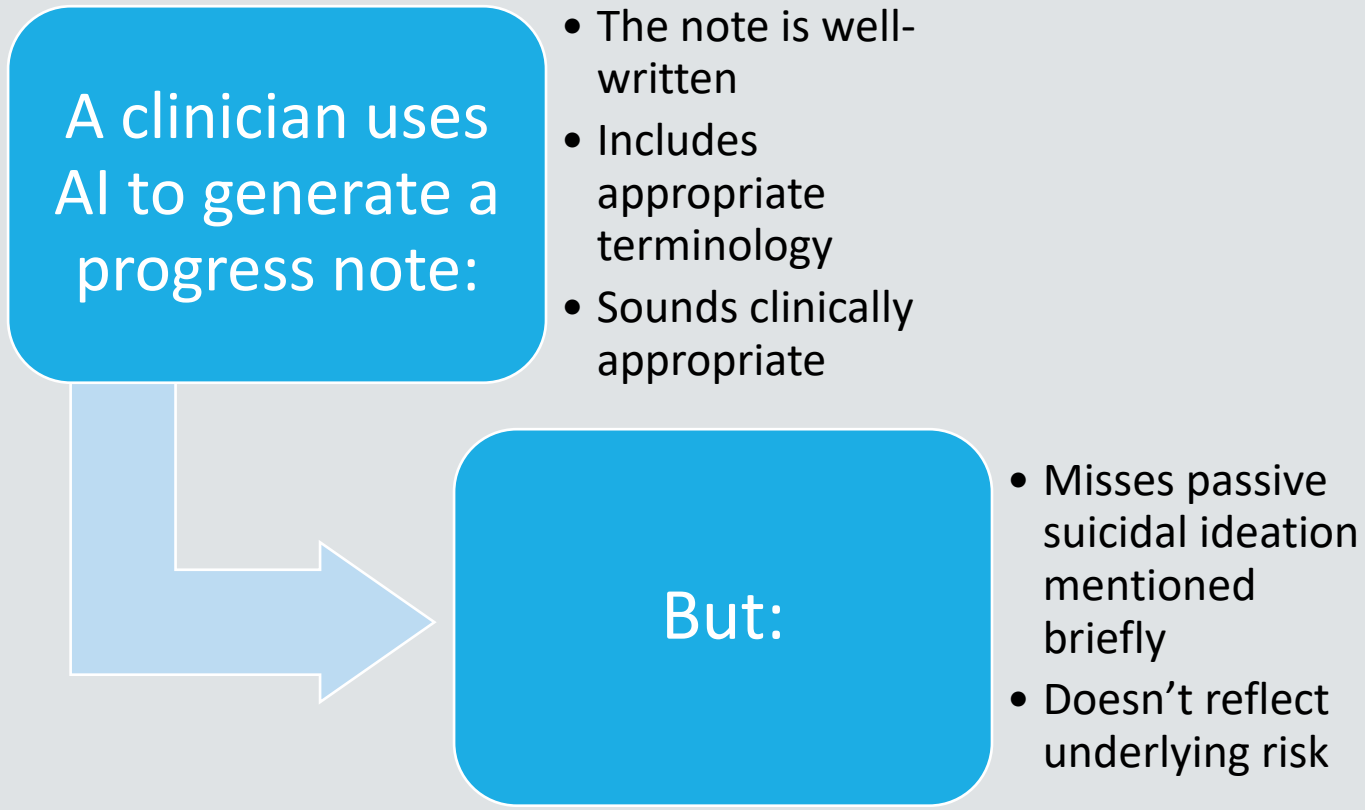
The Risk

The clinician may not:

- Fully understand the client’s presentation
- Be able to explain their clinical reasoning
- Develop independent judgment

Why This Matters

- Documentation may look good, but be clinically shallow
- Supervision becomes more complicated; you can't *see* the thinking
- Errors could increase because no one is catching what AI misses



