Lobby

Which of the following ethical concerns are currently on your mind as you consider using AI in your prevention work? (Select all that apply):

- **Data privacy:** Worries about how community or participant data is stored and used by AI companies
- **AI limitations:** Concern that AI might produce inaccurate or inappropriate content about my community
- **Loss of the ”human touch:"** Fear that automated content will feel inauthentic or damage community trust
- **Misinformation:** Concern that AI might "hallucinate" facts or provide incorrect health guidance
- **Digital divide:** Worry that AI tools will only benefit well-resourced organizations
- Other (please note in the chat)

PTTC HHS REGION 9
Pacific Southwest

Navigating AI in Prevention: A Practical Guide for Professionals

Part 2: Navigating AI Ethics: Decision-Making and Community Guardrails

Rebecca L. Cooney

Disclaimer

PTTC HHS REGION 9 Pacific Southwest

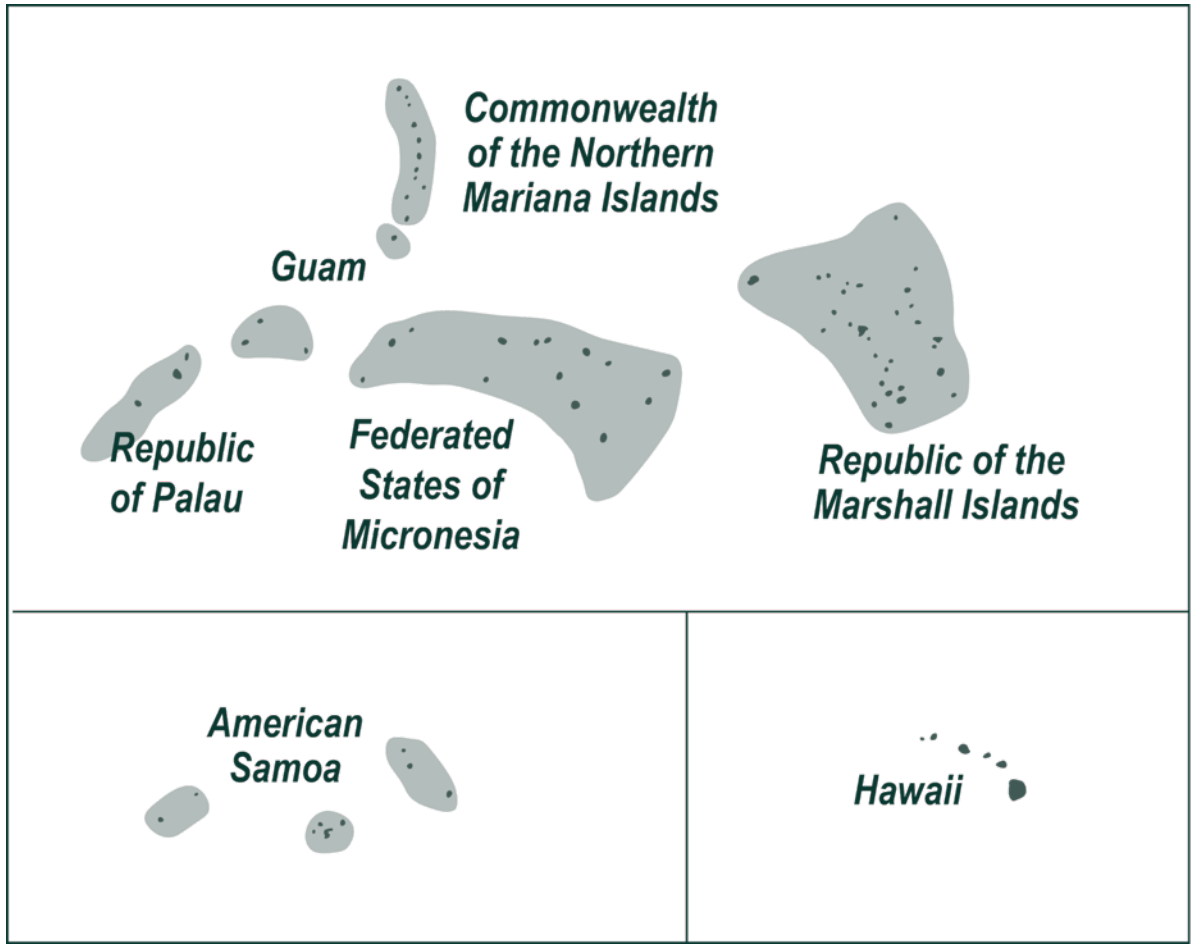
This presentation is supported by SAMHSA of the U.S. Department of Health and Human Services (HHS) through SAMHSA Cooperative Agreement #5H79SP084319-02. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by SAMHSA/HHS, or the U.S. Government.

Any mention of specific brand or company names is for illustrative purposes only and does not constitute an endorsement or recommendation by the funding agency or the PTTCs.

This webinar is being recorded and archived and will be available for viewing after the webinar. Please contact the webinar facilitator if you have any concerns or questions.

Pacific Southwest

University of Nevada, Reno
CASAT
School of Public Health



Housekeeping

- Chat feature
- Q&A feature
- Webinar recording and materials
- Certificates of attendance

Rebecca L. Cooney

Educator

Leader integrated marketing communication strategy

Specialist in multimodal curriculum development

Session 1 Recap: Navigating AI in Prevention

- Session 1 defined **artificial intelligence (AI) as a co-pilot**, not an auto-pilot, built for augmentation rather than professional replacement
- **Core Toolkit Applications:**
  - Distilling technical documents, such as 50+ page needs assessments, into actionable 1-page summaries
  - Generating campaign slogans, initial logic model drafts, and strategic agendas for community coalitions
  - Scaling a single prevention message into variations for multiple platforms like Instagram, Facebook, and local outreach
- **The Human-AI Workflow:**
  - **The "Heavy Lift":** AI manages the initial technical tasks of drafting, data synthesis, and design resizing
  - **The "Final Polish":** Prevention professionals apply a strategic communication filter to ensure every output remains stigma-free and culturally accurate

Today's Focus:

While Session 1 focused on **speed and scale**, Session 2 focuses on **safety and integrity -** establishing the guardrails necessary to ensure community safety and ethical application of the use of AI

Learning Outcomes

- Recognize key ethical considerations before integrating AI into prevention practice, with a focus on privacy, compliance, and digital dignity
- Analyze the risks of inaccuracies in AI when tailored to rural and priority service areas
- Apply a decision-making framework to determine the appropriateness of AI tools for specific community-led prevention tasks
- Describe practical next steps for piloting and evaluating AI applications within their own organizations and local programs

