Terre des Hommes

GROOMING AND DETECTING IT

WHAT IS GROOMING?

Grooming is the process of establishing/building a relationship with a child either in person or through the use of the Internet or other digital technologies to facilitate either online or offline sexual contact with that person.

There are many different ways in which someone can groom a child, however, grooming is generally done through manipulation: **emotional connection** and **feelings of trust** are used against children for the purpose of abuse. By slowly pushing emotional and/ or physical boundaries (e.g. discussing sexual topics in a non-age-appropriate way), child groomers **normalise manipulation and abuse**.

To reach a large number of potential victims and to avoid detection, perpetrators of online grooming use multiple channels (NCMEC). A common tactic used by perpetrators is called '**off-platforming**': from public platforms, they move the exchanges into applications that use end to end encryption (**E2EE**) or ones that don't have detection tools (WeProtect).

In 2022, IWF found that **78%**, of all CSAM identified, included **content created by offenders grooming** and encouraging the children to behave sexually over a webcam (193% increase compared to 2020, IWF).

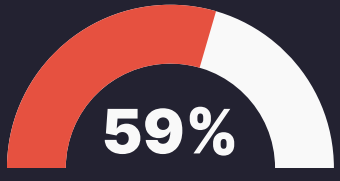
TYPES OF GROOMING

1. **Unwanted requests for sexual favours or images**
2. **Requests for sexual conversations**
3. **Manipulating children into sexual relationships**

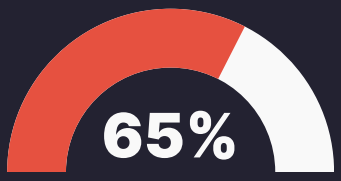
GROOMING AFFECTS CHILDREN FROM 8 TO 16 YEARS OLD, MOSTLY BETWEEN 11 AND 13

Almost 2 in 5 children were asked to do something sexually explicit online they were uncomfortable with or did not want to do. (WeProtect)

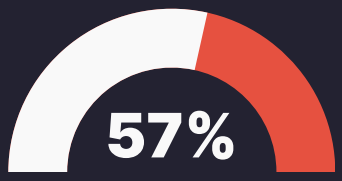
1 in 4 children had an adult they knew or someone they did not know ask them to keep their sexually explicit interactions online a secret. (WeProtect)



TRANSGENDER/NON-BINARY CHILDREN experienced online sexual harm



LGBTQ+ CHILDREN experienced online sexual harm



CHILDREN WITH DISABILITY experienced online sexual harm

Source: Global Threat Assessment 2021, WeProtect

GROOMING & GAMING

Despite that more than 70% of the platforms involved were Snapchat and Meta-owned applications (Facebook, Instagram and Whatsapp), online gaming platforms pose an ulterior challenge. There, interactions between adults and children are normalised. (UNICRI, WeProtect)
Key features of social gaming platforms, as well as metaverses, which can facilitate grooming include:

1. The player/user is embodied in a digital space through an avatar;
2. The player/user is able to interact with both friends and strangers;
3. The interactions can be multilayered, incorporating multiple ways to socialise, e.g. live communication audio chat or avatar gestures;
4. The **experiences are immersive through high levels of engagement** (VR and AR headsets).

"Sometimes I talk and play Fornite with people from other places and sometimes, I lie about my age for security reasons."

14 years old boy, Bolivia
(Child Safety by Design)

"It is like fooling yourself, it can lead you to have an older person talk to you as if you were their age, they can tell you things that you don't know or that you should not learn at your age."

15 years old boy, Bolivia
(Child Safety by Design)

PREVENTION & TOOLS

The focus often lies on identifying sexual predators or criminal material after the violence took place. While these efforts are vital for post-incident response, we need to shift our attention towards proactive measures.

By leveraging the power of computational risk detection, we can identify individuals or patterns that indicate potential risks, enabling us to intervene early and prevent harm.

Risk mitigation strategies from service providers, such as **warnings, prompts, nudges and intelligent coaches** are highly effective:

- Intelligent assistance or nudges play a crucial role in helping children make better decisions while navigating the online world
- Warning pop-ups and educational prompts can teach children how to navigate risk in real-time.

Standardising safety messaging across platforms and utilising privacy preserving human-centred technologies further enhances the effectiveness of prevention efforts.

AI & DETECTION

AI technology can be used to identify and stop online grooming of children at an **early stage** (e.g. in chat conversations) even if the language is coded or cryptic, as automated tools can help identify potential exploitation patterns.

Proposed systems can be based on natural language processing (NLP), for instance, a bag of words (BoW) approach, and use sets of classifiers to automatically distinguish conversations linked to child grooming. Chatbots using NLP can be deployed to engage and identify perpetrators online.

AI detecting of promoters of harmful video content and comments on platforms such as YouTube Kids, with a detection rate of 85.7% (Kaushal et al, 2016) or Microsoft's grooming detection with an accuracy rate of 88%, before human review (EC, 2023).

ANTI-GROOMING TECHNOLOGY

Anti-grooming technology is built for purpose and to minimise privacy invasion. Robust human review and oversight ensures that only criminal grooming conversations are flagged to law enforcement.

By combining pre-incident risk detection with post-incident risk mitigation strategies, we can create a comprehensive preventative approach, enhancing effectiveness and sending a strong message that we are committed proactively to safeguarding children from harm.

Terre des Hommes

For more information consult: Child Safety by Design 2022, IWF 2023, NCMEC 2016, NSPCC 2023, WeProtect - Global Threat Assessment 2021, WHO 2022