Using AI in Survivor Advocacy: A Harm-Reduction Guide for Shelters and Anti-Violence Workers



Table of Contents

Potential Risks in Using AI in Anti-Violence Work 3

Decision Prompts When Considering AI Use in Anti-Violence Work 6

High-Risk Uses of AI in Survivor Advocacy....8

Lower Risk Uses of AI....11

Organization Policies and Practices....11

Summary....14

Artificial intelligence (AI) is increasingly part of everyday life. For under-resourced anti-violence programs, these tools may offer practical benefits. Tools such as ChatGPT, Microsoft Copilot, writing assistants, transcription and meeting tools, and AI features built into email, word-processing programs, and other software can make some administrative tasks faster and easier.

At the same time, using AI in survivor-serving work can create privacy, confidentiality, safety, legal, ethical, and organizational risks. Survivor information is often highly sensitive. A disclosure of information that identifies a survivor, their location, children, circumstances, or safety concerns can have serious consequences. Once information is shared with a third-party AI tool, the organization may have less control over where it is stored, who can access it, how long it is retained, and whether it may be used for other purposes.

This guide provides a privacy-first, harm-reduction approach to using AI in Canadian women's shelters, transition houses, and anti-violence organizations. It is intended to help organizations and workers recognize risks and develop responsible AI policies and practices.

Important: This handout provides general information, not legal advice. Organizations should determine which privacy laws and other legal obligations apply to them and seek advice from their privacy officer, legal counsel, or relevant privacy regulator when necessary.

Potential Risks in Using AI in Anti-Violence Work

AI raises a wide range of concerns, including questions about ethics, environmental impact, labour impacts, bias, accessibility, and the broader social effects of these technologies. These are important issues, but they are beyond the scope of this guide.

This guide focuses on the risks most directly connected to survivor safety and anti-violence practice, including privacy, confidentiality, security, legal obligations, and organizational responsibility.

1. AI Tools are not necessarily private or confidential

You cannot assume that information entered into an AI tool is kept private and confidential. Information may be sent to a third-party service and become subject to that provider's terms, privacy practices, retention policies, training practices, security measures, and contractual arrangements. Depending on the tool and its settings, conversations may be stored, reviewed, retained for a period of time, sold, disclosed in a subpoena request, or used in other ways described in the company's privacy policies. This means other people may be able to access the information you enter without you knowing.

Canadian privacy obligations do not disappear because information is entered into an AI system. The law that applies to an organization can depend on factors such as its province or territory, whether it is a public or private organization, and the nature of its activities. For example, BC's Personal Information Protection Act (PIPA) applies to private organizations, including non-profit organizations, while other legislation can apply to public bodies.

Organizations should not assume that because a tool is widely used, free, or provided through Microsoft or Google, it is appropriate for sensitive information. Canadian privacy regulators have specifically identified privacy concerns related to generative AI, including overcollection of personal information, consent and transparency, inaccurate information, difficulties with access/correction/deletion, and accountability.

2. The information you provide might be stored longer than you think or permanently

Deleting a conversation from your screen does not necessarily mean that every copy of the information has immediately been deleted from the company's systems. Different AI services have different retention and deletion practices. Some may retain information for a certain period of time for security, legal, technical, or other purposes. There may also be circumstances where a company is legally required to preserve or disclose information.

For example, information held by a company could potentially be subject to a legal demand such as a court order or subpoena, depending on the circumstances and jurisdiction.

3. AI can provide inaccurate information

AI can produce information that is:

- Incorrect
- Outdated
- Incomplete
- Based on the wrong jurisdiction
- Presented without important context
- Entirely Fabricated

This matters when workers use AI to find information about legal processes, government programs, community services, or survivor support.

For example, AI may confidently provide:

- An incorrect court procedure
- An outdated legal requirement
- A nonexistent case or resource
- The wrong eligibility requirements
- Information about a service that no longer exists

4. AI can create a false sense of certainty or security

AI can produce recommendations, classifications, summaries, or rankings that appear objective. They are not necessarily objective, accurate, or appropriate.