The Inherent Worth of Digital Dignity

- Human dignity is a sacred foundation of ethical AI
- AI prompts should be designed to show respect toward the people of the community you serve
- AI applications must never objectify or subordinate local communities
- Use of AI should provide new means to advocate for and defend community rights, never to infringe upon them

Privacy: Beyond Legal Compliance

Privacy is essential to protecting human autonomy and individual agency

Data protection must be ensured throughout the entire lifecycle: collection, sharing, archiving, and deletion

Sensitive data requires valid legal bases and informed consent - especially when dealing with health or substance use trends

Accountability must be integrated into the initial system design to protect community information

The Ethical AI Life Cycle

1

Planning & Intent (Research & Design):

Set a clear, evidence-based goal for your program and ensuring the AI's purpose is both ethical and necessary for your community

2

Building the Safety Net (Development):

Construct the AI tool with privacy-first settings and verify that outputs accurately reflect the communities you serve

3

Community Launch (Deployment & Use):

Let the community know AI is in the mix; providing clear, "plain-language" explanations of how the tool works and what it does with data

4

Health & Wellness Check (Maintenance & Operation):

Regularly check in on the AI tool once it's live to make sure it is helping people and not causing accidental social or mental harm

5

The Safety Audit (Monitoring & Evaluation):

Review AI outputs to catch safety risks or errors that only show up after the tool has been out in the world for a while

6

Sunsetting (Close-Outs):

Know when to retire an AI tool because it has met its goal or is no longer accurately reflecting the communities you serve

Data Quality and Accuracy of Outputs

- AI algorithms can reproduce errors or inappropriate patterns from their training data, affecting reliability
- Poor quality training data can lead to inaccurate outputs about substance use and rural populations
- Inappropriate automated recommendations can undermine community trust in organizations
- Prevention professionals have an ethical responsibility to review and verify AI outputs for accuracy and appropriateness

Discussion

Think of a common **cliché or stereotype** people have about the community you serve. If you asked AI to 'write a social media caption' or 'create a flyer' about your town, how might it accidentally rely on those old clichés and get the story wrong?

Note answers in the chat

Case Study #1: Issues with Data Synthesis

The Scenario: An agency uses AI to scan statewide survey data to automatically predict which neighborhoods should receive priority for school-based prevention funding

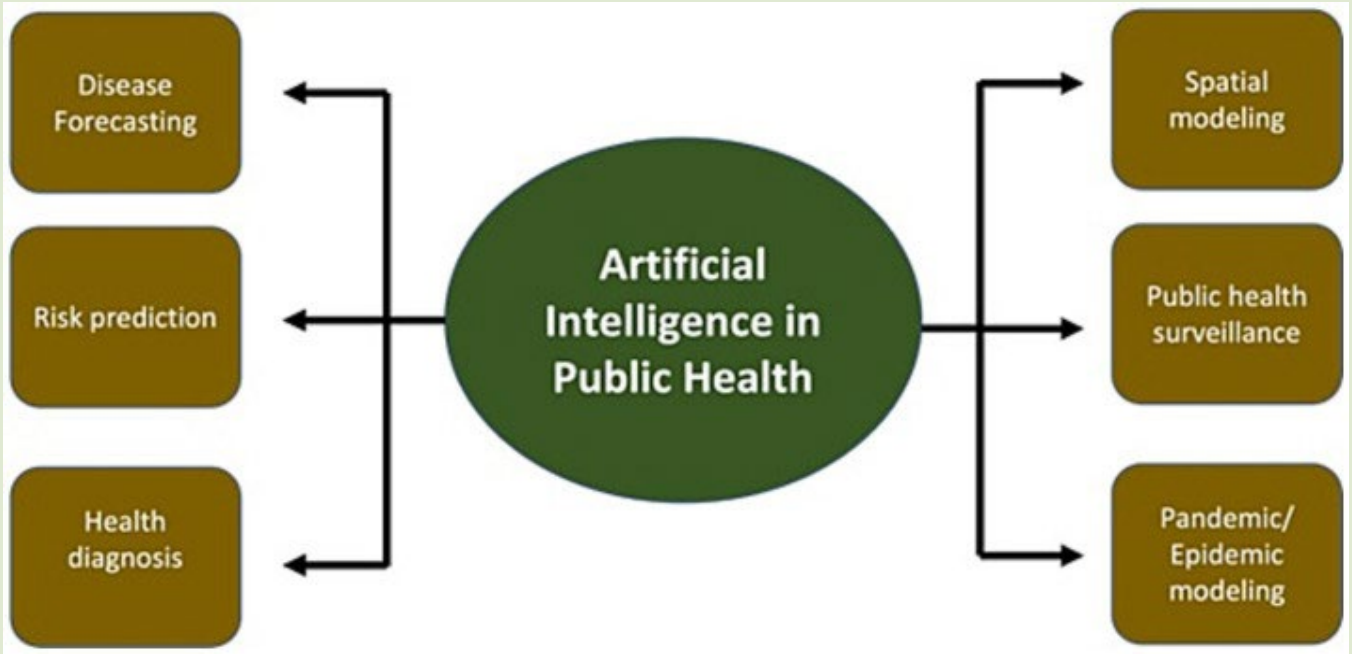
The Issue: The AI over-targets specific zip codes based on outdated historical data rather than current health needs

The Ethical Check: Applying the principle of fairness - is the algorithm producing appropriate and reliable results?

