Travailler avec des personnes survivantes qui utilisent l'IA: Guide pratique à l'intention du personnel antiviolence

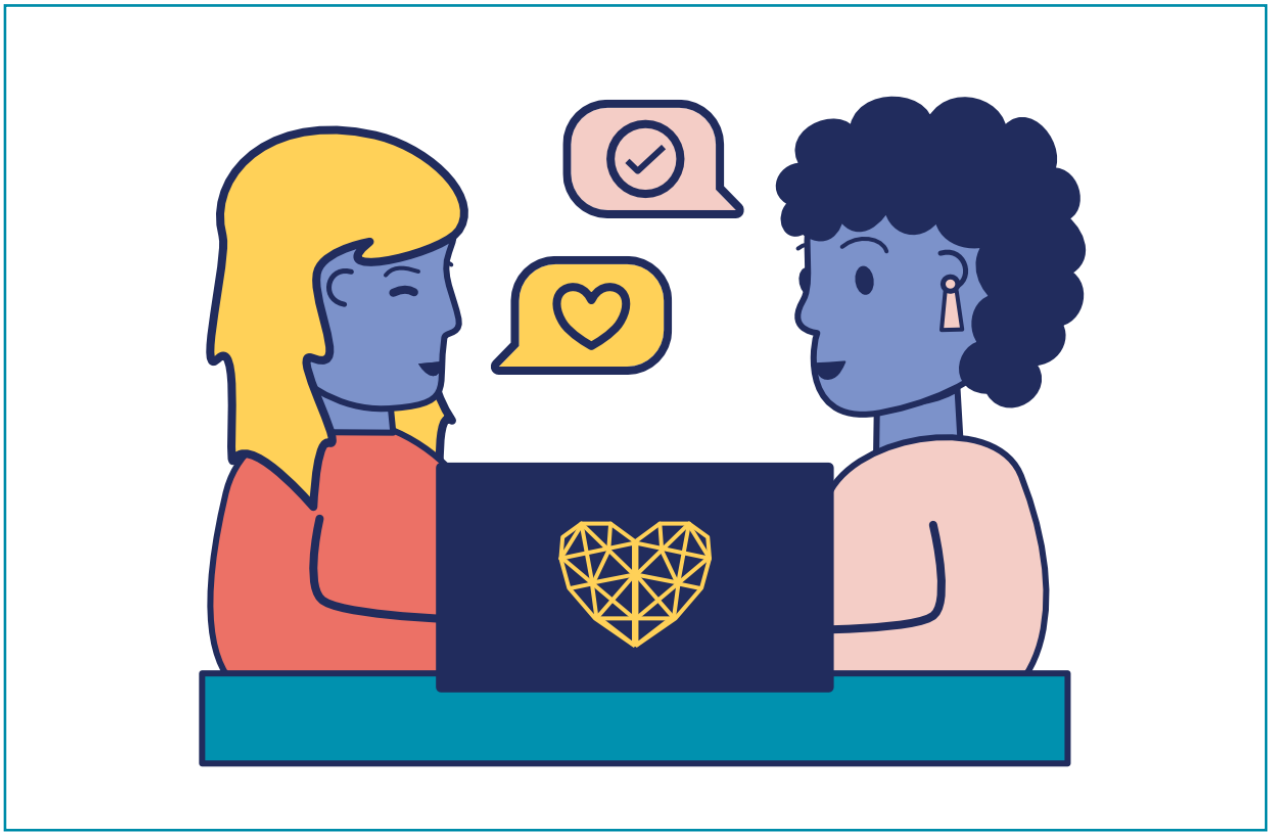


Table des matières

Comment l'IA peut faire défaut aux survivantes 3

1. L'IA peut fournir des informations inexactes ou inventées de toutes pièces 3
2. L'IA dispose de connaissances locales limitées 3
3. L'IA comporte des risques pour la vie privée 4

Sujets de conversation 4

1. Valider le récit de la survivante 4
2. Vérifiez ensemble les informations importantes 5
3. Concentrez-vous sur le risque, pas sur la technologie 6

Quand il faut redoubler de prudence 7

Les survivantes peuvent utiliser des chatbots basés sur l'IA pour obtenir des informations sur les procédures judiciaires, la planification de sécurité, les services disponibles, le logement, l'aide financière ou d'autres formes de soutiens. L'IA peut séduire, car elle est immédiatement disponible et fournit des réponses sous une forme conversationnelle et accessible.