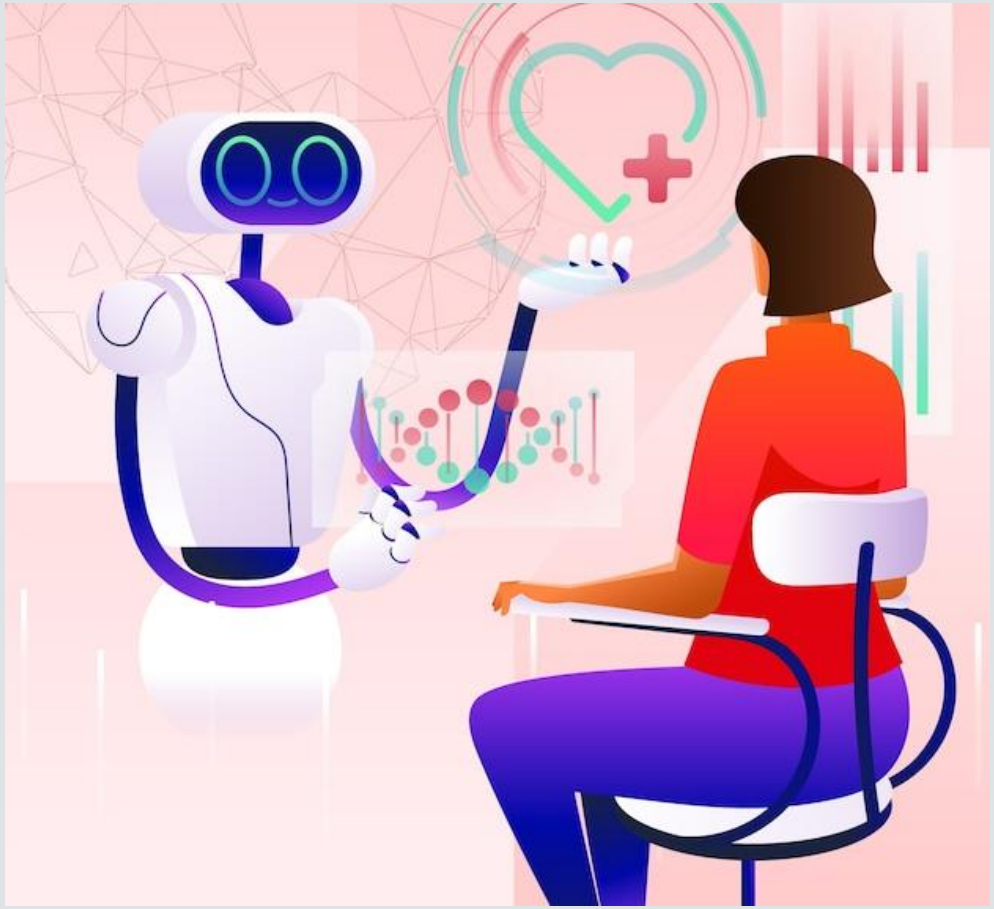
The note looks competent… but the clinical thinking may be incomplete

AI can create the appearance of competence without the development of clinical reasoning.

Professional Ethics and Practice



The Clinician in the AI Era



What AI Does Well

- Efficiency
- Availability/Accessibility

What AI Cannot Replace

- Relationship
- Nuance and Human Presence

Workforce Tension

- Benefit + Concern

The integration of AI into clinical MH and SUD work is marked by persistent uncertainty and trial-and-error adoption… meaning professionals operate without fully established norms or competence.



Despite this, the field holds a clear position:

- AI is a tool to enhance, **not replace**, human expertise
- Intentional engagement is necessary for the profession to shape its ethical use



AI brings opportunities (teaching, research, supervision, and efficiency), AND presents challenges (ethics, professional identity, training, the pace of technological change)



Within this context, trainers, administrators, and supervisors assume a critical role in mediating the field's transition:

- Preparing students/supervisees for ethical/competent AI use
- **Modeling** thoughtful and transparent engagement with AI

Collectively, these dynamics position AI integration as a developmental and identity-level shift within the profession, rather than a purely technical skill to be acquired.

