Social Inclusion
2025 • Volume 13 • Article 9998
https://doi.org/10.17645/si.9998

cogitatio

ARTICLE

Open Access Journal

Ethical Implications of AI-Driven Chatbots in Domestic Violence Support

Hanna Mielismäki and Marita Husso

Faculty of Social Sciences, Tampere University, Finland

Correspondence: Hanna Mielismäki (hanna.mielismaki@tuni.fi)

Submitted: 30 January 2025 **Accepted:** 26 June 2025 **Published:** 2 September 2025

Issue: This article is part of the issue "Fostering the Socially and Ecologically Sustainable Digitalisation of Welfare States" edited by Paula Saikkonen (Finnish Institute for Health and Welfare) and Marta Choroszewicz (University of Eastern Finland), fully open access at https://doi.org/10.17645/si.i514

Abstract

Our study explored the opportunities, challenges, and ethical considerations of using artificial intelligence (AI)-driven chatbots in domestic violence (DV) support. DV is a serious public health and social problem. Identifying it as early as possible is important in violence prevention. However, victim-survivors of DV are often reluctant to disclose violence, and service practitioners may lack the capacity or confidence to address violence-related issues. To tackle these challenges, the use of AI-driven chatbots presents opportunities to address DV by providing information and guiding users to appropriate services. However, interactions between humans and AI systems lie at the intersection of the human need for practical assistance and the risks inherent in digital communication—raising ethical considerations, particularly in the vulnerable context of DV. Semi-structured interviews with 25 victim-survivors, DV professionals, and criminal justice experts suggest that while the implementation of AI-driven chatbots can greatly enhance access to information, it also poses significant challenges related to safety and accountability. This is because interactions with chatbots lack essential elements for comprehensive situational assessment and documentation of DV cases, and for the establishment of a support network. These insights underscore the critical role of human interaction in addressing DV cases, while also highlighting the potential roles of chatbots as intermediate support systems for victim-survivors and as supplementary tools for welfare service practitioners in identifying different forms of DV and supporting the appropriate referral of cases. This study advances understanding of how AI-driven chatbots can be ethically and sustainably implemented in DV support systems.

Keywords

AI chatbots; digital ethics; domestic violence; gender-based violence; help-seeking; human-technology interaction; responsible AI; sustainability; technology-facilitated abuse; welfare services

1. Introduction: Current Challenges in Seeking Help for Domestic Violence

Domestic violence (DV) is a serious global public health and social problem. It may significantly damage the mental and physical health of victim-survivors of DV (Bellis et al., 2019; Ellsberg et al., 2015; Miller & McCaw, 2019; Stubbs & Szoeke, 2022). Furthermore, the widespread use of digital tools has introduced new forms of violence. For example, technology-facilitated abuse (TFA) involves controlling, monitoring, or harming victim-survivors through mobile phones and other digital devices (Boethius et al., 2023; Storer et al., 2023; Woodlock et al., 2020). This form of abuse enables DV perpetrators to continue exerting control even after separation, making them omnipresent in victim-survivors' lives. As a result, the abuse extends beyond physical encounters and intensifies the psychological distress experienced by victim-survivors (Afrouz, 2023; Harris & Woodlock, 2019). However, victim-survivors often hesitate to disclose their experiences to authorities—including social and healthcare practitioners, as well as the police—due to shame and the stigma associated with victimisation (Harris & Woodlock, 2019; Husso et al., 2021; Meyer, 2016; Piippo et al., 2021) or of being judged or not believed by service practitioners (Storer et al., 2022). In addition, victim-survivors are often unaware of the different forms of DV and which of those they are experiencing, their legal rights related to DV, and the procedures for disclosing DV to authorities (Decker et al., 2019; Saxton et al., 2021).

Research indicates that social welfare and healthcare service providers are in key positions to assess risks of violence, implement interventions, and organise support services (Husso et al., 2021; Miller & McCaw, 2019; Siltala et al., 2023). These practitioners assess individual circumstances, identify the issue's broader social context, and determine which other professionals or services are already involved (Trevithick, 2012). Collaborative, multidisciplinary approaches are particularly vital when working with clients facing complex challenges. The ultimate aim of these approaches is to empower and support the client—both materially and non-materially—in enhancing their own well-being (Trevithick, 2012). For example, victim-survivors typically need basic information about different forms of DV and guidance in accessing the necessary support services, including those related to housing and economic and mental recovery (Rogers et al., 2023).

Although DV, as a form of gender-based violence, has long been recognised as a human rights violation, there remains a pressing need for more effective interventions within the service system (Husso et al., 2021; Piippo et al., 2021; Turner et al., 2017; Wright et al., 2022). For example, in Finland, most victim-survivors remain unidentified by social and healthcare service providers despite using these services twice as often as individuals who have not experienced DV (Siltala et al., 2023). Practitioners typically consider DV-related problems either too challenging or too complicated to solve (Husso et al., 2021; Piippo et al., 2021). It is suggested that the insufficient actions of practitioners are due to their inadequate education on DV (Husso et al., 2021; Piippo et al., 2021; Siltala et al., 2023).

This gap in responding to DV appears to result from victim-survivors' reluctance to disclose their experiences of violence and service providers' limited capacity, confidence, or willingness to intervene in such situations. However, the physical, emotional, and social consequences of unaddressed and unrecognised DV may be serious for both victim-survivors and those witnessing the violence (Callaghan et al., 2018; Holt et al., 2008; Miller & McCaw, 2019; Stubbs & Szoeke, 2022). Early identification of DV is therefore essential for enhancing the safety of victim-survivors and preventing further harm (Siltala et al., 2023).

Artificial intelligence (AI) is increasingly being integrated into digitised welfare services, enabling novel approaches to addressing DV, such as the use of AI-driven chatbots to support cases that might otherwise remain unaddressed (Novitzky et al., 2023). These automated computer programs simulate human interaction through spoken, written, or visual communication and perform tasks that previously required a human agent (Henman, 2020). The interaction interfaces act as service assistants and are often integrated into service providers' websites or available as mobile applications. While general-purpose chatbots are designed to handle a wide range of topics, domain-specific chatbots are designed to meet specific needs, with their content typically curated by experts in the relevant field.

The benefits of these tools include handling routine activities, such as answering simple and repetitive questions and sharing information (Draughon Moret et al., 2022). The automation of routine tasks in welfare services has been justified by its potential to free up experts to focus on more demanding and critical responsibilities. Yet, distinguishing between tasks that require human judgment and those that can be automated remains challenging (Parviainen & Rantala, 2022). For example, Longoni et al. (2019) discovered that people perceive human practitioners as more capable of addressing their diverse needs and individual circumstances than machines. Nevertheless, the implementation of AI-driven chatbots has the potential to enhance service efficiency and end-user experience by providing instant responses to multiple user requests simultaneously (Nordheim et al., 2019). However, interactions between humans and AI systems lie at the intersection of the human need for practical assistance and the risks inherent in digital communication. This raises ethical considerations regarding the use of chatbots in the sensitive and vulnerable context of DV (Butterby & Lombard, 2025; Henry et al., 2024; McGreevey et al., 2020).

In this study, we explore the opportunities, challenges, and ethical considerations of using AI-driven chatbots in DV support from the perspectives of victim-survivors, DV professionals, and criminal justice experts. This research contributes to the ongoing discussion on the ethical and sustainable integration of chatbots into support systems for victim-survivors and to the broader agenda of sustainable development by addressing several of the United Nations (2015) SDGs. These goals include fostering societies free from fear and violence (SDG 16), advancing the use of technology to promote the empowerment of women (SDGs 5 and 9), enhancing the accessibility and reliability of welfare services (SDGs 10 and 16), and promoting health and well-being (SDG 3). Additionally, in 2024, the UN adopted a resolution to promote the safe, secure, and reliable development and use of AI systems that respect human rights and foster sustainable development, further emphasising the importance of ethical AI technology in achieving the SDGs (United Nations, 2024).

In studying AI-driven chatbots within the service system and DV support, an interdisciplinary approach is essential (Henry et al., 2024; Lindgren & Dignum, 2023). Our research draws on perspectives that are grounded in justice, equity, and the ethical use of technology in sensitive support contexts. We integrate insights from research on interpersonal violence and TFA with the normative ethical framework provided by the High-Level Expert Group on AI (AI HLEG, 2019), which outlines principles for the responsible development and use of AI. This combination enables a nuanced understanding of both the technological and human-centred dimensions of AI implementation in support services.

