Artificial Intelligence (AI) & Victim Services:

A Comprehensive Guide for Advocates

National Network to End Domestic Violence
SAFETY NET PROJECT

Artificial Intelligence & Victim Services: A Comprehensive Guide for Advocates

Introduction

This piece is the first of a two-part series: click here for our survivors' guide on considerations and best practices for AI use.

Artificial intelligence (AI) tools are becoming embedded in nearly every profession, and victim services are no exception. 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 when working with survivors of domestic violence, sexual assault, stalking, technology-facilitated abuse, and other forms of trauma.

At the National Network to End Domestic Violence (NNEDV), we recognize both the promise and the peril of these emerging technologies. This guide is designed to share practical guidance and help victim service providers (VSPs) and other advocates make informed, mission-aligned decisions about whether and how to use AI tools in the context of victim services.

> ***Important Note**: No two programs or situations are exactly alike. This guide offers general best practices and safety considerations, but individual programs and coalitions may need to adapt these recommendations to fit their specific legal obligations, funding requirements, and community needs. For questions, technical assistance, or additional support, please don't hesitate to contact us.*

Table of Contents

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 increasingly being integrated into nearly every form of technology that we interact with.

Generative AI is a newer form of artificial intelligence that goes a step further. 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, which hit headlines and went viral in late 2022.[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 human professional bound by confidentiality, most AI tools log and store what you tell them (often to improve the AI or for other business purposes). Without proper safeguards, anything you input could potentially be seen by third parties or even show up in another user's output through data leaks. This is a critical concern when dealing with sensitive information in victim services.

In summary, ***generative AI systems are powerful but fallible tools.*** They can assist with certain tasks, but they lack true understanding, empathy, and moral judgment. This is why it is extremely

[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.*

important to use them cautiously, especially in fields like victim services where trust, confidentiality, and safety are paramount.

Where AI Might Help

To be clear, AI does have some useful applications in victim service settings when used carefully for low-risk tasks. Some examples of where generative AI might help include:

- **Document Summaries**: AI can help advocates quickly summarize or review lengthy public documents or reports. For instance, an AI tool might condense a 50-page research paper into a one-page summary. However, **do NOT upload documents containing any personally identifying information (PII) or confidential case details** into an AI tool. Even for internal use, sensitive content should be removed before AI use.
- **Language Translation**: Some AI-based translation tools can improve communication when qualified human interpreters aren't available. They might help translate general information materials or non-confidential messages into languages spoken by survivors in your community. Always verify the translation's accuracy and clarity, especially for critical safety or legal information.
- **Internal Content Drafting**: AI can assist with brainstorming internal, non-confidential content like generic policy language, training outlines, or grant proposals. This can save time for busy staff. Crucially, ***survivor data must be excluded.*** AI might help brainstorm grant language or rephrase a public outreach paragraph, but you should never feed it real names, case specifics, or any details that could identify or endanger someone.

These use cases are internal and low-risk by design. They avoid putting sensitive survivor information into the AI. In these constrained scenarios, AI can be a helpful productivity tool to optimize an advocate's work. Even then, human oversight is needed to ensure the output is correct and appropriate.

Where AI Can Go Wrong

In many other scenarios, generative AI tools pose unacceptable risks in the victim services context. It's important to recognize the situations where AI use can go wrong or should be strictly avoided:

- **Chatbots for Survivor Support**: AI chatbots should not replace the survivor-centered expertise of human advocates. While a chatbot might be programmed to sound empathetic, it is ultimately a machine. Chatbots are often impersonal or inaccurate, and they lack the ability to truly understand context or emotion. There have been instances of chatbots giving inappropriate or even dangerous advice in sensitive situations. For example, one AI chatbot encouraged a user to commit suicide when the user expressed

fears about climate change.[4] Chatbots can often be misleading: some can claim to be therapists and even fabricated license numbers when prompted.[5] Chatbots simply cannot replicate the nuanced, relationship-based support that trained advocates provide. An "emotional support" chatbot might give a survivor false reassurance, which could delay them from seeking help from real people. At worst, a poorly designed bot might actively cause harm by normalizing abuse or encouraging self-harm.[6] Bottom line: Never use an unspecialized AI chatbot as a replacement for a human advocate or counselor when supporting a survivor.

