Emerging technologies: threats and opportunities for the protection of children against sexual exploitation and sexual abuse

Prepared by
Professor Victoria Baines,
independent consultant

Background paper for the Lanzarote Committee

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Emerging technologies: threats and opportunities for the protection of children against sexual exploitation and sexual abuse

Background paper for the Lanzarote Committee

Prepared by
Professor Victoria Baines,
independent consultant

Council of Europe

Cover design and layout: Documents and Publications Production Department (SPDP), Council of Europe.

Cover and layout: Documents and Publications Production Department (SPDP), Council of Europe.

This publication has not been copy-edited by the DPDP Editorial Unit to correct typographical and grammatical errors.

Printed at the Council of Europe

Secretariat of the Council of Europe
Convention on the Protection of Children against Sexual Exploitation and Sexual Abuse
(Lanzarote Convention)
F-67075 Strasbourg Cedex

lanzarote.committee@coe.int
www.coe.int/lanzarote

Contents

CONTEXT	**5**
ARTIFICIAL INTELLIGENCE (AI)	**7**
Generative AI	7
Machine Learning	8
IMMERSIVE ENVIRONMENTS – VIRTUAL REALITY (VR), AUGMENTED REALITY (AR), EXTENDED REALITY (XR)	**11**
The Internet of Things	12
Robotics	13
Live CSEA	13
CROSS-CUTTING THEMES AND QUESTIONS	**14**

Context

Child sexual exploitation and abuse (CSEA) is a constantly evolving phenomenon. Whether motivated by sexual interest or by the prospect of financial or other benefit, perpetrators make full use of the opportunities afforded to them. As Information Technology continues to develop and offer new tools to adults and children alike, there is a need for regular review of current trends in offending and victimisation, along with assessment of the potential for misuse presented by emerging technologies, to ensure that national and international responses to technology facilitated CSEA remain fit for purpose.

This paper invites the Lanzarote Committee to consider the impact of ongoing and near future technological developments on the nature and scale of CSEA. It does so by considering the threats and opportunities presented by the following technologies:

- Artificial Intelligence (both generative AI and Machine Learning)
- Extended Reality, comprising Virtual Reality, Augmented Reality, and other immersive digital environments
- Advances in robotics
- The Internet of Things
- Live video generation and broadcasting

Artificial Intelligence (AI)

Generative AI

Artificial Intelligence is the general term used to describe two types of advanced digital tools. The first is Generative AI, the best known of which is ChatGPT. Generative AI consists of Large Language Models (LLMs) that are trained to produce text and images by imitating patterns in pre-existing datasets. In text, this results in the creation of text that is natural and plausible, but not necessarily entirely factually accurate. The better and more comprehensive the training set, the more complete and persuasive the text generated – whether this is an answer to a request for factual information, or a task to produce generic content in a standard format, such as marketing material or an official letter. Generative AI can also be used to write and check code for computer programs. Increasingly, the text generation component is used to populate the conversations we have with bots, such as non-human customer service chat on websites and social media. Recently, we have seen the emergence of 'companion' apps that create an AI buddy or partner with whom the user can interact. Some of these are official extensions of a social media service, such as Snapchat's My AI.[1] Others are standalone apps that offer a customisable virtual girlfriend or boyfriend. Similar apps enable the user to recreate a digital version of a deceased loved one so that they can continue to chat and interact with them online.

While these apps may well reduce loneliness and provide comfort to some members of society, there is potential for their misuse to harm children. Virtual friend apps could engage children and young people in sexual conversation that is not age appropriate and that would be a criminal offence in some jurisdictions if committed by an adult human. Where apps or a potential victim's device have poor cybersecurity, those seeking to groom children could exploit these vulnerabilities to hijack an app and pose as a child's buddy or relative. This would not necessarily require a high level of technical skill but could, for instance, be achieved by persuading the child to share their account password. At a more general level, bots could be used to automate grooming approaches to a large number of children by one perpetrator.

In terms of image and video, Generative AI can be used to produce entirely new, synthetic content, or to alter existing material. Widely available tools such as DALL-E and Midjourney enable users to do both, and have safeguards that prevent the generation of sexual images. However, some 'face-swap' apps operated by less reputable companies allow users to upload a photo or video of a real person and then alter it in a way that transforms it into sexual content. States have already seen instances of young people using 'declothing apps' on their classmates and circulating the resulting material.[2] In the related field of cybercrime cloned audio recordings are already being used to dupe people into falling for scams.