The Professional Fix: The specialist resets the prompt parameters to focus on current service gaps and community-identified needs, manually weighting the results to ensure balance and consistency

Regional Balance and Fair Use

- AI tools should celebrate the community and avoid one-size-fits-all" translations or generic messaging
- Providing the same quality of AI access and participation for both rural and urban sectors is essential for fair service delivery
- Ensuring every staff member understands AI is a matter of fairness and professional growth
- If a tool produces an unfair or unrepresentative result, there must be a clear process to catch it, fix it, and protect the community



“Predictive models can find patterns and trends that can guide public health measures by examining data on previous epidemics and other pertinent elements, such as population and weather patterns.” – Olawade, D.B. et al

Transparency and The Right to Know

- Transparency is essential to ensure AI works for the common good and prevents harm
- Making the functioning of algorithms intelligible and traceable so stakeholders understand the *why* behind an output
- People have the right to be informed when a decision that affects them is made by or informed by AI
- Greater openness allows the community to check the work and decreases the risk of errors or unfair treatment

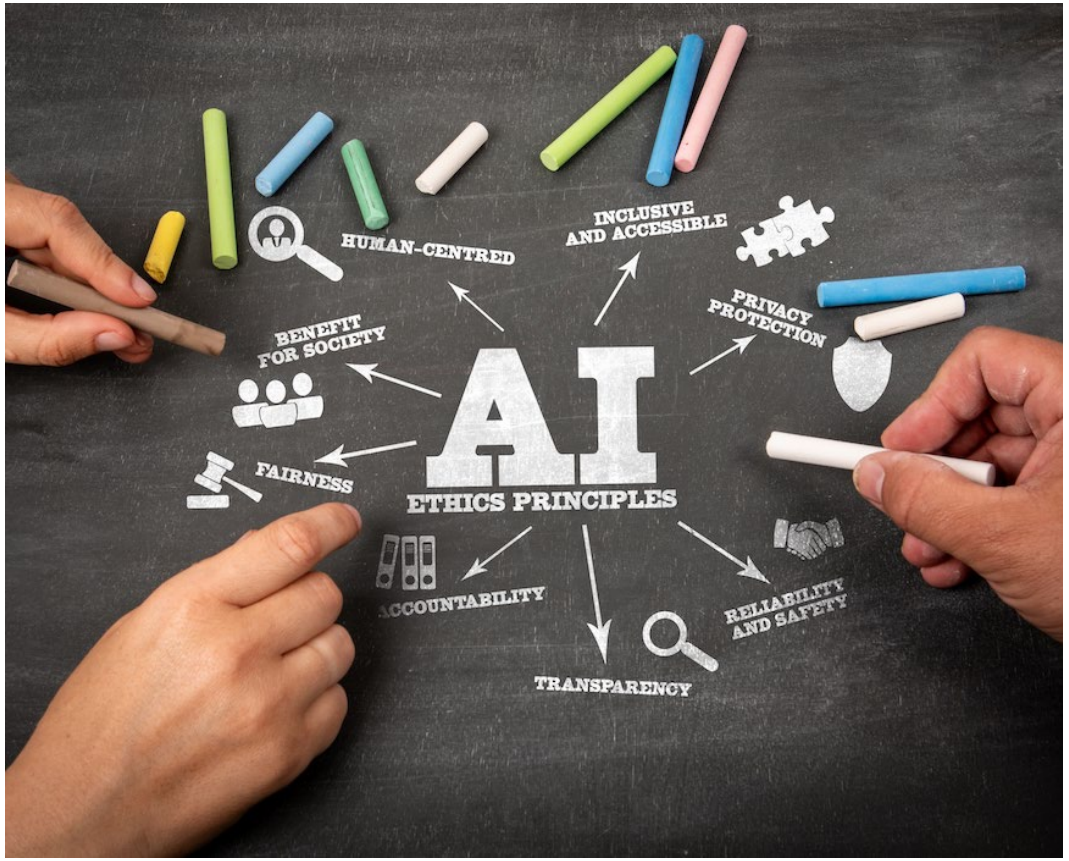
Case Study #2: Lost in Translation

- **The Scenario:** A prevention team in rural Arizona uses an AI tool to translate an underage drinking flyer into a local dialect for Spanish-speaking families
- **The Issue:** The AI generates a clinical Spanish translation that feels culturally alien, judgmental, or overly formal to the specific local community
- **The Ethical Check:** Can the choice of words be justified to the community?

The Professional Fix: A bilingual specialist reviews the draft, replacing clinical terms with regional verbiage and ensuring the tone remains supportive rather than lecture-based

Human Oversight and Responsibility

- AI is a tool for assistance, not a replacement for human professionals
- Final human determination should apply to all decisions that have a significant or irreversible impact
- Ethical and legal responsibility for an AI-informed action must always be attributable to a human or legal entity
- An AI system can never replace ultimate human responsibility and accountability



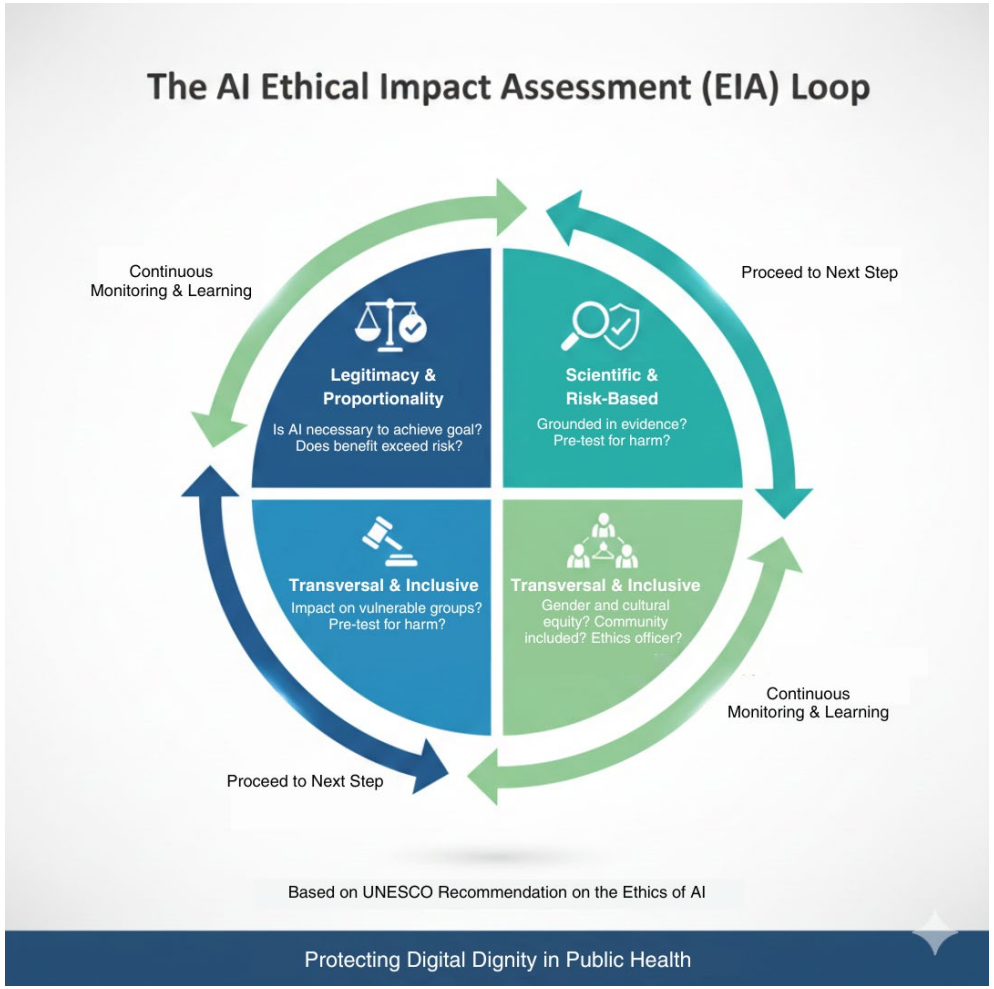
Discussion

Where do you personally draw the line? At what point does an AI-drafted message **stop being a useful starting point** and **start feeling like it could damage** your credibility with the community?

