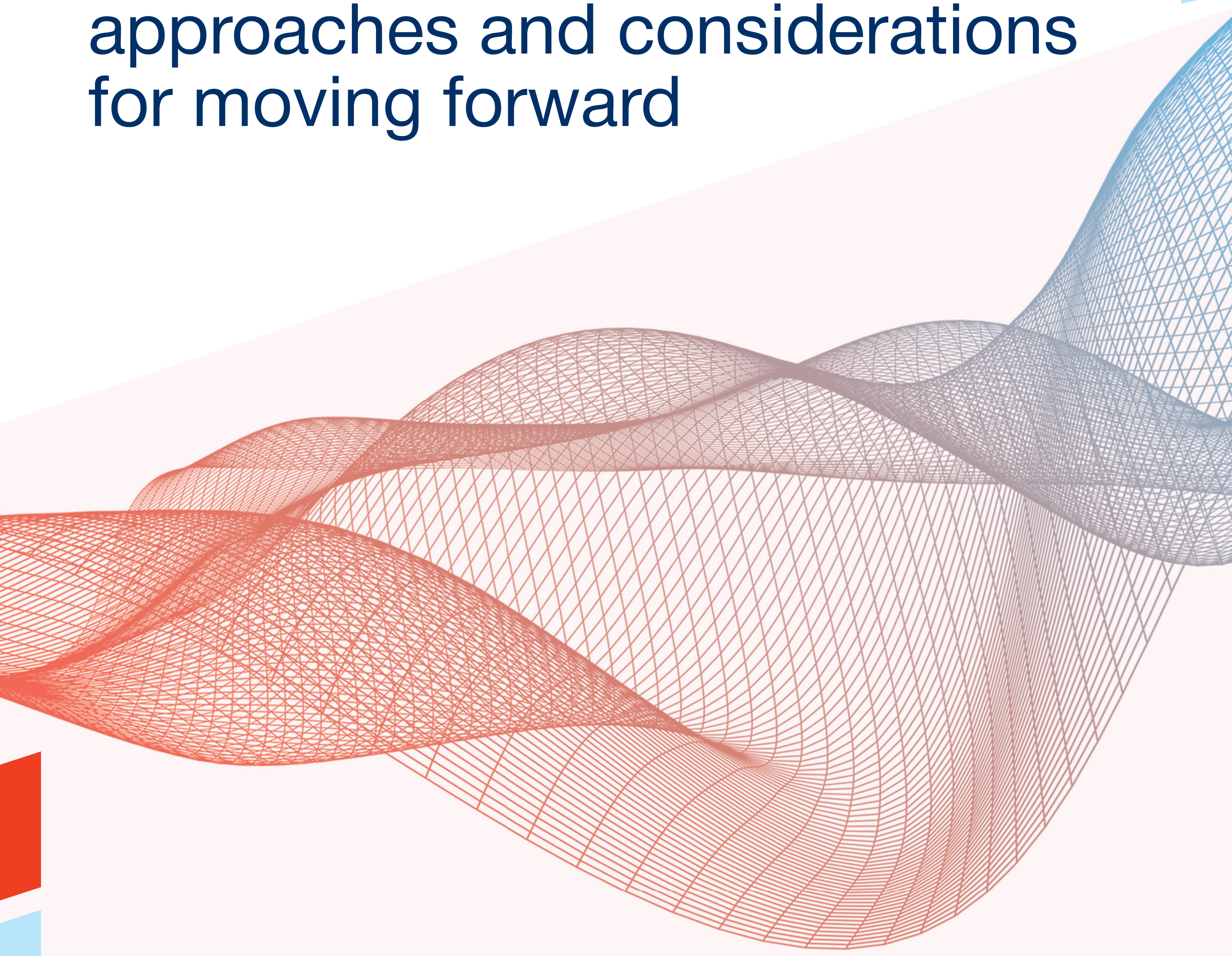
Office of the Special Representative and Co-ordinator
for Combating Trafficking in Human Beings

BRIEF

Policy responses to technology-facilitated trafficking in human beings:

Analysis of current approaches and considerations for moving forward



OSCE Organization for Security and
Co-operation in Europe

Introduction

Internet and communication technology (ICT) has led to the emergence and rapid expansion of technology-facilitated trafficking in human beings (THB).[1] The misuse of technology has become central to the business model of human traffickers and is present at each stage of the crime from grooming and recruitment, to control and coercion, to exploitation. At the most basic level, ICT – and the internet specifically – facilitates connectivity among perpetrators, between traffickers and their victims, as well as with users of goods and services extracted from victims.

