
open space

Developing a chatbot to support victim-survivors who are subjected to domestic abuse: considerations and ethical dilemmas

Kate Butterby, kate.butterby@gcu.ac.uk
Nancy Lombard, nancy.lombard@gcu.ac.uk
Glasgow Caledonian University, UK

This Open Space piece discusses the implementation of a chatbot for victim-survivors who are subjected to domestic abuse. Being deployed as part of the ISEDA project (Innovative Solutions to Eliminate Domestic Abuse), a Horizon Europe project involving 15 partners from nine European countries, the chatbot will aim to provide women who are subjected to domestic abuse with information about domestic abuse, local support services, links to emergency services and potentially act as a place to store evidence that can be used in court. We discuss the ethical considerations surrounding the implementation of the chatbot within the project, and use of technology to support women within the domestic abuse sector in general. We highlight some of the positives alongside pitfalls of this way of working, and outline some of the considerations surrounding longevity of the chatbot in light of empowering under-funded women's services.

Keywords technology • chatbot • domestic abuse • ethical considerations

Key messages

- Taking a critical approach when developing digital interventions to support victim-survivors is essential to ensure safe and ethical tools are produced.
- Chatbot developers should be cognisant of existing research, interventions, suggestions, and learn from best practice.

To cite this article: Butterby, K. and Lombard, N. (2024) Developing a chatbot to support victim-survivors who are subjected to domestic abuse: considerations and ethical dilemmas, *Journal of Gender-Based Violence*, XX(XX): 1–9, DOI: 10.1332/23986808Y2024D000000038

The use of mobile-based technology to support victim-survivors who are subjected to domestic abuse is becoming more prevalent, with the 24/7 availability and ease of use reported as positives. This thought-piece discusses the potential of a chatbot to support women who are subjected to domestic abuse. It argues that although technology may be useful for victim-survivors it should not take the place of the expertise provided by support services.

The authors are part of the Innovative Solutions to Eliminate Domestic Abuse (ISEDA) project alongside 15 partners from nine European Countries. ISEDA is a

Horizon Europe project which aims to tackle and eliminate domestic abuse using multi-sector expertise via modern technological tools and practices. One of the key tools to be deployed by the project is a chatbot for victim-survivors who are subjected to domestic abuse. Chatbots are a type of artificial intelligence (AI) with which users can engage to find answers to questions (Tzelios et al, 2022), and are defined as ‘A computer program designed to simulate conversation with a human user, usually over the internet’ (Oxford University Press, nd). The chatbot will provide victim-survivors with information about domestic abuse and legislation in their country, alongside information on how to report to police, should they wish to. It will be informed and developed by those working in the area of gender-based violence and technology. Though chatbots have come under some criticism, tools which have the potential to support those working with victim-survivors (and victim-survivors themselves) should be tested and considered.

ISEDA and chatbot development

It is envisaged that victim-survivors will use the chatbot to seek help and find support and information. It is also hoped the chatbot will have the capability to store and record evidence to be used in court (whether this occurs and the types of evidence collected would be dependent on individual countries and their legislative practices). The project is overseen by an ethics board, which works to ensure privacy, safety and security around use of the chatbot. The ethics around its use is ongoing (through the life of the project).

Digital interventions (rather than other sources of support such as helplines) may be preferred by some victim-survivors. For example, they may be looking for information if they are unsure whether their experience constitutes domestic abuse before deciding whether to speak to someone. Research by Storer et al (2022) suggests people find it easier to maintain confidentiality and/or remain anonymous when using digital interventions versus disclosing to a person. A chatbot would be available to victim-survivors 24/7, and potentially easier to access than support services, which may have specific opening times.

The ISEDA chatbot is designed to be used across Europe. Primarily, it will give women access to supportive technology in countries where services are not always commonplace, for example, countries where women’s services lack consistent funding, or where police training on domestic abuse is sporadic. The support provided to women via the chatbot will include information about domestic abuse and local support services, legislation, and their rights within the country within which they reside, in addition to (with victim-survivors’ permission) the storage of evidence. The initial pilot of the chatbot will be implemented in Bulgaria, Catalonia and Greece, and if successful, rolled out within other European countries in the future.