Note answers in the chat

Introduction: The Decision-Making Framework

- Using technology requires a step-by-step review to ensure the work remains fair and respectful of the community
- One option for organizations is to use a "Safety Check" process to find benefits and fix risks before any new tool is used
- A fair review must happen at every step: From the first idea and building the tool, to using it and eventually sunsetting it



UNESCO "Recommendations on the Ethics of AI" (2022)

A Practical Checklist for Ethical AI Use

This 4-step framework is the **Ethical Impact Assessment (EIA)** model. It is designed to be applied at every stage of the life cycle to ensure any AI tool is anchored in professional standards.

1

Step 1: The Right Tool Check (Legitimacy)

- Is AI truly necessary for this goal, and is it the best fit for our community's needs?

2

Step 2: The Reliability Check (Accuracy)

- Is the content backed by trusted prevention science? "Field-test" outputs with small groups to catch errors before they go public.

3

Step 3: The Community-Centered Check (Engagement)

- Does this respect everyone? Review every message to ensure the tone if appropriate and respectful for all in the community.

4

Step 4: The Final Say Check (Accountability)

- Who is in charge? Always maintain human oversight so we can fix errors and address community concerns immediately.

Applied from principles outline in UNESCO "Recommendations on the Ethics of AI" (2022)

The Reliability Check: Accuracy & Integrity

- AI outputs must be cross-referenced with prevention-approved research to ensure alignment with established science
- Professionals maintain the responsibility to verify health statistics and claims against trusted national sources like the CDC or SAMHSA
- Before a full community launch, pilot AI-drafted materials with a small advisory board or youth leaders to ensure the messaging is relatable and safe
- Consistent accuracy audits prevent "hallucinations" (fictional information) from damaging an organization’s community reputation

Human Oversight: The "Final Say"

- AI is designed for assistance, but the final determination of any message or strategy remains a human responsibility
- Ethical and legal responsibility for every AI-informed action must be attributable to a specific person or legal entity
- Professionals must not give up control of decisions that involve community safety, irreversible impacts, or high-stakes health outcomes
- Establish internal processes for correcting automated errors and addressing community concerns immediately

The Ethical AI Life Cycle

1

Planning & Intent (Research & Design):

Set a clear, evidence-based goal for your program and ensuring the AI's purpose is both ethical and necessary for your community

2

Building the Safety Net (Development):

Build private AI tools, auditing for unfair patterns and ensuring respectful, locally relevant language that avoids repeating harmful labels

3

Community Launch (Deployment & Use):

Let the community know AI is in the mix; providing clear, "plain-language" explanations of how the tool works and what it does with data

4

Health & Wellness Check (Maintenance & Operation):

Regularly check in on the AI tool once it's live to make sure it is helping people and not causing accidental social or mental harm

5

The Safety Audit (Monitoring & Evaluation):

Review AI outputs to catch safety risks or errors that only show up after the tool has been out in the world for a while

6

Sunsetting (Close-Outs):

Know when to retire an AI tool because it has met its goal or is no longer the safest way to serve the community; ensuring data is deleted securely

Case Study #3: The AI Chatbot Trap

- **The Scenario:** A youth coalition implements an AI chatbot to provide 24/7 answers to frequently asked questions about mental health and substance misuse
- **The Boundary Issue:** Youth begin using the bot for "synthetic empathy," sharing deep personal trauma and treating the algorithm as a human peer or therapist
- **The High-Stakes Risk:** AI lacks real-life experience and may give unsafe health advice, miss subtle signs of a crisis, or take advantage of a user's natural thinking patterns

The Professional Fix: Implement prominent "AI-Only" labeling, set clear boundaries on capabilities, and build in automatic hand-offs to human crisis intervention

Example: “AI-Only” Label Chatbot Setup

① ② ③

The Entry Point Disclaimer

Before the chat begins, a pop-up would be acknowledged. This fulfills the ethical requirement of "proactive notification.”

- **Label Text:** "Hi! I am the Prevention Partner AI. I am a computer program, not a human counselor. I can provide general facts and resources 24/7, but I cannot provide therapy or help in a crisis. Do you understand?"
- **Action:** User must click "I understand" to proceed.

The Persistent Identity Badge

While the chat is active, the interface would maintain a constant visual reminder of the bot's nature.

- **Visual Element:** A small, permanent tag at the top of the chat window that says 🤖 **“Automated Assistant"**
- **Avatar Choice:** Use a non-human icon (like a robot or abstract shape) rather than a human face to reduce the tendency for youth to attribute human feelings to the bot.

The Boundary Guardrail (Crisis Trigger)

Because AI may miss subtle cues, the "Professional Fix" involves hard-coded keywords that trigger human intervention.

- **The Scenario:** If a user types words like "hurt," "hopeless," or "emergency," the AI is programmed to stop generating text.
- **The Human Hand-off:** The bot immediately displays a highlighted box: *"It sounds like you’re going through a lot. I am just an AI and cannot help with this. Please call or text 988 now to speak with a real person who can help."*.