Equally, there has been an observed increase in the circulation on the dark and open web of Child Sexual Abuse Material (CSAM) that are realistic 'deepfake' images and videos of real children, which arguably can be addressed under Article 20 Lanzarote Convention.[3] This raises a number of considerations, among them the possibility for perpetrators to use manipulated images and video of a victim as leverage to coerce them into producing 'real' Child Self-Generated Sexual Images and Videos (CSGSIV), potentially addressed by Articles 21 and 23 Lanzarote Convention. Criminal justice and legislative responses may also be challenged where there is no evidence of a sexual offence against the child subject, even if the emotional impact on victims is documented.[4]

1. https://help.snapchat.com/hc/en-gb/articles/13266788358932-What-is-My-AI-on-Snapchat-and-how-do-I-use-it
2. https://www.bbc.co.uk/news/world-europe-66877718; https://www.bbc.co.uk/news/technology-67521226
3. https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/
4. https://www.bbc.co.uk/news/articles/cpdlpj9zn9go

Food for thought

- How should victim impact be evaluated, relative to i) offline CSEA, ii) CSAM obtained through solicitation, iii) CSGSIV that is misused, and iv) AI generated CSAM in which there is no real child?
- Should the production and distribution of CSAM that puts a real child in a fictional sexual scenario be judged a less serious offence than CSAM that is a record of a sexual offence, and on what basis?
- What is the appropriate regulatory (and societal) response to peer production and distribution of deepfake CSAM?
- Can technical solutions be mandated that prevent manipulation of images/video involving children?

These questions extend also to the production and distribution of CSAM that is entirely synthetic in that it does not feature the likeness of a real child, particularly relevant to Article 20 of the Lanzarote Convention which criminalises the production, offering or making available, distributing or transmitting, procuring, possessing and knowingly obtaining access to CSAM. Article 20 paragraph 3 provides a possibility for States to exclude the production and possession of CSAM "consisting exclusively of simulated representations or realistic images of a non-existent child" from the scope of these offences. The ability to create life-like images of non-existent people has been available for some years.[5] Historically, legislative and child safeguarding responses have prioritised enforcement against CSAM on the basis that it is a record of a sexual offence against a real child. Where there is no real child, can the same basis be invoked, or does there need to be an alternative basis such as obscenity? Is a different approach required that focuses on the risk to children of someone who has an interest in synthetic CSAM, or indeed its potential to corrupt children? It is, for instance, reasonable to assume that individuals who seek to groom children for sexual activity may share synthetic CSAM with potential victims to normalise sexual activity between adults and children.

Food for thought

- If there is no 'real' child in the material, what is the appropriate regulatory response and sanction?
- Does the offender present less of a risk to children than one who possesses 'real child' CSAM?
- Can technical solutions be mandated that prevent creation of images/video/chatbots that appear to be children?

Machine Learning

While there is less hype currently about Machine Learning than about Generative AI, its impact on society is just as great. Machine Learning denotes the development of algorithms that learn from data, make generalisations, and then perform tasks without explicit instruction. Predictive analytics, computer vision, natural language processing, and speech recognition are all examples of this technology.

Machine Learning is used extensively to combat CSEA, especially on online platforms, where companies have invested in the development of tools that flag suspicious or inauthentic behaviour based on analysis of very large sets of account metadata; also in reviewing and classifying content with a view to identifying illegal and reportable material. Because the tools learn, their accuracy and detection capability improve all the time, resulting in an increase in the number of accounts flagged for suspicious activity and volume of CSAM reported to law enforcement. Equally, it may enable platforms to identify previously unnoticed indicators for CSEA for which there may be no obvious legal basis to report to the authorities.

At the same time, content review and classification have until now relied on platforms being able to view unencrypted content. The challenge to content review posed by end-to-end encryption is ongoing. Solutions that propose scanning for illegal material when it is on a user's device ('client-side scanning') have been welcomed by child protection specialists but rejected by privacy advocates who are concerned that the technology will be used to scan for content other than CSAM.

Meanwhile, machine learning itself has the potential to assist perpetrators in their offending. AI is increasingly used to automate the gathering of open-source information.[6] The tools used to 'scrape' and analyse large datasets can be used legitimately or by criminal actors to select their targets. Since perpetrators engaged in

5. https://thispersondoesnotexist.com/
6. https://www.oii.ox.ac.uk/news-events/open-source-intelligence-osint-and-ai-the-informational-pivot-of-intelligence-analysis/

solicitation and sexual extortion of children are already known to research and engage with potential victims on multiple platforms, it is reasonable to expect that they will seek to automate this process to make it more efficient and scalable to a larger number of victims.