Cependant, les informations générées par l'IA peuvent s'avérer erronées, incomplètes, obsolètes ou ne pas correspondre à une situation particulière. Lorsqu'une survivante croit fermement aux informations fournies par l'IA, le personnel peut avoir du mal à corriger les erreurs sans nuire à la relation de confiance.

L'objectif n'est pas de convaincre les survivantes que l'IA est toujours néfaste ou que le personnel a toujours raison. Au contraire, le personnel peut aider les personnes ayant subi un traumatisme à prendre du recul, vérifier les informations importantes et prendre des décisions éclairées tout en préservant leur autonomie.

Comment l'IA peut faire défaut aux survivantes

1. L'IA peut fournir des informations inexactes ou inventées de toutes pièces

Les chatbots basés sur l'IA peuvent générer des informations qui semblent correctes, mais qui sont en réalité erronées. Ils peuvent fournir des informations inexactes concernant la législation, les procédures juridiques, les procédures judiciaires, les délais ou les services. Ils peuvent également citer des affaires, des lois ou des sources qui n'existent pas. Ils peuvent même contester vos corrections ou valider une réponse erronée. Une réponse détaillée ou convaincante ne constitue pas une preuve de l'exactitude de l'information.

Voici un exemple d'IA générant des informations erronées (en anglais).

2. L'IA dispose de connaissances locales limitées

Il se peut que l'IA ne sache pas ce qui se passe actuellement dans la communauté d'une survivante. Par exemple, l'IA ne peut pas nécessairement savoir si une maison d'hébergement dispose d'un lit aujourd'hui, si un programme local a une liste d'attente ou si un service a récemment modifié ses critères d'admissibilité.

3. L’IA comporte des risques pour la vie privée

Les survivantes peuvent saisir dans un outil d’IA des informations très personnelles concernant des abus, leurs enfants, leur situation financière, des questions juridiques ou leur localisation. Selon l’outil et ses paramètres, les informations partagées avec une IA peuvent être conservées pendant une longue période, et les entreprises peuvent avoir accès aux conversations. Cela signifie qu’une conversation pourrait être utilisée pour entraîner d’autres chatbots, vendue à une autre entreprise technologique ou divulguée en cas de réception d’une assignation à comparaître. Cela peut entraîner de graves risques en matière de vie privée et de confidentialité pour les survivantes.

Sujets de conversation

Voici quelques conseils pratiques lorsque vous soutenez une survivante qui utilise un outil d’IA ayant fourni des informations erronées.

1. Valider le récit de la survivante

Une survivante a peut-être eu recours à l’IA parce qu’elle avait besoin d’informations rapidement, en toute confidentialité, ou lorsque d’autres formes de soutien n’étaient pas disponibles. Vous pouvez reconnaitre que le chatbot lui a été utile sans pour autant assumer que les informations fournies sont exactes.

Au lieu de:

«On ne peut pas faire confiance à ChatGPT.»

Vous pourriez dire:

«On dirait que le chatbot vous a été utile quand vous en aviez besoin. Ça vous dit de regarder ensemble les informations qu'il vous a fournies?»

Si une survivante est fermement convaincue que la réponse de l'IA est correcte, lui répéter sans cesse qu'elle est erronée peut lui donner l'impression d'être ignorée ou la rendre moins disposée à s'engager. Restez curieuse et faites en sorte que la vérification des informations soit une démarche que vous menez **avec** la survivante, et non **pour** elle.

