CONCLUSION PAPER

RAN thematic event
24 April 2024, Zagreb, Croatia

Online P/CVE Interventions

Key outcomes

The **RAN thematic event** on **Online P/CVE Interventions** brought together practitioners with various backgrounds including exit workers, community police, youth workers, OSINT experts, and counter/alternative narrative experts to discuss the practice of P/CVE practitioners working online. We sought to answer the question: how can the internet provide solutions for intervention providers to connect with clients effectively?

We looked at the challenges practitioners face while working online, identified practical solutions to these, considered good practices in this field and neighbouring problem areas, and brainstormed about how to futureproof this work amid an evolving technological landscape.

Key outcomes:

- Practitioners need further training to scale their largely offline expertise to the online domain. Additional consideration must be given to develop new playbooks for doing this work effectively and safely.
- Given the nature of the threat, and the use of technology, no service should be exclusively offline. However, hybrid interventions may be possible, with the first contact made online. Investment should be given to this first moment of contact and then the necessary triage mechanisms to distribute cases to practitioners. A strong focus on trust building and positive user experience will lead to greater retention of clients.
- Advertising tools commercially can be used further for building brand awareness but may have limitations when it comes to reaching high-risk audiences. Rather, building strong multistakeholder partnerships with tech platforms is necessary to build stronger pathways for referrals.
- New and emerging technologies such as AI certainly provide opportunities, and practitioners should be empowered to innovate, with clear guidance about safety-by-design principles to mitigate risks. Any findings should be shared among practitioners to help further innovation and scaling.

Highlights of the discussion

There is consensus that violent extremism exists seamlessly across online and offline domains. Practitioners are interested to see how the emerging regulatory landscape affects the prevalence of extremist content online, and the extent to which it addresses former concerns that the internet was an "ungoverned space". However, they tend to agree that there is a gap when it comes to true online prevention and identify that supporting intervention providers to carry out their work online, where the problem is manifesting, and enabling online-offline referral pathways is the urgent solution required to bridge this gap.

Inspiring practices help to identify challenges

Inspiring practices were shared, each of which were recognised as partially filling the gap of true online prevention, but were neither scaled across the EU nor comprehensive in the solution they offered. For example, Moonshot presented their approach which uses commercially available advertising tools to redirect users towards alternative narratives but acknowledged the shortcomings of this approach in reaching high-risk users. Austrian digital streetwork approach Jamal al-Khatib was presented. They aim to be where the target audience is, and to communicate with them to encourage behaviour change but acknowledge the difficulties of scaling this approach, the need to keep pace with the different platforms used by the target audience, and the mixed results. The Dutch National Support Centre presented their web-based portal for individuals requiring support related to radicalisation but acknowledged the reliance on proactive outreach from the target audience to the service directly. Together, they nonetheless provided a jumping off point for inspiring discussions about future solutions.

Participants identified the challenges of:

- reaching the target audience, once it had been defined (branding and messaging to appeal to this target audience, incentivising engagement, the difficulties of advertising to this population);
- delivering and maintaining an online service to meet the needs of the target audience (providing this out of working hours, managing data protection, developing escalation pathways for imminent threats of violence, keeping your staff safe);
- retaining clients throughout the process, and converting them to long-term cases (e.g. building trust from the first contact, and maintaining this without face-to-face contact)
- training social workers and other intervention providers to take their work online, recognising the differences required to engage with clients in this setting (less control, more anonymity, shorter time).

Opportunities provided by emerging technologies

With the guidance of Violence Prevention Network, various opportunities to resolve these challenges were considered and further explored, in breakout sessions. These challenges include systemic shifts in the current terrorist and violent extremist landscape, and applications of technology, and include the following examples.