Food for thought

- As techniques for flagging suspicious activity improve, what are the considerations for law enforcement workloads around the world? Is there sufficient legal basis for platforms to share data on CSEA that does not meet the current evidential thresholds?
- As end-to-end encryption becomes standard for messaging services, how will platforms and authorities identify and evidence illegal content and activity?
- Can a technical solution be mandated that prevents children's data from being subject to mining by CSEA perpetrators?

Immersive Environments – Virtual Reality (VR), Augmented Reality (AR), Extended Reality (XR)

The line between our physical and digital worlds is now very blurred. Augmented Reality (AR) enhances our offline experiences by overlaying data on them – whether we are using a mapping app or playing a mobile game like Pokémon Go. In Virtual Reality (VR), hardware including headsets and controls with haptic (touch) feedback immerse us in a 360° experience in which we feel emotionally and, increasingly, physically present. The use of external hardware in social computing is not in itself new: joysticks, steering wheels, and even connected weapons have been available for decades. Likewise, online gamers have felt like they were 'really there' in virtual worlds for just as long. However, more recent improvements in Internet speeds and reliability, coupled with considerable investment from games companies, have resulted in the development of ever richer and more compelling virtual worlds, some of which allow for the integration of 'real world' video calling and social media into computer generated environments.

Given the established market for CSAM, it will come as no surprise that there is demand for immersive sexual experiences involving child-like partners, which may raise questions under Article 20 Lanzarote Convention. Platforms such as Virt-A-Mate enable a user to create their ideal sexual partner, and law enforcement reports that childlike avatars are being created for that purpose. Equally, data shows that law enforcement has already found CSAM on VR headsets.[7]

As noted in a recent roundtable hosted by the WeProtect Global Alliance, grooming approaches in immersive environments may also be more persuasive and therefore more successful.[8] As observed in other online gaming environments, in-game currencies and virtual items are highly desirable to children, and risk being used to recruit or coerce a child into producing CSGSIV or performing a sexual activity, potentially giving rise to offences under Articles 20 to 23 Lanzarote Convention.

The immersive quality of virtual worlds can prove challenging to criminal justice actors when gathering and presenting evidence. Law enforcement authorities that are used to requesting files containing text conversations, static images, and recorded video, often across international borders, may find that these are insufficient to evidence an adverse online experience in 360° from the point of view of the alleged victim. The sheer complexity and scale of immersive environments makes it impractical for platforms to record the entirety of users' experiences in such a way that they can presented in a court of law. Accordingly, some platform providers have recommended to users that they record all their activities and experiences as they happen, in case they need to provide evidence to the authorities of an alleged crime.

What is clear is that CSEA in immersive environments is no longer a hypothetical threat. In 2018, a case in the US came to worldwide attention in which a Roblox avatar belonging to a 7-year-old girl was 'gang raped' by two adult male avatars[9]. Regardless of the legal considerations, the description by the child's mother of the incident's impact clearly demonstrated that this was not simply experienced as a 'virtual crime' but as a traumatic event in emotional terms. Earlier this year, police in the UK announced that they were investigating an offence of sexual abuse of a girl under the age of 16 in Virtual Reality.[10] In the context of heightened sense of emotional and physical presence in immersive environments, this prompts the question whether the current practical distinction between online and offline CSEA offences in many countries' legislation gives due consideration to the impact on victims of abusive experiences in immersive environments and provides sufficient coverage for offences that are experienced as real, even where there may be no physical injury.

7. https://www.bbc.co.uk/news/uk-64734308
8. https://www.weprotect.org/library/beyond-the-headset-extended-reality/
9. https://www.bbc.co.uk/news/technology-44697788
10. https://news.sky.com/story/police-investigate-sexual-abuse-of-young-girls-avatar-in-the-metaverse-prompting-nspcc-warning-13041003

Food for thought

- Do existing criminal procedures allow for electronic evidence gathering from immersive virtual worlds, and what are the practical considerations?
- Can a technical solution be mandated whereby adult users are prevented from creating child-like avatars? How might that be enforced?
- Should exposure of children to immersive sexual content be a more serious offence than causing them to view static images and videos?
- Can a technical solution be mandated whereby children are prevented from coming into contact with sexual content or activity? How might that be enforced?