2. Vérifier ensemble les informations importantes

L'IA peut constituer un point de départ utile, mais il convient de vérifier les informations importantes avant qu'une survivante ne s'appuie sur celles-ci pour prendre une décision.

Si la survivante se sent à l'aise pour partager sa réponse, examinez ensemble les éléments suivants:

- Quelle question ou quelle information a été fournie à l'IA?
- Ce qu'a répondu l'IA
- Est-ce que l'IA a fourni une source?
- Pouvez-vous non retrouver la source originale?
- Est-ce que l'information est à jour?
- Est-ce que cela s'applique ou non à la province ou au territoire de la survivante?

- Est-ce que l'information provient d'une source officielle ou fiable, telle qu'un tribunal, une administration, un organisme juridique ou le site web d'un prestataire de services?
- Que pourrait-il se passer si les informations étaient erronées?

La manière dont une question est posée peut influencer la réponse. L'IA peut également donner l'impression d'être sûre d'elle ou d'approuver les hypothèses contenues dans une question, même lorsque les informations sont incomplètes ou erronées.

Une fois que vous aurez aidé la survivante à vérifier les informations importantes, celle-ci pourra décider de la suite à donner.

3. Se concentrer sur le risque, pas sur la technologie

Plutôt que de vous demander si un outil d'IA donné est «bon» ou «mauvais», concentrez-vous sur les conséquences possibles au cas où les informations fournies sont erronées ou incomplètes.

Par exemple:

«L'IA peut fournir des informations générales sur le droit canadien, mais elle ne connaît pas forcément les règles applicables dans votre province, devant votre tribunal ou dans votre situation particulière. Vérifions ces informations avant de s'y fier.»

«Un chatbot peut vous fournir des informations sur les maisons d'hébergement ou les services disponibles, mais il ne peut pas savoir si une place est disponible maintenant. Vérifions ça tout de suite.»

«Étant donné que ces informations pourraient avoir une incidence sur vos droits ou votre sécurité, vérifions-les avant que vous ne preniez une décision en vous basant sur la réponse de l'IA.»

Quand il faut redoubler de prudence

L'IA peut aider les survivantes à trouver les questions à poser. Cela ne devrait pas être leur seule source d'information pour tout ce qui pourrait avoir une incidence sur leur sécurité, leurs droits ou leur accès aux services.

Encouragez-les à vérifier les informations générées par l'IA avant d'agir lorsqu'elles concernent:

- Droits ou obligations juridiques
- Procédures judiciaires, dates limite ou documents
- Ordonnances de protection ou autres voies de recours
- Procédures policières ou pénales
- Immigration ou statut juridique
- Questions relatives à la protection de l'enfance ou à l'éducation des enfants
- Disponibilité de logements ou d'hébergements
- Aide financière ou prestations
- Sécurité personnelle
- Soins médicaux ou de santé mentale
- Partage de preuves ou d'informations personnelles
- Toute décision pour laquelle des informations erronées pourraient accroître le risque

Pour plus d'informations sur la manière d'aider les parties non représentées à utiliser l'IA en toute sécurité, consultez le document du Rise Women's Legal Centre intitulé AI for Unrepresented Parties: Minimizing Risk.

N'oubliez pas: Votre rôle n'est pas de rivaliser avec l'IA en matière de crédibilité. L'objectif est d'aider la survivante à prendre des décisions en toute connaissance de cause, tout en préservant la confiance, la sécurité et l'autonomie.

La violence fondée sur le genre facilitée par la technologie (VFGFT) s'inscrit dans un continuum qui peut se manifester aussi bien en ligne qu'en personne. Si vous ou quelqu'un que vous connaissez êtes la cible de VFGFT, sachez qu'il existe de l'aide. Vous pouvez consulter hebergementfemmes.ca pour trouver une maison d'hébergement près de chez vous afin d'explorer vos options et d'élaborer un plan de sécurité. Il n'est pas nécessaire de séjourner dans une maison d'hébergement pour bénéficier de services et d'un soutien gratuits et confidentiels.