2. AI-Driven Chatbots as Tools for DV Support

Finnish people have the most positive attitudes towards AI among the European countries surveyed (Bergdahl et al., 2023). Finland pioneered digitalisation in Europe in the mid-1990s through governmental initiatives. Current government policy prioritises digital assessment of service needs and utilisation of AI in service guidance. Key public services are already fully available online, and further advancement of service digitalisation and automation is targeted (Finnish Government, 2024). Currently, 97.61% of Finnish internet users utilise eGovernment services, and 82% of the population have at least basic digital skills (European Commission, 2024a, 2024b).

As chatbots become more integrated into support services, it is important to understand how their design and use may influence users' experiences and well-being in both online and offline settings. Researchers have noted that trust in the offered digital service might be compromised if it fails in safety (Kretzschmar et al., 2019; McGreevey et al., 2020; Xu et al., 2021). Trust is considered a key element in the adaptation of new technology (de Visser et al., 2018; Jobin et al., 2019). For victim-survivors of DV who engage with chatbot services, perceptions of safety and reliability may affect their willingness to use such tools and the extent to which they find them helpful.

The European Union AI Act aims to regulate the development, deployment, and use of AI systems, with a particular focus on safety and reliability (European Union, 2024). In parallel, the AI HLEG (2019), established by the European Commission, developed the *Ethics Guidelines for Trustworthy AI*. These guidelines are grounded in fundamental rights and European values and define "trustworthy AI" as lawful, ethical, and robust throughout its lifecycle. However, the guidelines remain relatively general in nature, as the implementation of AI systems presents diverse challenges and risks across various fields. Therefore, technological solutions should be examined as part of a broader sociotechnical system (Abdelnour-Nocera & Clemmensen, 2019) to address the practical challenges and conflicts that may arise during the implementation of AI systems.

In this study, we examined the use of AI-driven chatbots in the context of DV and the support system for victim-survivors. Drawing on our findings, we considered how the ethical principles outlined by the AI HLEG (2019) could inform the implementation of these systems in supporting victim-survivors. Our research question was: What are the opportunities, challenges, and ethical considerations of using AI-driven chatbots in DV support? We focused on the four main principles proposed by the AI HLEG: prevention of harm, fairness, explicability, and respect for human autonomy. These guidelines emphasise both technical robustness and the user's safety, well-being, and human rights.

Regarding the first principle, the AI HLEG (2019) emphasises that technological innovations are expected to contribute to individual and societal well-being, while also minimising the risk of harm, whether intentional or unintentional (see also Braunschweig & Ghallab, 2021; Floridi et al., 2018). This is of particular concern in this study, as AI-driven digital tools for help-seeking should not further disadvantage those who are already in a vulnerable position. One major challenge is that the same digital tools that enable victim-survivors to seek help may also allow perpetrators to exert harmful power and control—for example, by hacking devices, installing spyware, or monitoring victim-survivors' online activity via mobile phones and computers (Afrouz, 2023; Boethius et al., 2023; Woodlock et al., 2020). Closely related to these risks, privacy and data security

are relevant concerns, as victim-survivors may disclose sensitive personal information when interacting with chatbots (Butterby & Lombard, 2025; Henry et al., 2024).

The second principle, fairness, is outlined by the AI HLEG (2019) as emphasising the need for AI systems to operate in ways that support just, fair, and unbiased interactions with users. According to the international ethical principles of social work (International Federation of Social Workers, 2018), fairness involves a commitment to equity and inclusion, including efforts to reduce barriers to access and to ensure that support reaches vulnerable or marginalised populations. In the context of DV, fairness refers not only to equal treatment but also to equity, which requires recognising and addressing structural inequalities that may affect different user groups. For example, the data used to train chatbots may reflect existing cultural, gendered, or societal biases, which can inadvertently reinforce stereotypes or marginalise victim-survivors (Bailey & Burkell, 2021). To address these concerns, Kretzschmar et al. (2019) note that the training data and outputs should be evidence-based, inclusive, and accountable. Furthermore, scholars have drawn attention to the importance of ensuring that individuals with varying levels of digital literacy, language proficiency, or disabilities can access and benefit from these tools (Mishna et al., 2021; Taylor, 2017). This perspective aligns with the AI HLEG's (2019) principle of fairness, which includes non-discrimination and equal participation. Accordingly, the implementation of chatbots can promote more equal access to information and services, as they have been found to be especially useful for people who might not otherwise seek help (Kabacińska et al., 2022; Montagni et al., 2020; Torous et al., 2021). This is particularly relevant for victim-survivors who are isolated due to their abusers' coercive control, which involves restricting their movements and limiting their access to resources and support networks (Al-Alosi, 2020; Novitzky et al., 2023).

Regarding the third principle, explicability, the AI HLEG (2019) has pointed out that the requirement for it in AI systems is context-dependent and becomes increasingly critical when the consequences of erroneous or misleading outputs are severe. In the context of DV, automated support tools, such as AI-driven chatbots, may pose risks to victim-survivors' decision-making processes, particularly if they place undue trust in the system's outputs or perceive it as capable of providing optimal advice or solutions (AI HLEG, 2019; Henman, 2020). Users should be able to distinguish whether they are interacting with a human or an AI system, as this distinction is found important for supporting informed consent and fostering trust (AI HLEG, 2019; Xu et al., 2021).

Finally, the fourth principle—respect for human autonomy—emphasises the importance of fostering equitable interaction and empowering victim-survivors to make informed decisions with the support of AI systems, particularly in contexts involving vulnerable groups (AI HLEG, 2019). For instance, the ability to seek help anonymously through chatbots can reduce victim-survivors' fear of humiliation or judgment by service practitioners when they disclose deeply personal and emotionally sensitive experiences (Storer et al., 2023). This sense of anonymity has been shown to empower victim-survivors and increase their confidence in taking further steps to access support services (Al-Alosi, 2020; Novitzky et al., 2023). However, manipulative forms of guidance by AI systems have been identified as a potential threat to autonomous decision-making, as they may limit users' ability to act freely and independently (Devillers et al., 2021).

3. Methods

This study was conducted as part of a Horizon Europe project titled Innovative Solutions to Eliminate Domestic Abuse (ISEDA), which involved partners from nine European countries: Bulgaria, Catalonia (Spain), Cyprus, Greece, Finland, France, Germany, Italy, and Scotland. One of the key objectives of this project was to develop and pilot a chatbot for women victim-survivors of DV. Seven countries involved in the project conducted focus groups and/or individual interviews to better understand the challenges in seeking help for DV, as well as the positive and negative perceptions of chatbots.

Focus groups were considered suitable for this purpose, as they enabled participants to articulate their views through interaction with others in a shared setting. To support open and productive discussion, the groups were composed to be relatively homogeneous in terms of participants' background and familiarity with the topic, which is known to facilitate communication (Bryman, 2004; Fern, 2001). At the same time, differences in individual experiences contributed to a range of perspectives on the phenomenon (Kitzinger, 1994). First, we asked the participants about their perceptions of the challenges in seeking help for DV from public services, including social and healthcare services and the police. Then, we asked them to reflect on the opportunities, challenges, and risks associated with the use of AI-driven chatbots to seek help for DV. We defined a chatbot for them as a computer program designed to simulate conversations with humans.

In this study in Finland, we conducted 13 semi-structured interviews in 2023. In total, 25 informants were interviewed, either individually or in groups of 2 to 3. The details are outlined in Table 1. The interview format was selected based on participants' preferences and practical factors such as availability and comfort in discussing sensitive topics (Lazar et al., 2017).

Table 1. Interview formats and informant categories.

	Focus group interviews		Individual interviews
	Number of groups	Informants per group	
Victim-survivors	1	2	3
DV professionals	5	2–3	1
Criminal justice experts	1	3	2

The victim-survivors were women over 18 years of age whom violence prevention organisations selected for the interviews. These women acted as peer support in these organisations. The organisations ensured that these women were not under threat of violence before inviting them for an interview. The professionals and experts represented civil society organisations, as well as national and regional governmental administrations from different parts of Finland. Many of them had extensive experience in the field of DV and maintained a broad network within the relevant service system, including social and healthcare workers and the police.