Here, we ask whether technology (such as the chatbot) should be integrated within women’s services as a complementary tool available to support victim-survivors, and explore whether this can be done without adding to the workload of the already under-funded and overworked domestic abuse sector. We also highlight ethical issues surrounding the use of technology within the development of the chatbot specifically, and in the field of domestic abuse more broadly, and consider how these issues can be addressed.

Positives and pitfalls of existing technological interventions

Existing literature highlights that technology can be a useful tool to support women who are subjected to domestic abuse. Technological interventions may reduce barriers and be experienced as more private, less intimidating and mean victim-survivors feel less judged (Lucas et al, 2014; Glass et al, 2017; Storer et al, 2022).

As relevant to the ISEDA project, research has shown that people may feel more able to disclose sensitive information to a computer-generated avatar than an actual person (Lucas et al, 2014; Pickard et al, 2016; Bae Brandtzæg et al, 2021; Park and Lee, 2021). Technology also has the bonus of being more readily available than in-person support services, meaning that victim-survivors can use it in ways that meet their needs (such as choosing when and where to access), which is particularly positive for those who feel more isolated (Tarzia et al, 2017; Emezue et al, 2022). Research has also highlighted that technological tools are particularly suited to young people due to technology often being the preferred way of communicating for this demographic (Tarzia et al, 2017; Park and Lee, 2021; Storer et al, 2022). For example, Piccolo and Alani's, 2020 research with 11–17 year olds in the UK found that 82 per cent liked the idea of a chatbot for tackling online abuse, and it would be their chosen source of help beyond their families if they were to experience abuse. In Koulouri, Macredie and Olakitan's 2022 research exploring the acceptability of chatbots to support young adults' mental health, only 5 per cent had sought help from a helpline, compared to 12 per cent from a mental health app. Storer et al (2022) also found that digital technology can help to minimise barriers for young people in relation to seeking support from domestic abuse organisations. Digital interventions such as chatbots may also provide a level of privacy that speaking out loud or attending support in-person does not, which is likely to be particularly important for victim-survivors who could be overheard by the perpetrator.

Alongside the positives of technology, there are downfalls. When victim-survivors are accessing technology, safety and security should be paramount to ensure that they are not put at more risk (Eisenhut et al, 2020). Examples of when they do not feel safe include fear of being tracked by the perpetrator, personal data not being secure and private, and concerns that they will not be able to delete data off their device (Xu et al, 2021; Afrouz, 2023; Sabri et al, 2023).

The accessibility of technology should also be considered, such as whether some victim-survivors may not have access to it, for example, due to socio-economic constraints, and whether minoritised groups such as neurodiverse, deaf, older survivors, and survivors who struggle with literacy are able to access digital technologies in their current forms (see Emezue et al, 2022; Sabri et al, 2023; Sumra et al, 2023). Other research suggests that technology such as chatbots are unable to emulate 'real' human characteristics, such as empathy, reading of body language, or make eye-contact (Storer et al, 2022), with suggestions made that rather than replacing human contact, they can be used as a pathway to help people access these services (Tarzia et al, 2017; Xu et al, 2021). When considering the technology itself, there is also the possibility that it is not suitably maintained, such as information being outdated and with broken hyperlinks; problems which were found by Brignone and Edleson (2019) during their evaluation of existing domestic abuse applications. It is critical that the current project works to mitigate these potential pitfalls by learning from existing research in the area.

Ethical issues relating to chatbot development

As part of our work with ISEDA, we spoke with the developers and facilitators of FollowItApp, a mobile app designed to support victim-survivors who have been subjected to stalking in Scotland. Developed by the Media Co-op alongside the Scottish Women's Rights Centre (SWRC) and Rape Crisis Scotland (RCS), the app was originally released as a pilot in 2018 via women's services across Scotland such as RCS and Scottish Women's Aid, before being updated to enable centralised delivery through the SWRC. The creators worked collaboratively with the police, solicitors, women's organisations and survivors in its creation, to develop functions such as the recording of details of stalking incidents (including photos and videos, the location of the incident, what happened, and if there were any witnesses). During wider discussion with the FollowItApp developers and facilitators, consideration of the ethical issues of developing technological tools was deemed essential.

