02/12/2022

CONCLUSION PAPER

RAN cross-cutting event
27 September 2022, Berlin, Germany

The Online Dimension of Extremism and Improving Online P/CVE Efforts

Key outcomes

This RAN cross-cutting event brought together practitioners, policymakers and researchers to explore the online dimension of extremism and corresponding preventing and countering violent extremism (P/CVE) efforts, find a shared understanding of challenges across the stakeholder community, and take stock of available monitoring tools and P/CVE interventions in the digital world. This paper presents the input of and discussions between participants at the meeting, including inspiring practices and lessons learned from the different stakeholder groups' respective experiences, and it underlines the following main points.

1) The online landscape is constantly evolving. Therefore, **understanding which platforms the target group is using and what they are currently discussing** remains key. Recent examples and inspiring practices can be found in this paper.

2) When it comes to systematic monitoring, content classification and threat analysis online, there is **a continuous need for human oversight to put content into context**. Prevention with a human rights and gender lens also requires the reduction of algorithmic or automatic AI-tracking systems and their replacement with more nuanced mechanisms or combining their application with human or manual monitoring.

3) With regard to **online P/CVE interventions**, several developments can be identified and leveraged, such as:
   a) using online alternative narratives instead of counternarratives;
   b) creating referral processes, including digital opportunities for people to seek support ('self-referral');
   c) using online advertising to redirect users away from extremist content; and
   d) the need to keep adapting media literacy approaches to new target groups as a non-confrontational and non-securitised approach focused on critical thinking, cognitive biases, communication skills and emotional resilience.

4) Various challenges in the field relate specifically to the EU General Data Protection Regulation (GDPR). **Various privacy as well as ethical issues** have raised concerns among practitioners as they seek to develop P/CVE interventions and activities online. Due to a missing legal and ethical framework, guidelines and trainings on ethical, legal and GDPR boundaries as well as on the mental well-being of practitioners are necessary.

Highlights of the discussion

With the development of new and emerging technologies, accelerated by the COVID-19 pandemic, people's lives both online and offline should be understood along a continuum. The concept of 'onlife' has been used to highlight the absence of a hard divide between the online and offline dimensions of an individual's life. Similarly, efforts taken by extremist groups to attract individuals can take place in both the physical and digital spheres simultaneously. A recent example is the racist terrorist attack in Buffalo (New York, United States), where the perpetrator mentioned having spent lots of time online due to the pandemic. He stated that he got inspired by previous violent right-wing extremist attacks and the extremist content he found online. In addition, the perpetrator livestreamed the shooting and posted a 180-page manifesto, which was shared widely online. To enable the development of P/CVE activities and interventions, either online or offline, that are able to address this phenomenon, it is important to grasp its breadth and multifaceted nature, which remains a common challenge identified during previous RAN activities and events (see more information under 'Further Reading').

Evolving narratives in the online sphere

The online landscape is constantly evolving and affecting behaviour in the offline world. Therefore, understanding where the target group of a given P/CVE intervention is located online and what they are currently discussing remains key. Since the COVID-19 pandemic, **conspiracy narratives** have increasingly been shared online and evolved around the following ideas: the established 'elites' are hiding the truth, COVID-19 was created to control the people, traditional media and institutions cannot be trusted, democracy does not work and is actually authoritarianism, and the people need to protect themselves as the authorities cannot be trusted. Staying updated on the constantly evolving narratives and the platforms on which they are discussed is important for the P/CVE field, as most conspiracy narratives are being spread online and can lead to "the belief that an in-group's success or survival can never be separated from the need of hostile action against an out-group" ([1]). The context of COVID-19 has also allowed for **violent right-wing extremist movements** across the world to successfully capitalise on shared anxieties, fuelling conspiracy narratives centring minority groups and authorities. More than 2 years into the pandemic, other newer topics have emerged:

- online content related to monkeypox recontextualising some of the same narratives seen around COVID-19;
- pro-Russian narratives connected to the Russian war in Ukraine stipulating that Russian president Vladimir Putin is fighting the 'deep state' and/or the World Economic Forum, that the war is a pretext to bring in more immigrants (Great Replacement narrative) or that President Putin is liberating imprisoned children in Ukraine (QAnon);
- a movement comprised of Dutch farmers claiming that their government is part of a larger scheme to increase population control, destroy traditional ways of living, and replace 'native groups' with immigrant groups (again, Great Reset, Great Replacement narratives);
- counter-intuitively to the concept of nationalism, a venturing beyond nationalism to the concept of whiteness can be observed as a general trend; and

([1]) Berger, J. M. (2018). *Extremism*. MIT Press.

- post-truth, a term that refers to a phenomenon whereby emotions and personal beliefs play a bigger role in shaping public opinion than objective facts.

Also, the number of incidents and cases of **xenophobia** from 2020 onwards increased across the world, with data indicating that this increase may be directly related to the COVID-19 pandemic. In addition, the 'incel' movement (an abbreviation of involuntary celibate), that is to a large extent active online, believes they are unable to access sex, mainly due to genetic factors, evolutionarily predetermined processes of mate selection and societal structures. The incel ideology can be seen as part of the online 'manosphere', a collection of online spaces promoting masculinity and misogyny, and opposing feminism. Lastly, **the confidence and trust in liberal democracy** has diminished and misinformation and fake news increasingly normalised. This normalisation of misinformation and sometimes extremist narratives has partly been a consequence of the lack of communication (due to lack of data) around centralising and imposing measures related to COVID-19. Also, disinformation (although not normalised) is increasingly being used online. Trusted community actors can play a crucial role in communication, as can independent fact-checkers (e.g. Maldita.es, Faktograf; more information under 'Inspiring digital practices'), and P/CVE actors should ask the question: 'Who has an audience that is willing to listen?'

Systematic monitoring, content classification and threat analysis online

When monitoring specific individuals, one of the first steps practitioners should take is to map their digital neighbourhood, which is as important as the offline component of their neighbourhood. Social media landscapes are personalised and depend on certain identity factors such as gender, age, language and (sub)culture. Further, the personal motivation to follow certain accounts is connected to a feeling of belonging to a certain in-group. Apart from forming the basis for the development of individualised P/CVE interventions at the individual level, such mappings are also crucial for monitoring activities and identifying trends and trajectories within local environments and they support the development of such tailored interventions.

Coded language and the human factor

When analysing threats there is a continuous need for human oversight to put certain data into a specific cultural context in order to be able to understand its meaning. New trends come and go so fast that AI alone is not enough to keep track. Furthermore, prevention with a human rights and gender lens also requires improvement of algorithmic or automatic AI-tracking systems and their replacement with more nuanced mechanisms or with human or manual monitoring. Some points that practitioners, policymakers and researchers should be aware of are:

- Stay up to date with the newest trends when it comes to **coded language** to be able to understand dog whistles and how to capture and identify the meaning of the coded language ([2]).
    - Certain emoticons or signs that have been created for general purposes are given a **different meaning** by extremist groups to disseminate their ideology. These 'trends' or 'operations' are created on platforms like 4chan and then disseminated on mainstream platforms like Twitter and Facebook ([3]).
- **Multi-agency groups** working on the national level are needed to address these constantly evolving challenges.

Leveraging technology for P/CVE purposes

Leveraging technologies in targeted P/CVE interventions can be challenging. This is partly because of a lack of understanding of the fast-evolving online ecosystem, but also because practitioners and researchers may not be

([2]) An example is the 'OK' hand sign used to signify 'white power'; see: https://www.nytimes.com/2019/12/15/us/ok-sign-white-power.html
([3]) Other examples are 'Operation Blue the Jew', 'Operation Jogger' and 'Operation Jewish Claps' whereby images, words or emoticons are discussed on platforms like 4chan and given a different meaning than the usual, mainly connected to right-wing narratives. Subsequently these 'codes' are disseminated on mainstream media to spread these narratives.

able to access online spaces where extremists operate. It is therefore important to explore how a more systematic use of technology can be used and further explore cooperation with security actors within national and European legal frameworks, specifically the GDPR.