The Internet of Things

Related to the notion of sexual abuse in immersive digital environments is the ongoing development of technologies that connect objects to the Internet and to each other, most often known as the Internet of Things (IoT).

In addition to smart home technology and use in critical infrastructure, connectivity is increasingly present in items such as adult sex toys, the retail market for which expanded considerably at the height of the COVID pandemic. Devices on the open market now include adult toys that can be controlled remotely by a partner using an app, and that can be connected to performances by adult entertainers on platforms such as OnlyFans. The capability already exists to integrate these devices into Virtual Reality (VR) sexual experiences. Pornography is already about much more than video: it, too, is becoming more about experience and direct participation.

Consequently, it is reasonable to expect that children will be exposed to device enabled pornography (Article 22 Lanzarote Convention). Given the emerging indicators above of child sexual abuse in VR, authorities should also prepare for the occurrence of child sexual abuse facilitated by connected sex toys. Where a child is subject to physical sensation delivered remotely by an adult, or where they are induced to deliver that sensation to themselves or other children. There is arguably a strong case for considering this to be child sexual abuse (Article 18 Lanzarote Convention) rather than offences of CSAM (Article 20 Lanzarote Convention), pornographic performance (Article 21 Lanzarote Convention), or corruption of children (Article 22 Lanzarote Convention), that attract lesser sentences in certain jurisdictions.

At a more general level, research has demonstrated that poor cybersecurity in connected devices such as children's toys and baby monitors could enable access to children in the home by unknown adults. Absent additional security measures, cameras, speakers and microphones embedded in devices can be remotely activated to converse with and capture footage of children[11]. In the last few years, legislation has been introduced in many – but not all – countries to improve IoT security by mandating unique passwords, security issue reporting, and security updates for consumer-connected devices. These regulatory solutions do not, however, exclude the possibility of remote access to children being achieved through social engineering – for instance, by persuading the child or a caregiver to disclose the password, which needs to be addressed through awareness raising and education.

Food for thought

- Is there sufficient legislative coverage to ensure safety and security in connected home devices and children's toys? Is this effective in deterring lax practices and preventing harm to children, including CSEA?
- Is there sufficient public awareness of the need to secure connected devices with good cyber hygiene? How might this be improved?
- Can a technical solution be mandated whereby children are prevented from using connected adult sex toys? How might this be enforced?

11. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3407264; https://www.pentestpartners.com/security-blog/making-childrens-toys-swear/

Robotics

Several companies have now developed companion robots for offline sexual activity. Owners can specify the precise look and dimensions of their partner, customise their expressions, voice, verbal and physical responses, and personalities. The most sophisticated sex robots combine animatronic movement with Machine Learning and Generative AI. Considerable time, effort, and resources are expended on making them sufficiently lifelike and anatomically correct that people are willing to spend thousands of dollars on them. For now, at least, their high price points mean they are beyond the reach of all but the wealthiest in society, and the leading companies involved in their development prohibit orders for child-like robots. Such advances in robotics need to be born in mind when considering how to regulate child-like robots, in a similar way to regulation of child-like sex dolls. [12]

Food for thought

- Would images/video of an adult engaged in sexual activity with a child-like doll constitute CSAM?
- Should the manufacture, import/export of child-like sex dolls be prohibited, and how might that be enforced?

Live CSEA

A dominant trend in recent years is the speed with which CSEA offences are now committed. The transition from 'studio' production of CSAM to solicitation of Child Self-generated Sexual Images and Videos (CSGSIV) now extends into live video distribution using widely available platforms and broadcast technologies.

Perpetrators of the commercial sexual exploitation of children have for some years live-streamed abuse. Due to improved connection quality and speed, increasingly it has become possible for paying viewers in one State to direct the abuse remotely in another country, with specific instructions and requests. Some State Parties have therefore begun to prosecute those who remotely direct the sexual abuse of a child either for involvement in or its commission or inciting the offence. In 2013, a 52-year-old man was sentenced in Sweden to eight years' imprisonment for eleven cases of instigating gross rape of a child in the Philippines. In the United Kingdom, the application of extra-territorial jurisdiction to CSEA legislation has enabled the authorities to prosecute offenders in the UK for remotely inciting the sexual abuse of children in the Philippines.