AI should not independently determine:

- Whether a survivor is at risk
- Whether someone qualifies for a service
- Who should receive a shelter bed
- Who should be prioritized on a waitlist
- Whether a survivor's account is credible
- Whether a safety plan is adequate
- What intervention a survivor should receive

Survivor-serving decisions require human judgment, context, lived experience, cultural understanding, professional expertise, and survivor autonomy.

Decision Prompts When Considering AI Use in Anti-Violence Work

Before using an AI service or introducing an AI tool, organizations should understand its:

- Data retention practices
- Deletion procedures
- Backup practices
- Data storage locations
- Access controls
- Secondary uses of information
- Training or model-improvement practices
- Sub-processors and third-party service providers

When reviewing the policies and information listed above, consider:

- What privacy legislation applies to us?
- What personal information would be entered into the tool?
- Is using AI necessary for this purpose?
- What authority do we have to collect, use, or disclose information?
- Where is the information stored or processed?
- Who can access it?
- How long is it retained?
- Is it used to train, improve, or otherwise develop the AI system?
- Can the information be permanently deleted, not just deleted for us while still being stored on a server?
- What happens to backups?
- What happens if the provider experiences a data breach?
- What contractual protections are in place?
- Can we retract information we have entered?
- Have we completed an appropriate privacy and security assessment>
- Does the tool meet our confidentiality and security requirements?

Organizations should also consider whether AI-generated content or prompts could become part of organizational records and potentially be subject to applicable access, retention, legal, or disclosure obligations.

Ask yourself, “Is this particular use of this particular tool appropriate, necessary, authorized, and consistent with our organization's legal and confidentiality obligations?”

Remember: “anonymous” does not always mean anonymous.

High-Risk Uses of AI in Survivor Advocacy

As a general rule, do not enter survivor, client, or case information into public or unapproved AI tools. This includes information about:

- Survivors
- Children and youth
- Current or former partners
- Family members
- Friends or other people connected to a survivor
- Volunteers or staff when the information is confidential

Personal information can include much more than a person's name.

Avoid entering:

- Addresses or locations
- Workplaces or schools
- Phone numbers or email addresses
- Dates of birth
- Court or file numbers
- Case histories
- Information about abuse or violence
- Safety concerns
- Information about children
- Health or mental health information
- Financial information
- Immigration or legal information

- Details that could identify a person even when their name has been removed

A description may still identify someone when combined with other details.

For example: "Write a case note for my appointment with Jane today. She is struggling with mental health and housing for her three children "should not be entered into an AI tool.

But removing “Jane” does not automatically make the prompt safe. For example, "write a case note for a 37-year-old survivor in a small northern community who has three children, is experiencing depression, and is currently staying at [specific shelter].”

The combination of details may still identify the person.

Organizations should prohibit or carefully control AI use for activities involving confidential survivor information, including:

- **Case documentation**
    - Case notes
    - Client files
    - Incident reports
    - Case summaries
- **Survivor communications**
    - Emails
    - Text messages
    - Letters containing personal information
- **Safety and risk**

    - Safety plans
    - Risk assessments
    - Danger assessments
    - Threat assessments
- **Service decisions**
    - Eligibility decisions
    - Prioritizing access to services
    - Determining risk levels
    - Deciding who receives limited resources
- **Meetings and conversations with or about survivors**
    - AI meeting transcription or note-taking
    - AI-generated meeting summaries
- **Legal or court-related information**
    - Survivor case histories
    - Court documents containing personal information
    - Legal correspondence
    - Summaries of a survivor's legal situation

A simple rule: If you would not share the information with an unrelated third party or want the AI company to be subpoenaed for it, do not put it into an AI tool unless your organization has specifically assessed and authorized that use.