Westmarland et al (2013) highlighted that many apps assume that a victim-survivor has private access to their phone, which may not be the case, with research finding that perpetrators use technology to abuse victim-survivors (Afrouz, 2023). With the development of new technology such as the chatbot, there needs to be safeguards in place in order to mitigate malicious dual reporting by perpetrators (see also Brooks-Hay and Kyle, 2015), and also to prevent them from accessing data input by victim-survivors which could be used to further control them. One recommendation is that security measures should be set up within apps that can distinguish between a victim-survivor using an app, and a perpetrator. This could be, for example, based on keystroke or contextual cues (see Freed et al, 2018). Freed et al also illustrate further security mechanisms, such as the app having a quick exit button, and data being deleted if the password is entered incorrectly. Brignone and Edleson (2019) highlight similar, in that apps should have passwords, hidden areas and the option to disable push-notifications. Location services are used by many existing apps in order to, for example, find local support services for women. However, this function can also be used maliciously by perpetrators to track victim-survivors (Brignone and Edleson, 2019; Draughon-Moret et al, 2022) and chatbot developers should be cognisant of how this could jeopardise women's safety.

Women may be more likely to reach out for support when they feel most unsafe. The chatbot therefore should be programmed in such a way that it is able to quickly identify if a woman needs immediate assistance (such as police or ambulance) and be able to guide them to seek this support. This is an area that will need detailed consideration – in the case of face-to-face support, it is perhaps simpler (though not always easy by any means) to ascertain how safe somebody is, by asking appropriate questions, listening to tone of voice and picking up on body language. A chatbot, however, will need to be programmed to establish, quickly, a woman's perception of her level of safety. If a woman is determined to be in immediate danger, the chatbot will not be a suitable tool to be used. Currently, text is being developed to direct users to call emergency services in this situation. Whatever the solution, the woman's safety should be the first and foremost consideration.

In order to most successfully support women who have been subjected to domestic abuse, the chatbot needs to be trauma-informed, both in terms of language used, and imagery; particularly as research suggests that some existing chatbots are insensitive and inaccurate (Park and Lee, 2021). Current research highlights the importance

of empathy being shown within digital technologies, alongside language being reaffirming, validating, sensitive, non-judgemental, positive and friendly (Tarzia et al, 2017; Park and Lee, 2020; O'Campo et al, 2021; Sabri et al, 2023). At the same time, language should not normalise dangerous behaviours (O'Campo et al, 2021), that is, a chatbot should enable victim-survivors to feel listened to and reassured, but without the language being used giving the impression that the abusive behaviours they are describing or being subjected to are normal/acceptable. The language should also be accessible. Developing the chatbot in conjunction with professionals from women's services, those who work in the area of gender-based violence and alongside victim-survivors themselves will help to ensure that the chatbot is fit-for-purpose in this regard.

Platform and data ownership and funding

A key area to be considered when developing the chatbot is who will own/host, fund and update the platform, to ensure that the chatbot continues to be maintained sufficiently.

The issue of ownership of the chatbot is one that needs to be carefully considered. As is shown within existing research, women have concerns over the security and privacy of their data when using technology (Sumra et al, 2023). Who owns the platform is likely to influence the level of trust that a victim-survivor feels when using it to access support or to store sensitive information. For the FollowItApp, Media Co-op, a not-for-profit feminist co-operative, who work in an ethical, sustainable and collaborative way, own the intellectual property of the app. However, victim-survivors who use the app own and control their own data. For the ISEDA project, it is vital therefore that data ownership is carefully considered and communicated to users of the chatbot so that they can make an informed decision about whether and how they wish to use it.

