The Safety Net Survivor's Guide to Artificial Intelligence (AI):

Understanding Artificial Intelligence, Protecting Your Autonomy, and Making Empowered Choices

National Network to End Domestic Violence
SAFETY NET PROJECT

The Safety Net Survivor's Guide to AI: Understanding Artificial Intelligence, Protecting Your Autonomy, and Making Empowered Choices

Introduction

This piece is the second of a two-part series: click here for our guide on considerations and best practices for AI use by victim service providers and advocates.

Technology is evolving quickly, and artificial intelligence (AI) tools are becoming part of more and more everyday experiences. While AI does have the potential to save time and expand access to resources, these same tools can also introduce serious safety, privacy, and trust risks, especially for survivors of domestic violence, sexual assault, stalking, technology-facilitated abuse, and other forms of trauma. Some AI tools may offer advice or comfort that sounds human, but is actually misleading or unsafe, and most of them quietly store–and sometimes even share–what you tell them.

At the National Network to End Domestic Violence (NNEDV), we believe survivors deserve clear, trustworthy information about how AI tools work, what risks they carry, and what options may exist for survivors and their loved ones who engage with these technologies. This guide was created to support you in making informed, confident choices about if or how you use AI, on your own terms.

> ***Important Note**: Every survivor's situation, needs, comfort levels, and boundaries are different. This guide shares general safety tips and things to keep in mind, but not every recommendation will apply to everyone. We encourage you to take what's useful to you, leave what's not, and know that your safety and autonomy come first. For questions, technical assistance, or additional support, please don't hesitate to contact us.*

Table of Contents

Introduction 2

What Is AI, and Why Does It Matter? 3

Guidance for Survivors Using AI Tools 4

Understanding the Limits: AI Isn't Confidential or Human 4

Privacy Tips: Protect Your Personal Information 5

Emotional Cautions: AI Is Not a Therapist, Advocate, or Companion 6

Safer Ways to Use AI (if You Choose To) 7

Always Remember These Safety Rules when Using AI: 8

Conclusion & Final Reminders 8

What Is AI, and Why Does It Matter?

Artificial intelligence (AI) is a type of technology that mimics human thinking to perform tasks like writing text, transcribing audio, translating language, or spotting patterns in data. Many of us interact with AI every day without realizing it – autocomplete in texts and emails, facial recognition in photos, traffic updates in maps, and shopping recommendations all rely on AI systems in order to function. AI is being added to more and more apps, websites, and services, including some that survivors may come across when looking for information or support.

Generative AI is a newer kind of AI that became widely used in late 2022. Unlike earlier AI tools that simply followed pre-programmed rules, generative AI can create entirely new content – like emails, images, summaries, or code – based on user input (known as "prompts" or "queries"). Tools like ChatGPT, Microsoft Copilot, and Google Gemini are some of the most well-known examples. Generative AI's mainstream breakthrough came with the public launch of ChatGPT.[1]

At their core, these systems don't actually "understand" your request the way a person does. They are essentially sophisticated guessers: they predict what words (or images or phrases) are most likely to come next based on patterns in the massive datasets they were trained on. This means they can sound confident even when they're wrong, and that's where many serious risks begin. In other words, a chatbot's response might sound eloquent and firm, while being misleading or even completely fabricated.

Generative AI can also carry the biases and gaps of its training data. If the data used to train the AI contain harmful stereotypes or lack diverse perspectives, the AI's outputs can reflect those same biases. For example, AI chatbots trained on Western-centric data have misinterpreted cultural expressions of distress.[2] One user reported that a mental health chatbot told them their anxiety was "irrational" when they described discrimination at work.[3] Examples like these demonstrate how these tools cannot necessarily simulate cultural competence or empathy.

Finally, unlike a counselor, advocate, or therapist – who is legally and ethically bound to keep what you share confidential – **most AI tools store the things you tell them and keep logs of your conversations**, often to improve the AI or for other business purposes. Without proper safeguards, anything you input could potentially be seen by others, leak into another user's results, or even be discoverable in court. This is a serious risk, especially if you're sharing personal or sensitive information with an AI tool.