Several developments can be identified with regard to using technology for P/CVE purposes, such as:

- using **natural language processing**: a branch of AI concerned with the programming of computers to process and analyse large amounts of language data to understand text and spoken words in a similar way as humans do;
- **scraping** selected profiles of individuals or groups from platforms and then checking the data for extremist content;
- **advertising** technologies: inserting referral mechanisms based on individual searches by placing ads leading to offline counselling on websites with extremist content; and
- utilising **social media companies' data** for research purposes and building trusted relationships to harness opportunities for collaboration.

Digital P/CVE interventions

P/CVE campaigns

When talking about digital P/CVE approaches, digital counter- and alternative narratives are one of the first approaches that come to mind ([4]). Some points resulting from discussions to be aware of (and that can cause negative effects) are:

- Research into counternarratives and counter-memes has shown that they can cause **'backfire effects'**, with some experts noting that counternarratives might even accelerate radicalisation.
- If alternative messaging aimed at individuals on the pathway of radicalisation comes from **a source that is not trusted** (e.g. government-led campaigns), it often does not work.
- The flagging or banning of accounts spreading harmful content with radicalising purposes tends to be gendered or determined by **assumptions about gender-based signs of radicalisation** and engagement in violent narratives.
- Extremist or harmful narratives can be easily **intertwined with mainstream discourse** and so pave the way towards endangering fundamental human rights.

Rather than counternarratives, the integration of positive conversations about identity, reconciliation and tolerance into the public space through media assistance is a relevant option. Another option is looking at ways to offer counselling to individuals who are attracted to extremist content. An educational and pedagogical constructive approach is needed beyond framing things as P/CVE or counterterrorism, as security language can have a negative connotation. In terms of the right organisation to carry out this work, it should be noted that for NGOs it can be easier to communicate online as opposed to state actors, as they might enjoy more trust. For this, however, collaboration between state actors and NGOs is needed.

([4]) The alternative narrative is one of several 'soft' approaches to preventing and suppressing radicalisation to violence (based on the United Nations Security Council Resolution 2354 (2017)). Distinct from its sister strategy, the counternarrative, an alternative narrative is not intended to directly challenge the content of violent extremist rhetoric. Instead, it serves to undermine its 'predominant assumptions' (Gavin, 2005). Ultimately, an alternative narrative tells a different story, focusing on what society is 'for' rather than 'against', whilst remaining completely distinct from the 'discourse and influence' of a dominant, problematic narrative (Adame & Knudson, 2007). For more information, see this RAN paper: Lessons Learned from Alternative Narrative Campaigns (2021).

Transferring from offline to online: digital 'street work'

When it comes to online youth or street work, practitioners are increasingly present in the online sphere. For example, replying in comments and via direct messages, or DMs, is a way of getting in contact online (see also 'Inspiring digital practices'). Some reflections that were shared during the meeting are:

- 'Online street work' usually includes three aspects: **empowerment, emotional reflection and a question**. In this way, people feel seen and heard.
- There are pros and cons of using **an organisational versus a private account**. On the one hand, an organisational account can impose barriers in establishing contact/rapport with target groups as it is more impersonal compared to a private account; on the other hand, it better safeguards practitioners' privacy.
- P/CVE practitioners indicate that **interpretation online misses body language indicators**: it is difficult to estimate someone's true feelings.
- In addition, **trust** is even harder to build online than offline.