- **AI Meeting Assistants in Survivor Meetings**: AI-powered note takers or virtual assistants (like those that transcribe or summarize meetings) should never be used in survivor meetings or support groups unless the survivor is fully informed, consenting, and in control. These tools often upload data to third-party servers. If used improperly, they could record sensitive conversations without true consent. In some cases, these notetakers can generate completely inaccurate information that can become part of an electronic health record and be difficult to correct once generated.[7] If, in a rare case, a survivor wants to use an AI assistant, they must opt in freely with informed, voluntary consent and a clear understanding of how the tool works and where the data will go. In practice, it is safest to exclude AI from any confidential or therapeutic conversation. Human note-taking or other non-digital means are preferable to protect confidentiality.
- **Automated Risk Assessments or Decision-Making**: Using AI to evaluate a survivor's risk level, determine eligibility for services, or make decisions about urgency and resource allocation is both dangerous and unethical. For instance, AI prediction models in social services and law enforcement have a documented history of bias and error. The frequency of flawed or one-size-fits-all predictions means an AI should never determine something as critical as whether or how a survivor gets help. Every survivor's situation is unique, and these decisions require human judgment, cultural competence, and empathy, factors an algorithm simply doesn't have.
- **Other Misuses**: In general, any AI use that involves processing survivor-provided information (especially identifiable information) is problematic unless strict conditions are met (as discussed in the next section). Even seemingly minor uses, like plugging a survivor's story into a chatbot to draft a safety plan or asking an image-generation AI to visualize evidence, could lead to breaches of privacy or create distorted, untrustworthy results. When in doubt, err on the side of not using AI for survivor-related matters. There is usually a safer, human-driven alternative available.

---

[4] '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/

[5] Instagram's AI Chatbots Lie About Being Licensed Therapists, https://www.404media.co/instagram-ai-studio-therapy-chatbots-lie-about-being-licensed-therapists/

[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

[7] What to know about an AI transcription tool that 'hallucinates' medical interactions, https://www.pbs.org/newshour/show/what-to-know-about-an-ai-transcription-tool-that-hallucinates-medical-interactions

While generative AI mistakes can be troublesome in any field, the consequences of those mistakes in victim services can be uniquely severe. Misinformation, privacy violations, or a loss of trust can literally be life-threatening in the context of domestic violence and trauma. This is why it's critical for both advocates and survivors to approach these tools with extreme caution, set clear boundaries, and prioritize human safety over speed or convenience.

The next section provides tailored guidance for victim service providers. It contains information on privacy, confidentiality, and how to safely evaluate or use AI in victim services. Our AI guidance for survivors can be found here: it contains information on protecting personal privacy, understanding the limits of AI "support," and how to maximize safety when using these tools.

Guidance for Victim Service Providers and Advocates

For advocates and organizations serving survivors, few responsibilities are more important than safeguarding survivors' privacy and maintaining confidentiality. Introducing AI tools into your work must be done, if at all, with utmost care and adherence to existing confidentiality obligations. This section details what advocates need to know about the risks of AI, strategies to mitigate harm, and how to make an informed decision about using (or not using) AI in victim services.

Privacy vs. Confidentiality: What's the Difference?

In discussions of technology and survivor data, the terms privacy and confidentiality are both crucial, but they have distinct meanings:

Privacy refers to an individual's right and ability to control *their own personal information*. It's about a survivor's choice over who gets to know what about their life. For example, a survivor might choose not to share their new phone number or address with anyone for safety reasons, or may opt out of location tracking by an app. In essence, privacy is an individual right to decide how much information to share with others.[8]

Confidentiality refers to the ethical or legal duty *of someone else to protect the survivor's information* once it's been shared in confidence. In victim services, when a survivor discloses something to an advocate, the professional is bound to keep that information secure and secret (with a few exceptions, like mandatory reporting). Confidentiality is grounded in laws and ethical standards that prohibit sharing someone else's information against their will.[9]

[8] Primer on Privilege — Safety Net Project
https://www.techsafety.org/privilege-primer
[9] *Id.*

In short, privacy is controlled by the survivor (their choice to disclose or not), while confidentiality is the responsibility of the professional or service (to protect what is disclosed).[10] Both are absolutely critical to our work.