In summary, AI tools can be useful for small, simple tasks, but they're not human, usually not confidential, and not always accurate. They don't truly understand, and they can't offer care or

[1] Generative AI: What Is It, Tools, Models, Applications and Use Cases https://www.gartner.com/en/topics/generative-ai
[2] AI Therapists Are Biased—And It's Putting Lives at Risk | Psychology Today https://www.psychologytoday.com/us/blog/the-human-algorithm/202504/ai-therapists-are-biased-and-its-putting-lives-at-risk
[3] *Id.*

judgment the way a trained person can. This is why it's so important to use AI cautiously, especially if you are navigating an unsafe situation.

This next section provides tailored guidance for survivors. It contains information on protecting your privacy and maximizing your safety if or when engaging with AI tools. Our AI guidance for victim service providers and advocates can be found here.

Guidance for Survivors Using AI Tools

If you are a survivor of trauma or abuse, you might come across AI tools like chatbots or voice assistants that claim to offer help, or you might simply consider using general AI apps (like a chatbot that comes with your phone or a website) for information or support. It's important to approach these tools with caution. This section highlights potential risks and suggests ways to maximize safety and control if you do decide to use AI tools.

Important Note: *No AI chatbot or app can replace the human support of a trained advocate, counselor, or trusted friend. But if you choose to interact with AI for help or information, being informed about the potential risks can help you protect your safety, privacy, and well-being.*

Understanding the Limits: AI Isn't Confidential or Human

When you talk to a domestic violence advocate or counselor, there are ethical rules and often laws that protect your privacy: what you share is confidential, and the advocate isn't allowed to divulge it without your permission (barring extreme exceptions like immediate danger or court orders). This is not the case with AI services.

Most AI chatbots are run by companies that typically store your conversations on their servers. What you tell an AI could potentially be seen by tech employees, used to "train" the AI further, or even handed over in legal situations. For instance, one court order forced an AI company to save all user chats, even those users thought they deleted.[4] So, assume that anything you type into an AI could be saved or shared. It's not a private diary or a confidential chat, even when it feels one-on-one.

Separately, we encourage you to always remember that AI is not a person. It may sound obvious, but when you're upset or lonely, it can be easy to see a chatbot as a kind of friend or support. AI language models are designed to produce smooth, conversational replies, and some can even mimic a warm, nurturing tone. Despite that, the AI has no real understanding or empathy, because

[4] Tech Policy Press – Belle Torek, "For Survivors Using Chatbots, 'Delete' Doesn't Always Mean Deleted," June 10, 2025. https://www.techpolicy.press/for-survivors-using-chatbots-delete-doesnt-always-mean-deleted/

it's generating responses based on patterns. It might say validating things, but it might also give dangerously misguided advice because it doesn't truly grasp your situation.

Real-life examples show how wrong this can go. In one case, a man in Belgium became very emotionally attached to an AI chatbot (even seeing it as a confidante). The chatbot encouraged him in harmful ways, and sadly, he died by suicide after following the AI's disturbing suggestions.[5] In another case, a teenager was using a chatbot that pretended to be a famous fictional character; the AI "normalized" suicidal thoughts and even gave encouraging messages about them, and the teen later took his own life.[6] These are extreme tragedies, but they highlight that an AI cannot be trusted with your mental health or safety planning. An AI might miss red flags or fail to give you crucial help that a human would provide, because it doesn't truly understand consequences. It also won't proactively call for help if you're in danger.

Bottom line: Think of AI as a public tool, not a private confidant, and as a robot, not a human friend. Anything very personal, like details of abuse, identifying information, or intense emotional struggles, is safer shared with a real person who is trained to help. If you wouldn't post it publicly or tell it to a stranger, you probably shouldn't tell it to a chatbot.