- Social media advertising tools are not set up for P/CVE practitioners to reach a very small target audience indicating a propensity to violence, therefore direct behaviour change in secondary/tertiary prevention is difficult.
    - However, they might be more effective in reaching bystander audiences such as family members; or could be used to raise general awareness of the support services available among the general public, or to build a base level of trust in the practice.
    - Social media integration may be a fruitful approach for web-based services, to improve the possibility of "meeting clients where they are", and could use existing tech such as messenger services, rather than requiring them to visit a third-party website that they have never been to before.
    - Digital streetwork could be scaled up on emerging or alternative platforms where advertising is more difficult, user numbers are smaller, but where there is a recognised problem with extremism. Engaging in comments and responding to target audience behaviour, as well as implementing secondary prevention approaches on-platform, or referrals to support services, may be a worthwhile approach.
    - There is still a space for counter- and alternative-narratives, and more broadly social and behaviour change communication interventions, and there is an emerging body of literature about how to mitigate risk (including backfire effect) and increase impact. For example, they are more likely to be effective if they promote action such as engagement with a support service.
- Direct integration with tech platforms might be a more useful way to reach this identified target audience, as tech platforms have the right data, which they currently use for content moderation and scaled enforcement practices.
    - Building consensus among tech platforms for a pivot towards a user-orientated prevention-focused approach and setting up systems to enable online-offline referrals, will be a larger upfront investment, but could compound the potential impact of online interventions through greater precision of targeting. This could be achieved through integration with law enforcement referral pathways, which already exist between tech platforms and police services for incidents of imminent threat to life; or they could be linked to scaled user/content enforcement systems, where there is an opportunity to communicate with users when they are deplatformed or content is removed. Alternatively, this could be operationalised through regulation-required transparency notices, in-app intervention experiences (like comment covers or public service announcements) and would likely look different on all platforms.
    - Participants identified some precedents for proactive safety interventions with at-risk users in neighbouring harms like child safety. We agreed that building safeguards into such interventions would be essential to mitigate risk and enable scalability, but otherwise recognised the need for greater multistakeholder integration between practitioner support services and the private sector.
- Chatbots and other AI-powered services could be a good supplement to practitioner-led online services, but never a replacement. They could help provide an "always-on" service to ensure full coverage for individuals who need support at times when a practitioner is not available.
    - If Large Language Models were trained using scripts, playbooks, and even conversations from practitioners, they could do a reasonable job in this scenario. This could be further boosted through training data from formers, or family members.

- They could match the communication style of the client to improve engagement and retention, aim to build trust and keep the conversation going, which may in some cases be less off-putting than a "real-life" conversation and could signpost resources.

- We agreed, however, that transparency about the use of AI in this service was non-negotiable, and that other safeguards should be built before experimenting with this practice.

Recommendations

1. Coordinate interactions with tech platforms to make clear, evidence-based, achievable requests for engagement with multistakeholder solutions in P/CVE. Prioritise small changes that can have big impacts and leverage existing bodies such as the EU Internet Forum or the Global Internet Forum for Counterterrorism.

2. When innovating with new technology, build in safety-by-design principles to mitigate risks, and prioritise managing the controllables, reducing assumptions, measuring robustly, and sharing information with fellow practitioners to help inform evidence-based solutions.

3. Consider the value of centralising online-offline referral pathways with wider distribution to a network of support services, to reduce the costs for developing new platforms. Invest in training P/CVE practitioners to receive new cases from online pathways, and in training non-P/CVE practitioners in the basics of P/CVE so as to scale up existing services rather than create new ones.

Inspiring practices

- The Dutch National Support Centre for Extremism is an organisation specialised in handling people and their network who are (possibly) radicalised, like family members. They offer consultation and advice, support, knowledge and experience. As a part of their support and consultation services, they have launched a chat option on their website. This chat box offers the opportunity to anonymously share their own concerns or concern for a loved one. Professionals will answer questions and provide extra information and support where needed.

- Online streetwork project Jamal Al-Khatib aims to deliver authentic alternative narratives to jihadist and Islamist-extremist online propaganda. They make short videos on the internet to counter extremist narratives online. They take on a narrative biographical approach and carry out online streetwork.

Further reading

RAN papers:

- RAN C&N (2022): Digital frontrunners: Key challenges and recommendations for online P/CVE work
- RAN FC&S (2022): Hybrid social work and digital awareness for family support
- RAN (2022): Hybrid youth and social work
- RAN REHAB (2022): Exploring hybrid and digital rehab work
- RAN Y&E (2022): Integrating the online dimension into offline pedagogical practices
- RAN (2022): The Online Dimension of Extremism and Improving Online P/CVE Efforts

Other:

- European Commission. (2022). Guidelines for teachers and educators on tackling disinformation and promoting digital literacy through education and training. Publications Office of the European Union.
- Evans, A. T., & Williams, H. J. (2022). *How extremism operates online: A primer*. RAND Corporation.
- Winter, C., Neumann, P., Meleagrou-Hitchens, A., Ranstorp, M., Vidino, L., & Fürst, J. (2020). Online extremism: Research trends in internet activism, radicalization, and counter-strategies. *International Journal of Conflict and Violence, 14*(2).