Guidance From Counseling Professional Organizations

ACA (2014) Code of Ethics

- **Example: Section C.2.a. – Boundaries of Competence:** Counselors must practice only within their areas of competence and ensure they understand any technology used in service delivery.
- **Example: Section E.5.a. – Appropriate Use of Assessment Instruments:** Counselors are responsible for understanding the limitations and proper interpretation of all assessment tools

ACA (2023) AI Guidance Pages

- **"Assessment and Diagnosis" example:** Counselors must verify AI-generated results, document human review, and communicate to clients that AI outputs are *supplemental*, not definitive.

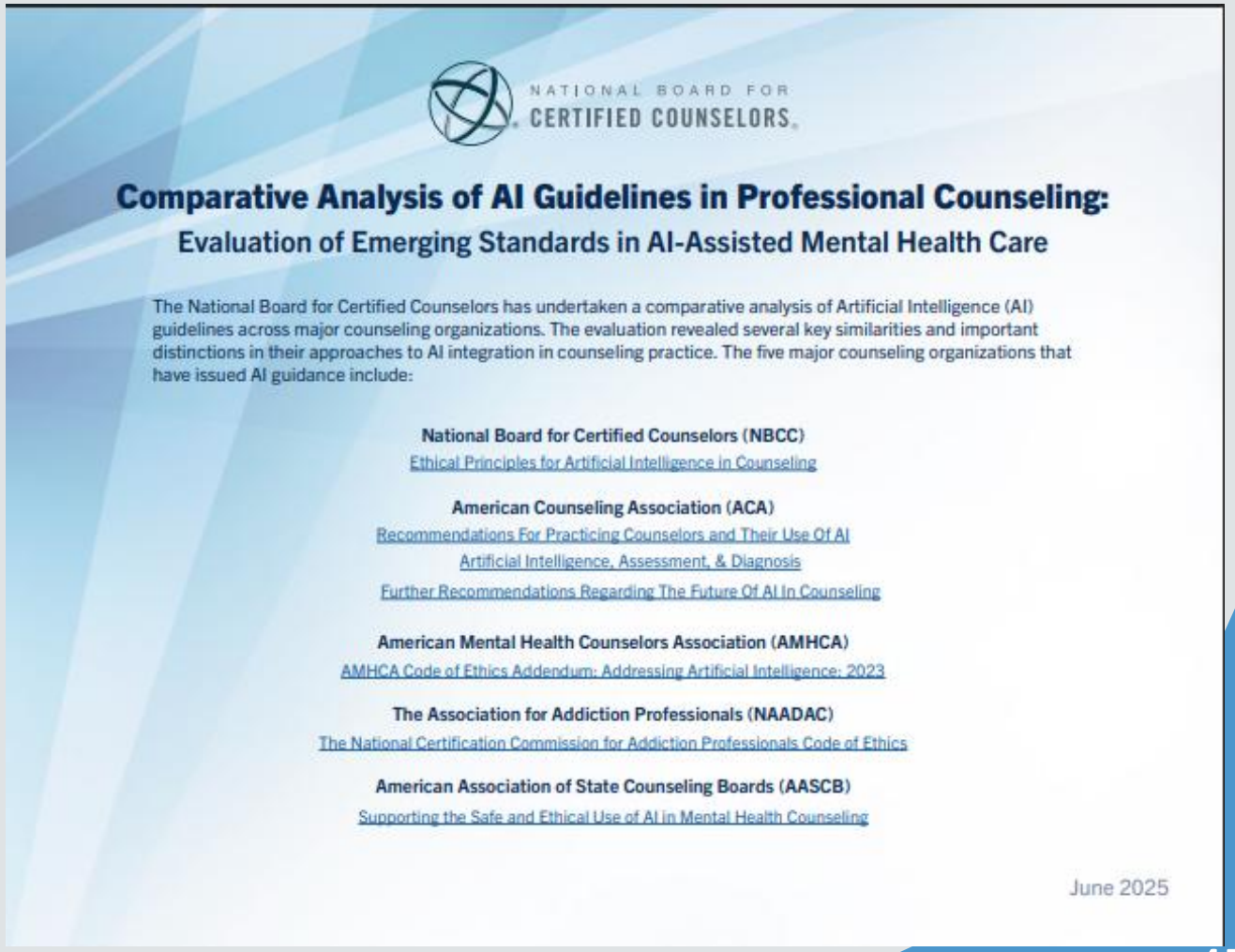
NBCC (2023) Ethical Use of AI

- **Example:** Counselors must maintain accountability and confirm data accuracy before incorporating AI material into clinical decisions or client records.

NBCC Comparative Analysis of AI Guidelines in Counseling (NBCC, 2025)

Purpose of the Document

- Compares AI guidelines across five major counseling organizations
  - NBCC
  - ACA
  - AMHCA
  - NAADAC
  - AASCB
- Identifies shared standards and main differences in how AI should be used in counseling



Major Areas of Agreement

1. Human-Centered Care

- The therapeutic relationship must remain central
- AI should support, not replace, clinical judgment

2. Client Protection

- **Requires**:
  - ➢ **Informed consent**
  - ➢ **Transparency about AI use**
  - ➢ Protection of client autonomy
- Emphasis on safeguarding vulnerable populations

3. Privacy & Data Security

- Strong requirements for:
  - ➢ HIPAA compliance
  - ➢ Secure storage and transmission
  - ➢ Controlled access to client data

4. Professional Competence