Digital/media literacy for practitioners and the general public

Besides education on digital literacy for first-line practitioners, wider media literacy awareness as a way of early prevention should be sustained and become more systematic. This entails developing a broader and more in-depth understanding of dis- and misinformation and better identification of fake news. This should be combined with a non-confrontational and non-securitised approach focused on critical thinking, cognitive biases, communication skills and emotional resilience. Practical experiences shared during the meeting indicate that critical thinking and information assessment skills are most effectively developed when also emotional intelligence, empathy and social skills are being enhanced in parallel ([5]). There is a need to incorporate this more in the school curricula and explore how this can also be adapted for an older target audience with different needs.

Legal and ethical concerns

Various challenges relating specifically to the GDPR, privacy and various ethical issues have raised concerns among practitioners as they seek to develop P/CVE interventions and activities online. There is a lack of understanding of the legal boundaries within which they can operate as well as of the possible judiciary implication they could face in situations of infringement. This challenge is likely to hamper the development of digital P/CVE interventions. Specific issues and questions that were raised are:

- **Ethical concerns:**
  - Practitioners are in vulnerable positions especially if a situation escalates online and your real name is being used. If you use a fake name, however, this brings up ethical issues and could harm the relationship.
  - Monitoring an open versus a closed group: whereas a closed group requires creating alternative identities, this falls into the sphere of intelligence agencies.
  - What are the indicators along the online radicalisation pathway that warrant intervention from law enforcement agencies and intelligence?
  - How can the mental well-being of practitioners be protected?

- **Legal concerns:**
  - Because a lot of information is available online, what can we monitor and what are the legal boundaries for digital monitoring?
  - Can practitioners approach individuals online without consent?

([5]) Bulgarian Safer Internet Centre (2021). Guide for development of critical thinking and fostering resilience. Also: Centre for the Study of Democracy. (2021). Overcoming youth vulnerabilities to far-right narratives. Policy Brief No 102.

- Does monitoring and approaching individuals need to be documented in the same way as offline cases? Reporting can be challenging, especially if it is not certain whether the reports will be published.
- Especially small civil society organisations (CSOs) are struggling with the GDPR. Differences between national legal (and organisational) and ethical frameworks in the EU make it challenging to exchange good practices.

Recommendations

Recommendations for policymakers, practitioners and researchers with regard to ethical and legal frameworks:

- Develop **clear ethical guidelines and a legal manual** for CSOs engaging in online P/CVE and raise awareness about this among practitioners and policymakers.
- Develop **trainings** for practitioners on dealing with ethical and legal issues and inform them about GDPR rules and the implications these have for their work (see more background information under 'Legal and ethical concerns').
- Take into account **guidelines for the mental well-being of practitioners** who are engaged in online P/CVE work.
- The transnationality of the online threat has rendered national laws, policies and measures against this threat significantly less effective — explore and develop **transnational approaches to P/CVE** in spite of the lack of common legal frameworks and definitions.

Recommendations for practitioners and researchers with regard to strengthening understanding of the online dimension:

- Invest in **better understanding** trends in online extremism and keeping an overview of existing digital and hybrid P/CVE projects.
- Create an **inventory of problematic and coded language** accessible for practitioners and updated regularly based on the most recent research and monitoring findings.
- Include **digital skills in the curricula** for P/CVE practitioners.
- Train **monitoring skills** in relation to practitioners' and security agency staff's respective profiles, responsibilities and needs.
- Enhance digital literacy for the wider public. Specific angles can be:
  - educating **parents** on early signs of radicalisation online;
  - enhancing online literacy of **older generations**;
  - improving **cultural literacy**;
  - supporting **societal role models** (community, religious, etc.) in engaging online as well.