We provided all the informants with detailed information about the purpose of this study and obtained their verbal informed consent to participate prior to conducting the interviews. Each interview lasted 53 to 82 minutes. We recorded a total of 14 hours of audio material and transcribed it for analysis. We removed all identifying information before completing the analysis and identified the informants only by their interview numbers (e.g., G3 for Group 3) and participant numbers (e.g., R2 for Respondent 2). The original interview data are in the Finnish language and the interview extracts used in this article are English translations of them.

We analysed the data using theory-driven thematic content analysis (e.g., Braun & Clarke, 2006), not directly based on theory but guided by concepts and findings on the topic from previous research. This approach enabled us to identify and examine patterns in the informants' understanding and experience of the process of seeking help for DV, and their perceptions related to it. The goal was to identify relevant ideas that repeatedly arose from the data.

In the first phase of the analysis, we repeatedly read the entire dataset to understand all aspects of the data. In the second phase, we systematically coded the dataset by identifying and labelling facilitating and challenging elements related to help-seeking for DV, interactions with authorities, and the use of chatbots. In the third phase, we began organising these codes into potential themes and sub-themes, using tables to support the sorting and comparison process. In the fourth phase, we reviewed the coded interview extracts for each theme to ensure their coherence and consistency with the research question. In the fifth phase, we examined the themes based on both the literature and the informants' perspectives and then deepened the analysis to identify the opportunities and risks associated with the use of AI-driven chatbots. Finally, in the sixth phase, we reflected on the findings through the lens of the AI HLEG's (2019) four ethical principles: prevention of harm, fairness, explicability, and respect for human autonomy.

This entire study complied with the ethical guidelines of the Finnish National Board on Research Integrity TENK (2023). The Ethics Committee for Human Sciences in the Tampere Region issued a statement on this study's ethical acceptability. Additionally, the necessary research permits were obtained from the interviewees' respective organisations.

4. Results: Opportunities and Challenges of AI-Driven Chatbots in DV Support

This section presents the interviewees' perceptions of AI-driven chatbots in the context of DV support, highlighting both opportunities and challenges associated with their use. The identified opportunities are (a) providing information and guidance during the early stages of DV, (b) offering a non-judgmental space for sharing experiences, and (c) reducing the emotional strain of violence disclosure. The identified challenges are (a) absence of case documentation, (b) lack of holistic situational assessment, and (c) concerns related to accountability and safety.

4.1. Providing Information and Guidance During the Early Stages of DV

All the interviewees emphasised the importance of knowledge about the different forms of DV and its multifaceted impacts on victim-survivors, including its physical, psychological, and social consequences. However, they expressed disappointment in what they perceived as the authorities' insufficient knowledge of DV. In contrast, they viewed AI-driven chatbots as more consistent in providing standardised and timely information, which they considered more reliable in certain situations. Accordingly, participants viewed these systems as low-threshold support tools during the early stages of DV, when abuse is just beginning to occur, to help prevent violence from becoming normalised in victim-survivors' lives. The following criminal justice expert pointed out an important aspect of the challenges facing those who disclose violence:

> A computer-based channel like this, where a person can reflect on their own experiences safely and there is no threshold for whether they are believed or whether their case is too trivial, or whether they

> are burdening these professionals unnecessarily, will probably be able to dispel human doubts about whether the person's situation is serious enough to warrant help. (Criminal justice expert, G1R1)

Victim-survivors often consider themselves not "ideal" victims worthy of support (Meyer, 2016) and feel that they have the burden of proof of being targets of violence (Piippo et al., 2021). The interviewees also expressed that victim-survivors need information on support services, including legal, medical, and psychological support. Accordingly, they perceived that the use of AI-driven chatbots in DV support could lower the threshold for seeking help by providing accessible information and guidance. This, in turn, was seen as potentially encouraging victim-survivors to reflect on their situations and make more informed decisions about how to proceed. The victim-survivors expressed that knowledge about DV is essential for them to prevent violence from becoming normalised, which Piippo et al. (2021) found to be the case for many victim-survivors. For example, the following interviewee described how she needed guidance in these matters instead of being belittled by service practitioners:

> I feel that I was not taken seriously. I was hushed up and told that it would be alright. And now, years later…it's not OK. If someone had taken me seriously and told me about my rights and the rights of children, what's right and what's wrong, that would certainly have helped. (Victim-survivor, G10R1)

> We victims are terribly skilled at not seeing the problem as big as it actually is, especially in the early phase. At that phase, AI could be useful. (Victim-survivor, G10R2)

These quotes highlight an important aspect of experiencing violence: victim-survivors may question their experiences and thereby prolong their suffering. For example, abusers may use coercive control, distort their targets' perceptions of reality, and erode their autonomy (Stark, 2007). The institutional practices of bypassing and silencing problems related to DV that G10R1 described may enforce the normalisation of violence.

Accordingly, most of the participants mentioned that the use of chatbots could facilitate the disclosure of DV by guiding victim-survivors on how to proceed with their cases. In addition to supporting victim-survivors directly, both the DV professionals and the criminal justice experts considered chatbots as potential tools to assist the social and healthcare practitioners and the police in facilitating the disclosure of DV and proceeding systematically with DV cases. The following professional described the systemic gap in which victim-survivors hesitate to disclose violence, while service practitioners may not be able to properly identify a case of DV—highlighting how chatbots could also support professionals in recognising and responding to such cases:

> A person who repeatedly visits health care services with various symptoms but is never asked about violence and does not disclose it themselves—we have considered this a situation where AI could assist: to guide, explore, and ask about the theme of violence. (DV professional, G3R3)

> I believe AI can certainly ask specific questions, perhaps even better, and will remember to ask all of them, provided the settings ensure that all necessary questions are included. In this way, human omission can be avoided. (Criminal justice expert, G2R1)

G2R1 described the mechanical feature of the AI systems as a benefit. This insight illustrates the diverse needs of different users in different contexts. This expert referred to the preliminary questioning, where contextual

matters, such as the hectic hours of a weekend, may affect a police officer's ability to routinely conduct all necessary questions.

4.2. Offering a Non-judgmental Space for Sharing Experiences

The interviewees perceived AI-driven chatbots as providing a supportive space for victim-survivors to share their experiences without the fear of being judged by the other party. The victim-survivors noted that the other party's supportive interaction and acknowledgement of their feelings are essential in vulnerable situations, as these show that their experiences are taken seriously. They described many ways in which victim-survivors are often ignored by authorities, including social and healthcare workers and the police. These approaches were associated with negative interactions, such as belittling and victim-blaming, which were perceived as negatively affecting the likelihood of them getting the help they need. In contrast, the interviewees perceived AI systems as lacking such attitudinal attributes:

> There is, of course, a risk if AI also ends up belittling the seriousness of the problem. That is, what kinds of questions and response options are created for early-stage support—and what kinds of paths the AI then leads the user into—there are enormous risks in that. But somehow, I see that these risks are not as big…as if you call the family counselling centre or the deaconess of a church and then they say: "Nonsense, try to get along." So, in a way, the bot would never respond like that. It would aim to resolve the situation. (Victim-survivor, G10R1)

> It significantly lowers the threshold when you do not have to disclose violence to potentially contemptuous and arrogant male police officers. Instead, if this process could be neutral, most people would likely prefer to tell an AI what has happened rather than fear being belittled for their experiences, which creates a very high threshold. (DV professional, G7R1)

The fear of being belittled by service practitioners hinders victim-survivors from disclosing violence, as seen in these quotes. G10R1 implied the importance of actions and the need to solve problematic situations. Acknowledging victim-survivors' experiences is essential for empowering them and helping them make sense of their situations (Keeling & Fisher, 2015; Wood et al., 2022). Without validation of their feelings, they may question their experiences and become accustomed to living in violent situations (Piippo et al., 2021).

However, the victim-survivors also expressed the need for genuine human connections when facing life challenges and being in vulnerable situations. The following quotes highlight the contrast between programmed interaction and human interaction:

> The victim has a need to be heard, and with a bot, they do not necessarily feel that they are genuinely heard, as it is only a program that answers as it is programmed to. (Victim-survivor, G8R1)

> At that moment, the victims long especially for humanity from these social and healthcare services. The conversational bot is anything…but humane….Typically, the victim feels alone in their situation and as if they are fighting against windmills. Especially in these situations, they long for human contact and to be seen as human. (Victim-survivor, G8R1)

These quotes emphasise the importance of distinguishing between hearing and listening. Whereas listening is often perceived as an easy or passive task, it is, in fact, a complex skill that requires conscious effort and practice to build trust and understanding (Trevithick, 2012).