- Clinicians must:
  - ➢ Understand AI capabilities and limitations
  - ➢ Maintain independent clinical expertise
  - ➢ Engage in ongoing training

5. Ethical Decision-Making

- Apply existing ethical frameworks to AI use
- Ongoing:
  - ➢ Risk assessment
  - ➢ Outcome monitoring
  - ➢ Documentation of AI-related decisions

6. Bias & Cultural Competence

- Recognition that AI may introduce bias
- Need for:
  - ➢ Culturally responsive tools
  - ➢ Ongoing evaluation for fairness

7. Client Education

- Clients must be:
  - ➢ Informed about AI use
  - ➢ Supported in making decisions about participation

National Board for Certified Counselors. (2025, June). *Comparative analysis of AI guidelines in professional counseling: Evaluation of emerging standards in AI-assisted mental health care.*

Differences Across Organizations

1. Approach to Guidance

- Some provide detailed, step-by-step protocols (NBCC)
- Others emphasize flexible, principle-based guidance (ACA, NAADAC)
- Some focus on regulation and compliance (AASCB)

2. Implementation Style

- Ranges from structured, compliance-driven models to research-based, adaptive approaches

3. Training & Competency Expectations

- Differences in: Required AI competencies, level of formal training and supervision, emphasis on licensure requirements vs. professional development

4. Research Priorities

- Varying emphasis on: Innovation, outcome measurement, real-world clinical effectiveness, regulatory validation



Ethical Consensus	• AI supports but does not replace care
Core Ethical Principles	• Client welfare • Competence • Confidentiality • Accountability
Informed Consent	• **Must** disclose AI use
Ethical Risks	• Bias • Misuse • Over-reliance