General recommendations for national authorities and policymakers:

- When it comes to the list of recommendations above, pay attention to **financing these types of projects** and pushing them at the policy level.
- **Include the online dimension** in national P/CVE strategies and prioritise the online dimension in curricula for P/CVE practitioners, such as teachers, youth workers, social workers and psychologists.
- **De-securitise** language. Avoid security language in educational responses: for example, do not use terms such as 'prevention of radicalisation' and focus on positive elements such as fostering democratic life and strengthening critical thinking and emotional intelligence.
- Work on a **transnational P/CVE approach**, because the online threat also does not stay within national boundaries. This can focus on tech companies or NGOs sharing knowledge and new developments, but also on an ethical framework beyond national boundaries.
- Governments need to work **bottom-up** and collaborate with local NGOs and communities.
- State actors should build **sustainable partnerships** with the private sector, like social media companies, to harness opportunities for collaboration with regard to the leveraging of technology for P/CVE purposes.
- In security agencies, there should be **semi-automatic monitoring** (automatic detection of keywords based on a dictionary/taxonomy), followed by content analysis. The keywords the semi-automatic monitoring is based on should be regularly updated by practitioners.

Inspiring digital practices

Leveraging technology:

- The **Redirect Method** is an open-source methodology that uses targeted advertising to connect people searching online for harmful content with constructive alternative messages.
- **CrowdTangle** is a tool with which you can follow, analyse and report on what's happening with public content on social media.
- The **hash-sharing database** of the Global Internet Forum to Counter Terrorism (GIFCT) is a safe and secure database of images and videos produced by terrorist groups that members of GIFCT had removed from their platforms.
- The **OSINT Framework** breaks down different areas of interest and connects them to tools that could help you gather the information you need.
- **Maltego** is a tool for graphical link analyses that offers real-time data mining and information gathering, as well as the representation of this information on a node-based graph, making patterns and multiple order connections between said information easily identifiable.

Monitoring and detecting early signs and interventions:

- The **City of Mechelen** is building an online branch of their local approach to P/CVE — one that complements and strengthens their current offline approach.
- **Faktograf.hr** is a non-profit media organisation specialised in fact-checking, i.e. checking the factual accuracy of claims in the public space.
- **Maldita.es** is a non-profit organisation that aims to control disinformation and public discourse with fact-checking and data journalism techniques.
- The **European Observatory of Online Hate** is a standardised framework for monitoring more than 20 social media platforms.

Youth work in the digital space:

- The project **streetwork@online** seeks to establish a dialogue with young people from Berlin between the ages of 16 and 27. This will strengthen their capacity for critical reflection and support them in their identity formation.

Early prevention through media literacy and building online resilience:

- The online campaign **Find Another Way** aims at building resilience among Bulgarian youth for far-right narratives online.
- The Bulgarian Safer Internet Centre created a **guide for teachers** with guidelines on critical thinking and fostering tolerance among students from the age of 14 to 19.
- The **Safer Internet Centres** possesses many resources on media literacy and critical thinking.

Working evidence-based:

- The **INDEED** project aims to use evidence-based approaches to strengthen first-line practitioners' and policymakers' knowledge, capabilities and skills for designing, planning, implementing and evaluating preventing violent extremism, countering violent extremism and deradicalisation initiatives.

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 Ad hoc paper (2022). Hybrid youth and social work
- RAN REHAB (2022). Exploring hybrid and digital rehab work
- RAN LOCAL (2022). How to deal with the local impact of online (extremist) activities
- RAN Y&E (2022). Integrating the online dimension into offline pedagogical practices
- RAN LOCAL (2021). An online P/CVE approach for local authorities: challenges, tips & tricks

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.
- European Commission. (2022). *Final report of the Commission expert group 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.
- Molas, B. (2022). *Dutch flags and maple leaves: How conspiracy theories created a transnational far-right.* International Centre for Counter-Terrorism.
- Szmania, S., & Fincher, P. (2017). Countering violent extremism online and offline. *Criminology & Public Policy, 16*(1), 119-125.
- Valentini, D., Lorusso, A. M., & Stephan, A. (2020). Onlife extremism: Dynamic integration of digital and physical spaces in radicalization. *Frontiers in Psychology, 11, Article 524.*
- 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).