Privacy harms in the context of AI might include a survivor's data being accessed by unauthorized parties, a survivor losing control over personal information they entered into a tool, or corporate entities retaining data that the survivor intended to keep private. Confidentiality harms involve breaches of the trust or legal obligations we have as advocates: for example, if an advocate inadvertently shares a survivor's story with an AI service, that could violate funder requirements or laws like VAWA that require keeping that information confidential.

When considering AI, advocates must protect both privacy and confidentiality: ensure survivors' personal data doesn't leak or get misused (privacy), and uphold all legal and ethical commitments not to reveal what survivors share in confidence (confidentiality). Unfortunately, most AI tools on the market today pose challenges on both fronts.

Privacy Risks of AI Tools: Data Exposure and Third-Party Access

Most widely available AI tools were not designed with survivor safety or privacy in mind. In fact, many generative AI services (such as popular chatbots, AI writing assistants, voice transcription tools, etc.) explicitly "learn from" and repurpose the information you share with them for future training and product improvement. This means that anything an advocate or survivor types into such a tool might be stored on remote servers, analyzed by the AI company, and even used to refine the model or appear in responses to other users. If you are using a general-purpose AI chatbot, you must assume that what you enter could be recorded or reused in some way.

Here are some common characteristics of commercial AI platforms that pose serious privacy concerns for survivor data:

- **Data Storage**: They often store user inputs on third-party servers (the AI company's cloud) without giving users control over how long that data is kept or who can access it. Users often cannot easily delete their data permanently.
- **Data Logging**: Conversations and interactions are frequently logged for "quality assurance" or analytics. Even if the AI interface feels like a private chat, the content might be reviewed by developers or used to train the AI.
- **Transparency Challenges**: AI tools usually lack transparency about their data handling. It's often unclear where the data are stored, who can see it, or how long it's kept. Privacy policies may be vague or hard to find.
- **No True Consent Mechanism**: Most AI apps don't offer a meaningful opt-in consent for data use. Simply by using the service, you often "agree" that the company can use your input data. This is especially problematic in trauma-informed work where survivors should

[10] *Id.*

have a choice about how their information is used. The concept of informed consent is usually absent from these tools' design.

- **No Emergency Protocols**: Chatbots and AI assistants are not equipped to handle crisis situations or mandatory reporting in real time. They can't easily escalate to a human or call emergency services if a user is in danger (and some terms of service even disclaim responsibility for that). This is more of a safety functionality issue, but it means if a survivor reveals abuse or suicidal thoughts to an AI, the tool won't respond with the same protocol a trained advocate would.

All of the above create significant risks to survivor privacy. For example, if a survivor shares personal details or abuse experiences with a generative AI chatbot on a website, those details could end up in a future AI model update, or even appear in another user's output through a phenomenon known as "data leakage." There have been cases where private information inadvertently surfaced in unrelated users' AI queries because it was retained in the model's memory. Even describing a survivor's situation (without using names) to an AI could result in that scenario being stored and learned by the system, effectively handing sensitive info to a third party.

A stark illustration of privacy risk exists in a recent legal case, where a federal court ordered OpenAI to preserve all ChatGPT output logs, including chats that users had deleted or thought were temporary.[11] In other words, "delete" didn't mean deleted, and content users believed to be gone was now required to be saved and could potentially be examined in court.[12] The data under this recent order include conversations that users would have erased or run in "temporary chat" modes. Even if an AI platform promises privacy or gives a deletion option, those assurances can be overridden by legal demands. Content a survivor shares with a chatbot could be retained or even disclosed in legal proceedings, without the survivor's knowledge or consent.