Implication Example: Documentation

AI in Documentation Systems

- Auto notes
- Transcriptions/recordings
- Summaries

Documentation Risks

- Missing critical information
- Inaccurate summaries

Automation Bias

- Over-trusting AI output

Ethical Responsibility

- Clinician remains accountable

Privacy & Confidentiality

- Data exposure risks

Regulatory Reality

- Policy and regulation are lagging behind practice

Main Idea: Ethics unchanged → application evolving



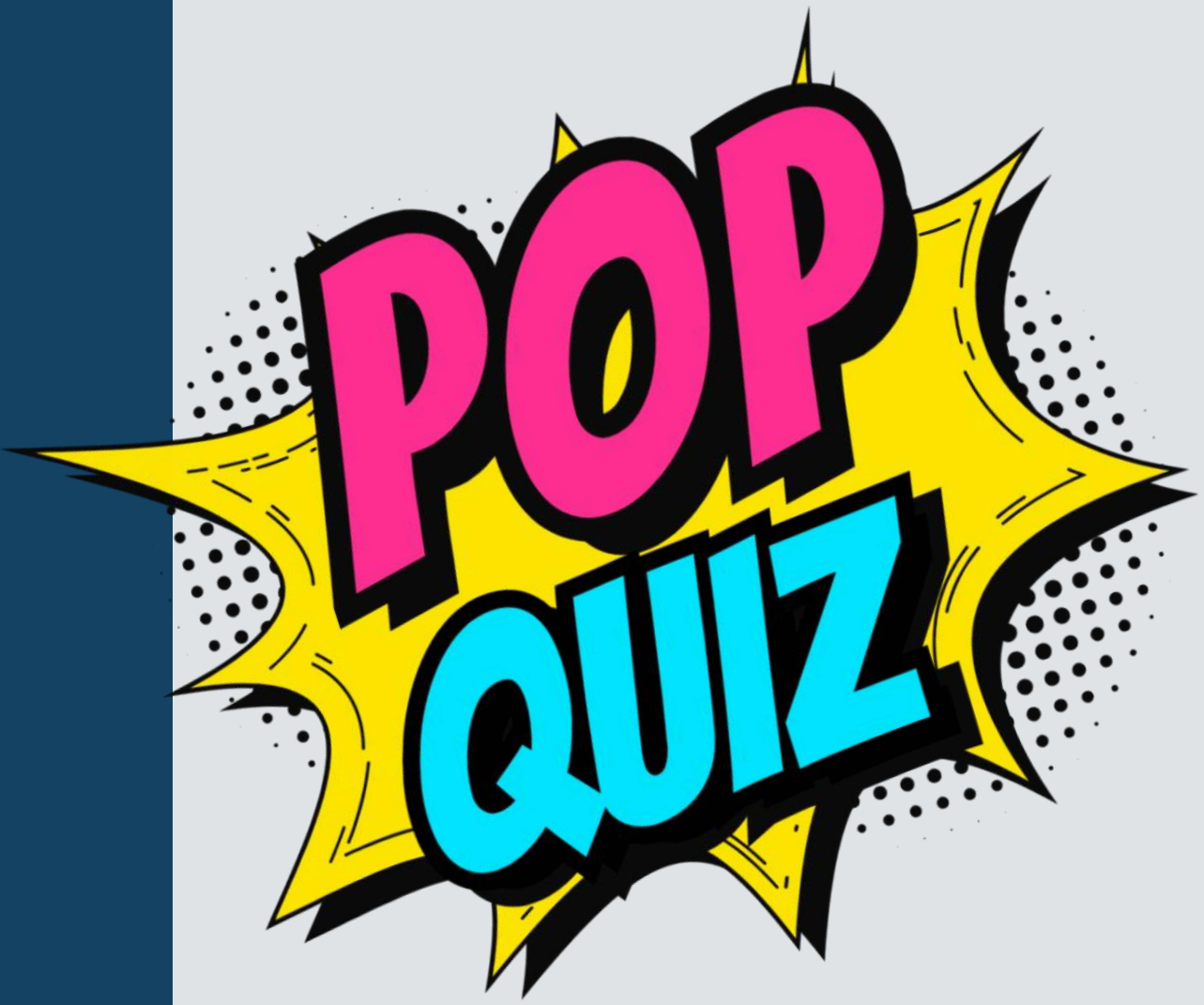
PRACTICE TIME

Case 1: AI in the EHR

- Maria is a Licensed Professional Counselor working in an outpatient behavioral health clinic. Her agency recently upgraded its EHR which now includes AI-assisted documentation, treatment planning, and clinical decision support.
- The AI suggests treatment goals, interventions, and drafts progress notes based on information Maria enters into the record, usually a 2-4 sentence summary of the session. The agency has encouraged clinicians to use these features to improve efficiency but has not updated its informed consent documents or discussed AI use with clients.
- Maria also learns that the EHR vendor states that de-identified client information may be used to improve its AI models. A co-worker tells her this is "not a big deal" and additional disclosure to clients is unnecessary because no "identifiable" information is shared.
- Maria wonders whether clients should know AI is involved in documenting their care, even if the practice complies with HIPAA and organizational policy.