Why this works for Prevention Specialists:

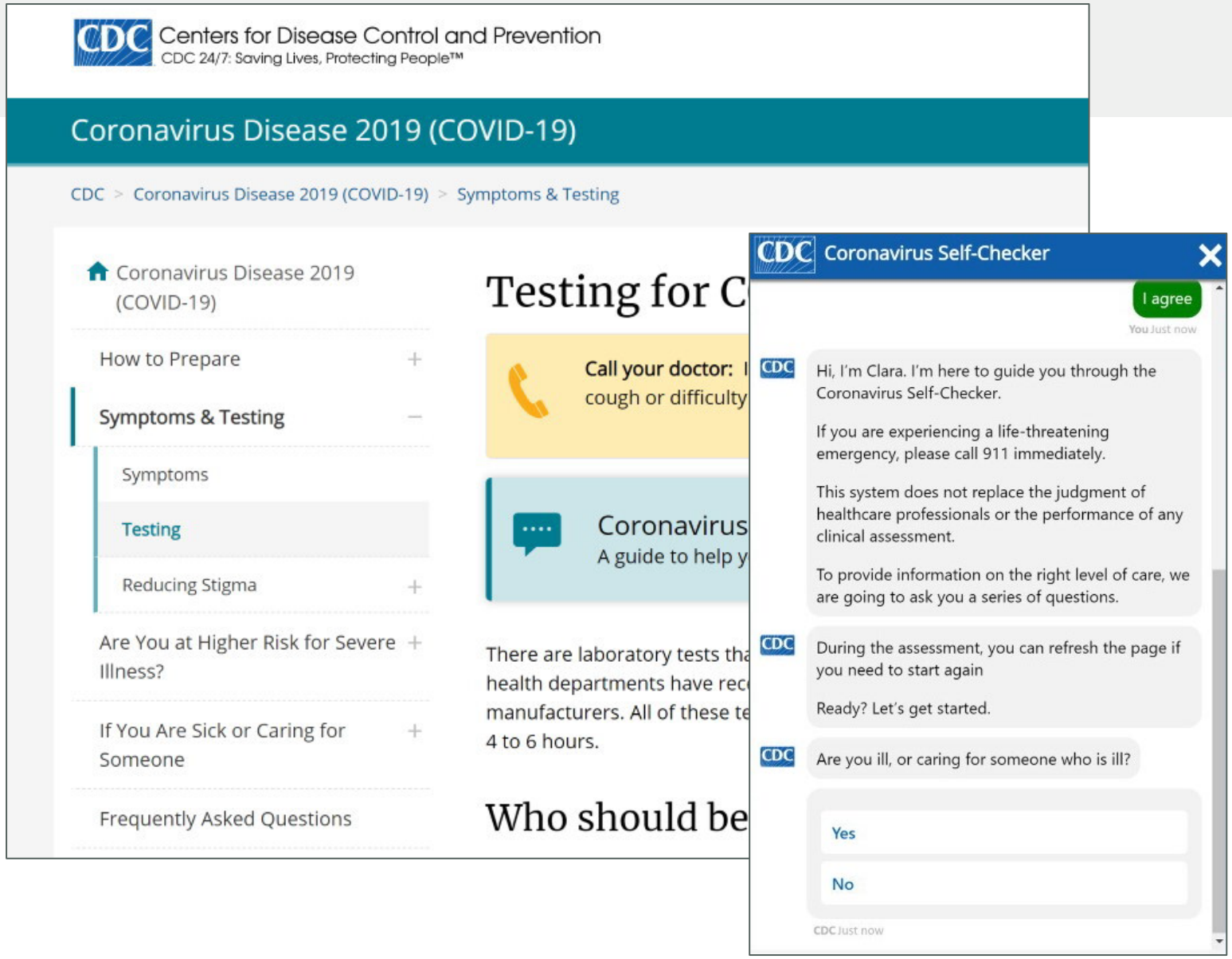
- It ensures that life-and-death decisions are never ceded to an algorithm
- By being transparent about the bot's limitations, the agency respects the user's autonomy and right to know
- Establishing clear, attributable accountability for the bot’s boundaries is a core pillar of ethical AI governance

Example AI Chatbot:

CDC Covid-19 Self-Checker “Clara”

[no longer active]

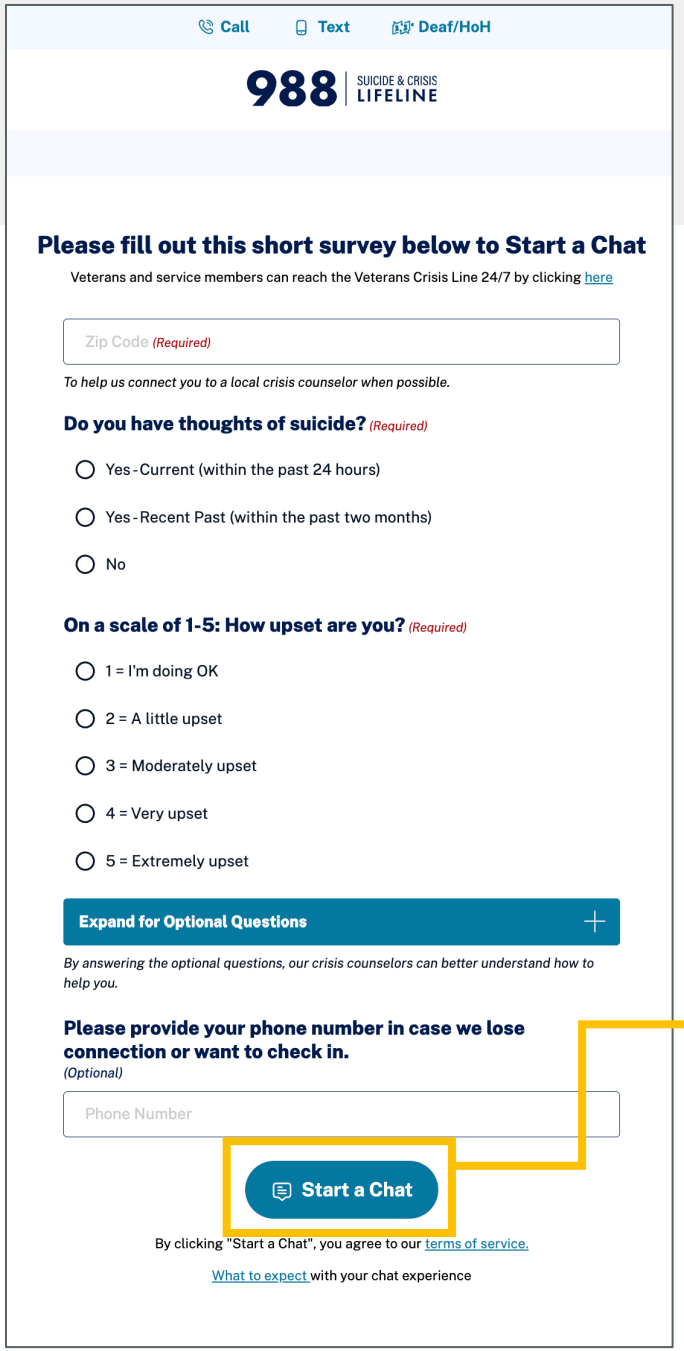
- **Purpose:** Automated triage to assess COVID-19 risk and manage home care
- **Efficiency:** Answered over 1 million messages daily, freeing up staff for critical cases
- **Clear Boundaries:** Explicitly stated it could not provide a medical diagnosis or testing locations
- **Ethics in Action:** Used ”legitimacy" by only handling information-heavy tasks, not clinical care