4.3. Reducing the Emotional Strain of Violence Disclosure

According to the interviewees, fear, as a human emotion, can act as a barrier to accessing support, as, in their experience, it was linked to practitioners' reluctance to engage with DV issues. In contrast, the interviewees believed that the use of AI-driven chatbots could facilitate help-seeking by automatically presenting the necessary questions about violence, raising hopes for a more systematic handling of cases. For example, Husso et al. (2012) and Piippo et al. (2021) found that social workers may fear asking clients about DV because they worry about their clients' reactions or the potential emotional burden it places on the practitioners themselves, such as anxiety, frustration, or sadness:

> People might fear hearing about violence, which can prevent them from wanting to get involved. And I think that's quite common, at least, based on my experience….AI doesn't have that need to protect itself, which can lead to better solutions. (Victim-survivor, G9R1)

> There could be something that removes the human element in such a way that it doesn't always depend on someone's opinion whether the matter progresses or not. For example, some kind of entirely computer-based, AI-driven form system that would automatically score risks. (Victim-survivor, G9R1)

When the victim-survivors discussed interactions with human practitioners and AI, they often pondered the meaning of humanity and the non-human attributes of AI in interactions. The presence or absence of humanity seemed to both facilitate and hinder access to support. While the human practitioners' empathetic approach was considered a fundamental requirement for successful encounters, human emotions were seen as a burden to the successful progression of processes. As seen in the quotes, the victim-survivors regarded AI as a tool that can be designed to minimise subjective variability in interactions and to handle their cases more systematically.

The interviewees acknowledged that DV is deeply institutionalised and reflected in the attitudes of service practitioners. However, the interviewees also acknowledged the risk of human biases being embedded in AI systems through coding and training processes, as one DV professional noted:

> You have to be very careful about what database a chatbot uses so that the general biases of people are not transferred to it. (DV professional, G6R1)

This quote implies the variety of chatbots and the challenge of knowing which of them is reliable (Kabacińska et al., 2022; Montagni et al., 2020). Draughon Moret et al. (2022) found that the quality of apps varies significantly. Nevertheless, some of the interviewees saw more potential for eliminating negative bias from AI systems than from human practitioners.

4.4. Absence of Case Documentation

The victim-survivors regarded anonymous interaction with chatbots as beneficial for those who are not yet ready to take further steps in their current situation. This allows them to gain knowledge anonymously and decide whether to proceed with their cases, thereby avoiding unwanted consequences from authorities, such as child welfare notifications or the child being taken into care by social services. The victim-survivors explained that they might fear disclosing violence due to concerns about case documentation, as one interviewee expressed:

> The church deacon is the only person whom you have the courage to approach and talk to,…not the authorities, because they always document and save everything. And the consequences are much worse than…if you simply do not seek help and just stay quiet….if only you can just…seek help if you are in a violent relationship without having to fear that if you…disclose violence, they will take your children away or you will end up in a terrible hassle. (Victim-survivor, G5R1)

This quote shows a tension between the need to seek help and the fear of consequences. It also illustrates the pressure victim-survivors face when they have to stay in a violent relationship due to the fear of disclosing violence. This shows that victim-survivors need more decision-making power regarding how their cases are handled by authorities. However, avoiding documentation contradicts the professionals' views, which highlight the importance of documentation, as described in the following quote:

> How does the information transfer and get recorded or documented if the conversation with the AI doesn't lead to anything at that moment, but it leads to something two months later, or if the same person seeks help multiple times?...One would always hope to know at what point a human should get involved….Banks already operate like this. Just yesterday, I dealt with a bank bot first, and the bot concluded that a human was needed. (DV professional, G12R2)

The professionals noted that repetitive incidents of violence would remain undocumented if victim-survivors interacted only with chatbots. DV is typically an ongoing process rather than a single incident (see Stark, 2007); however, individual contacts with chatbots remain separate and do not form a client case that allows for the tracking of history or progression. The professionals acknowledged this nature of the phenomenon of DV by pointing out the need to regard DV cases as a process. Furthermore, because violence affects the entire family (Callaghan et al., 2018; Holt et al., 2008), the professionals were especially concerned about potential emergency child protection needs and the limitations of chatbots in identifying the criticality of situations:

> In these relationships, there are, on average, at least two to three underage children, especially among those that have lasted a long time. So, how does the situation unfold, considering that the need for child protection can be immediate, especially in urgent situations like these? (DV professional, G12R1)

Failing to recognise immediate risks and the need for instant reaction could endanger the safety and well-being of victim-survivors and their children. According to national statistics in Finland, in 65% of cases where violence has been recurring, children have been repeatedly exposed to violence in their families (Attila et al., 2023). However, the need for child protection may remain undiscovered with the use of chatbots, as expressed by DV professionals.

4.5. Lack of Holistic Situational Assessment

DV professionals expressed scepticism regarding the capacity of chatbots to adequately respond to crises and potential traumas caused by violence, which they viewed as requiring a dynamic and multidimensional approach. For example, one professional referred to the seriousness of DV, with potential lethal consequences, and raised concerns about the extent to which chatbots can execute the required critical situational assessments, whereas human practitioners were considered to have gained deep empirical knowledge through long experience in assisting victim-survivors:

> The assessment of the risk of violence is very dynamic and highly multidimensional. It includes a strong qualitative element, and although there are certain systematic aspects that can be considered, there is also a significant amount of experiential knowledge that develops only through direct encounters and over the years. So, when we talk about threats to life and health, to what extent can a chatbot reliably make such an assessment? (DV professional, G12R2)

As shown in this quote, recognising and responding to violence-related problems requires profound interaction and holistic observation of situations. According to Trevithick (2012), the initial meeting with a client should emphasise the use of open-ended questions to gain a deeper understanding of the client's situation and circumstances, including the broader social context. While chatbots are considered beneficial in answering simple and repetitive questions (Draughon Moret et al., 2022), in DV cases, there are no simple answers in diverse situations, as the following professional noted:

> At least, from the perspective of harassment, these cases are so terribly complex, and I don't know if there are any straightforward answers in the end. The amount of data that would need to be input is quite considerable. Even though I have had such a long work history in this field, almost every day, I encounter questions that make me think: "Wow, I've never heard of that before" and "I don't really know what advice to give or what to do here; let's think about it together," as there aren't really any good solutions. (DV professional, G7R2)

This quote highlights another essential element of situational assessment: Practitioners also use creative problem-solving to assess various solution options in complex DV cases where potential risks need to be assessed from various angles. Considering the amount of data required to address all nuanced DV scenarios, doubts arose about the capability of AI-driven chatbot systems to perform these kinds of tasks. Furthermore, the DV professional referred to the overwhelmingly complex nature of stalking cases, as digital devices have enabled the continuation of violence even after separation.

Accordingly, a substantial element of holistic assessment includes observing signs of TFA, about which most of the interviewees raised concerns. The intimate matters and vulnerabilities of victim-survivors may be exposed to hacking, misuse of information, or the installation of spyware by a current or ex-partner:

> What evidence do women have of this? If it's not bruises on their bodies, then it's in their stories or, just like R2 said, they should be able to show if there are message threads, or if a tracking program has been installed on their phone, or cameras at home, like last time, when cameras were eventually found in every room upon investigation. But how does the bot verify these? (DV professional, G3R3)

This form of abuse can significantly impact the safety and privacy of victim-survivors, making it essential to recognise the signs of abuse also in the environment where victim-survivors live. Besides determining the criticality of victim-survivors' circumstances, the situational assessment must succeed in preventing the escalation of violence. The DV professionals highlighted the need for more holistic observations than a mechanical assessment provided by AI, as it may not capture important signs of abuse, including nonverbal cues:

> How can victim-survivors recognise their own situation and the related issues if they have lived in it for a long time? Of course, the professional also makes other observations besides just ticking boxes. All of that can easily be missed if it's entirely done by AI, replacing the human. I see more risks in that. (DV professional, G3R1)

G3R1 referred to the tendency of violence to become normalised in victim-survivors' lives (see Husso et al., 2012; Piippo et al., 2021). However, violence typically becomes more serious over time, making early identification of DV essential and potentially life-saving. Such identification often relies on holistic observations, including listening, body language, and gaze direction (Reith-Hall & Montgomery, 2022). Digital technologies are not considered capable of emulating human characteristics in this sense, as they cannot read body language or perceive hurt and pain (Storer et al., 2023).