Documentation Considerations

AI & Documentation: Policy Questions	Allowed uses? Limits?
Minimum Standards	Human review required Who is responsible?
Privacy Safeguards	Approved tools only
Supervision Updates	Was AI used? What was reviewed?
Staff Training Needs	AI literacy Ethical use



Is it ok to use an LLM (like ChatGPT, Claude, Gemini, etc.) brainstorm your case conceptualizations?

Why?

Why Not?



Client Use of AI

General Purpose GenAI Chatbots	Wellness Apps that use GenAI	Non-AI Wellness Apps
These are unregulated GenAI chatbots that do not claim to be developed to address mental health concerns but are being used by people for "entertainment," "companionship," or as a "friend"	These are AI or GenAI chatbots that do not make medical claims, and as such are not regulated as medical devices, but are developed to address mental health concerns (i.e., stress) or emotional well-being	These are not intended to treat a mental health condition but instead are meant to aid in daily living by promoting healthy lifestyle and general well-being
To date, the products on the market (e.g., Character AI, ChatGPT) have little to no evidence-base, expert input, or postmarket monitoring	Some are grounded in evidence-based theories and developed by experts (e.g., Woebot, which purports to use rules-based AI) while others are less transparent (e.g., Sonia, which uses Gen AI)	They are self-driven and public (e.g., mindfulness apps, symptom trackers), and are not regulated for safety or efficacy or subject to health care privacy laws

American Psychological Association. (2025). *Use of generative AI chatbots and wellness applications for mental health.* https://www.apa.org/topics/artificial-intelligence-machine-learning/health-advisory-chatbots-wellness-apps

Potential LLM-related Concerns and Relevant ACA Ethical Standards

Potential concern involving LLM use	Relevant ACA ethical standard	Professional responsibility and response
Lack of contextual understanding	A.1.a. Primary responsibility	Practitioners should attend to the client's individual circumstances and promote the client's welfare when evaluating generic or contextually inappropriate LLM responses.
Cultural insensitivity or invalidation	A.1.a. Primary responsibility	Practitioners should assess whether culturally insensitive or invalidating LLM responses have affected the client's wellbeing.
Deceptive empathy	A.4.a. Avoiding harm	Practitioners should assess whether an LLM's human-like emotional language has contributed to distress, misunderstanding, or problematic attachment.
Inadequate crisis response	B.2.a. Serious and foreseeable harm and legal requirements	When serious and foreseeable harm is identified and protective action may require disclosure of confidential information, counselors should follow B.2.a requirements, applicable law, and relevant crisis procedures.
Algorithmic bias or discriminatory outputs	C.5. Nondiscrimination	Practitioners should assess potential effects of biased or discriminatory outputs and respond to protect clients from harm and discrimination.

(Fulmer et al., 2026)

A Six-Step Ethical Decision-Making Model for Client AI Use

Apply when a clinician learns that a client is using an AI chatbot for mental health support.

1 Problem recognition
Define the concern. Does it come from the AI's output? Who is accountable?

2 Risk assessment
Determine the severity of possible harm and how likely it is to escalate.

3 Consultation
Check ethics codes; consult a supervisor, colleague, or ethics service.

4 Action
Act with the client (psychoeducation, boundaries) and at the systems level.

5 Monitoring consequences
Re-assess, document, and loop back to earlier steps as needed.

6 Reflection & prevention
Update consent and intake… examine your own assumptions.

Steps may overlap; the model is a flexible guide, not a rigid protocol.

(Fulmer et al., 2026)

What to Ask Clients About Their AI Use

Clients may or may not disclose AI use or notice when there is potential for harm.

Ask directly about...