Privacy Tips: Protect Your Personal Information

If you do decide to use an AI chatbot or similar tool for any reason, be very careful about what information you share. Here are some concrete privacy tips for interacting with AI as a survivor:

- **Stay Anonymous**: Do not share your name, address, phone number, email, or any details that could identify you. Avoid mentioning other people's names (like those of family members) or specific locations (your workplace or your city, especially if it's small, etc.). For example, rather than saying "My ex John Doe who lives at 123 Main Street did XYZ," you could speak generally: "my ex-partner did XYZ." The less personal data, the better. This way, even if the conversation were somehow exposed, it's harder to link it to you specifically. Even **metadata or the tiniest geographic identifiers in uploaded photos** can ultimately be personally identifying.
- **No Sensitive Identifiers**: Similarly, don't enter things like your social media handles, specific court case numbers, or any account info/passwords into an AI. Some people might think an AI can help fill out forms or write letters, but don't provide things like your Social Security number or bank info in the prompt. (AI should never be used for that kind of task).
- **Don't Rely on Deletion**: If the app or site has a feature to delete your chat history or go "incognito," understand that it might not truly delete everything. As noted, companies may

[5] 'He Would Still Be Here': Man Dies by Suicide After Talking with AI Chatbot, Widow Says https://www.vice.com/en/article/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says/
[6] AI Therapists Are Biased—And It's Putting Lives at Risk | Psychology Today https://www.psychologytoday.com/us/blog/the-human-algorithm/202504/ai-therapists-are-biased-and-its-putting-lives-at-risk

retain data on the back-end. Use those features (they can reduce what's visible on your device), but assume the data still exist somewhere. It's like clearing a chat on your phone: the other party (in this case, the AI company) might still have a copy.

- **Beware of Linked Accounts**: Some AI services let you log in with Google, Facebook, or other accounts. If you're concerned about privacy, consider using a burner email or an account not linked to your real identity. This adds a layer of anonymity. Also, check if the chatbot is public (some AI "companions" allow others to view conversations), and make sure it's a private session.
- **Read and Understand Privacy Policies**: This is a big ask, but if you can, skim the AI tool's privacy policy or FAQs. Look for statements about data use. If it says data may be shared with third parties or used to improve the service, that means your chats are not confidential. Almost all freely available AI tools will have such clauses. Knowing this can reinforce that you should keep sensitive information out of it.

In summary, treat the AI like a stranger on the internet: never give out info that could be used to locate you, contact you off-platform, or impersonate you.

Emotional Cautions: AI Is Not a Therapist, Advocate, or Companion

Many survivors have understandable reasons for turning to online resources or even chatbots for support. Some AI tools advertise themselves as offering a non-judgmental ear or coping advice. However, be very cautious about using AI for emotional support or crisis help. Here's why:

- **False Reassurance and Delayed Help**: A chatbot might say things that make you feel a bit better momentarily ("I'm here for you," "That sounds very tough, you're so strong for coping"). While that can feel supportive, it might also delay you from reaching out to a real person. If you spend hours venting to a bot, that's hours you weren't talking to a counselor, friend, or hotline who could offer real assistance. A chatbot's comfort is superficial: it won't follow up with you, and it won't notice subtle signs of escalating risk like a human would.
- **Unreliable or Dangerous Advice**: AI is not an expert, even when it sounds like one. It might give incorrect advice about what to do (for example, wrong info about legal options or medical facts). It could even suggest something harmful. There have been reports of chatbots encouraging violence or self-harm, or normalizing abuse by failing to challenge harmful statements. If you told a chatbot, "I feel like I'm to blame for the abuse," it might not effectively counter that myth the way a trained professional would. Be wary: any advice or "facts" from an AI should be double-checked with a trusted source.
- **Emotional Attachment to AI**: It's surprisingly easy to start feeling emotionally attached to a chatbot, because it always replies and can be programmed to be polite or caring. Survivors who feel very isolated or misunderstood might pour their hearts out to an AI. But note: the relationship is an illusion. The AI doesn't truly understand or empathize, so the "support" has limits. If you find yourself getting very dependent on chatting with it, it may be helpful to take a step back. Consider dialing back usage or talking with a real counselor

about how you're feeling. Over-reliance on AI can increase feelings of isolation in the long run.