The lesson for advocates: assume that anything put into a commercial AI tool could be saved or accessed by others. If that information would pose a risk to a survivor if leaked, it does not belong in the AI. This includes not just obvious identifiers like names or addresses, but also any detailed personal circumstances, court case details, or narratives of abuse. Err on the side of caution: it is far better to forego an AI's convenience than to expose a survivor to new harms.

Confidentiality and Compliance Concerns

In addition to general privacy risks, advocates must consider legal confidentiality requirements and funding rules that apply to their work. Many victim service programs are bound by laws such as the Violence Against Women Act (VAWA), the Family Violence Prevention and Services Act (FVPSA), the Victims of Crime Act (VOCA), as well as state-level advocate privilege laws and professional ethics. These laws typically prohibit sharing a survivor's personally identifying information without specific, time-limited, informed consent from the survivor. Violating

[11] The New York Times Company v. Microsoft Corporation et al, Filing 551, https://docs.justia.com/cases/federal/district-courts/new-york/nysdce/1:2023cv11195/612697/551

[12] For Survivors Using Chatbots, 'Delete' Doesn't Always Mean Deleted, https://www.techpolicy.press/for-survivors-using-chatbots-delete-doesnt-always-mean-deleted/

confidentiality can not only endanger survivors, but jeopardize an organization's funding or even lead to legal liability.

Mainstream AI chatbots and tools are not designed to be compliant with these confidentiality obligations. As our recent analysis on AI and survivor privacy in *Tech Policy Press* notes, widely available platforms like ChatGPT are generally not operating within frameworks that trigger confidentiality protections.[13] Additionally, using these tools generally does not meet the standard of a protected or privileged communication (since the AI company is a third party, not a confidential advocate or attorney). Attorney-client privilege, for instance, is highly unlikely to cover a client's disclosures to a consumer AI chatbot. The confidentiality that applies in an advocate-survivor conversation does not extend to survivor-AI conversations.

To clarify how AI usage intersects with key laws and regulations, consider the following:

- **VAWA**: Applies to VAWA-funded victim service providers. It prohibits disclosing personally identifying information (PII) about survivors without the survivor's time-limited, informed, written consent (with some narrow exceptions). If an advocate were to input survivor PII into an AI tool that stores data, that could be considered an unauthorized disclosure, since the information is effectively shared with the AI company (a third party) and possibly used beyond the survivor's intent. This likely violates VAWA confidentiality rules.
- **VOCA and FVPSA**: VOCA-funded programs and FVPSA-funded programs have similar confidentiality provisions as VAWA. They generally require keeping survivor information confidential and not disclosing it without consent. Using a chatbot that logs data could conflict with these requirements, unless you had consent (and even then, many funders frown on practices that risk exposure of survivor data).
- **HIPAA**: If you are a covered entity or working in a context involving health information (e.g., a sexual assault program associated with a healthcare provider), HIPAA requires safeguarding Protected Health Information (PHI). Any AI tool used in a healthcare context would need a business associate agreement and strong security. Mainstream commercial chatbots are not HIPAA-compliant out of the box. For instance, if a survivor asks a health-related question via a general AI chat service, that could inadvertently create a HIPAA violation if PHI is shared.
- **Attorney-Client Privilege**: While not applicable to most advocates unless you're in a legal role, it's worth noting that communications with legal counsel are privileged only if kept confidential. If an attorney or client puts information into a public AI tool, it may constitute a waiver of privilege. In the same vein, advocate-survivor privileged communications (in states where that exists) could be jeopardized if an advocate were to share those details with an AI service, since privilege generally requires no unnecessary third-party disclosure.

To summarize: No popular AI chatbot or generative AI platform currently meets the strict confidentiality requirements for direct survivor data. These laws mandate a level of protection and

[13] *Id.*

control that these AI services do not provide. Because AI platforms aren't typically built with those safeguards, the onus is on advocates to not put sensitive data into them in the first place.

Best practice: Unless an AI tool can demonstrably guarantee no data retention or third-party access (which is not the case for general AI tools), do not input any survivor-identifying or case-specific information into the AI. Focus on using AI, if at all, in ways that do not involve client data.