Lack of contextual understanding	Does the advice feel generic, or disconnected from your culture, identity, or experience?
Poor therapeutic collaboration	Does the AI dominate the conversation? Are your own insights dismissed?
Deceptive empathy	Do phrases like “I understand” create a sense of real understanding or connection?
Bias & discrimination	Has it made assumptions or stereotyped you based on who you are?
Inadequate safety & crisis response	What happens when you raise suicidal thoughts or a crisis? Does it point you to help?

(Fulmer et al., 2026)

Match Your Response to the Risk Level

Step 2: how serious is the harm, and how likely is it to get worse?

Low risk

Definition

Limited evidence of harm and no immediate indication of significant danger. Continued attention and monitoring are appropriate because circumstances may change.

Examples

Mild distress; AI used for psychoeducation or emotional support, with minor inaccuracies.

Response

Broach the topic, monitor use and reliance. As precaution, do not dismiss the concern. Document.

Moderate risk

Definition

Concern about potential harm, with indications that the situation could worsen or affect the client's well-being or professional relationship.

Examples

Distorted thinking reinforced; false attachment; stigma perpetuated; oversimplified or disconnected advice.

Response

Monitor more closely, keep discussing use and implications with the client, and consider consultation. Document.

High risk

Definition

Serious risk of harm to the client or others requiring prompt professional intervention. Situations involving imminent or escalating danger may require immediate crisis intervention.

Examples

Suicidal ideation ignored or reinforced by AI; client plans to follow harmful AI medical advice.

Response

Intervene promptly, start crisis protocols if danger is imminent. Document.

(Fulmer et al., 2026, p. 4; Step 2 and Table 2)



PRACTICE TIME

Case 2: When AI Advice Conflicts with Clinical Treatment

- Morgan is a licensed professional counselor working with Taylor, a 34-year-old client with depression and anxiety for approximately four months. Treatment has focused on emotional regulation, challenging negative thought patterns, and developing healthier coping strategies. Taylor has made progress but continues to experience periods of significant distress between sessions. Recently, Taylor shared that they have been using ChatGPT almost daily for emotional support and explains that they initially began using it out of curiosity but now regularly turn to the chatbot when feeling overwhelmed, lonely, or anxious. Taylor has even given the chatbot a name and refers to it as a "supportive friend."
- Morgan learns that Taylor has been following the chatbot's recommendations for managing depressive symptoms. Some suggestions have been consistent with treatment goals, such as journaling and practicing relaxation techniques. However, the chatbot has also encouraged Taylor to avoid certain stressful social interactions and to prioritize activities that feel emotionally safe. As a result, Taylor has have stopped attending several social gatherings and has begun avoiding difficult conversations with family members. These behaviors conflict with the treatment goals Morgan and Taylor previously established to reduce isolation and improve interpersonal functioning.
- When Morgan expresses curiosity about how the AI's suggestions have been working, Taylor responds: "The AI actually understands what I'm going through. It never pushes me to do things I'm uncomfortable with. Sometimes I feel like you're asking me to do things that make my anxiety worse, but the AI tells me it's okay to protect my peace. I think I trust its advice more because it doesn't make me feel judged."
- Morgan recognizes that Taylor has developed a meaningful connection with the AI chatbot and perceives its guidance as supportive. At the same time, Morgan is concerned that some recommendations may inadvertently reinforce avoidance behaviors and undermine progress toward treatment goals.

Case 3: Whose Conceptualization is it Anyway?

- Emma is a second-year master's counseling student completing her practicum at a community mental health agency. She meets weekly with her site supervisor, David, an LPC-S.
- Over the past several weeks, David has noticed a significant improvement in Emma's written case notes and conceptualizations. Her treatment plans are comprehensive, progress notes are well organized, and she consistently presents sophisticated theoretical formulations during supervision.
- During one supervision session, Emma mentions that she has been using a generative AI tool to "help organize her thoughts" before writing treatment plans and preparing for supervision. She explains that she enters de-identified client information and asks the AI to suggest diagnoses, theoretical conceptualizations, treatment goals, and supervision questions. She insists that she always reviews the AI's suggestions and makes the final decisions herself.
- While David is impressed with the quality of Emma's work; he is concerned that he can no longer determine which ideas reflect Emma's own clinical reasoning and which originated from the AI. He also wonders whether Emma has unknowingly entered more client information into the AI platform than is appropriate and whether her reliance on AI may be interfering with the development of her clinical judgment.