Currently, within the ISEDA project, funding is available for the setup of a pilot chatbot. It is, however, unclear about what will happen to the tool when the project ceases in terms of continued funding. The owners of the FollowItApp submit frequent funding applications due to the app needing to be updated on a regular basis, and it is envisaged that the ISEDA chatbot will require similar, if not more regular, updates. There is also the question of who has the capacity, experience and knowledge to own and update the platform, going forward. For the pilot, there is input from practitioners working with victim-survivors and perpetrators, police, academics and technology experts, which ensures rounded and informed input. When the chatbot pilot is deployed, victim-survivors will form an integral part of this by providing feedback on early versions. Following project completion, ideally this level of expertise will continue, though funding constraints are unlikely to allow this. To that end, a dilemma ensues. Women's services are equipped with the knowledge around gender-based violence and trauma, work closely with victim-survivors and have the ability to design sensitive resources and messaging. They are, however, under-funded and already working at capacity, in addition to potentially lacking the technological expertise needed in order to update and maintain a chatbot (there will of course be individuals working in these services who do possess this knowledge, but it is not women's services' area of expertise). On the other hand, technology companies possess the technological expertise needed to develop and update the chatbot, but may not have gender-based violence knowledge.

While chatbot setup is currently funded by European Union (EU) funding, who should take responsibility moving forwards? Is it the EU's responsibility to continue to fund what they have initiated? Should funds be provided to third-sector/non-governmental women's organisations so that they can be empowered to take ownership and control of technology that they use? These are difficult questions to answer, and questions that require a collaborative response from all sectors involved with ISEDA. Our current position is that women's services in Bulgaria, Greece and Catalonia should be provided with funds to recruit more staff in order to manage the chatbot and work collaboratively with technology experts to provide a joined-up, sustainable service. This funding should not be at the expense, however, of continuing to provide funds to women's services to cover costs for the existing vital work they undertake.

Evaluation of the chatbot

The purpose of an evaluation is to help to identify criteria for successes and challenges, lessons learned, areas for improvement, and new goals (Zarinpoush, 2006) as well as to understand the progress, success, and effectiveness of a project (Zarinpoush, 2006). In order that digital technologies continue to serve their purpose, it has been suggested that evaluation should be built into an app itself (Westmarland et al, 2013), as well as consideration being given as to how apps can be used in harmful ways. When discussing chatbots, Xu et al (2021) state that they should be reviewed, quality checked and evaluated regularly.

The chatbot developed for ISEDA needs to be evaluated in the pilot stage of this project to ensure that it is fit-for-purpose. To do so, groups of victim-survivors, professionals from third-sector/non-governmental women's organisations and police will be asked to participate in this initial evaluation. Evaluation questions for each group have been developed to assess the technological elements of the chatbot, how well it meets the needs of domestic abuse survivors and the extent to which victim-survivors feel safe while using/as a result of using the chatbot. The findings from these questions will be actioned to develop the chatbot further. The evaluation will be led by the authors of this article, supported by members of the pilot sites in Bulgaria, Greece and Catalonia in conjunction with the chatbot creators. Further to this, ongoing evaluation of the chatbot is key to ensure it remains up-to-date, accurate, trauma-informed and is technologically sound.

Conclusion

With the increase in reliance on technology within society, it is sensible that its use should be reflected upon within the domestic abuse support sector. Technology enables many women who are subjected to domestic abuse more convenient ways of accessing support, anonymity and potentially the confidence to be more open than when speaking to a person face-to-face. However, it is also prone to glitches, out-of-date information, possible issues with data security and the potential of it being abused by perpetrators. The chatbot currently being developed by ISEDA aims to give more women across several European countries access to information about domestic abuse support services in their local area and assist them to make a report to the police if they choose to. It is currently not clear, however, once the tool is developed, who will host and update this platform, and how it will be funded

in future years. There is an argument for specialist third-sector/non-governmental women's organisations to integrate the chatbot within their existing services and manage its deployment going forward. However, the women's sector lack funding and are overworked, meaning doing so without a stringent plan would increase their workload to unsustainable levels. We recommend that to support chatbot longevity, the women's sector is allocated funds to take the chatbot forward, by being able to recruit/work alongside more staff (including those with technology expertise) to continue to provide a trauma-informed, sensitive tool to support women. These funds, should not, however, take away vital funding from women's services in general, who provide much-needed support to diverse groups of women, including those who do not use or have access to technology for various reasons. Whatever decisions are made in relation to hosting and providing the chatbot service, it is imperative that the tool is hosted and deployed within all countries in an ethical way, with the safety of women at the forefront of all decisions. We look forward to grappling with the issues outlined in this thought-piece as we progress within the ISEDA project.