While the interviewees highlighted that chatbots may lack the capacity to appropriately process the complex and nuanced situations of victim-survivors, this challenge was described as even more pronounced for victim-survivors from immigrant backgrounds, who often face linguistic barriers and cultural differences in the meanings of concepts. Concerns were raised about chatbots potentially misinterpreting victim-survivors or providing incorrect instructions, especially when the language used was not standard or contained errors:

> If the client writes something incorrectly, will they get the right kind of help or the right answers? Or will they get something completely different from what they need? Also, if it's the first contact and you send a message and it doesn't work and feels frustrating, will you then seek help from anywhere else, or will you think, "I just can't do this anymore—I'm done"? (DV professional, G4R2)

> Whenever I've tried using a chatbot, I've never gotten the right answer through it. Especially the fact that you can't have any typos if you're trying to write in Finnish. It doesn't understand if you write something incorrectly. (DV professional, G4R1)

These reflections illustrate how easily misunderstandings and misguidance can occur in chatbot interactions, particularly for users with limited language proficiency. The inability of chatbots to process non-standard language or accommodate linguistic variation was identified as a significant barrier to effective support. Thus, special attention should be paid to language, imagery, and wording in the chatbot development process (Butterby & Lombard, 2025). Ensuring that these elements are carefully considered can help mitigate some of the risks associated with chatbot interactions.

4.6. Concerns Related to Accountability and Safety

The DV professionals were concerned that the responsibility for initiating contact with support services might fall on the victim-survivors once they had been provided with contact information by a chatbot. This was of special importance because the participants acknowledged the vulnerable position of victim-survivors due to the threat of violence and their potentially lowered functional capacity due to violence. Hence, there is a contradiction between their vulnerability and the expected strong agency to take action to advance their case.

> It seems that when interacting with an AI chatbot, much of the responsibility—including assessing the situation and maintaining the connection—is left to the person seeking help. (DV professional, G3R2)

> After certain programmed answers or questions, the person should be directed forward very quickly—because currently, seeking help and reporting violence are different matters. (Criminal justice expert, G2R1)

The DV professional pointed out the safety aspect: The need for help may not become apparent to service providers if interactions are limited to a chatbot. The criminal justice expert stated that these tools should be designed to detect indicators of emergencies and initiate the necessary predefined steps to ensure that the situations are reported and victim-survivors are guided forward. Many DV professionals pointed out that, for example, social work practitioners have a responsibility to ensure that victim-survivors have a safety network and that their children are protected from violence. Hence, there seems to be a lack of agency in the process of seeking help via chatbots.

> I strongly advocate for a human touch alongside this, as receiving emotional and psychosocial support is equally important. How can we ensure that this information flows while also connecting the help to the specific person? (DV professional, G3R3)

The interviewees highlighted that victim-survivors typically need multiprofessional support to secure safe and lasting disengagement from their abusers. The DV professionals were concerned that no one would be held accountable for victim-survivors' safety. On the whole, the interviewees considered DV too severe and complex to be entrusted to chatbots and expressed the need for a comprehensive assessment of the criticality of situations and concrete measures to ensure the safety of victim-survivors and their children.

Long-standing structural barriers to addressing DV were highlighted in the following comment, in which the respondent questions whether chatbots in DV support could overcome the challenges embedded in social structures and institutional practices:

> I do not see any added value a chatbot would have compared to what already exists. This is such a serious matter, and I feel that it requires a human touch. This problem is so deeply rooted in Finnish society. (Victim-survivor, G10R2)

This comment underscores the importance of critically assessing which specific issues technology can effectively address. On a similar note, the comment below acknowledges that while chatbot technology may help victim-survivors recognise DV and access information about support services, the broader service

system continues to pose structural challenges—particularly related to resources. Disclosing violence requires that practitioners have both the capacity and resources to respond to DV, offer support services, and handle a potential increase in criminal reports from victim-survivors:

> At the moment, the resources are probably insufficient, and knowing how long the criminal proceedings currently are, if there's even a little more pressure, it just means the process will be extended for everyone involved. (Criminal justice expert, G11R2)

Criminal processes are already lengthy and mentally exhausting, often leaving victim-survivors without adequate support before, during, and after legal proceedings. If these processes are further prolonged, their safety may be compromised. Several victim-survivors pointed out that one reason for not reporting violence is the lack of safety guarantees—even when their primary hope is simply for the violence to stop.

5. Ethical and Sustainable Integration of AI-Driven Chatbots in Support Services

To understand the ethical implications of using chatbots in DV support services, the findings were reflected on considering the ethical principles of prevention of harm, fairness, explicability, and respect for human autonomy. The results indicate that the implementation and use of AI-driven chatbots both align with these ethical principles and face challenges regarding them, illustrating how abstract guidelines can encounter various practical challenges depending on the context in which these AI systems are deployed.

Considering the ethical principle of preventing harm, the results suggest that the implementation of AI-driven chatbots may be consistent with this principle in certain contexts, while also revealing inconsistencies or limitations in practice. The challenges lie particularly in the absence of a responsible party to take the necessary concrete actions for organising multidisciplinary support and establishing a safety network for victim-survivors and their children. The interviewees pointed out that current AI-driven chatbot systems have limitations in situational assessment, particularly in processing contextual information holistically and in identifying the criticality of situations or the potential for digital abuse. These concerns resonate with findings from previous studies. For example, Palanica et al. (2019), who explored the use of healthcare chatbots, referred to the need for personalised knowledge in complex cases—a need that was consistently highlighted across all focus groups in the present study. Similarly, Emezue et al. (2022) identified the problem of accountability in digital interventions. They highlighted the need to critically consider the demand for agency that is placed on victim-survivors, as they may be living under constant threat of violence. In such circumstances, the expectation of victim-survivors' agency may seem unreasonable, considering how significantly the violence may impair their ability to function (see Miller & McCaw, 2019; Stubbs & Szoeke, 2022).

Expanding the discussion on harm prevention, all focus groups acknowledged the risks of being exposed to TFA when using a chatbot to seek help, in line with recent research (Afrouz, 2023; Boethius et al., 2023; Storer et al., 2023; Woodlock et al., 2020). As Afrouz (2023) noted, victim-survivors may struggle to identify TFA, especially if they lack technical skills. However, as shown by Boethius et al. (2023) and Novitzky et al. (2023), service practitioners also face challenges in addressing TFA. Similarly, the present study raised concerns about the extent to which AI-driven chatbot systems can support the detection of potential misuse of the device on which they are deployed. Hence, the results indicate that the responsibility for assessing the safety of a

chatbot, verifying the validity of its responses, and taking further actions, including reaching out to support services, falls on victim-survivors.

These concerns underscore the importance of establishing clear criteria for assessing the reliability of digital tools (see Kabacińska et al., 2022; Montagni et al., 2020), given the variability in app quality that Draughon Moret et al. (2022) found in their study. As Boethius et al. (2023) noted, victim-survivors should not be forced to limit their online presence, as technology can also empower them in various ways, including by providing them with access to support and information. Therefore, chatbots could, for example, be integrated into existing welfare service websites, which citizens are familiar with and may rely on more than apps from different sources. This would also address the structural challenges of scattered services (see Husso et al., 2021) and prevent the recurrence of systemic deficiencies in the digital dimension.

While the ethical principle of preventing harm raised concerns in some respects, the integration of AI-driven chatbots may also support it by lowering barriers to help-seeking—a dynamic also observed in prior studies on digital support tools (Al-Alosi, 2020; Kabacińska et al., 2022; Montagni et al., 2020; Novitzky et al., 2023; Torous et al., 2021). The interviewees perceived the automated functions of AI systems—such as systematically asking about experiences of DV—as encouraging victim-survivors to speak about issues that they might otherwise consider too emotionally exhausting for service practitioners to address. This perception resonates with earlier findings (Husso et al., 2012; Piippo et al., 2021; Virkki et al., 2015), which suggest that practitioners' own emotions can hinder them from intervening in DV.