Important Note: Keeping information confidential is not just a legal requirement, but a core part of building and maintaining trust with survivors. Many survivors seek services precisely because they are guaranteed a confidential, judgment-free space. If we introduce tools that compromise that space, even unintentionally, we risk eroding the very foundation of our support.

Questions to Ask AI Vendors (for Programs Considering New Tools)

If your program or coalition is considering adopting any AI-powered software, especially those that might process or store survivor data (for example, an AI case management add-on, a chatbot on your website, or an AI transcription service), due diligence is essential. Before using the tool in any capacity, get clear, concrete answers from the vendor about their data practices and capabilities. Here's a checklist of questions to ask vendors:

- **Data Use for Training**: "Is any user input data used to train or improve the AI model?" You need to know if the content you input will be fed back into the AI's learning process. If yes, that means the data is being retained and analyzed. Ideally, the answer should be "No, or we offer a secure opt-out so your data are not used."
- **Data Storage and Retention**: "Where are user data stored, and for how long? Who can access it? Can it be deleted permanently upon request?" If the vendor cannot tell you exactly what happens to the data (e.g., stored for 30 days then deleted, accessible only to certain security personnel, etc.), that's a red flag. We should seek minimal retention and the ability to delete data on demand.
- **Opt-Out and Control**: "Does the tool allow us to opt out of data collection or turn off the AI features when needed?" Some platforms might have a "private mode" or an option not to save histories. Others might allow an on-premise installation where data stay locally. If no meaningful controls are available, assume all data are collected.
- **Security and Compliance**: "Is the tool compliant with relevant privacy laws, and can it meet requirements like VAWA confidentiality or HIPAA if needed?" Most likely the answer will be that the base product is not VAWA/HIPAA compliant, but listen for whether they have any compliance offerings. If a vendor markets specifically to healthcare or government, they might have a secure version. If they can't speak to compliance at all, that's a sign the tool is consumer-grade and not meant for sensitive information.
- **Memory/Logging**: "Does the AI have a memory of past inputs? If so, can that feature be disabled or cleared?" For example, some chatbots let you wipe conversation history. If the AI "remembers" previous interactions to inform responses, that means data are stored.

Ideally, you want no persistent memory when working with sensitive topics, or at least the ability to clear it.

If the vendor cannot or will not answer these questions clearly and in writing, the tool is likely not appropriate for survivor-facing work. Remember that even if a tool seems useful, it must first do no harm. An evasive or uninformed vendor is not someone you want to bet your clients' safety on.

On the other hand, if the vendor provides solid answers–for instance, "No, we do not train on your data, yes we can delete data on request within 24 hours, here's our encryption and compliance detail..." you will still need to get informed consent from your organization's leadership, legal counsel, and survivors (more on consent below) before using it.

Informed Consent: Survivors Must Always Have a Say

Survivor autonomy is central to trauma-informed care, and any introduction of AI into your service delivery should respect that autonomy. This means survivors must always be informed and have a choice in whether AI tools are used in handling their information or their service provision.

If you as an advocate are considering using an AI tool in any part of a survivor's case or personal data, you should ensure the survivor:

- Knows that an AI tool is being used (or considered) to assist in the process. Never use an AI "behind the scenes" on survivor data without disclosure. For example, if you're using a translation AI or drafting a letter with AI help, let the survivor know that this technology is involved.
- Has the option to opt out and say "No, I prefer you not use that tool with my information." There should be no negative repercussions for the survivor if they choose not to have AI involved. Always be prepared to use a human-only alternative if a survivor isn't comfortable with AI.
- Understands how their data will be used. This means you should be able to explain in plain language what the AI does, where the information goes, and what safeguards are (or aren't) in place. For instance: "We could use a computer program to translate this document. It might save time, but I want you to know that the program will send the text to a company's server to do the translation. They say they don't store it long-term, but I cannot guarantee complete privacy. Would you rather use this tool, or should we find a human translator? It's your choice."

In many cases, once survivors learn about the uncertainties of AI, they may opt out, and that is completely valid. If a survivor doesn't give clear, informed, voluntary consent for an AI-related process, do not use the AI for that survivor's information or services.