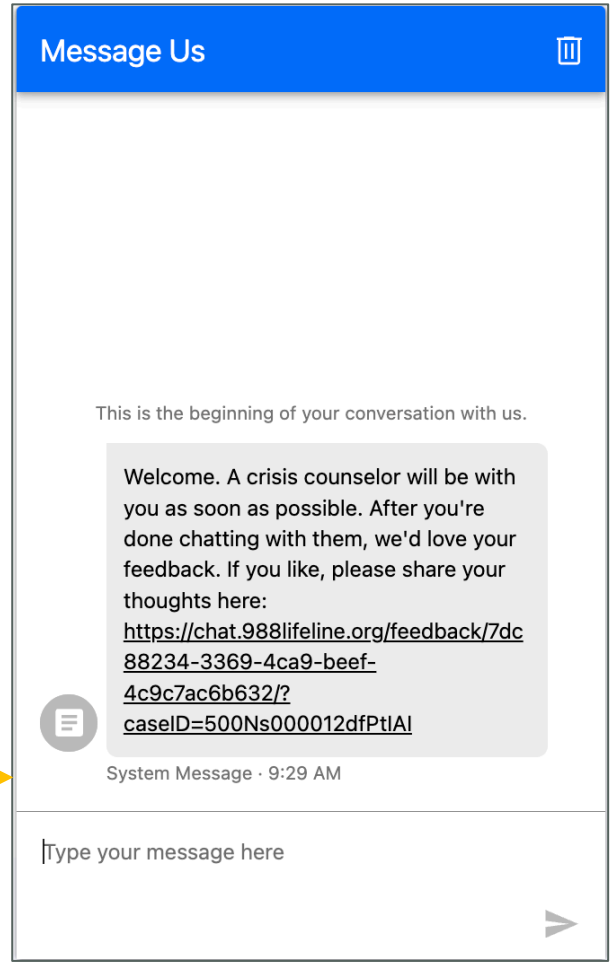


Source: PC Mag (Mar 2020)

Example AI Chatbot:
Suicide Hotline “988 Lifeline”

- **Crisis Response:** 24/7 support for mental health and substance use via call, text, or chat
- **Broadening Reach:** Serves speakers of multiple languages
- **Privacy First:** Users can remain anonymous; georouting connects people to local care without pinpointing precise GPS locations
- **Ethics in Action:** Demonstrates accountability by routing users to local experts while maintaining strict human-led support





Source: SAMHSA (2026)

Roadmap: Establishing AI Policies and Best Practices

- Organizations move through stages of readiness; start with internal behind-the-scenes tasks to build technical fluency
- Begin with low-stakes tasks, such as summarizing internal meeting notes or drafting staff memos
- Develop agency-specific policies that define where AI is - and is not - appropriate for community use
- Strengthen institutional expertise by collaborating with technical partners and regional prevention networks

Sample Outline: Agency-Specific AI Policy

I. Purpose and Values (define primary goal and commit to human-in-the-loop oversight)

II. Approved Use Cases (Green Zone) (include examples of appropriate use)

III. Restricted Use Cases (Red Zone) (outline examples of prohibited use like confidential data and automated counseling)

IV. Mandatory Verification Steps (fact-checking, cultural review)

V. Transparency Requirements (requirements for labeling of chatbots and AI-use disclosures)

➡️ Visit CDC's Vision for Using Artificial Intelligence in Public Health

Training and AI Literacy

- AI literacy is an essential skill to protect staff and community members from undue technological influence
- Training should focus on questioning the output, treating AI as a thought partner rather than an absolute source of truth
- A variety of voices within the organization must be part of the decision-making process for selecting and using AI tools
- Effective AI training combines technical use with ethics, community perspectives, and social sciences

IDEAS: Practical AI Training for Prevention Staff

- **AI 101: The "Co-Pilot" Basics:** Demystify AI terms and establish that these tools are meant to assist, not replace, human professionals.
- **Prompting for Prevention:** Learn to write clear instructions to generate better drafts for flyers, logic models, and grant summaries
- **Spotting "Hallucinations":** Practice fact-checking AI outputs against trusted sources like the CDC or SAMHSA to catch "plausible" but false information
- **The Stereotype Audit:** Review AI-generated content to ensure it doesn't reinforce harmful clichés about rural life or substance use
- **Privacy-First Practices:** A simple guide on which data is safe to enter into AI tools and what must remain confidential (e.g., participant names)
- **The "Human Edge" Focus:** Identify the skills AI cannot replicate, such as empathy and community relationship building

The Prevention Specialist's AI Toolkit

Strategy and Content Development: The Co-Pilot

Gemini Google AI Studio	**Gemini / AI Studio:** Best for drafting logic models, grant narratives, and brainstorming campaign slogans
perplexity	**Perplexity:** A "search AI" that provides citations, making it excellent for finding vetted prevention stats or local health data
Hero AI TRANSFORM CARE. SAVE LIVES	**Headline Hero:** Useful for creating high-impact, stigma-free titles for community flyers and newsletters

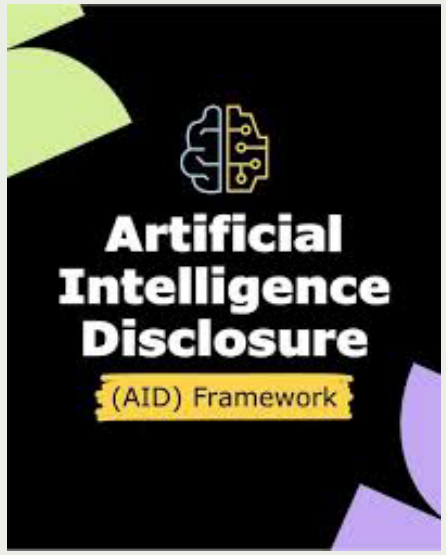
Visuals and Community Outreach

Canva Adobe Firefly	**Canva / Adobe Firefly:** Essential for generating culturally appropriate images and resizing social media assets for local reach
Napkin	**Napkin AI:** Automatically turns your program data or prevention steps into visual diagrams—perfect for "explaining the science" to community members
guidde.	**Guidde:** Creates how-to video guides; great for showing community partners how to access resources or navigate your agency's website