Reflecting this, many interviewees saw chatbots as potentially useful tools for supporting practitioners in their work with DV cases by systematically asking the necessary questions to enhance the disclosure of violence. Hence, this could partly address the systemic gap whereby victim-survivors are reluctant to disclose violence (see Harris & Woodlock, 2019; Husso et al., 2021; Meyer, 2016; Storer et al., 2022) and service practitioners may struggle to effectively identify and intervene in DV cases (see Siltala et al., 2023; Turner et al., 2017; Wright et al., 2022). However, practitioners' motivation to use chatbots as tools could be questioned, given the insufficient use of tools specifically developed to map experiences of violence. For example, in Finland, over 50% of police officers and social and healthcare service workers use no tools at all for mapping DV (Niklander et al., 2019).

Turning to the ethical principle of fairness, the results show that the implementation of AI-driven chatbots can enhance access to information, thereby increasing awareness of DV and the legal rights of victim-survivors. These findings align with recent studies that suggested chatbots' potential to enhance confidentiality and to encourage victim-survivors to make an initial outreach for help from support services (Butterby & Lombard, 2025; Novitzky et al., 2023; Storer et al., 2022). Receiving information on a low threshold is both fair and empowering, aside from respecting victim-survivors' autonomy to make informed decisions about their situations. However, the results also indicate unequal opportunities to access information, especially for victim-survivors from immigrant backgrounds who face linguistic barriers. Additionally, Emezue et al. (2022) highlighted that victim-survivors with diverse ethnic, cultural, and linguistic needs may face obstacles in using DV apps. This shows how vulnerabilities may intersect and accumulate, and how, to prevent intersectional inequalities, it is important to perceive the diversity of capabilities to utilise digital services (AI HLEG, 2019; Bailey & Burkell, 2021; Mishna et al., 2021).

With regard to the ethical principle of explicability, the findings highlight challenges in ensuring that AI systems are understandable and interpretable to all users. As the interviewees noted, misunderstandings and frustration caused by unsuccessful interactions can ultimately deter victim-survivors from seeking help. Earlier research suggests that trust and perceived safety are central to users' willingness to engage with digital services (de Visser et al., 2018; Jobin et al., 2019; Kretzschmar et al., 2019; McGreevey et al., 2020). If a chatbot fails to communicate effectively due to language or cultural barriers, its functioning may become opaque to the user, thereby undermining both the user's trust and their ability to make informed decisions—echoing findings by Mishna et al. (2021) and Xu et al. (2021). Furthermore, in the present study, DV professionals expressed significant concerns about the reliability of chatbot-generated responses in complex and nuanced situations, particularly in critical cases requiring immediate action to protect victim-survivors and their children, a concern also raised by Butterby and Lombard (2025).

Finally, reflecting on the ethical principle of respect for human autonomy, the results indicate challenges in balancing victim-survivors' self-determination and institutional responsibilities, particularly in relation to anonymous user interactions with chatbots. Based on the victim-survivor interviews, these interactions were perceived as retaining victim-survivors' decision-making power over the progression of their cases and as possibly helping them avoid unwanted consequences by authorities, such as child protection interventions. However, this presents difficulties in ensuring comprehensive support and accountability. While anonymous interaction may support autonomy and reduce fear of judgment (see Al-Alosi, 2020; Novitzky et al., 2023; Storer et al., 2023), DV professionals emphasised the importance of intervention and documentation of cases. They highlighted that emergency situations requiring child protection or cases of recurring violence would remain undocumented if only chatbots are interacted with. This concern is even greater when considering how underreported violence has been within the service system (see Husso et al., 2021; Piippo et al., 2021; Siltala et al., 2023), potentially creating tensions between respect for victim-survivors' autonomy and the legal obligation of authorities to protect children. This tension highlights that the balance between different ethical principles may vary depending on the end-users' needs.

6. Conclusion

The findings from the present study suggest that while the implementation of AI-driven chatbots can significantly enhance access to information on DV and the legal rights of victim-survivors, it also raises ethically significant concerns regarding safety and accountability. Interactions with chatbots lack essential elements for comprehensive situational assessment, documentation of violence, and the formation of a support network. Hence, the results highlight the importance of human interaction in DV cases, aligning with recent studies (Butterby & Lombard, 2025; Domingo-Cabarrubias et al., 2023; Henry et al., 2024). However, the need for human interaction also requires that service practitioners possess the essential professional skills and knowledge to address DV. This requirement remains a structural deficiency yet to be resolved (Husso et al., 2021; Niklander et al., 2019; Piippo et al., 2021).

The results further emphasise the need to consider the vulnerabilities of victim-survivors, ensuring their safety and acknowledging the severity of DV, while questioning which specific issues technology can effectively address. Attempting to address problems that are structural and deeply embedded in institutional practices with technological solutions may divert attention from efforts to strengthen existing service systems and practices (see Hunt et al., 2020; Lindgren & Dignum, 2023). Therefore, it is essential to ensure

that the implemented technology reinforces the service system while supporting and strengthening practitioners' expertise in addressing DV (see Emezue et al., 2022; Hunt et al., 2020; Inkster et al., 2018). This study thereby suggests that AI-driven chatbots could serve victim-survivors as an intermediate support system, aligning with findings by Henry et al. (2024). In addition, they may function as a supplementary tool for welfare service practitioners to identify DV and proceed with cases—an area that remains underexplored in existing research.

DV causes a wide range of social and health problems and significant human suffering. Developing effective intervention practices requires a broad understanding of the distinctions and interconnections between online and offline domains as well as the opportunities and limitations of AI-driven solutions. Meeting these challenges demands comprehensive ethical consideration of safety, accountability, and the potential impacts of automated interventions.

While this study contributes to the ethical and practical discourse on AI-driven interventions in DV support, it is not without limitations. First, the conceptual distinction between online and offline domains often fails to capture the experiences of victim-survivors, who navigate violence across both spheres, where digital and physical realities are deeply interwoven. This ambiguity presents challenges not only for victim-survivors but also for practitioners and researchers seeking adequate language to describe AI-mediated interactions. Second, additional research is needed to examine the user experiences of AI-driven chatbots specifically designed for victim-survivors. Third, while many interviewed DV professionals had relevant education and experience, social and healthcare service practitioners were not directly interviewed. Furthermore, this study identified several areas requiring further exploration, including the tension between respecting human autonomy and the duty of authorities to document DV cases, as well as the potential for practitioners to incorporate AI tools into violence prevention and intervention strategies.

Acknowledgments

We would like to thank our interviewees for sharing their time and expertise, and we would also like to thank the anonymous reviewers for their valuable feedback throughout the process.

Funding

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101073922 (ISEDA). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them. Publication of this article in open access was made possible through the institutional membership agreement between Tampere University and Cogitatio Press.

Conflict of Interests

The authors declare that there is no conflict of interest.

References

Abdelnour-Nocera, J., & Clemmensen, T. (2019). Theorizing about socio-technical approaches to HCI. In B. R. Barricelli, V. Roto, T. Clemmensen, P. Campos, A. Lopes, F. Gonçalves, & J. Abdelnour-Nocera (Eds.), *Human work interaction design. Designing engaging automation* (Vol. 544, pp. 242–262). Springer. https://doi.org/10.1007/978-3-030-05297-3_17

Afrouz, R. (2023). The nature, patterns and consequences of technology-facilitated domestic abuse: A scoping review. *Trauma, Violence, & Abuse*, *24*(2), 913–927. https://doi.org/10.1177/15248380211046752

Al-Alosi, H. (2020). Fighting fire with fire: Exploring the potential of technology to help victims combat intimate partner violence. *Aggression and Violent Behavior*, *52*, Article 101376. https://doi.org/10.1016/j.avb.2020.101376

Attila, H., Keski-Petäjä, M., Pietiläinen, M., Lipasti, L., Saari, J., & Haapakangas, K. (2023). *Sukupuolistunut väkivalta ja lähisuhdeväkivalta Suomessa 2021: Loppuraportti*. Statistics Finland. https://urn.fi/URN:ISBN:978-952-244-717-3

Bailey, J., & Burkell, J. (2021). Tech-facilitated violence: Thinking structurally and intersectionally. *Journal of Gender-Based Violence*, *5*(3), 531–542. https://doi.org/10.1332/239868021X16286662118554

Bellis, M. A., Hughes, K., Ford, K., Ramos Rodriguez, G., Sethi, D., & Passmore, J. (2019). Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: A systematic review and meta-analysis. *The Lancet Public Health*, *4*(10), e517–e528. https://doi.org/10.1016/S2468-2667(19)30145-8

Bergdahl, J., Latikka, R., Celuch, M., Savolainen, I., Soares Mantere, E., Savela, N., & Oksanen, A. (2023). Self-determination and attitudes toward artificial intelligence: Cross-national and longitudinal perspectives. *Telematics and Informatics*, *82*, Article 102013. https://doi.org/10.1016/j.tele.2023.102013

Boethius, S., Åkerström, M., & Hydén, M. (2023). The double-edged sword—Abused women's experiences of digital technology. *European Journal of Social Work*, *26*(3), 506–518. https://doi.org/10.1080/13691457.2022.2040437

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braunschweig, B., & Ghallab, M. (2021). Reflections on AI for humanity: Introduction. In B. Braunschweig & M. Ghallab (Eds.), *Reflections on artificial intelligence for humanity* (pp. 1–12). Springer. https://doi.org/10.1007/978-3-030-69128-8_1

Bryman, A. (2004). *Social research methods*. Oxford University Press.