Finally, advocates should document consent when possible. If a survivor agrees in writing (even an email or text saying "Yes, you can use the translation app for my document"), that's ideal. But

even a noted verbal consent is better than nothing. This protects both the survivor's intent and the organization.

Important Note: *Offering real choice may also mean offering to undo something. If you try an AI tool with a survivor's permission but then the survivor feels uneasy, you should be willing to stop using it, delete any data you can, and continue without it. The survivor's comfort and trust in the process are more important than whatever benefit the AI tool is providing.*

Using AI for Content Creation and Outreach: Risks and Best Practices

Some programs are experimenting with generative AI to help create content, such as brainstorming training materials, revising outreach blog posts, or even generating social media posts and brochures. At first glance, this might seem low-risk since it doesn't involve personal survivor data. However, there are still important concerns to keep in mind when using AI for any kind of resource creation or public-facing content:

- **Plagiarism and Copyright Issues**: Many generative AI models were trained on massive datasets scraped from the internet, including copyrighted academic articles, news stories, books, and non-profit publications. These tools often do not cite their sources, and they can sometimes produce text that closely mimics or even directly copies phrases from published works. If an advocate uses AI to draft an article or report, there is a risk of inadvertent plagiarism or unintentional copyright infringement if the AI regurgitates existing content. This could lead to credibility issues or even legal problems for your organization. Always review AI-generated text critically: run it through a plagiarism checker if possible, and ensure that any facts or unique phrasing aren't lifted verbatim from elsewhere without attribution.
- **Accuracy and "Hallucinations"**: Generative AI is notorious for producing inaccurate or fabricated information (often called "hallucinations"). The AI may present incorrect statements as if they were factual. For example, a chatbot might confidently provide a statistic or a quote that is completely made-up, or misstate the law on a certain issue. Recent studies have found disturbingly high rates of such false outputs: one analysis showed ChatGPT-3.5 generated wrong or made-up references about 40% of the time, and the more advanced GPT-4 model still did so about 29% of the time.[14] The onus is on the human user to fact-check every AI-generated claim. *Never assume the AI is correct.* If you use AI to draft a section of a grant or a fact sheet about domestic violence, you must verify all facts, statistics, quotations, and definitions against trusted sources before publishing. AI cannot be trusted as an authoritative source: treat it like an intern who needs everything reviewed.
- **Bias and Tone**: As mentioned earlier, AI outputs can reflect biases. When creating materials for your program, consider whether the AI might be introducing subtly biased language or perspectives. For instance, if you ask an AI to write a brochure about domestic

[14] Journal of Medical Internet Research - Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis
https://www.jmir.org/2024/1/e53164/

violence, and it was trained on data with gender stereotypes, it might produce content that inadvertently blames victims or only highlights certain types of abuse. Always review the tone and framing. Ensure the content aligns with trauma-informed, inclusive principles.

- **Automation Bias (Overreliance on AI)**: One of the greatest dangers in using AI for writing or research is actually a human error: the tendency to give AI outputs more credit than they deserve. This is known as automation bias, when people often trust a computer-generated answer simply because it looks professional or authoritative, even if they would question the same statement coming from an unknown person. Studies show that users often over-rely on automated systems and fail to double-check them, which can lead to critical errors.[15] As advocates using AI, we must consciously resist this bias. If something the AI produces doesn't seem entirely right, or even if it does, we need to verify it. Always apply your own expertise and judgment.
- **Reputation and Quality Control**: Finally, consider the audience's trust. If AI helps draft a blog post, and that post ends up containing a mistake or a strangely phrased section, readers might lose trust in your organization's information. Maintaining a high standard of accuracy and clarity is crucial in victim services outreach. It can be embarrassing (or worse, actively harmful) to share content that later turns out to be wrong or plagiarized. Any AI-generated content must be heavily vetted and edited by knowledgeable staff before it carries your organization's name.