- **Triggering or Inappropriate Responses**: Because AI isn't actually sensitive to your well-being, it might respond in ways that are upsetting or triggering. If you encounter a response from an AI that makes you feel worse, remember that you're interacting with a flawed machine, not a person who intended to hurt you. Take a break, and reach out to a human if you need emotional first aid after a bad bot experience.

In extreme cases, dependency on AI "companions" has been linked to negative mental health outcomes. The examples of the Belgian man and the Florida teen show how far off track it can go when AI starts to fill a role it shouldn't.[7] Lawsuits are ongoing in those cases, raising questions about accountability for AI platforms that exploit vulnerable users with poor safeguards. While those are extreme, they highlight a simple truth: no chatbot can fully substitute for human care. If you are in crisis, feeling unsafe, overwhelmed, or considering self-harm, please reach out to a trained crisis counselor or someone you trust.

Safer Ways to Use AI (if You Choose To)

All that said, you might still find some AI tools useful in lower-stakes ways. The key is to use them with clear boundaries. Here are a few safer-use tips and examples for survivors:

- **Information Gathering**: You could use AI to ask general questions or get definitions. For example, "What are common signs of gaslighting?" or "Explain the cycle of abuse." These are broad questions that don't reveal personal info. The AI might give a helpful overview (though double-check any critical info with a quick web search or a reputable site). This can be like using it as a fancy search engine. Just be cautious that sometimes AI gives incorrect info. Verify anything important via trusted resources (like womenslaw.org or techsafety.org).
- **Idea Generation**: Maybe you want ideas for self-care or for how to phrase something (like writing a statement or a journal entry). You could ask, "Can you give me some ideas for calming activities when I feel anxious at night?" The AI might list generic things like breathing exercises, music, etc. If nothing else, it's brainstorming. Again, ensure the suggestions are safe and sensible – most will be, but if something odd comes up, use your judgment or consult a professional.
- **Language Translation or Drafting (with caution)**: If English isn't your first language, you might use AI to help translate a question you want to ask an advocate, without including identifying details. For instance, you could type in Spanish, "How do I get a protective order?" and have it translate to English, then take that translation to a legal aid chat (not that you must, most lines have interpreters too). Or if you're drafting an email to a landlord, you can ask the AI for a general template, but don't paste in your exact address or full story.

[7] *See* footnotes 4, 5, and 6.

- **Creative Outlets**: Some survivors use writing as an outlet. If you treat the AI as a creative tool (like to write a poem about healing, or to role-play a fictional empowering scenario), that might be a safe use, since it's more about expression than advice or data. Even then, keep it fictional or generalized, so you're not feeding it your personal details.

Always Remember These Safety Rules when Using AI:

- **No Personal Identifiers**: (Worth repeating) Don't give away who you are.
- **Double-Check Critical Info**: If the AI tells you something that could impact your decisions (legal steps, safety actions), verify it with a human expert or authoritative source.
- **Listen to Your Feelings**: If an AI conversation is making you uncomfortable or just doesn't feel right, you can stop. You are never obligated to continue or to follow its lead.
- **Prioritize Authentic Human Support**: Use AI as a supplement for low-level tasks or curiosity, not as a replacement for getting real help. If you have access to an advocate, counselor, support group, or even a trusted friend, those should remain your primary supports. AI is just a tool – and sometimes a faulty one at that.
- Finally, **you have the right to say "no" to AI.** If an organization or anyone ever wants to involve an AI in your care (like using a screening chatbot or recording session notes with AI), you can absolutely decline. You should feel empowered to ask: "Is this tool secure? What happens to my information? Can we do this without the AI?" Advocates should respect your wishes. Your comfort and safety come first.