Funding

This work was supported by the European Union under Grant Number 101073922: ISEDA. Funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [10039269]. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or UKRI. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgements

This work is based upon a literature review we wrote as part of one of the work package directives within the ISEDA project. We would like to thank Rocio Vicente and Carmen Vives from the University of Alicante for their literature suggestions and comments on the report submission.

Conflict of interest

The authors declare that there is no conflict of interest.

References

- Afrouz, R. (2023) The nature, patterns and consequences of technology-facilitated domestic abuse: a scoping review, *Trauma, Violence & Abuse*, 24(2): 913–27. doi: [10.1177/15248380211046752](https://doi.org/10.1177/15248380211046752)
- Bae Brandtzæg, P.B., Skjuve, M., Kristoffer Dysthe, K.K. and Følstad, A. (2021) When the social becomes non-human: young people's perception of social support in chatbots, in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21), New York, Association for Computing Machinery, 257: 1–13. doi: [10.1145/3411764.3445318](https://doi.org/10.1145/3411764.3445318)
- Brignone, L. and Edleson, J.L. (2019) The dating and domestic violence app rubric: synthesizing clinical best practices and digital health app standards for relationship violence prevention smartphone apps, *International Journal of Human-Computer Interaction*, 35(19): 1859–69. doi: [10.1080/10447318.2019.1574100](https://doi.org/10.1080/10447318.2019.1574100)
- Brooks, O. and Kyle, D. (2015) *Dual Reports of Domestic Abuse Made to the Police in Scotland: A Summary of Findings from a Pilot Research Study, Project Report*, Dundee: Scottish Institute for Policing Research.

- Draughon-Moret, J., Todd, A., Rose, L., Pollitt, E. and 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): e28959. doi: [10.2196/28959](https://doi.org/10.2196/28959).
- Eisenhut, K., Sauerborn, E., García-Moreno, C. and Wild, V. (2020) Mobile applications addressing violence against women: a systematic review, *BMJ Global Health*, 5(4): e001954. doi: [10.1136/bmjgh-2019-001954](https://doi.org/10.1136/bmjgh-2019-001954)
- Emezue, C., Chase, J.D., Udmuangpia, T. and Bloom, T.L. (2022) Technology-based and digital interventions for intimate partner violence: a meta-analysis and systematic review, *Campbell Systematic Reviews*, 18(3): e1271. doi: [10.1002/cl2.1271](https://doi.org/10.1002/cl2.1271)
- Freed, D., Palmer, J., Minchala, D., Levy, K., Ristenpart, T. and Dell, N. (2018) ‘A Stalker’s paradise’: how intimate partner abusers exploit technology, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 667: 1–13. doi: [10.1145/3173574.3174241](https://doi.org/10.1145/3173574.3174241)
- Glass, N.E., Perrin, N.A., Hanson, G.C., Bloom, T.L., Messing, J.T., Clough, A.S., Campbell, J.C., Gielen, A.C., Case, J. and Eden, K.B. (2017) The longitudinal impact of an internet safety decision aid for abused women, *American Journal of Preventive Medicine*, 52(5): 606–15. doi: [10.1016/j.amepre.2016.12.014](https://doi.org/10.1016/j.amepre.2016.12.014)
- Koulouri, T., Macredie, R.D. and Olakitan, D. (2022) Chatbots to support young adults’ mental health: an exploratory study of acceptability, *ACM Trans. Interact. Intell. Syst.*, 12(2): 1–39. doi: [10.1145/3485874](https://doi.org/10.1145/3485874)
- Lucas, G.M., Gratch, J., King, A. and Morency, L.P. (2014) It’s only a computer: virtual humans increase willingness to disclose, *Computers in Human Behavior*, 37: 94–100. doi: [10.1016/j.chb.2014.04.043](https://doi.org/10.1016/j.chb.2014.04.043)
- Media Co-op (2023) Followitapp, 6 November, <https://mediaco-op.net/projects/follow-it-app/>.
- O’Campo, P., Velonis, A., Buhariwala, P., Kamalanathan, J., Hassan, M.A. and Metheny, N. (2021) Design and development of a suite of intimate partner violence screening and safety planning web apps: user-centered approach, *Journal of Medical Internet Research*, 23(12): e24114. doi: [10.2196/24114](https://doi.org/10.2196/24114)
- Oxford University Press (nd) Definition of Chatbot, 8 August. <https://www.oed.com/search/dictionary/?scope=Entries&q=chatbot>.
- Park, L. and Lee, J. (2020) Can a conversational agent lower sexual violence victims’ burden of self-disclosure?, in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA ’20)*, New York, Association for Computing Machinery, 1–8.
- Park, L. and Lee, J. (2021) Designing a conversational agent for sexual assault survivors: defining burden of self-disclosure and envisioning survivor-centered solutions, in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI ’21)* 634, New York, Association for Computer Machinery, 1–17.
- Piccolo, L.S.G. and Alani, H. (2020). *Prospecting Chatbots for Tackling Online Abuse*. Centre for Policing Research and Learning, Milton Keynes: The Open University.
- Pickard, M.D., Roster, C.A. and Chen, Y. (2016) Revealing sensitive information in personal interviews: is self-disclosure easier with humans or avatars and under what conditions?, *Computers In Human Behaviour*, 65: 23–30. doi: [10.1016/j.chb.2016.08.004](https://doi.org/10.1016/j.chb.2016.08.004)