As technology becomes ever more central to both licit and illicit marketplaces – a situation accelerated by the COVID19 pandemic – the challenge posed by technology-facilitated THB is set to increase. Effective and comprehensive responses are therefore urgently required. In particular, measures that foster safety and counter the harms – including substantive human rights violations – facilitated by technology are needed.

Until now, the primary policy response of governments to this challenge has been to allow the technology industry to self-regulate and voluntarily enact safety measures. This approach has not succeeded; in fact, the problem has grown significantly worse. Current efforts are almost entirely reactive and focus on removing previously-identified exploitation materials featuring children; measures to identify content featuring exploitation of adults are virtually non-existent.

In light of the growth of technology-facilitated THB as well as limited policy action by governments, the OSCE Office of the Special Representative and Co-ordinator for Combating Trafficking in Human beings developed the occasional paper ***Policy Responses to Technology-Facilitated Trafficking in Human Beings: Analysis of the Current Situation and Considerations for Moving Forward***. This Brief summarizes the findings of the Occasional Paper, drawing attention to different policy approaches taken by OSCE participating States to tackle technology-facilitated THB, including the successes and failures of self-regulation. It examines how technology-facilitated THB is addressed in criminal justice frameworks; the policy approaches taken towards online platforms; and the policy challenges specific to combating technology-facilitated THB. Finally, it offers recommendations for policy makers on developing policies and legislation to combat technology-facilitated THB.

Addressing technology-facilitated THB in criminal justice legal frameworks

Two threshold questions must be considered when examining the response to technology-facilitated THB: 1) whether technology-facilitated THB is covered in the definition of THB in national legislation; and 2) whether criminal procedures cover the investigation and prosecution of technology-facilitated THB, including the collection and use of electronic evidence in court.

First, with regard to statutes criminalizing THB, there is an ongoing debate as to whether technology-facilitated trafficking should be explicitly recognized in international and national legal frameworks, or whether existing definitions are sufficiently flexible and do not require amendment.

Currently, the predominant approach among the OSCE participating States is to apply THB frameworks originally crafted for “offline” contexts to technology-facilitated THB offences without express reference to technology in the statutory definition of the crime. Many practitioners do not believe an express reference to technology in legislation is critical. However, some stakeholders argue that incorporating an explicit reference to technology would be a valuable tool to ensure that traffickers do not escape justice for technology-facilitated crimes, as well as help raise awareness and mobilize resources to address these crimes.

[1] “Technology-facilitated trafficking” is understood for the purposes of this report as human trafficking offences (defined in line with the UN Protocol on Trafficking of Persons) that use digital technologies during any element of the offence.

As a middle-ground approach, in many jurisdictions policymakers could adopt interpretive guidance to clarify that the legal definitions of THB include technology-facilitated THB. This would extend application of the law to such offences and ensure laws are applied coherently.

Second, while inclusion of references to technology is generally not seen as an urgent matter by many practitioners, there is little ambiguity about the critical need to reflect technology-facilitated THB in national codes of criminal procedure, which impacts the collection and storage of online evidence, access to electronic devices and the collection of evidence using artificial intelligence.

Here, deficiencies in legislative approaches are numerous. For example, with regard to child sexual exploitation (CSE), both the Council of Europe Convention on Cybercrime (the Budapest Convention) and the Council of Europe Convention on Protection of Children against Sexual Exploitation and Sexual Abuse (also known as the "Lanzarote Convention") emphasize the need for procedural reform enabling effective investigation and prosecution of CSE facilitated by ICT. However, while a number of countries have responded by enacting legislation that regulates criminal procedure, there remain a number of OSCE participating States that lack regulatory frameworks governing the collection and use of digital evidence, or that have frameworks premised on voluntary data sharing. Moreover, the Council of Europe conventions are not comprehensive in addressing human trafficking: the Budapest Convention does not expressly reference THB and the Lanzarote Convention only applies to exploitation of children, not adults.

Exacerbating these challenges is the fact that online platforms commonly delete unlawful content that is reported to them or that they identify in their own investigations. The Committee of Ministers of the Council of Europe has recommended that intermediaries should retain such material to facilitate criminal investigations, however, this recommendation has not yet been implemented within many national codes of criminal procedure, further hampering prosecutions of technology-facilitated THB offences.

Another barrier to the collection of evidence is the general inability of law enforcement to covertly access devices as part of an investigation. Historically, States have not provided sufficient legal protocols on how to carry out such procedures at the national level, limiting proactive investigations. However, a growing number of OSCE participating States have introduced legislation allowing law enforcement to access suspects' computers when investigating technology-facilitated offences, including THB. A second wave of countries are currently in the process of drafting and enacting legislation to harmonize procedures for online investigative techniques for technology-facilitated criminal offences, the collection of electronic evidence, and the use of electronic evidence in prosecutions.