Knowledge Management and Data Triage

NotebookLM	**NotebookLM:** A specialized tool where you upload your own documents to ask questions and get accurate summaries without hallucinations
tally*	**Tally:** An AI-powered survey tool to collect community feedback in a user-friendly format
MUCKROCK	**MuckRock FOIA Requests:** Helps specialists track public records and local health policies across different jurisdictions

Knowledge Management and Data Triage



AI Disclosure Tools: Generate "AI-Only" labels and disclosure statements needed to maintain community trust (e.g., tools that generate visual labels and badges, automated AI use statements such as The AID (Artificial Intelligence Disclosure) Framework, and interactive disclosure features

👋 Jan

Jan: A "private AI" that runs on your own computer without sending data to the cloud - ideal for analyzing sensitive internal reports while protecting privacy

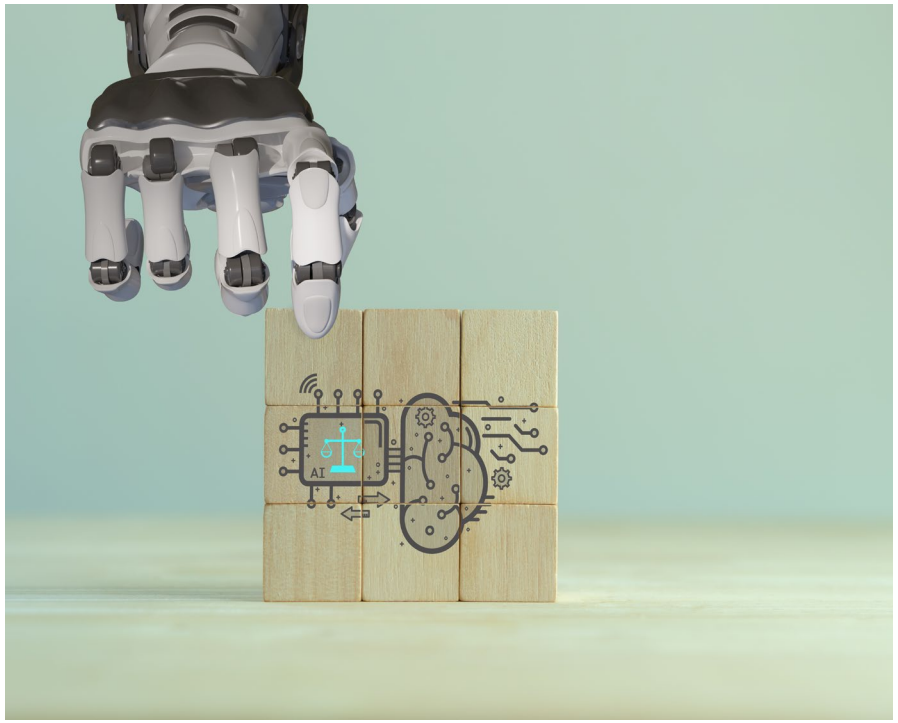
Discussion

Looking at the checklist we covered today, what is one small **safety check** you could add to your team's workflow next week to make sure your AI use stays ethical?