As the Committee is aware, the incorporation of cameras into mobile devices has increased the risk of children and young people being solicited, exploited or coerced into producing and sharing CSGSIV, some of which has been detected on adult platforms. This gives rise to the question of whether these providers are doing enough to keep sexual images and videos of children off their services. Innovative voluntary partnerships exist, such as StopNCII.org, which enables people (including children) who have shared intimate images of themselves to send digital fingerprints for them to leading global platforms, who can then prevent them being uploaded by someone else.[13]

When CSGSIV is published or otherwise shared without the consent of the subject – whether obtained by peers or as a result of solicitation/coercion by an adult – a race against time can ensue to ensure that the material is removed from circulation as quickly as possible to prevent re-victimisation. Where live video and sexual extortion are involved, responding to CSEA increasingly requires emergency intervention to safeguard a child in crisis. Criminal justice structures are often not designed to be that agile, especially when international data disclosure may be required. Emerging technologies may be able to provide solutions to enhance rapid removal and prevent resharing or republishing of such materials once reported by the victim.

Food for thought

- Is there sufficient regulatory coverage to combat the phenomenon of children being solicited, exploited or coerced into producing 'live' CSGSIV?
- Can a technical solution enhance prevention efforts in this regard?
- Are procedural powers sufficiently agile to allow law enforcement/criminal justice to respond immediately to ongoing, live abuse and/or coercion?

12. https://www.bbc.co.uk/news/uk-41203239; https://journals.sagepub.com/doi/10.1177/1023263X231176908; https://www.cps.gov.uk/legal-guidance/sex-dolls-childlike
13. https://stopncii.org/partners/industry-partners/

Cross-cutting themes and questions

Some consistent themes and questions emerge from the analysis above. The first concerns who should be responsible for preventing and combating technology facilitated CSEA. While a criminal justice response is clearly still required, online services have exclusive access to privileged information on the misuse of their platforms. As outlined above, their cooperation is often essential to removing a child's feeling of crisis when they are being sexually extorted or when CSGSIV has been published without their consent. **Corporate liability (LC Article 26)** is naturally a greater focus in current legislative efforts to prevent harm to children online, with the UK Online Safety Act and the EU's Digital Services Act among recent instruments aimed at setting baselines of compliance, achieving greater transparency, and ensuring that platforms' terms of service are effectively enforced.

Absent a global ban on declothing apps and tools for creating deepfakes, **education for children (LC Article 6)** on the appropriate use of technology, respect for other people's digital content, and the issue of consent is urgently needed. The phenomenon of non-consensual sharing of CSGSIV by peers identified as an emerging issue by the Committee in its 2019 *Opinion* has become more pressing in light of young people's ability to create sexualised images and video of their peers.[14]