Best Practice for Content Creation: AI can be a useful brainstorming partner or first-draft generator, but it should never be the final word. You might use AI to get a rough outline or to suggest some phrasing, but a trained advocate or subject matter expert should review and refine the material thoroughly. Think of AI as a starting point for your writing process, not a shortcut to skip human review. If you don't have time to verify and thoughtfully edit what the AI produced, it's better not to use it. Also, for important public-facing documents (reports, educational materials), you might decide it's safest to write them the old-fashioned way, or only use AI in minor ways (like strengthening transitions between paragraphs) to avoid any risk of error.

By keeping these guidelines in mind, advocates can prevent the misuse of AI in content creation. The goal is to harness any potential efficiency gains without sacrificing accuracy, originality, or the trust of your audience.

Environmental Considerations (AI's Carbon & Water Footprint)

Many victim service organizations are committed not only to survivor safety, but also to broader principles of community well-being, which for many includes environmental sustainability. It's worth noting that generative AI use has environmental impacts that advocates might care about, especially if your program or coalition values eco-conscious practices.

[15] What is Automation Bias? How to Avoid its Pitfalls | Informa TechTarget
https://www.techtarget.com/searchitoperations/definition/What-is-automation-bias

Generative AI models are computationally intensive. This means they require significantly more energy than traditional software or simple web searches. In fact, a single query to an AI-powered chatbot can use up to ten times as much energy as a standard Google search.[16] One analysis estimated that generally a generative AI system may use 20–30 times more computing power to accomplish a task than regular programming would.[17] All that energy consumption contributes to carbon emissions if the electricity is not from renewable sources.

Beyond electricity, significant amounts of water are needed to cool the servers that train and run these AI models.[18] Tech companies operate massive data centers, and cooling the equipment (often using water in cooling towers) can strain local water resources and ecosystems. The rapid growth of generative AI has led to a surge in data center construction and energy use worldwide. If your organization has committed to reducing its carbon footprint or is part of a community that stresses sustainability, you might decide to limit AI use for that reason as well.

This doesn't mean never use AI (the environmental impact of a few queries is small), but it is another factor to be aware of. If staff start using AI as a support for everything (including things a simple search or a bit of human effort could have done), it could multiply the energy consumption unnecessarily.

Practical tip: If you do use AI tools, use them efficiently. Keep prompts and sessions as concise as possible. Don't run dozens of needless queries when one targeted query would do. Avoid using AI in place of a normal search for simple facts (since a regular search engine query is far less resource-intensive). For example, to find a hotline number or a definition, a quick web search is "greener" than asking a chatbot. Also, logging out of or shutting down AI tools when not in use can ensure you're not unintentionally sending data.

Of course, this environmental perspective is more of an added consideration: survivor safety and privacy still come first. Still, it aligns with the idea that we should be intentional and purposeful in using technology, and consider not just the direct benefits and risks to survivors, but also indirect impacts on the world they live in.

Conclusion & Final Reminders

As AI tools continue to evolve and integrate into everyday life, the pressure to adopt them quickly and broadly can feel overwhelming. But in victim services, speed and convenience must never

[16] Ai's environmental impact: Energy consumption and water use – Planet Detroit https://planetdetroit.org/2024/10/ai-energy-carbon-emissions/

[17] *Id.*

[18] Explained: Generative AI's environmental impact | MIT News | Massachusetts Institute of Technology https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117

come at the expense of survivor safety, trust, or autonomy. Both advocates and survivors should feel confident not using a technology if it isn't safe or doesn't align with their needs and values.

Generative AI can offer meaningful support in limited, controlled contexts: for example, helping staff with internal tasks that don't involve private data, or giving survivors easy access to general information. 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. This means the burden falls on us, as advocates, to approach these technologies with vigilance, care, and survivor-centered practices.

Whether someone is 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 encourage you to use this guide as a living resource: discuss it in staff meetings, use it to guide your program's AI use policies, and consider this information within the context of the evolving technological landscape. The conversation around AI and victim services is just beginning, and your experiences and voices are crucial in shaping how these tools are approached in the future. With care, compassion, and informed choices, we can uphold the values of our work while engaging with new technology on our own terms.

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.