A further emerging issue in the area of e-evidence is the generation of evidence with the use of artificial intelligence (AI) tools where the human factor is minimal or absent. Examples of e-evidence in THB cases generated by software without human intervention already exist, including a number of projects that use chatbots to engage with sex buyers attempting to procure "services" of THB victims. Although this practice is already used in some OSCE participating States, there is no consensus on how policymakers and magistrates in the OSCE region treat evidence gathered by an AI system, highlighting the need for clear policy guidance in criminal justice proceedings.

Policy approaches to online platforms

To date, governments have generally allowed the technology sector to self-regulate on the topic of combating exploitation and THB. However, regulatory approaches in this area are currently in a period of transition at both the national and regional levels. Several key topics dominate policy discussions. First is the degree to which the operations and practices of the technology sector should be self-regulated, co-regulated or government-regulated. Second - and closely related to the self-regulation / government-regulation debate - is whether compliance with industry standards should be voluntary or mandatory.

A primary example of **self-regulation** is the evolution of Terms of Use. Online platforms have responded to growing public critique of widespread misuse of their services by making their Terms of Use increasingly stringent. A number of social media companies, including some of the largest like Facebook, VKontakte and Youtube, commit in their Terms of Use to removing content that "facilitates or coordinates the exploitation of humans, including human trafficking."[2] Given the increasing volume of exploitative content online, however, it seems clear that such Terms of Use are not broadly effective as a deterrent.

A second example of self-regulation has been attempts to harmonize responses across the technology industry, often through the establishment of multi-stakeholder initiatives involving different types of organizations and stakeholders.

Nonetheless, despite promising initiatives such as the Voluntary Principles to Counter Online Child Sexual Exploitation, recent history confirms there are many shortcomings in self-regulatory approaches, particularly when compliance is **voluntary**, including: limited or non-existent industry standards; inconsistent and inadequate adoption and application of voluntary principles; lack of incentives for compliance at scale in self-regulatory frameworks; and broadly worded rules lacking in clear indicators of compliance or breach.

In short, there is a need - and indeed a growing push - to move away from total reliance on self-regulation and toward robust **State-led regulatory frameworks**, including those which combine aspects of self-regulation with enhanced State powers and oversight. Such frameworks can foster harmonization and a level playing field, while avoiding safe havens for criminals and challenging impunity.

Specific issues related to regulating the technology industry on THB

Beyond the overarching questions of self-regulation versus State-led regulation, and voluntary versus mandatory compliance, a number of key policy issues relevant to THB arise in regulating the technology industry across the entire arc of technology development from design to development to consumer use: prevention of harm including through age verification; monitoring content online; liability for harmful content; removal of prohibited content; and the blocking of online platforms.

First, prevention efforts to curb misuse of technology are fundamental. Among promising prevention initiatives, the use and development of **age verification** guidelines - including those that focus on verifying the age of individuals depicted in uploaded material, individuals uploading material, and individuals viewing explicit materials - deserves increased adoption and implementation. These tools can help ensure that content being uploaded or shared on online platforms does not feature minors, for example in sexual service advertisements.

Second, monitoring of content online is critical to identifying exploitative content and removing it. Consistent with the traditional self-regulation and voluntary approaches used in most countries, the **monitoring of content** on platforms has been guided by the bedrock principle that online platform companies have no obligation to monitor third-party content. The principle of no obligation to monitor is often been linked with a second, equally fundamental principle - **no liability for third-party content** – and the two principles related to monitoring and liability are typically enshrined together in national legislation.

However, the long-accepted inviolability of these two principles is currently being challenged on a number of fronts. Critics of the

[2] See Facebook, Facebook Terms of Service. Available at: www.facebook.com/terms.php (accessed 21 October 2021).

protections afforded by these principles to online platforms argue that they were designed to enable the growth of the internet, and that they are no longer required in an age in which online platforms have profits exceeding the GDP of many States. Moreover, there is now a corresponding need to prioritize the safety of users online and those exploited via online platforms over economic concerns or the right to privacy.

States are increasingly exploring a variety of measures to encourage or mandate monitoring by companies of their services or platforms. Key questions in this respect involve what content companies should monitor for, the degree to which monitoring is conducted by human moderators or tech tools and what reporting obligations they have to authorities. Moreover, the evolution of monitoring is being directly challenged by privacy restrictions and encryption.

The principle of no liability is also being reconsidered by recent jurisprudence and new legislation holding online platforms accountable – either from a civil or criminal perspective. One example is the United States' 2018 enactment of the FOSTA-SESTA package. The passage of the Online Safety Act 2021 in Australia, an OSCE Partner for Co-operation, has also changed the approach to the principles of monitoring and liability of online platforms. Central to these policy initiatives is the desire both to hold technology companies accountable for harm that they knew – or should have known – about, and to give victims of harm an avenue to seek redress.