Note answers in the chat

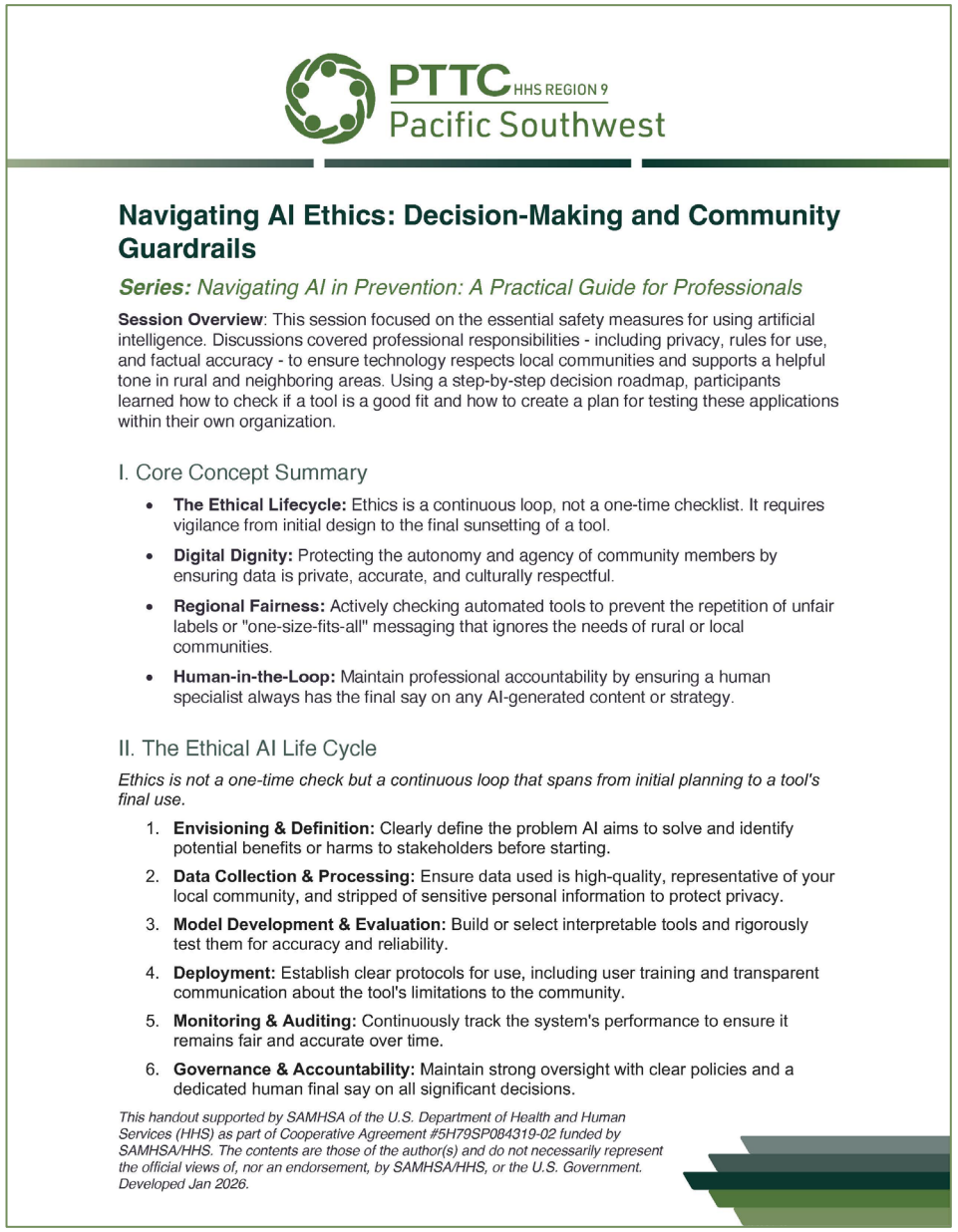
Summary: An Ethics-First Approach

- **A Unified Shield for Your Work:** Ethical principles like digital dignity, fairness, and transparency are not extra tasks; they work together to protect the integrity of your programs and the trust of your community
- **Building a Foundation of Trust:** Using established ethical standards helps local agencies create a professional environment that is fair, predictable, and trustworthy for everyone
- **Supporting the Whole Community:** Using AI ethically is about more than just technology - it supports long-term community wellness, your own professional growth, and the health of the environment
- **Your Professional Compass:** While tools will change, a commitment to human rights and community respect remains the permanent guide for navigating new technology in prevention



Final Takeaways and Ethical Application

- **The Continuous Loop:** Ethics is not a one-time checklist but an ongoing Life Cycle from initial planning to tool retirement.
- **Human-in-the-Loop:** A human professional must maintain ultimate responsibility and have the final say on every AI-assisted output.
- **Zone-Based Governance:** *Green Zone*: Use AI for low-stakes administrative tasks and brainstorming. *Red Zone*: Never enter identifiable data or use AI for direct counseling.
- **Reliability Protocols:** Mandatory verification of statistics against vetted sources like CDC or SAMHSA.
- **Digital Dignity Check:** Prioritize regional audits and community privacy before any AI-assisted launch



PTTC HHS REGION 9
Pacific Southwest

Navigating AI Ethics: Decision-Making and Community Guardrails

Series: *Navigating AI in Prevention: A Practical Guide for Professionals*

Session Overview: This session focused on the essential safety measures for using artificial intelligence. Discussions covered professional responsibilities - including privacy, rules for use, and factual accuracy - to ensure technology respects local communities and supports a helpful tone in rural and neighboring areas. Using a step-by-step decision roadmap, participants learned how to check if a tool is a good fit and how to create a plan for testing these applications within their own organization.

I. Core Concept Summary

- **The Ethical Lifecycle:** Ethics is a continuous loop, not a one-time checklist. It requires vigilance from initial design to the final sunsetting of a tool.
- **Digital Dignity:** Protecting the autonomy and agency of community members by ensuring data is private, accurate, and culturally respectful.
- **Regional Fairness:** Actively checking automated tools to prevent the repetition of unfair labels or "one-size-fits-all" messaging that ignores the needs of rural or local communities.
- **Human-in-the-Loop:** Maintain professional accountability by ensuring a human specialist always has the final say on any AI-generated content or strategy.

II. The Ethical AI Life Cycle

Ethics is not a one-time check but a continuous loop that spans from initial planning to a tool's final use.

1. **Envisioning & Definition:** Clearly define the problem AI aims to solve and identify potential benefits or harms to stakeholders before starting.
2. **Data Collection & Processing:** Ensure data used is high-quality, representative of your local community, and stripped of sensitive personal information to protect privacy.
3. **Model Development & Evaluation:** Build or select interpretable tools and rigorously test them for accuracy and reliability.
4. **Deployment:** Establish clear protocols for use, including user training and transparent communication about the tool's limitations to the community.
5. **Monitoring & Auditing:** Continuously track the system's performance to ensure it remains fair and accurate over time.
6. **Governance & Accountability:** Maintain strong oversight with clear policies and a dedicated human final say on all significant decisions.

This handout supported by SAMHSA of the U.S. Department of Health and Human Services (HHS) as part of Cooperative Agreement #5H79SP084319-02 funded by SAMHSA/HHS. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by SAMHSA/HHS, or the U.S. Government. Developed Jan 2026.

Q&A

References

- American Psychological Association (Jul 2025). *Ethical Guidance for AI in the Professional Practice of Health Service Psychology*
- Annenberg Public Policy Center (2025, Jul 22). *AI Used for Real-Time Selection of Actionable Messages for Government and Public Health Campaigns*.
- Centers for Disease Control (CDC) (2025, Aug 22). *CDC's Vision for Using Artificial Intelligence in Public Health*.
- Olawade, D.B., Wada O.J. et al (2023, Oct 26). *Using artificial intelligence to improve public health: a narrative review*. Frontiers in Public Health, PMID 37954052
- Pham, T. (2025, May 14). *Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use*. Royal Society Open Science, PMID 40370601
- Stark, K. (2025, Aug 11). *AI Breakthrough May Transform Public Health Campaigns*. Penn Leonard Davis Institute of Health Economics
- UNESCO (2021). *Recommendation on the Ethics of Artificial Intelligence*

Let's connect.

Rebecca L. Cooney

rebecca-cooney.com
@rebeccacooney (LinkedIn)

Post-Webinar Feedback

Please click on the link in the chat to complete a very brief online feedback form!

Thank you!

International Certification & Reciprocity Consortium (IC&RC)

This webinar is under these domains:

- Planning & Evaluation
- Communication
- Public Policy and Environmental Change
- Professional Growth & Responsibility

Learn more about IC&RC: https://internationalcredentialing.org/

Connect with us!

Join our mailing list:

- www.pspttc.org
- http://eepurl.com/gIssWD

Find us on LinkedIn:

- https://linkedin.com/company/pacific-southwest-pttc

Email with general questions:

- pspttc-info@casat.org

Like us on Facebook:

- https://tinyurl.com/PSPTTC-Facebook

Pacific Southwest PTTC Contact Information

Clarissa Lam Yuen

Senior Project Coordinator

clamyuen@casat.org

Thank You!

University of Nevada, Reno
CASAT
School of Public Health

PTTC HHS REGION 9
Pacific Southwest