Butterby, K., & Lombard, N. (2025). Developing a chatbot to support victim-survivors who are subjected to domestic abuse: Considerations and ethical dilemmas. *Journal of Gender-Based Violence*, *9*(1), 153–161. https://doi.org/10.1332/23986808Y2024D000000038

Callaghan, J. E. M., Alexander, J. H., Sixsmith, J., & Fellin, L. C. (2018). Beyond "witnessing": Children's experiences of coercive control in domestic violence and abuse. *Journal of Interpersonal Violence*, *33*(10), 1551–1581. https://doi.org/10.1177/0886260515618946

Decker, M. R., Holliday, C. N., Hameeduddin, Z., Shah, R., Miller, J., Dantzler, J., & Goodmark, L. (2019). "You do not think of me as a human being": Race and gender inequities intersect to discourage police reporting of violence against women. *Journal of Urban Health*, *96*(5), 772–783. https://doi.org/10.1007/s11524-019-00359-z

Devillers, L., Fogelman-Soulié, F., & Baeza-Yates, R. (2021). AI & human values: Inequalities, biases, fairness, nudge, and feedback loops. In B. Braunschweig & M. Ghallab (Eds.), *Reflections on artificial intelligence for humanity* (pp. 76–89). Springer. https://doi.org/10.1007/978-3-030-69128-8_6

de Visser, E. J., Pak, R., & Shaw, T. H. (2018). From "automation" to "autonomy": The importance of trust repair in human-machine interaction. *Ergonomics*, *61*(10), 1409–1427. https://doi.org/10.1080/00140139.2018.1457725

Domingo-Cabarrubias, L., Woodlock, D., Alexander, C., Sato, M., Grant, G., & Weinberg, J. (2023). The role of technology in improving access to justice for victims of family violence: Challenges and opportunities. *Law, Technology and Humans*, *5*(1), 1–10. https://doi.org/10.5204/lthj.2469

Draughon Moret, J., Todd, A., Rose, L., Pollitt, E., & Anderson, J. (2022). Mobile phone apps for intimate partner and sexual violence prevention and response: Systematic search on app stores. *JMIR Formative Research*, *6*(2), Article 28959. https://doi.org/10.2196/28959

Ellsberg, M., Arango, D. J., Morton, M., Gennari, F., Kiplesund, S., Contreras, M., & Watts, C. (2015). Prevention of violence against women and girls: What does the evidence say? *The Lancet*, *385*(9977), 1555–1566. https://doi.org/10.1016/S0140-*6736*(14)61703-7

Emezue, C., Chase, J. D., Udmuangpia, T., & Bloom, T. L. (2022). Technology-based and digital interventions for intimate partner violence: A systematic review and meta-analysis. *Campbell Systematic Review*, *18*(3), Article 1271. https://doi.org/10.1002/cl2.1271

European Commission. (2024a). *DESI indicators—eGovernment usage: Finland*. https://digital-decade-desi.digital-strategy.ec.europa.eu/datasets/desi/charts/desi-indicators?indicator=desi_egov_users_anys&indicatorGroup=desi2023-4&breakdown=ind_total&period=desi_2024&unit=pc_ind_ilt12

European Commission. (2024b). *Digital decade country report 2024: Finland*. https://digital-strategy.ec.europa.eu/en/factpages/finland-2024-digital-decade-country-report

European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Fern, E. F. (2001). *Advanced focus group research*. Sage. https://doi.org/10.4135/9781412990028

Finnish Government. (2024). *Finland's national roadmap: EU Digital Decade Policy Programme 2030* (Publication of the Finnish Government 2024:7). https://urn.fi/URN:ISBN:978-952-383-743-0

Finnish National Board on Research Integrity TENK. (2023). *The Finnish code of conduct for research integrity and procedures for handling alleged violations of research integrity in Finland 2023*. https://tenk.fi/sites/default/files/2023-05/RI_Guidelines_2023.pdf

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Harris, B. A., & Woodlock, D. (2019). Digital coercive control: Insights from two landmark domestic violence studies. *British Journal of Criminology*, *59*(3), 530–550. https://doi.org/10.1093/bjc/azy052

Henman, P. (2020). Improving public services using artificial intelligence: Possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration*, *42*(4), 209–221. https://doi.org/10.1080/23276665.2020.1816188

Henry, N., Witt, A., & Vasil, S. (2024). A "design justice" approach to developing digital tools for addressing gender-based violence: Exploring the possibilities and limits of feminist chatbots. *Information, Communication & Society*. Advance online publication. https://doi.org/10.1080/1369118X.2024.2363900

High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. European Commission. https://doi.org/10.2759/346720

Holt, S., Buckley, H., & Whelan, S. (2008). The impact of exposure to domestic violence on children and young people: A review of the literature. *Child Abuse & Neglect*, *32*(8), 797–810. https://doi.org/10.1016/j.chiabu.2008.02.004

Hunt, X., Tomlinson, M., Sikander, S., Skeen, S., Marlow, M., du Toit, S., & Eisner, M. (2020). Artificial intelligence, big data, and mHealth: The frontiers of the prevention of violence against children. *Frontiers in Artificial Intelligence*, *3*, Article 543305. https://doi.org/10.3389/frai.2020.543305

Husso, M., Notko, M., Virkki, T., Holma, J., Laitila, A., & Siltala, H. (2021). Domestic violence interventions

in social and health care settings: Challenges of temporary projects and short-term solutions. *Journal of Interpersonal Violence*, *36*(23/24), Article 0886260519898438. https://doi.org/10.1177/0886260519898438

Husso, M., Virkki, T., Notko, M., Holma, J., Laitila, A., & Mäntysaari, M. (2012). Making sense of domestic violence intervention in professional health care. *Health and Social Care in the Community*, *20*(4), 347–355. https://doi.org/10.1111/j.1365-2524.2011.01034.x

Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR Mhealth and Uhealth*, *6*(11), Article 12106. https://doi.org/10.2196/12106

International Federation of Social Workers. (2018). *Global social work statement of ethical principles*. https://www.ifsw.org/global-social-work-statement-of-ethical-principles

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kabacińska, K., McLeod, K., MacKenzie, A., Vu, K., Cianfrone, M., Tugwell, A., & Robillard, J. M. (2022). What criteria are young people using to select mobile mental health applications? A nominal group study. *Digital Health*, *8*, Article 20552076221102775. https://doi.org/10.1177/20552076221102775

Keeling, J., & Fisher, C. (2015). Health professionals' responses to women's disclosure of domestic violence. *Journal of Interpersonal Violence*, *30*(13), 2363–2378. https://doi.org/10.1177/0886260514552449

Kitzinger, J. (1994). The methodology of focus groups: The importance of interaction between research participants. *Sociology of Health & Illness*, *16*(1), 103–121. https://doi.org/10.1111/1467-9566.ep11347023

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., & Singh, I. (2019). Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical Informatics Insights*, *11*, Article 1178222619829083. https://doi.org/10.1177/1178222619829083

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction* (2nd ed.). Morgan Kaufmann.