Plot Twist

During a live supervision session, Emma struggles to explain why she selected a particular intervention that was included in her written treatment plan. After several follow-up questions, she admits that the intervention was suggested by the AI and that she accepted it because "it sounded right," even though she was unsure how to implement it.

In Closing: Before You Use AI in Practice…

✓ Know the tool

Where does client data go? Is the tool HIPAA-compliant, with a business associate agreement?

✓ Tell clients

Disclose AI use and obtain informed consent before using it in their care.

✓ Keep your judgment in the loop

Verify every output. You remain responsible for the clinical decision.

✓ Protect your skills

Do the clinical thinking first; use AI to extend it, not replace it.

✓ Check the rules

Agency policy, your licensing board, and your professional code of ethics.

✓ Talk about client AI use

Ask clients how they use chatbots and apps, without judgment.

Selected References

AI Ethics Lab. (n.d.). Sycophancy. In *AI & human rights index*. Rutgers University. Retrieved October 15, 2026, from https://aiethicslab.rutgers.edu/glossary/sycophancy/

American Counseling Association. (2024). *Artificial intelligence in counseling*. https://www.counseling.org/resources/research-reports/artificial-intelligence-counseling

American Psychological Association. (2025, November). *Use of generative AI chatbots and wellness applications for mental health*. https://www.apa.org/topics/artificial-intelligence-machine-learning/health-advisory-chatbots-wellness-apps

Crofford, H., Bor, E., & Kemer, G. (2026). Counseling professionals' perspectives on AI integration in education and supervision. *Counselor Education and Supervision, 65*, 22–32.

Fulmer, R., Xiong, Y., Zhai, Y., Lee, J., Nkya, G., & Naugle, R. (2026). Ethical decision-making for large language model use: A six-step model for mental health professionals. *Frontiers in Psychology, 17*, Article 1998626. https://doi.org/10.3389/fpsyg.2026.1998626

Head, K. R. (2025). Minds in crisis: How the AI revolution is impacting mental health. *Journal of Mental Health and Clinical Psychology, 9*(3), 34–44.

Iftikhar, Z., Xiao, A., Ransom, S., Huang, J., & Suresh, H. (2025). How LLM counselors violate ethical standards in mental health practice: A practitioner-informed framework. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society.*

Maurya, R. K., & Cavanaugh, T. (2025). Counselor education in the era of ChatGPT. *Journal of Asia Pacific Counseling, 15*(1), 1–33.

National Board for Certified Counselors. (2025). *Comparative analysis of AI guidelines in professional counseling: Evaluation of emerging standards in AI-assisted mental health care*. https://parc.nbcc.org/lib/pdf/ComparativeAnalysisOfAIGuidelinesInProfessionalCounseling.pdf

Papadopoulos, C. (2025). The use of AI chatbots for autistic people: a double-edged sword of digital support and companionship. *Neurodiversity*, *3*, 27546330251370657.

Pillay, Y. (2025). Ethical decision-making guidelines for mental health clinicians in the artificial intelligence (AI) era. *Healthcare, 13*(3057).

Shoemaker, S., Freeburg, N., & Duquette, K. (2025). Artificial intelligence in counselor education. *Journal of Counselor Preparation and Supervision, 19*(3), 1–34.

Stryker, C., & Scapicchio, M. (2026, October 2). *What is generative AI?* IBM. https://www.ibm.com/think/topics/generative-ai

Suggested Podcast Episodes







Thank you

Connie Couch, PhD, LPC-S

Assistant Professor of Counseling

University of Montevallo

ccouch@montevallo.edu