If AI is ever used in a process that could significantly affect a survivor, the organization should have clear safeguards, meaningful human oversight, and a way to identify and address errors. Canadian privacy guidance emphasizes assessing the necessity, proportionality, validity, and reliability of generative AI uses, especially when they may significantly impact people.

Lower Risk Uses of AI

A harm-reduction approach does not necessarily require organizations to ban every AI-enabled feature.

Lower risk uses may include tasks that involve information that is non-confidential, non-personal, and does not identify a survivor or their circumstances.

For example:

- Editing a public-facing document such as your website or a brochure
- Improving grammar
- Shortening a non-confidential document
- Brainstorming general ideas
- Translating non-identifying information
- Drafting general professional communications
- Creating general professional communications
- Creating general training content
- Brainstorming grant language that contains no survivor information

Organization Policies and Practices

AI adoption should be an organizational decision. Every organization using AI should establish clear rules for staff.

Policies should address:

- **The process for reviewing and approving AI tools**
    - Require an appropriate privacy and security assessment before adopting new AI tools and specify who at the organization will conduct this assessment.
    - Before adopting an AI tool, review relevant vendor terms, privacy agreements, data processing arrangements, retention practices, and security safeguards using the prompts listed in earlier sections of this document.
- **Approved tools & uses**
    - Identify which AI tools, if any, are approved for organizational use.
    - Identify what kinds of tasks these tools can be used for, and what to avoid.
- **Prohibited information & tasks**
    - Clearly identify information that must never be entered into AI systems.
    - Prohibit using AI to make decisions about survivor eligibility, risk, prioritization, or services
- **Training**
    - Ensure workers understand both the benefits and risks of AI and know what to do when they are unsure.
- **Incident response**
    - Have a clear process for reporting and addressing suspected privacy breaches, accidental disclosures, or inappropriate AI use.
- **Regular review**
    - AI tools and their terms can change. Review and approved tools and organization policies regularly.

1. Be careful with AI features built into everyday software

AI is no longer limited to standalone chatbots. It may be built into tools we already use in our work, such as:

- Email
- Word-processing software
- Cloud storage
- Video-conferencing platforms
- Transcription tools
- Scheduling software
- Customer service systems
- Search engines
- Phones and operating systems

This can make AI use easy to overlook. “It’s already part of our software” does not mean it is automatically approved for confidential information. Organizations should identify where AI is embedded in the technology they already use and assess those features as part of their privacy and security practices.

2. What to do if confidential information is accidentally entered

Do not assume that nothing can be done. If a worker accidentally enters confidential or personal information into an AI tool:

1. Stop using the tool for that information.
2. Do not try to cover up the mistake.

3. Notify the appropriate person within your organization, such as your privacy officer, supervisor, executive director, or designated privacy contact.
4. Follow your organization's privacy breach and incident-response procedures.
5. Document what happened, including what information was entered, which tool was used, when it happened, and what actions were taken.
6. Get appropriate privacy or legal advice where necessary.

Privacy regulators provide guidance and resources for organizations responding to privacy breaches. The appropriate response may depend on the information involved, the organization, the applicable law, and the specific AI service.

Summary

AI can be a useful workplace tool, but convenience is not a sufficient reason to expose survivor information to any third-party technology system.

Anti-violence organizations have a responsibility to consider privacy, confidentiality, security, legal obligations, survivor safety, and organizational accountability when deciding whether and how to use AI.

A privacy-first approach does not require workers to fear technology. It requires organizations to understand the technology being used, assess the risks, establish clear boundaries, and keep people, not technology, at the centre of survivor advocacy.

Technology-Facilitated Gender-Based Violence (TFGBV) is part of a continuum of violence that can be both online and in-person. If you or someone you know is experiencing TFGBV, you are not alone. You can use ShelterSafe.ca to find a shelter/transition house near you to discuss options and create a safety plan. You don't need to stay in a shelter to access free, confidential services and support.