- Sabri, B., Saha, J., Lee, J. and Murray, S. (2023) Conducting digital intervention research among immigrant survivors of intimate partner violence: methodological, safety and ethical considerations, *Journal of Family Violence*, 38(3): 447–62. doi: [10.1007/s10896-022-00405-6](https://doi.org/10.1007/s10896-022-00405-6)
- Storer, H.L., Nyerges, E. and 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 organizations, *Children & Youth Services Review*, 139.
- Sumra, M., Asghar, S., Khan, K.S., Fernández-Luna, J.M., Huete, J.F and Bueno-Cavanillas, A. (2023) Smartphone apps for domestic violence prevention: a systematic review, *International Journal of Environmental Research and Public Health*, 20(7): 5246. doi: [10.3390/ijerph20075246](https://doi.org/10.3390/ijerph20075246)
- Tarzia, L., Lyser, D., Thrower, E. and Hegarty, K. (2017) ‘Technology doesn’t judge you’: young Australian women’s views on using the internet and smartphones to address intimate partner violence, *Journal of Technology in Human Services*, 35(3): 199–218. doi: [10.1080/15228835.2017.1350616](https://doi.org/10.1080/15228835.2017.1350616)
- Tzelios, C., Contreras, C., Istenes, B., Astupillo, A., Lecca, L., Ramos, K., et al. (2022) Using digital chatbots to close gaps in healthcare access during the COVID-19 Pandemic, *Public Health Action*, 12(4): 180–85. doi: [10.5588/pha.22.0046](https://doi.org/10.5588/pha.22.0046)
- Westmarland, N., Hardey, M., Bows, H., Branley, D., Chowdhury, M., Wheatley, K. and Wistow, R. (2013) *Protecting Women’s Safety? The Use of Smartphone ‘Apps’ in Relation to Domestic and Sexual Violence*, Durham: Durham Centre for Research into Violence and Abuse, Durham University.
- Xu, L., Sanders, L., Li, K. and Chow, J.C.L. (2021) Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review, *JMIR Cancer*, 7(4): e27850. doi: [10.2196/27850](https://doi.org/10.2196/27850)
- Zarinpoush, F. (2006) *Project Evaluation Guide for Nonprofit Organizations: Fundamental Methods and Steps for Conducting Project Evaluation*, Toronto: Imagine Canada.