Third, and closely related to the topic of monitoring is the issue of how to regulate the **removal of illicit or harmful content** once it has been identified. Although a significant proportion of content-removal by online platforms is voluntary based on the company's individual terms of service, a number of OSCE participating States have now enacted regulatory frameworks requiring online platforms to remove certain content that has been detected and empowering State authorities to compel online platforms to block or remove content. Central to this conversation is defining the content to be removed. Some approaches focus on requiring removal illegal content only, such as child sexual abuse material. However, there is increasing support for broader approaches that extend removal to content which is not per se illegal but that, for example, causes psychological harm, is done without the consent of the depicted person or is otherwise exploitative.

A more substantive form of content removal has been the development of legal instruments that allow State authorities to **take down or block entire websites** where prohibited content resides. This approach is most commonly done with regard to THB for sexual exploitation. The blocking of websites known to host content related to THB is occasionally at odds with law enforcement practitioners, who argue that such sites can provide valuable sources of information for investigations. Again, however, the issue of scale becomes dispositive since the intelligence gains from such sites are typically far outstripped by the harms caused to victims from their continued operation.

Finally, a key topic in policy development is the establishment of **transparency** obligations on technology companies, in particular public reporting on the volume of illicit activity on their platforms and the steps taken by the companies to prevent or mitigate the misuse of their services. While some companies currently report such information, the voluntary, self-regulation framework to date has resulted in fragmented and inconsistent transparency, obfuscating the scale of the crime and the response to it. Harmonization and clarity on standards is needed in order to provide policy makers, and the public, with the information needed to understand online exploitation and develop appropriate policy responses.

Permeating these key topics are debates on the appropriate relationship between combating exploitation and fostering safety online on one hand, and upholding other rights or principles such as freedom of expression and privacy on the other. Historically, safety online has taken a back seat to other considerations, however, policy makers can no longer ignore the profound need for action to improve safety for everyone.

Conclusions and recommendations

The misuse of technology by traffickers has been facilitated by inadequate protections across the technology sector. These issues primarily manifest in widespread grooming and recruitment, power and control over victims, and exploitation through depictions or advertisements.

While some companies have developed measures or tools to respond, the reliance on self-regulation has resulted in fragmented and inadequate adoption of safety measures, inconsistent and slow reporting to authorities, lack of redress for victims, and impunity for traffickers. Moreover, current efforts have focused primarily on reactive identification of previously known child exploitation material; actions to proactively prevent the dissemination of new material, to prevent grooming and exploitation, and to implement default safety measures have been minimal. Initiatives to address online exploitation of adults are almost completely absent.

Technology-facilitated THB requires strong legislative action by governments to establish mandatory industry standards, harmonize approaches and support enforcement. Policy development should involve input from the technology sector and civil society and take into account the unique characteristics of different platforms, but State-led intervention is critical. Thus, OSCE participating States should:

1. Ensure that technology-facilitated THB is covered by national legislation criminalizing THB and by relevant codes of criminal procedure, thereby ensuring that investigators and prosecutors have the necessary procedural tools to investigate, collect evidence, share information, bring indictments and present evidence in court.

2. Enhance State-led regulatory frameworks. Such regulation should prioritize safety and include robust, mandatory obligations on core responsibilities including:

a. Strong prevention measures including:

i. “Safety-by-design” principles in the design, development, and distribution of products and systems;
ii. Age-verification for persons depicted in, persons uploading, and persons viewing sexually explicit material. Consent verification should also be explored for any sexually explicit content prior to its distribution;
iii. High-visibility content removal request mechanism;

b. Due diligence obligations for their operations and systems to identify risks of misuse and mitigate them, including:

i. Undertake proactive monitoring for exploitative or harmful materials (not only illegal) and for misuse of platforms, and establish mechanisms to allow for direct
reporting by the public to companies;
ii. Remove prohibited content expeditiously, preserving it safely for possible use in investigations/prosecutions;
iii. Report illegal content to appropriate/designated authorities;
iv. Enforcement mechanism for failure to comply with the above;

c. Liability for harm caused by content on the platforms or exploitation occurring through the platform. Liability should be based on a “should have known” standard.

d. Transparency standards regarding the reporting of platform misuse, the steps taken to mitigate misuse and the outcomes of such efforts.

3. Strengthen cooperation between States, the private sector and civil society with the aim of improving data gathering and sharing between law enforcement, anti-trafficking actors and other relevant stakeholders.