The world of AI is new and evolving, and it's completely understandable to be unsure about it. When in doubt, lean on personal connections and trust your instincts. If something feels off with an AI interaction, it probably is. You deserve genuine, compassionate support, and while technology can sometimes assist, it should never compromise your privacy or well-being.

Conclusion & Final Reminders

AI tools are emerging everywhere, and it can feel like you're expected to use them just to keep up with the rest of the world. But you should never feel pressured to use a new tool, especially one that might put your well-being at risk, just because it's popular. No survivor should have to compromise their safety, privacy, or autonomy in order to engage with technology. Both advocates and survivors should feel confident not using a technology if it doesn't align with their needs and values.

Some AI tools can be helpful in small, low-risk ways. However, most tools on the market today are not designed with trauma, confidentiality, or legal compliance in mind, and the legal landscape around AI is still very much evolving. For now, this means that the burden falls on each of us to approach these technologies with vigilance and care.

Whether you are a victim service provider or a survivor, here are some shared best practices and final reminders:

- **Protect Personal Data**: Personally identifying survivor information should never be input into an AI tool unless you are absolutely certain the tool is secure, private, and compliant; and realistically, such certainty is hard to come by. For advocates, this is a professional mandate; for survivors, it's a self-protection strategy. When in doubt, keep the details out.
- **Respect and Demand Consent**: Survivors must always have an informed, voluntary choice in whether their information is processed by AI. Advocates must get consent every time, and be prepared with alternatives if consent is not given. Survivors must have the right to ask questions and say no.
- **AI as Assistive, not Authoritative**: AI should assist, not replace, human expertise, empathy, and judgment. No chatbot, no matter how advanced, can replicate the nuanced understanding and empathy of a trained advocate or the lived experience of survivors. AI can be a tool, but human brains and hearts must remain at the center of any decision-making or support. If an AI output doesn't seem right, double-check with outside sources and trust the human judgment to override it.
- **Double-Check AI Outputs**: The stakes are higher in victim services. When generative AI makes mistakes or hallucinations, the consequences can be life-altering. Always verify critical information that comes from an AI. If an AI helped draft something, review it carefully (and ideally have someone else review it too). If you're a survivor using AI for info, cross-reference with a known reliable source. It's an extra step, but it can prevent serious harm.
- **Prioritize Safety and Trust**: In this field, trust is everything. Survivors need to know their information is safe and their agency is respected. Advocates need to maintain credibility and abide by their ethical duties. Any technology that undermines trust – by leaking info, giving bad advice, or just creating confusion – should be set aside. It's okay to be slow or cautious in adopting new tech. What matters is that survivors feel secure and supported.
- **Stay Educated and Adapt**: Technology will continue to change rapidly. AI might become safer or new tools might emerge that are specifically designed for confidential survivor services. It's wise for both advocates and survivors to stay informed about these developments. Be open to learning (as you are by reading this guide), and be ready to adapt policies as needed. But also remember that our fundamental responsibility to survivors remains constant, no matter the tech.

In closing, generative AI is a powerful innovation with many potential benefits, but also many pitfalls. By clearly distinguishing between what is acceptable and what is not, setting distinct boundaries for advocates and survivors, and taking time to understand the overlap, we can navigate this space more safely.

We invite you to treat this guide as a living resource that you can return to, reflect on, and share as the technology around us continues to evolve. We encourage you to consider talking through these ideas with people you trust, and to think about what feels right for you. You don't have to make these decisions all at once, and you are always allowed to change your mind.

While the conversation around AI and survivor safety is still in its early stages, your experiences and insights are essential in shaping what comes next. As the conversation continues, your voice, experience, and choices will help shape a safer, more thoughtful future.

This project was supported by **Grant #15JOVW-23-GK-05170-MUMU** awarded by the Office on Violence Against Women, U.S. Department of Justice. The opinions, findings, conclusions, and recommendations expressed in this publication/program/exhibition are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Justice.

© 2025 National Network to End Domestic Violence, Safety Net Project.