Lindgren, S., & Dignum, V. (2023). Beyond AI solutionism: Toward a multi-disciplinary approach to artificial intelligence in society. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 163–172). Edward Elgar Publishing. https://doi.org/10.4337/9781803928562.00019

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *The Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

McGreevey, J. D., III, Hanson, C. W., III, & Koppel, R. (2020). Clinical, legal, and ethical aspects of artificial intelligence–assisted conversational agents in health care. *JAMA*, *324*(6), 552–553. https://doi.org/10.1001/jama.2020.2724

Meyer, S. (2016). Still blaming the victim of intimate partner violence? Women's narratives of victim desistance and redemption when seeking support. *Theoretical Criminology*, *20*(1), 75–90. https://doi.org/10.1177/1362480615585399

Miller, E., & McCaw, B. (2019). Intimate partner violence. *New England Journal of Medicine*, *380*(9), 850–857. https://doi.org/10.1056/NEJMra1807166

Mishna, F., Milne, E., Bogo, M., & Pereira, L. F. (2021). Responding to COVID-19: New trends in social workers' use of information and communication technology. *Clinical Social Work Journal*, *49*(4), 484–494. https://doi.org/10.1007/s10615-020-00780-x

Montagni, I., Tzourio, C., Cousin, T., Sagara, J. A., Bada-Alonzi, J., & Horgan, A. (2020). Mental health-related

digital use by university students: A systematic review. *Telemedicine and E-Health*, *26*(2), 131–146. https://doi.org/10.1089/tmj.2018.0316

Niklander, E., Notko, M., & Husso, M. (2019). *Intervening in domestic violence and training of professionals in social services and health care and the police: Evaluation of the EPRAS project*. National Institute for Health and Welfare (THL). https://urn.fi/URN:ISBN:978-952-343-413-4

Nordheim, C. B., Følstad, A., & Bjørkli, C. A. (2019). An initial model of trust in chatbots for customer service—Findings from a questionnaire study. *Interacting With Computers*, *31*(3), 317–335. https://doi.org/10.1093/iwc/iwz022

Novitzky, P., Janssen, J., & Kokkeler, B. (2023). A systematic review of ethical challenges and opportunities of addressing domestic violence with AI-technologies and online tools. *Heliyon*, *9*(6), Article 17140. https://doi.org/10.1016/j.heliyon.2023.e17140

Palanica, A., Flaschner, P., Thommandram, A., Li, M., & Fossat, Y. (2019). Physicians' perceptions of chatbots in health care: Cross-sectional web-based survey. *Journal of Medical Internet Research*, *21*(4), Article 12887. https://doi.org/10.2196/12887

Parviainen, J., & Rantala, J. (2022). Chatbot breakthrough in the 2020s? An ethical reflection on the trend of automated consultations in health care. *Medicine, Health Care and Philosophy*, *25*(1), 61–71. https://doi.org/10.1007/s11019-021-10049-w

Piippo, S., Husso, M., Hirvonen, P., Notko, M., & Glumbíková, K. (2021). Institutional and affective practices of domestic violence interventions in social work: Malignant positioning of victims. In M. Husso, S. Karkulehto, T. Saresma, A. Laitila, J. Eilola, & H. Siltala (Eds.), *Violence, gender and affect: Interpersonal, institutional and ideological practices* (pp. 113–133). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-56930-3_6

Reith-Hall, E., & Montgomery, P. (2022). The teaching and learning of communication skills in social work education. *Research on Social Work Practice*, *32*(7), 793–813. https://doi.org/10.1177/10497315221088285

Rogers, M. M., Fisher, C., Ali, P., Allmark, P., & Fontes, L. (2023). Technology-facilitated abuse in intimate relationships: A scoping review. *Trauma, Violence, & Abuse*, *24*(4), 2210–2226. https://doi.org/10.1177/15248380221090218

Saxton, M. D., Olszowy, L., MacGregor, J. C. D., MacQuarrie, B. J., & Wathen, C. N. (2021). Experiences of intimate partner violence victims with police and the justice system in Canada. *Journal of Interpersonal Violence*, *36*(3/4), NP2029–NP2055. https://doi.org/10.1177/0886260518758330

Siltala, H., Hisasue, T., Hietamäki, J., Saari, J., Laajasalo, T., October, M., Laitinen, H., & Raitanen, J. (2023). *Domestic violence-related use of services and the resulting costs in health, social and legal services*. Prime Minister's Office. http://urn.fi/URN:ISBN:978-952-383-317-3

Stark, E. (2007). *Coercive control: The entrapment of women in personal life*. Oxford University Press.

Storer, H. L., Nyerges, E. X., & Hamby, S. (2022). Technology "feels less threatening": The processes by which digital technologies can facilitate youth's access to services at domestic violence and sexual assault organisations. *Children and Youth Services Review*, *139*, Article 106573. https://doi.org/10.1016/j.childyouth.2022.106573

Storer, H. L., Scott, C. F., Rodriguez, M., & Nyerges, E. X. (2023). Technology is a "blessing and a curse": The perceived risks and benefits of digital technology adoption at domestic violence organisations that serve teens. *Journal of Technology in Human Services*, *41*(1), 96–124. https://doi.org/10.1080/15228835.2023.2179158

Stubbs, A., & Szoeke, C. (2022). The effect of intimate partner violence on the physical health and health-related behaviors of women: A systematic review of the literature. *Trauma, Violence, & Abuse*, *23*(4), 1157–1172. https://doi.org/10.1177/1524838020985541

Taylor, A. (2017). Social work and digitalisation: Bridging the knowledge gaps. *Social Work Education*, *36*(8), 869–879. https://doi.org/10.1080/02615479.2017.1361924

Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A. F., Keshavan, M., Linardon, J., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*, *20*(3), 318–335. https://doi.org/10.1002/wps.20883

Trevithick, P. (2012). *Social work skills and knowledge: A practice handbook* (3rd ed.). McGraw-Hill Education; Open University Press.

Turner, W., Hester, M., Broad, J., Szilassy, E., Feder, G., Drinkwater, J., Firth, A., & Stanley, N. (2017). Interventions to improve the response of professionals to children exposed to domestic violence and abuse: A systematic review. *Child Abuse Review*, *26*(1), 19–39. https://doi.org/10.1002/car.2385

United Nations. (2015, October 21). *Transforming our world: The 2030 agenda for sustainable development* (A/RES/70/1). https://digitallibrary.un.org/record/3923923

United Nations. (2024, March 21). *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development* (A/RES/78/265). United Nations Digital Library. https://digitallibrary.un.org/record/4043244

Virkki, T., Husso, M., Notko, M., Holma, J., Laitila, A., & Mäntysaari, M. (2015). Possibilities for intervention in domestic violence: Frame analysis of health care professionals' attitudes. *Journal of Social Service Research*, *41*(1), 6–24. https://doi.org/10.1080/01488376.2014.917449

Wood, L., Hairston, D., Schrag, R. V., Clark, E., Parra-Cardona, R., & Temple, J. R. (2022). Creating a digital trauma informed space: Chat and text advocacy for survivors of violence. *Journal of Interpersonal Violence*, *37*(19/20), NP18960–NP18987. https://doi.org/10.1177/08862605211043573

Woodlock, D., McKenzie, M., Western, D., & Harris, B. (2020). Technology as a weapon in domestic violence: Responding to digital coercive control. *Australian Social Work*, *73*(3), 368–380. https://doi.org/10.1080/0312407X.2019.1607510

Wright, E. N., Anderson, J., Phillips, K., & Miyamoto, S. (2022). Help-seeking and barriers to care in intimate partner sexual violence: A systematic review. *Trauma, Violence, & Abuse*, *23*(5), 1510–1528. https://doi.org/10.1177/1524838021998305

Xu, L., Sanders, L., Li, K., & Chow, J. C. L. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. *JMIR Cancer*, *7*(4), Article 27850. https://doi.org/10.2196/27850

About the Authors

Hanna Mielismäki is a doctoral researcher in social sciences at Tampere University. Her research focuses on domestic violence and technology-facilitated abuse, responsible AI, digital ethics, and the digitalisation of welfare services. She explores the ethical and societal implications of AI-driven technologies in support services.

Marita Husso is a professor of social policy at Tampere University. Her research interests include gender-based violence, digitalisation, service systems, care work, and gendered agency. She is the first editor of *Interpersonal Violence—Differences and Connections* (Routledge, 2017) and *Violence, Gender and Affect—Interpersonal, Institutional and Ideological Practices* (Palgrave, 2021).