

Sociotechnical Imaginaries of Responsible Design: an AI Training Intervention for Gender-based Online Harm

Azize Naji

Goggleminds, UK **Correspondence:** azize@goggleminds.co.uk

Shima YekkehbashHeidari

Cardiff Business School, Cardiff University, UK

This position statement reflects on the design of AVA, an AI-powered, trauma-informed conversation simulator for professionals responding to violence against women and girls (VAWG), and its relevance to the Socio-technical Imaginaries for Responsible Design workshop. We examine how sociotechnical imaginaries of gender-based online harm shape the representation of technology-facilitated abuse, professional preparedness for online harm disclosures, and the design of AVA as a responsible AI-powered training system.

CCS CONCEPTS • Human-centred computing • Human computer interaction (HCI) • Collaborative and social computing

Keywords: Trauma-informed AI, Technology-facilitated abuse, Gender-based online harm, Responsible design

1 CONTEXT AND POSITION

Our multidisciplinary team at Goggleminds deliver immersive, AI powered simulation training across health, policing, education, and related sectors, collaborating with NHS Trusts, universities, and public bodies. We are currently developing AVA, an AI powered, trauma informed conversation simulator that is designed to mimic the voice of a survivor and enables professionals working in VAWG contexts to practice realistic conversations safely and at scale, without re-traumatising real survivors. AVA supports learning through structured feedback to build communication, critical thinking, and decision-making skills. It also provides adaptive modules with wrap-around materials including empathy, listening, and adaptability. To ensure authenticity and relevance to practice, AVA is co-designed with survivors and multi-agency professionals across policing, health, social care, government, and academia via focus groups, interviews, and pre/post-session surveys.

AVA aligns closely with the aims of the Socio-technical Imaginaries for Responsible Design workshop [1], particularly in exploring how responsible digital design can address and mitigate gender-based online harm. Below, we outline priority areas guiding AVA's development, informed by participants' feedback and survivor-centred design principles.

1.1 Realistic representation of online and technology-enabled abuse

AVA's scenarios should realistically and accurately reflect the realities of gender-based online and technology-enabled abuse patterns and include diverse, marginalized voices. This includes digital coercive control, online stalking, monitoring, harassment, surveillance, and the misuse of everyday technologies (such as phones, social media, smart home devices) [2]. With a rapidly changing technology landscape, including new forms of digital abuse being facilitated through the use of AI, it is more critical than ever that training is updated to reflect these contemporary challenges to better enable practitioners to support women and girls experiencing these.

1.2 Preparing professionals for complex disclosures of online abuse

Frontline professionals responding to VAWG often face complex disclosures of online and technology-enabled abuse, involving high information load, uncertainty, and fragmented, layered narratives with overlapping harms

and subtle risk indicators [3]. Survivors may describe constant contact, pressure to share passwords, threats involving intimate images, or fears of digital monitoring without naming these experiences as abuse or recognising escalating risk [4]. Traditional training struggles to replicate this complexity safely and at scale, while live role-play risks re-traumatisation [5]. AVA addresses this by simulating realistic disclosures and providing structured, trauma-informed feedback on recognising online risk cues and responding appropriately, including pacing, validation, risk and safety questions, and signposting. Pilot use in practice and education settings will explore whether this approach improves confidence and decision-making when supporting survivors of gender-based online harm.

1.3 AVA as a responsible, safe learning environment

A core consideration in AVA's design is ensuring that technology itself does not become a harmful, extractive digital space, particularly given that many learners have lived experience of trauma or violence [6]. For example, Figure 1 shows an animated survivor character within learning modules, allowing realistic engagement while avoiding the use of video or audio from real survivors.

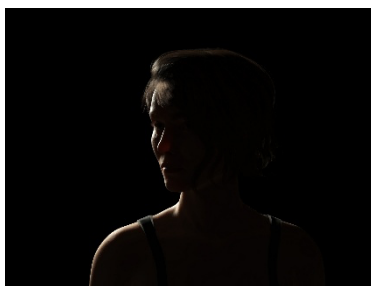


Figure 1: Example of an animated survivor character used in AVA's learning modules

Key principles guiding this work include preventing re-traumatisation despite the realism of scenarios and animated learning content; ensuring learner confidentiality and data privacy; delivering non-judgemental, supportive, trauma-informed feedback balanced between strengths and development; incorporating trigger warnings, break reminders, and guidance on seeking support; designing for accessibility and inclusion; and providing modules on learner wellbeing and safety, including lone-working protocols, awareness of vicarious and secondary trauma, and personal safety planning. We therefore approach AVA not only as a training intervention, but as a responsible sociotechnical intervention.

2 CONTRIBUTION TO THE WORKSHOP

Our priority development areas align with responsible design imperatives to avoid harm, do good, and govern the design process [7]. The *Socio-technical Imaginaries for Responsible Design* workshop is therefore directly relevant, offering a space to critically examine how gender-based online harm is imagined, represented, and operationalised. Engagement will support refinement of AVA's conversational model to reflect lived realities while avoiding simplification, stereotyping, or harm normalisation, and will enable critical reflection on our design choices within a broader discourse on responsible AI and sociotechnical imaginaries, while sharing and learning from others' experiences of navigating responsibility in technology design across varied organisational contexts. We see this workshop as a valuable space to contribute practitioner-grounded insights from AVA's development while deepening understanding of how responsible design frameworks can shape future training technologies addressing gender-based online harm.

REFERENCES

- [1] Min Zhang, A.K.B., Gordon Rugg, Irum Rauf, Dilrukshi Gamage, Bashar Nuseibeh, Silvia Podesta, Wanling Cai, Sarah Robinson. Sociotechnical Imaginaries of Responsible Design: A Case for Mitigating Gender-based Online Harm. in CHI Conference on Human Factors in Computing Systems. 2026. ACM.
- [2] Woodlock, D., The abuse of technology in domestic violence and stalking. *Violence against women*, 2017. 23(5): p. 584-602.
- [3] Freed, D., et al., Digital technologies and intimate partner violence: A qualitative analysis with multiple stakeholders. *Proceedings of the ACM on human-computer interaction*, 2017. 1(CSCW): p. 1-22.
- [4] Leader, C., A Socio-Technical Approach to Researching Technologically Facilitated Intimate Abuse, in *Tackling Domestic Abuse and Sexual Violence: A Systems Approach*. 2024, Springer. p. 75-92.
- [5] Turner, W., et al., Interventions to improve the response of professionals to children exposed to domestic violence and abuse: a systematic review. *Child abuse review*, 2017. 26(1): p. 19-39.
- [6] Hegarty, K., et al., Health practitioners' readiness to address domestic violence and abuse: A qualitative meta-synthesis. *PLoS one*, 2020. 15(6): p. e0234067.
- [7] Boehnert, J., M. Sinclair, and E. Dewberry, Sustainable and responsible design education: Tensions in transitions. *Sustainability*, 2022. 14(11): p. 6397.