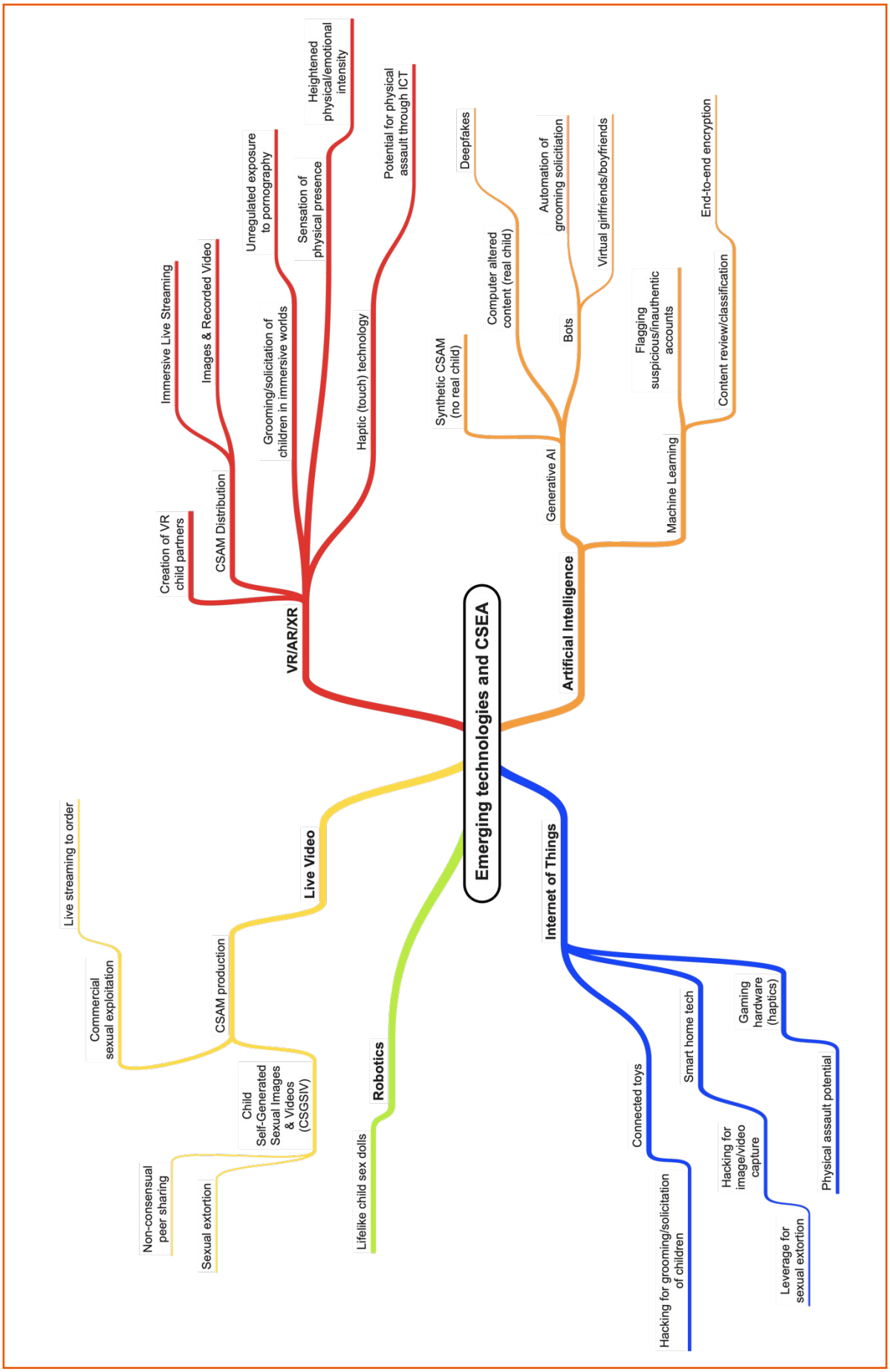
Equally, the emergence of immersive CSEA with heightened emotional and physical impact highlights the importance of ongoing psycho-social and therapeutic **assistance to victims (LC Article 14) preventive intervention programmes (LC Article 7) and intervention programmes (LC Chapter V) for perpetrators**. Increasing evidence of the emotional and physical impact of online CSEA may also give us cause to review the sentences for offences where there is no offline physical contact between perpetrator and victim, and to consider whether legal and conceptual distinctions between online and offline offences are still appropriate. This may be especially relevant where the provision of victim support or perpetrator intervention is dependent on the gravity of the offence as established in internal law.

Developments in ICT continue to provide perpetrators of CSEA with perceived 'safe spaces' in which to congregate, share knowledge, and commit offences. As the current controversy surrounding end-to-end encryption illustrates, there is a need for renewed focus on alternative approaches to detection, investigation, and evidence gathering. These include ensuring that law enforcement has the authority and capacity to conduct **covert operations (LC Article 30)** for all CSEA offences, and that there is adequate resourcing for **units specialised in CSEA investigation (Article 34)**.

We might also consider how these technologies can be used for good, to prevent and combat CSEA rather than simply contributing to its proliferation. AI-powered chatbots are already in use by several child helplines and Safer Internet Centres to provide live advice to children and young people. A number of initiatives use chatbots to intervene with perpetrators and those at risk of offending – for instance, when searching for CSAM on adult websites.[15] In other policy areas, Virtual Reality (VR) is used to generate empathy with, for example, the experience of dementia sufferers, or young refugees.[16] How might stakeholders in the fight against CSEA turn technological developments to their and children's advantage?

14. https://rm.coe.int/opinion-of-the-lanzarote-committee-on-child-sexually-suggestive-or-exp/168094e72c
15. https://www.iwf.org.uk/news-media/news/pioneering-chatbot-reduces-searches-for-illegal-sexual-images-of-children/
16. https://www.awalkthroughdementia.org/; https://www.with.in/watch/clouds-over-sidra/

Quick Reference Mind Map



PREMS 170424

ENG

www.coe.int

The Council of Europe is the continent's leading human rights organisation. It comprises 46 member states, including all members of the European Union. All Council of Europe member states have signed up to the European Convention on Human Rights, a treaty designed to protect human rights, democracy and the rule of law. The European Court of Human Rights oversees the implementation of the Convention in the member states.

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE