Jutta Oberlin,
Sarah von
Hoyningen-Huene

Navigating Digital Safety for Minors in Europe – A Legal Analysis and Comparison

Jutta Oberlin
Sarah von Hoyningen-Huene

Navigating Digital Safety for Minors in Europe – A Legal Analysis and Comparison

EIZ Publishing

Publisher: EIZ Publishing (eizpublishing.ch), Bellerivestrasse 49, 8008 Zürich,
eizpublishing@europa-institut.ch
Authors: Jutta Oberlin, Sarah von Hoyningen-Huene
ISBN:
978-3-03994-075-2 (Print – Softcover)
978-3-03994-076-9 (Print – Hardcover)
978-3-03994-077-6 (ePub)
DOI: https://doi.org/10.36862/70S3-AC9R
Version: 1.00 – 20260225

This work is available in print and various digital formats in **OpenAccess**. Additional information is available at: https://eizpublishing.ch/publikationen/navigating-digital-safety-for-minors-in-europe/.

Acknowledgements

It is with profound gratitude and sincere appreciation that we present this book, *Navigating Digital Safety for Minors in Europe: A Legal Analysis and Comparison*. The journey toward its completion has been shaped and supported by numerous individuals and institutions, without whom this work would not have been possible.

Our deepest thanks go first and foremost to the University of Zurich Library, whose generous sponsorship and financial support have made the publication of this book a reality. Their commitment to fostering academic inquiry and supporting emerging legal scholarship has been invaluable throughout this process. We are particularly grateful to Ms. Margit Dellatorre of the University of Zurich Library, who facilitated the funding process with remarkable simplicity and efficiency, and who supported us from the very beginning.

We, the authors, Sarah von Hoyningen-Huene and Jutta Oberlin, would also like to express our heartfelt gratitude to the Europa Institut at the University of Zurich for their exceptional guidance, unwavering support, and highly professional management of the publication process. In particular, we are indebted to Prof. Andreas Kellerhals and Dr. Tobias Baumgartner, whose expertise, encouragement, and thoughtful feedback have accompanied us from the earliest stages of this project to its final form. Their dedication to academic excellence has been both motivating and inspiring.

A special note of appreciation goes to our dear friend, Dr. Sandra Husi, whose professional insight and personal support have been indispensable. She has been a trusted advisor, a critical sounding board, and a compassionate listener throughout this intense and meaningful undertaking. Her willingness to stand by us with both expertise and friendship is something we value deeply.

The most personally, we – as two mothers – extend our deepest love and gratitude to our nine children — Maximilian, Carl, Eric, Sophia, Gregor, Charlotte, Arabella, Filippa, and Cecilia. Without you, this book would never have taken shape in the way it has. You are the reason our passion for this topic exists. It is through observing your interactions with the digital world that we have come to understand, more strongly than ever, the urgent need to protect children's rights, safety, and dignity in online environments. This book is dedicated to you and to all children, in the hope that our work may contribute, even in a small way, to building a safer and more empowering digital future.

We also wish to express our heartfelt gratitude to our beloved mothers, whose love and guidance have shaped our paths in profound ways. Though Sarah's mother is no longer here to witness her professional journey as it continues to unfold, she played a pivotal role in shaping who Sarah became, instilling in her, both as a mother and as a former professor at the University of Zurich, a passion for academic work and accompanying her university path with wisdom, encouragement, and unwavering support. Jutta's mother, through her extraordinary and loving care for Jutta's four children over many years, has made it possible for her to dedicate herself to demanding publication projects, and has taught her from early on that anything truly important must be pursued with love and devotion.

Lastly and above all, we are profoundly grateful for the friendship that has grown between us as co-authors, a bond that has accompanied us since our very first day of lectures at the University of Zurich. What began as a wonderful personal relationship has, over the years, naturally extended into our professional lives. Together, we now pursue, with passion and conviction, the topics closest to our hearts, especially the protection of children. We look forward to many more years of this remarkable friendship, which continues to grow and mature through nine children, numerous professional chapters, countless publications, podcasts, various speaking engagements and our non-profit-organization "Kinderrechte.Digital".

With heartfelt thanks,

Jutta Oberlin & Sarah von Hoyningen-Huene

Foreword

In recent decades, information and communication technologies have become deeply embedded in everyday life, fundamentally shaping how we connect, learn, and express ourselves. For children and young people growing up in this digital era, the internet and online platforms are not just tools but essential environments where much of their social and educational experiences unfold. The widespread daily use of digital technology among Europe's youth reflects this shift, with almost universal internet access and a strong engagement with social media. However, this digital immersion is a double-edged sword. While it offers unprecedented opportunities for creativity and connection, it also exposes young users to significant risks, from cyberbullying and privacy breaches to harmful content and exploitation. The vulnerability of children and adolescents online demands a careful and thorough examination of the legal and regulatory frameworks designed to protect them.

We, the authors of this publication, have been deeply engaged with this topic for many years, both professionally and personally as mothers. Over time, we have gathered extensive experience in this field through various roles, including employment with the public prosecutor's office, a major technology corporation, and active involvement in non-profit organizations. This diverse background enriches our perspective and informs the comprehensive analysis presented here.

We decided to write this book because we see an urgent need for action. The current situation, both from a regulatory and practical standpoint, is not sustainable and cannot adequately protect children and young people in the digital space. Without effective intervention, the risks will continue to grow, leaving minors increasingly vulnerable.

In this publication, we provide a comprehensive overview of how legal measures and policies across Europe and, by way of comparison, the United States, address the protection of minors in the digital environment. We highlight current challenges, identify gaps in existing safeguards, and evaluate the effectiveness of various approaches. By examining the roles of legislators, service providers, and non-profit organizations, we aim to foster a clearer understanding of how we can collectively ensure a safer online space for children and young people.

Only through coordinated, robust, and adaptive legal frameworks can we guarantee that the digital world remains a place of opportunity and growth, rather than harm, for the younger generation.

Table of Contents

I. Introduction 1

II. Chances in the Digital Age 5

1. Digital Skills 5
2. Connecting people through Social Media Platforms 5
3. Promoting Equal Opportunities for Children via Social Media 6
4. From Help to Empowerment 6

III. Internet Dangers and Risks – Threats in the Digital Age 9

1. Rise in Online Child Exploitation Fuels Industry Growth and Abuse 10
2. AI-CSAM 11
   A. CSAM 11
   B. A familiar phenomenon presented in a new guise 12
   C. Revictimization, Secondary Victimization and Retraumatization 12
3. Cybergrooming 13
4. Sexting/Sextortion 14
5. Cybermobbing 14
6. Hate Speech 16
7. Exposure to inappropriate and illegal content 17
8. Addiction 20
9. Extremism & Radicalization 21
10. Algorithms 22
11. Risky Online Behavior and Data Breaches 23
12. The threat to the 'digital' welfare of children 25

IV. International Guidelines and treaties on child online safety 29

1. OECD – Protecting Children Online 29
2. Council of Europe: Guidelines and advocating for respect, the protection and realization of the rights of the child in the digital environment 31
3. EU: A European strategy for a better internet for kids (BIK+) 35
   A. European Commission: A report from the consultation with children and young people 36
   B. Inclusion, Accessibility, and Digital Skills as Cornerstones of Child Empowerment 37
4. EU's Guide to age assurance 38
5. UN: General comment No. 25 (2021) on children's rights in relation to the digital environment 39
6. The 24/7 Network established under the Convention on Cybercrime (known as the Budapest Convention) 40
7. AEPD: Decalogue of principles for age verification and protection of minors from inappropriate content 41

V. Regulatory Approaches to Child Online Safety 45

1. Regulatory approaches to protect minors on the internet from a Youth Protection Perspective 45
   A. Switzerland 45

B. Germany 47
C. Regulatory Gaps 52
2. Regulatory approaches to protect minors on the internet from a Data Privacy Law Perspective 54
A. EU – General Data Protection Regulation (GDPR) 54
B. Switzerland 57
C. Excursus: USA 59
D. Regulatory Gaps 63
3. Regulatory approaches to protect minors on the internet from a Digital Law Perspective 64
A. EU 65
B. Germany; The Netzwerkdurchsetzungsgesetz (NetzDG) 71
C. Switzerland – Preliminary Regulation Framework for Communication Platforms 73
D. France – Kidfluencer Law & a new legislative proposal 76
E. The Netherlands 77
F. Spain 78
G. Ireland – Online Safety Act 78
H. Greece 79
I. Austria 81
J. UK 81
K. Regulatory Gaps 84
4. Regulatory approaches to protect minors on the internet from a Criminal Law Perspective 87
A. International Regulation 87
B. Regulating Crimes Against Children: A Comparative Overview of National Legal Approaches 91
C. Regulatory Gaps 97

VI. Institutional Responses to the Challenges of the Digital World for Children 99

1. The Watchdog Organizations 99
2. UNICEF: Children’s rights in the digital sphere – “promotion of media literacy instead of bans” 101
A. UNICEF: Policy guidance on AI for children 102
B. UNICEF: Protecting Children from Violence and Exploitation in relation to the digital environment – Policy Brief 103
3. CNIL’s 8 recommendations 104
4. German Safer Internet Centre (SID) / Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM) 106
5. Swiss Foundation for the Protection of Children (Kinderschutz Schweiz) 106

VII. Platforms and their actions/activities 109

1. Initiatives by the platforms 109
A. Instagram 109
B. TIKTOK 113
C. Youtube 116
D. Google 121
2. On the Sufficiency of Provider Measures in Child Online Safety 123
A. Instagram 123

B. TIK TOK 124
C. YouTube 126
D. Google 128
3. Enhancing Online Safety for Minors: A Provider Perspective 129
A. Platform Providers ensuring an age-appropriate digital environment 130
B. Expanding the definition of inappropriate content 130
C. Presence of police on the social media platforms 132
D. Promoting Media Literacy and Parental Awareness for Safer Social Media Use Among Minors 133
E. Digital Safety Risk Assessment Framework & Risk Classifications 134
F. Classifying Generative AI Models Capable of Producing Harmful Content as High-Risk Systems 139
G. Legal Requirement for Mandatory Social Media Training to Protect Minor Users 139
H. The EU law on chat control: The ultimate solution? 140
I. Social Media Ban for minors 145

VIII. Evaluation and Outlook 147

1. The Effectiveness of Existing Laws in Protecting Minors online 147
2. Outlook 149

IX. Conclusion 151

References 155

I. Introduction

Information and communication technologies (ICT)[1] have become an integral part of daily life in recent decades.[2] For children and young people who are growing up immersed in the internet and digital platforms, this environment offers numerous opportunities for connection, learning, and creative expression. Children and young people seem to be more connected than ever before. In 2024, the vast majority of young people across the European Union used the internet daily, with national rates ranging from 93% to 100%. On average, 97% of young individuals in the EU reported daily internet use, highlighting the near-universal integration of digital technology into the lives of Europe's youth. In 2024, social media platforms were a central part of daily life for young people in the European Union, with 88% of individuals aged 16 to 29 actively using them. This stands in stark contrast to the overall EU population, where only 65% engaged with social networks, underscoring the significant generational divide in digital communication habits and online social interaction.[3] They encounter the digital world at an increasingly young age, whether through social media, online games, or educational platforms. This greater connectivity opens a multitude of new opportunities, but it also presents well-known[4] and significant risks.[5]

Recent studies highlight both the benefits and the dangers of this digital immersion. A survey by UNICEF Switzerland and Liechtenstein[6] revealed that nearly one in three children felt unsafe in the digital world, underscoring the vulnerability of young users. In Switzerland, one in five children viewed influencers as role models, and in Liechtenstein, this figure was one in four. Concerns about privacy were also evident, with one in ten children reporting that their privacy had been violated by the sharing of photos or videos without consent. Similarly, the MIKE study[7] in Switzerland found that while children

[1] Information and communication technology, abbreviated as ICT, covers all technical means used to handle information and aid communication. This includes both computer and network hardware, as well as their software, glossary accessed under: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Information_and_communication_technology_(ICT).

[2] Eurostat, n. d.

[3] Eurostat, n. d.

[4] Steinbacher et al., 2024.

[5] UNICEF, 2021.

[6] UNICEF, 2021.

[7] Suter et al., 2023

aged 6 to 13 engaged with phones mostly for entertainment, the prevalence of social media use increased with age, exposing young users to platforms like WhatsApp, TikTok, and Snapchat. Despite the positive experiences many children have online, about half reported encountering frightening or unsuitable content. The JAMES study,[8] which surveyed young people aged 12 to 19, further confirmed that the digital space can be a source of harm, with a significant portion of young people experiencing cyberbullying, unwanted sexual approaches, or abusive behavior online.

While the digital transformation offers many opportunities, it also presents considerable security threats.[9] Children and adolescents are particularly vulnerable in the online space, facing dangers such as cyberbullying, data misuse, and exposure to inappropriate content such as deepfake child sexual abuse material and DeepNudes.[10] The digital world can be a place of growth and knowledge for young people, but it can also be a place of harm if the necessary protection is absent. A significant portion of child sexual abuse material (CSAM) globally is hosted in the European Union.[11] The EU has been a growing hub for CSAM since 2016, with nearly 90% of child abuse URLs hosted in Europe by 2019.[12] The Netherlands, in particular, has seen a sharp rise in the amount of CSAM hosted there, with Dutch authorities removing tens of thousands of webpages annually. There have been efforts to close legal loopholes[13] in the Netherlands to force hosting companies to act against illegal content more effectively.[14] This publication explores how the protection of minors in the digital world can be ensured through legal frameworks. It provides a comprehensive overview of existing laws and regulations concerning online safety for minors, alongside a comparative legal analysis that examines how different countries address these issues and how effective these regulations are. Special attention is given to the challenges and risks posed by insufficient regulation, and it highlights why it is crucial to implement legal measures that safeguard the safety and privacy of minors in the digital space. Only through effective, internationally coordinated regulation can we prevent the digital world from becoming a place where young people are deprived of their protection and rights.[15]

[8] Külling-Knecht et al., 2024

[9] Oberlin et al., 2024.

[10] Steinbacher et al., 2024.

[11] Internet Watch Foundation, 2022; Internet Watch Foundation, 2023.

[12] European Parliament, 2020.

[13] Government of the Netherlands, 2020, June 9.

[14] Government of the Netherlands, 2024, June 5; European Parliament, 2020.

[15] Oberlin et al., 2024.

This publication offers an overview of key measures and initiatives across Europe with a brief excursus to the United States that operate at multiple levels (legislation, policy frameworks, industry practices, and the role of non-profit organizations) to strengthen the protection of children in the digital environment. Both the legal provisions and the political measures that regulate the handling of specific dangers on the Internet are analyzed. It also examines the responsibility of internet providers; the measures they have taken so far, and how non-profit organizations contribute to education and prevention. This detailed stocktaking is followed by an in-depth evaluation focusing on the question of whether existing laws, measures and guidelines are sufficient to effectively protect children in the digital space. The gaps and challenges that still exist and what additional steps are needed to strengthen the protection of children and young people online will be critically scrutinized.

II. Chances in the Digital Age

1. Digital Skills

In today's increasingly digital world, ICT skills are essential for young people, not only to access and benefit from the vast range of internet-based opportunities, but also to participate actively and responsibly in the digital society. By the time they complete compulsory education, most young people in the EU have already used computers and the internet for various purposes. However, beyond mere access, it is crucial that they develop strong digital competencies to become empowered, informed, and responsible users.[16]

Digital skills cover five key areas: information and data literacy, communication and collaboration, content creation, safety, and problem-solving. To be considered as having at least basic digital skills, a person must have carried out at least one activity in each of these areas in the three months prior to the following survey. These might include actions such as engaging in social networks, taking an online course, writing code, or participating in online civic initiatives. According to 2023 data, the level of digital skills among young people aged 16 to 29 varies widely across Europe. While countries like Finland (94%), Malta (92%), and Czechia and the Netherlands (both 90%) report very high shares of youth with at least basic digital skills, the figures are much lower in others—such as Romania (46%) and Bulgaria (52%). The EU average stands at 71%, revealing a significant digital divide between member states. Equipping young people with digital skills is not just a matter of personal development; it's a foundation for future employability, democratic participation, and social inclusion. In a rapidly evolving digital economy, these competencies are becoming as fundamental as traditional literacy and numeracy, shaping the way the next generation works, learns, communicates, and engages with the world.[17]

2. Connecting people through Social Media Platforms

As of February 2025, approximately 5.56 billion people around the world were using the internet, representing 67.9% of the global population[18], serving as a

[16] Eurostat, n. d.

[17] Eurostat, n. d.

[18] Statista, 2023.

powerful tool for communication, self-expression, and global engagement. Social media, which 5.24 billion, or 63.9 percent of the world's population uses[19] has become a powerful tool for young people to connect, share, and build communities across borders. Platforms like Instagram, TikTok and Snapchat allow young users to stay in touch with friends, meet new people who share similar interests, and engage in global conversations. Whether through group chats, comment sections, or collaborative content creation, social media provides spaces where young people can express themselves, find support, peek into different worlds, and feel a sense of belonging. In a world that is increasingly digital, these platforms help bridge distances and create networks that go beyond geographic and cultural boundaries. They also offer opportunities for activism, learning, and collective action, empowering youth to have a voice in shaping their world. When used mindfully, Social Media can foster meaningful connections and strengthen young people's sense of identity and community. Therefore, Social Media can empower individuals to share their voices, exchange ideas, and build communities across borders in real time.[20]

3. Promoting Equal Opportunities for Children via Social Media

Digital inclusion during childhood is more than just access to the internet. It's about giving all children the chance to use digital tools in a way that supports their learning, creativity and voice. When done right and with professional guidance, technology can reduce educational inequalities and offer new ways for children to grow, express themselves and participate. For many children, especially those from underserved communities, digital tools can be a form of empowerment. Whether it's learning apps, creative platforms, or access to safe online spaces, technology can help level the playing field. But real equity in education requires more than devices: Studies have shown that internet access alone is not sufficient; guided educational content is essential to translate digital access into meaningful academic benefits.[21]

4. From Help to Empowerment

The digital space offers young people new ways to access support and build resilience, especially in times of personal struggle. Anonymous online services

[19] Statista, 2023.

[20] United Nations, n. d.

[21] Malamud et al., 2019.

and peer-support communities provide safe, low-threshold spaces to talk about issues like mental health, bullying, or discrimination without fear of stigma. A systematic review found that online peer-to-peer support groups significantly benefit youth facing mental health issues. Participants reported enhanced coping strategies, social connectedness, reduced isolation, and lower depressive symptoms, particularly when groups were moderated by peers or professionals. Studies have shown that two of the randomized controlled trials were associated with a significant positive outcome in comparison to the control group at post-intervention. In the remaining four studies, peer-to-peer support was not found to be effective.[22] At the same time, social media, online resources, and activist platforms empower teens to educate themselves and speak up for their rights, whether around LGBTQ+ identities[23], racial and gender equality (e.g. #BlackLivesMatters)[24], climate justice (e.g. #FridaysForFuture)[25], body positivity[26], or mental wellbeing.[27] Through hashtags, storytelling, and creative content, young people raise awareness, mobilize communities, and drive social change on issues that matter to them.

[22] Ali et al., 2015.

[23] Berger et al., 2022.

[24] Jackson et al., 2020; Burdick & Sandlin, 2017.

[25] Tait, 2019.

[26] Bennett, 2019.

[27] Martinez Sainz & Hanna, 2023.

III. Internet Dangers and Risks – Threats in the Digital Age

To evaluate the effectiveness of current regulations, it is essential to first understand the range of risks that minors encounter in the digital environment. While the online world offers children and adolescents vast opportunities for learning, communication, and creative expression, it also presents a multifaceted set of risks that can threaten their safety, well-being, and healthy development. These risks are grouped into five key dimensions by the authors: content-related, contact-related, conduct-related, contractual, and systemic risks.

Content risks refer to the exposure of children to harmful or age-inappropriate material online, including violent imagery, pornography, self-harm content, or extremist ideologies. Such exposure can be distressing and may lead to confusion, fear, or even trauma, particularly for younger or more vulnerable children.[28]

Contact risks arise when children interact with others in online spaces, such as chat rooms, social media platforms, or gaming communities, and become victims of harmful behavior. This includes cyberbullying, harassment, threats, or manipulation. A particularly dangerous form of contact risk is online grooming, where adults with sexual or abusive intent establish emotional connections with minors to exploit or abuse them. These encounters can leave deep psychological scars and often go undetected until significant harm has occurred.[29]

Conduct risks focus on children's own behavior in digital environments. This includes instances where minors engage in problematic activities themselves, such as spreading hate, participating in online shaming or bullying, or accessing illegal services like gambling or age-restricted content. These behaviors not only affect others but also put the children engaging in them at risk; socially, emotionally, and legally.[30]

Contract risks relate to the often-invisible legal and commercial mechanisms that children become entangled in online. Many digital platforms operate under complex terms of service, which young users routinely accept without understanding. This can lead to unintended consent to data collection, in-app

[28] Deutsches Kinderhilfswerk, n. d.

[29] Deutsches Kinderhilfswerk, n. d.

[30] Deutsches Kinderhilfswerk, n. d.

purchases, or binding agreements that are neither age-appropriate nor transparent.[31]

Finally, systemic risks represent a broader, more structural category. These stem from the way digital platforms and technologies are designed and governed. For example, opaque algorithms may push harmful content to children, engagement-driven design can foster screen addiction, and monetization models often prioritize profit over child safety. Systemic risks are especially concerning because they operate at scale, impacting not just individuals but entire generations of children by shaping their digital experiences in ways that are often invisible, unregulated, and difficult to counter.

1. Rise in Online Child Exploitation Fuels Industry Growth and Abuse

The prevalence of online child abuse has increased dramatically during the pandemic,[32] as both children and child sex offenders spend more time online. Digital technologies, while enabling rapid communication and access, have unfortunately also facilitated the spread of CSAM.[33] Offenders now have easier access to children through webcams, connected devices, and chat rooms, especially in social media and video games.[34] Anonymity provided by technologies like cloud computing, the dark web[35], end-to-end encryption[36], and livestreaming makes it more difficult to track these criminal activities. Online child abuse has worsened since the Covid-19 pandemic.[37] In 2024 over 300 million children were victims of online sexual exploitation[38] and abuse, and the United Nations once estimated that over 750,000 child sex offenders were searching for CSAM online at any given moment[39]. Today, this number exceeds 1 million.[40] The Internet Watch Foundation (IWF) reports a significant rise in the number of CSAM reports, from about 1 million in 2010 to 17 million in 2019, which includes millions of images and videos.[41] Most reports come from Elec-

[31] Deutsches Kinderhilfswerk, n. d.

[32] Ramaswamy & Seshadri, 2020.

[33] UNICEF Office of Research – Innocenti, n. d.

[34] UNODC, 2014

[35] Adel & Norouzifard, 2024; Saleem et al., 2022.

[36] See Kriechbaumer & Nath, 2015, and also related to this: Sayyed & Paul, 2025.

[37] Negreiro Achiaga, 2020.

[38] Childlight Global Child Safety Institute, 2024.

[39] IJM, 2020.

[40] Internet Watch Foundation, 2020, May 1.

[41] Internet Watch Foundation, 2021.

tronic Service Providers (ESPs) or from the users itself.[42] In 2023, the CyberTipline[43] received over 35.9 million reports related to suspected CSAM incidents.[44]

Technological advancements have made it easier for offenders to access, share, and distribute CSAM. While in the early days of the internet, slower connections limited the spread of CSAM, the proliferation of high-speed internet and mobile devices has facilitated a significant increase in online child exploitation. Today, offenders can quickly find and share CSAM on mainstream platforms, even using search engines. Research shows that such material can be found in as few as three clicks on major search engines. Additionally, the anonymity of the internet has allowed offenders to form online communities, making it easier to share their crimes and find new victims. The scale of CSAM crimes has grown, with younger children becoming the targets of abuse. Reports show that nearly half of CSAM images involve children under 10, and over 40% fall into the most severe categories of abuse.[45]

2. AI-CSAM

A. CSAM

The abbreviation CSAM (Child Sexual Abuse Material) refers to visual representations of the sexual abuse of children. This includes all images involving minors in sexual activities, such as photos, videos, and computer-generated images. It is important to recognize that not all images of naked children depict sexual abuse. Additionally, there are other types of obscene or exploitative images involving children, such as drawn depictions (cartoons) of child sexual abuse or attempts to portray child sexual exploitation humorously, which may also be illegal under certain laws.

[42] Negreiro Achiaga, 2020.

[43] NCMEC's CyberTipline is a reporting platform for the public and electronic service providers (ESPs) to report suspected child sexual exploitation, with child sexual abuse material (CSAM) being the most commonly reported issue, more under: https://www.missingkids.org/gethelpnow/cybertipline.

[44] National Center for Missing & Exploited Children, 2024.

[45] Negreiro Achiaga, 2020.

B. A familiar phenomenon presented in a new guise.

One particularly alarming phenomenon that has become increasingly important in recent years is the use of modern artificial intelligence to produce child pornography. Access to depictions of child sexual abuse is often a precursor to actual abuse, whether the images depict real exploitation or are deceptively realistic.[46] This technology enables the digital manipulation of images and videos of abused children, allowing for the creation of new, often more explicit representations.[47]

C. Revictimization, Secondary Victimization and Retraumatization

The revictimization of children who have experienced sexual abuse is one of the most serious and complex challenges in the field of child and youth protection research.[48] As a result, an additional layer of abuse and trauma is inflicted on the victim. The impact of such digital revictimization on the mental health of victims is profound[49] – especially the fear of re-surfacing.[50] Pre-existing psychological distress from the physical sexual abuse, such as post-traumatic stress disorder (PTSD), depression and anxiety disorders[51], is further exacerbated by the digital repetition and amplification of the abuse. In particular, the idea that the child is repeatedly and uncontrollably further exposed through the internet and in the media leads to a loss of control and may create a feeling of permanent threat.

Children whose images are manipulated and disseminated can be additionally burdened by the constant fear that further content will be created and disseminated without their knowledge or consent. This leads to a long-term, chronic feeling of shame, guilt and fear. These forms of digital revictimization not only hinder the healing process but may also reinforce children's sense of entrapment in their victim role, significantly impairing their development.[52]

[46] Insoll et al., 2021.

[47] For an extensive analysis on this subject, see Oberlin & von Hoyningen-Huene, 2025, August.

[48] Papalia et al., 2021.

[49] Papalia et al., 2021.

[50] Joleby et al., 2020.

[51] Maniglio, 2013.

[52] Parti & Szabó, 2024.

3. Cybergrooming

The term “Cybergrooming” combines “cyber” and “to groom”. It represents the online counterpart to traditional grooming, which is a process where an adult builds trust with a minor and their social environment through manipulation and deception to prepare them for sexual abuse.[53] Grooming is often gradual and can be difficult for the child to distinguish from normal adult interaction.[54] Cybergrooming differs primarily because the internet allows offenders greater, more immediate, and constant access to victims[55], reducing their reliance on involving the victim’s social circle.[56] This shift favors extrafamilial abuse, though the offender may still consider the victim’s social environment. The rise of interactive social media has transformed grooming strategies into the digital realm.[57] A challenge in understanding Cybergrooming is the lack of a uniform definition.[58] The anonymity[59] and multiplicity of identities possible online contribute to an “online disinhibition effect”,[60] changing the scope and nature of grooming. For some offenders, the goal is not a physical meeting but sexual victimization through online contact alone.[61]

The goal may be to obtain sexual images or videos, engage in explicit conversations, or even arrange in-person meetings.[62] Cybergrooming is complex, involving various tactics such as sexualized verbal and written messages, exchanging pornographic material, flattery, bribery, or urging sexual acts.[63] Even in the absence of physical contact, Cybergrooming constitutes a serious form of sexual violence and is criminalized in certain jurisdictions. The perpetrator often uses a fake profile to inspire trust to obtain intimate images that can later be used to blackmail the minor.[64]

[53] Wachs, 2014; Alexiou, 2018.
[54] Stelzmann et al., 2020.
[55] Oberlin et al., 2024.
[56] Müller-Johnson, 2018.
[57] Wachs, 2014.
[58] Wachs, 2017; Alexiou, 2018; Rüdiger, 2020
[59] Oberlin et al., 2024.
[60] Alexiou, 2018.
[61] Bergmann & Baier, 2016; Rüdiger, 2020; Kattenberg, 2024.
[62] Oberlin et al., 2024.
[63] Wachs, 2017.
[64] Oberlin et al., 2024.

4. Sexting/Sextortion

The term Sexting refers to the consensual exchange of intimate, often pornographic, images in the digital space. This exchange typically occurs between two individuals who willingly and mutually share the content. It usually takes place through messaging services, social networks, or other digital communication channels and can include photos, videos, or messages containing personal, sexual content. However, Sexting carries significant risks, particularly when the involved parties share their own intimate images or videos with others. The biggest risk of Sexting is the unauthorized forwarding of these private, often very personal materials to third parties.[65] This can have severe consequences, such as extortion or the public release of the images, leading to enormous emotional and psychological distress for the individuals involved[66].

In the worst-case scenario, the person may be pressured to share more intimate content or other personal information to prevent the original images from being made public.This phenomenon is known as Sextortion, a term that combines the English words “sex” and “extortion.” Sextortion is a form of blackmail where the perpetrators use intimate content, which has been shared without consent, to extort the victim – sometimes to obtain more intimate materials or to force other actions such as payment or compliance.[67]

5. Cybermobbing

Cyberbullying or Cybermobbing is an increasingly widespread phenomenon that manifests itself on digital platforms and involves targeted bullying that often includes threatening and coercive elements. It goes far beyond simple insults or negative comments and can take various forms. Perpetrators not only use language, for example by writing hurtful messages or comments, but also use visual materials that they have either created themselves or taken from digital sources. Images or photomontages are often created with the aim of emotionally hurting and humiliating the victim. According to the JIM Study 2018[68], one-third of the surveyed adolescents have already encountered cyberbullying.[69] The effects of cyberbullying can be devastating, and in many tragic cases, have led to serious psychological and physical consequences – in

[65] Döring, 2005.

[66] Gassó et al, 2020.

[67] For more information; Schweizerische Kriminalprävention, n. d., and Oberlin et al., 2024.

[68] Feierabend et al., 2023.

[69] BSI, 2023.

extreme cases even to the suicide of the victim.[70] The rapid spread of such attacks in the digital world makes the situation even more dangerous. What was previously only accessible to a few people can now be made accessible to an unlimited number of people within seconds.[71] The internet provides a fertile ground for cyberbullying, largely because it removes the direct, human feedback that would normally act as an emotional and moral brake in face-to-face interactions. In traditional bullying (for instance, on the school playground) aggressors are confronted with the victim's immediate reaction: tears, fear, or withdrawal. These visible signs of suffering often trigger a natural sense of empathy or shame that can bring the harassment to an end. Online, however, this social corrective mechanism is absent. Perpetrators do not witness the victim's pain; the digital interface creates emotional distance and anonymity, which lower empathy and increase the likelihood of escalation. The screen acts as a psychological barrier, making it easier to dehumanize the target and perceive the act as harmless or even justified. Consequently, what might once have stopped when someone "was on the ground or started crying" now continues unchecked — amplified by the speed, reach, and permanence of online communication. This detachment from real-world consequences is one of the key factors that make cyberbullying more pervasive and psychologically damaging than traditional forms of bullying.

One particularly worrying aspect of cyberbullying is the use of artificial intelligence to create new, manipulated content from existing images. These so-called deepfakes; computer-generated, often deceptively real images or videos,[72] have the potential to portray people in extremely damaging or embarrassing scenarios. What is particularly alarming is that these manipulations can now be carried out with relatively simple means, without the need for in-depth technical knowledge or large financial resources.[73] Deepfakes can be used to create fake images or videos that have a sexual component,[74] for example, and thus devastatingly destroy the reputation of the people concerned. The deception is now so precise[75] that it is almost impossible for the victim to defend themselves against such false representations. The increasing prevalence of cyberbullying and the opportunities that modern technology offers for such attacks pose a serious threat to the mental health and well-being of

[70] Kaye, 2010, October 7.
[71] Beran & Li, 2005.
[72] Oberlin & von Hoyningen-Huene, 2025.
[73] Oberlin et al., 2024.
[74] da Silva Gioia, 2023, December 28.
[75] Oberlin & von Hoyningen-Huene, 2025.

those affected. It is vital that both society and legislators are made more aware of this problem to take effective action to protect victims and hold perpetrators to account.[76]

6. Hate Speech

Hate speech is another worrying phenomenon that is increasingly common on social media and gaming platforms. This form of verbal and non-verbal violence[77] is particularly frequently directed by young people against their peers.[78] Verbal and non-verbal contempt is communicated with the aim of humiliating and emotionally hurting the victim.[79] This type of communication manifests itself in violent abuse, slander, and insults aimed at devaluing the victim and damaging their psychological integrity.[80] The hate is often directed at people because of their origin, appearance, sexual orientation, or other personal characteristics, which become the target of marginalization[81].

What makes hate speech particularly dangerous is the fact that it is often accompanied by an apparent anonymity and distance in the digital world. This anonymity tempts people to say or write things that they would never dare to say in a direct, face-to-face conversation.[82] In the safe environment of the internet, many people find it easier to shed their inhibitions and make hostile, hurtful statements without experiencing the immediate consequences and direct suffering of the victim.[83]

Although there are only a few studies to date that explicitly address the damage caused by hate speech, it is generally recognized that the impact on the young people affected can be far-reaching.[84] The psychological consequences are often profound and can lead to lasting traumatization. Young people who are victims of hate speech are at an increased risk of psychological problems,[85]

[76] Oberlin et al., 2024.

[77] Oberlin et al., 2024.

[78] UK Safer Internet Centre, 2016.

[79] Oberlin et al., 2024.

[80] May, 2018.

[81] Oberlin et al., 2024.

[82] Weber, 2012.

[83] Oberlin et al., 2024.

[84] Wachs et al., 2023.

[85] This can manifest itself in mental illnesses such as anxiety or eating disorders, depression, self-harming behaviour, suicidal tendencies or low self-esteem. These young people may engage in increased risk behaviour in relation to drugs, risky sexual behaviour, poor academic performance and frequent absences from school or training, in: Oberlin et al., 2024.

as persistent verbal attacks can have a massive impact on self-esteem and trust in interpersonal relationships.[86]

The damage caused by hate speech not only affects the immediate well-being of the young people concerned but can also have a long-term impact on their mental health and social integration. Particularly during the formative developmental phase of adolescence, when self-image and identity are often still being formed, the psychological scars caused by repeated verbal attacks can have a particularly serious impact on a person's entire life.[87]

7. Exposure to inappropriate and illegal content

Studies suggest a significant link between the use of social media and the development of eating disorders.[88] Particularly concerning is the constant exposure to idealized body images on platforms like Instagram, which can negatively impact adolescents' self-image.[89] The pressure for perfection and the ubiquitous presence of retouched, often unrealistic portrayals of beauty amplify feelings of insecurity and comparison, significantly increasing the risk of eating disorders.

Another equally worrying aspect of social media is the spread of content related to suicide and self-harm. Platforms like Instagram are unfortunately also a space where topics like suicide and self-destructive behavior are prevalent, often fostering imitation effects. This phenomenon, known as the "Werther Effect," describes how the media portrayal of suicides or self-harming behaviors can trigger imitation of these actions in vulnerable individuals.[90] In 2022, Research by the Center for Countering Digital Hate (CCDH) reveals that TikTok's recommendation algorithm promotes harmful content,[91] such as self-harm, eating disorders, and pro-suicide videos, to teenagers within minutes of them expressing interest in these topics. The study involved creating accounts for users aged 13 in the US, UK, Canada, and Australia, including both "standard" and "vulnerable" accounts. The vulnerable accounts, which contained the term "loseweight" in their usernames, were specifically designed to reflect patterns seen in users seeking eating disorder-related content. After briefly engaging with content on body image and mental health, the algorithm

[86] Benner et al., 2018; Oberlin et al., 2024.
[87] Wachs et al., 2022; Madriaza et al., 2025.
[88] Fatt & Fardouly, 2023.
[89] Schweitzer, 2024; Oberlin et al., 2024.
[90] Arendt et al., 2019; Oberlin et al., 2024.
[91] Center for Countering Digital Hate, 2022.

quickly recommended suicide-related videos within three minutes and eating disorder content within eight minutes. Vulnerable accounts received significantly more harmful recommendations, with self-harm and suicide-related content shown 12 times more often than on standard accounts. These findings highlight the alarming speed and scale at which TikTok exposes young users to harmful content, raising concerns about the impact on their physical and mental health.[92] A study conducted by Amnesty International, in collaboration with the Algorithmic Transparency Institute, simulated the digital experience of children and young people struggling with mental health issues like anxiety and depression. They set up an automated TikTok account, programmed to interact with the app's "For You" feed. Over time, as the account scrolled through content, increasingly harmful videos were recommended. These included clips featuring children crying, expressing depressive thoughts, or describing self-harm and suicidal tendencies. One video featured an AI-generated voice reciting disturbing thoughts about overdosing and escaping pain. The experiment highlighted the alarming nature of TikTok's recommendation algorithm, which continuously exposes vulnerable users to distressing and potentially harmful content, raising concerns about its impact on mental health.[93] In 2024, self-harm on Instagram remained a significant issue. A study by Digitalt Ansvar, a Danish organization, found that while Instagram's AI could identify 38% of self-harm images and 88% of the most severe cases, the platform still failed to implement this technology effectively. This raised concerns that Instagram was not complying with EU regulations, such as the Digital Services Act, which mandates platforms to address risks to mental health. A survey by stem4 revealed that nearly half of children and teens had experienced withdrawal, excessive exercise, social isolation, or self-harm due to online bullying. Despite removing over 12 million pieces of self-harm and suicide-related content in 2024, Instagram's algorithm was found to unintentionally connect users to self-harm groups. Experts, including psychologist Lotte Rubæk, criticized Meta for failing to remove explicit content, contributing to the harm and suicide risk among vulnerable youth.[94] Such forbidden content also increasingly poses a significant problem on social media platforms. This includes a wide range of harmful and illegal materials, such as pornography, child pornography, AI-generated pornographic content, and violent depictions (Violent or graphic content refers to material that shows violent acts like murder, torture, and rape, as well as content that encourages or incites violence[95]).

[92] Hern & Milmo, 2022.

[93] Amnesty International, 2023.

[94] Bryant, 2024.

[95] World Economic Forum, 2023.

These types of content not only pose a direct threat to the mental health and well-being of users but also present a danger to society as a whole. For example, studies have found that between 19% of boys aged 10-12 and up to 84% of boys and 60% of girls aged 16-17 have been unintentionally exposed to pornography.[96]

While many social media platforms have clear guidelines against the distribution of problematic content, the problem persists.[97] In some cases, such content finds its way onto platforms through circumvented upload mechanisms or the use of encrypted networks. This issue is compounded by the anonymity of the internet and the difficulty in automatically detecting such content. Despite efforts by platforms like Facebook, Instagram, and TikTok to remove such material, new ways of hiding and disseminating it continue to emerge.[98] Particularly troubling is the distribution of child pornography, which carries severe legal, ethical, and moral implications.

A particularly new and concerning development is the use of artificial AI to generate pornographic images or videos. This presents an additional risk to children by enabling the creation of "deep fakes," which are manipulated images, audio, or videos that appear real. Alarmingly, publicly available images, including those of children, can be used to create deep fake pornography, known as "AI CSAM." AI image generators, originally designed for art and design, are increasingly used in pornography, creating unforeseen legal challenges. Older versions of AI software like Stable Diffusion, which had security flaws, were often used to create abusive content due to their training on existing abuse material. While newer updates prevent such content, older versions are still available on the dark web. The rapid advancement of AI means that generated images and videos are becoming increasingly realistic. Online tools allow users to alter the apparent age, body shape, and activities of depicted individuals. Open-source models for creating pornography are particularly concerning, as they allow for easy customization. Disturbingly, AI-generated abusive content is no longer confined to the dark web but is also appearing on the clear web, signaling its widespread availability.[99]

Violence on social media is another significant issue, particularly when it is depicted in explicit forms. This can include videos of physical violence in various

[96] Mitchell at al., 2003.

[97] Gorwa & Emmert, 2021.

[98] Ngo et al., 2024.

[99] Oberlin & von Hoyningen-Huene, 2025.

variations[100], terrorist attacks, criminally relevant calls for violence[101] or other extreme acts of aggression. In some cases, social media platforms promote such content by recommending videos to users who have engaged with similar material. Others are unwillingly confronted with such material.[102] These depictions can be disturbing and traumatizing, particularly for younger or vulnerable users. They can also contribute to the normalization of violence, being less sensitive to the suffering of others and, in extreme cases, even incite real-life violent acts.[103]

8. Addiction

Social media platforms carry a high risk of fostering addictive behavior among users.[104] Content tailored to the individual generates positive affirmations for the user in the form of likes, which can subsequently lead to a release of dopamine, which in turn favors the potential for addiction.[105]

This mechanism is grounded in the principles of operant conditioning, as formulated by B.F. Skinner, where variable rewards, such as irregular likes and comments, reinforce user behavior and increase engagement time.[106] Studies in neuroscience have confirmed that receiving social rewards, like a "like," activates the same neural circuits involved in drug addiction, particularly the mesolimbic dopamine system.[107] Moreover, neuroimaging studies have shown increased activity in the nucleus accumbens, a core area of the brain's reward system, when individuals receive positive social feedback on social media. This activation mirrors patterns found in other behavioral addictions, such as gambling.[108] The algorithmic structure of platforms like Instagram and TikTok also contributes significantly to this addictive potential. These platforms use machine learning to optimize content delivery based on past user behavior, a form of "persuasive technology".[109] This personalization enhances the perceived relevance of content and makes it harder for users to disengage. Re-

[100] UNICEF, n. d.

[101] Sempach, 2023.

[102] Klicksafe, 2024, June 25.

[103] American Psychological Association, 2013.

[104] Pellegrino et al., 2022.

[105] DAK Forschung, 2018, S. 18 ff.; Oberlin et al., 2024.

[106] Skinner, 1953; Eyal & Hoover, 2014.

[107] Meshi et al., 2013

[108] Sherman et al., 2016; Brand et al., 2019.

[109] Fogg, 2003.

search by Montag et al. notes that such algorithmic curation can trap users in a feedback loop, reinforcing behaviors through neurochemical rewards.[110]

Importantly, repeated engagement with these platforms can lead to behavioral addiction, characterized by symptoms such as mood modification, tolerance, withdrawal, and conflict, as defined by Griffiths.[111] A meta-analysis by Andreassen further supports the claim that excessive use of social media is associated with symptoms of addiction, especially in adolescents and young adults.[112]

9. Extremism & Radicalization

In the digital age, young people are increasingly confronted with various forms of digital violence, radical content, and extremist rhetoric. They can unknowingly become victims of propaganda or share problematic extremist content from others without realizing that these messages often hide anti-democratic and extremist strategies.[113] “Extremist content” refers to material that promotes involvement in, or aims to recruit individuals to, violent extremist groups. This includes terrorist organizations, hate groups, criminal organizations, and other non-state armed groups targeting civilians. It can feature names, symbols, logos, flags, slogans, uniforms, gestures, salutes, illustrations, portraits, songs, music, lyrics, or other items associated with violent extremist groups or individuals.[114]

What begins online can also lead to real-world violence, a risk that cannot be ignored. It is crucial for adolescents to find guidance to develop a clear stance against extremism and radicalization. Parents can play a key role by helping their children reflect on their views on religion and democracy, teaching them a respectful and constructive debate culture, and encouraging them to critically question the content they encounter. Given the growing threat of extremist ideologies in the digital world, studies show that a significant portion of young people in Switzerland are at risk of leaning towards various extremist directions. A sample survey revealed that 5.9% of the adolescents surveyed could be classified as right-wing extremists. Additionally, 7% of 17- and 18-year-olds were identified as left-wing extremists, while 2.7% of Muslim

[110] Montag et al., 2020.

[111] Griffiths, 2005.

[112] Andreassen, 2015.

[113] more information under: https://www.klicksafe.de/rechtsextremismus.

[114] World Economic Forum, 2023.

17- and 18-year-olds were labeled as Islamist extremists. These figures highlight the urgent need not only to raise awareness about the risks of extremism in the digital space but also to equip young people with the tools and strategies necessary to counter such influences.[115]

The study[116] of the Zurich University of Applied Sciences (ZHAW) and the Haute école de travail social Fribourg (HETS-FR), which examined the behavior of young people with extremist opinions, also shed light on their media behavior. 21% of Muslim young people had visited a website with radical Islamic content in the past year, and 23% watched jihadist propaganda songs. 9% of young people consumed left-wing extremist online content, while almost 15% listened to left-wing music. In the right-wing extremist area, 9% visited corresponding websites and 12% listened to right-wing extremist music. These results make it clear that extremist attitudes and behavior are widespread among young people in Switzerland and that the internet plays a central role in the spread of such ideologies. This is based on the fact that social networks, blogs, and comment sections offer numerous opportunities to express opinions, engage in public discourse, and spread propaganda. However, these platforms are increasingly being used to disseminate extremist views and recruit members for radical groups. The term "radicalization online" has become established in today's communication society.[117]

10. Algorithms

A particularly pressing aspect of systemic risks lies in the influence of algorithmic systems used by social media platforms. These algorithms, which determine what content users see in their feeds, are typically designed to maximize engagement by promoting emotionally charged, sensational, or provocative material.[118] For children and adolescents, whose cognitive and emotional regulation is still developing, this can lead to repeated exposure to potentially harmful or distressing content, such as violence, misinformation, or unhealthy beauty ideals. Moreover, algorithmic personalization can create so-called filter bubbles or echo chambers, reinforcing harmful beliefs or behaviors and limiting access to diverse perspectives.[119] Because these systems operate mainly without transparency or meaningful oversight, children are often unknowingly

[115] Jugend und Medien, n. d.

[116] Mazoni et al., 2019.

[117] Jugend und Medien, n. d.

[118] Sui et al., 2023.

[119] Knobloch-Westerwick & Westerwick, 2023.

subjected to curated digital experiences that may conflict with their best interests. The growing awareness of such algorithmic risks highlights the urgent need for child-centered platform design and stronger regulatory frameworks to ensure that automated content curation does not compromise the well-being of young users.[120]

11. Risky Online Behavior and Data Breaches

The degree to which children engage in risky online behaviors and neglect fundamental privacy and safety measures significantly influences their susceptibility to exploitation. Empirical evidence underscores that minors who partake in hazardous activities both on- and offline are exposed to substantially heightened risks. For instance, the seminal 2009 European Union Kids Online study revealed that 50% of young internet users shared personal information with strangers, 40% accessed pornographic material, 30% encountered violent or hateful content, and 10% consented to in-person meetings with online acquaintances.[121] A 2012 survey further indicated that 44% of teenagers admitted to viewing online content disapproved of by their parents,[122] while a 2011 study found that 42% of teens routinely deleted their browsing history to conceal their activities.[123] Risk of online solicitation and grooming correlates strongly with risky behaviors. Young people engaging in aggressive online conduct, such as posting offensive comments, frequenting pornographic websites, or interacting with unsolicited material via peer-to-peer networks, are more frequently targeted by explicit solicitations.[124] Moreover, willingness to discuss sexual topics with strangers online substantially amplifies vulnerability.[125] Notably, a psychosocial profile marked by multiple risky behaviors is a stronger predictor of solicitation risk than the specific technological platform employed.[126] Furthermore, the digital environment affords minors the opportunity to construct alternate identities, which may encourage risk-taking behaviors not exhibited offline. A relatively emergent channel of victim access is online gaming, which, by its immersive and often violent nature, fosters intense engagement and emotional attachment to virtual achievements. Disturbingly, cases have emerged of children committing

[120] Saldías, 2024.
[121] Livingstone & Haddon, 2009.
[122] Cox Communications, 2012.
[123] GFI Software, 2011.
[124] Berkman Center, 2008; Wolak et al., 2008.
[125] Berkman Center, 2008.
[126] Berkman Center, 2008.

suicide following the loss of in-game assets.[127] The virtual anonymity and absence of real-world social cues in gaming environments render young players particularly susceptible to offenders who cultivate trust and manipulate social dynamics to facilitate exploitation. Offenders may exploit gaming relationships by offering assistance or companionship as a pretext for initiating sexual advances or future abuse.[128]

The excesses to which risky online behavior can lead are illustrated by a recent case in France. On August 18, 2025, a popular online streamer died during a marathon livestream that had lasted nearly 12 days, broadcast on the platform Kick. His death occurred live on screen, during a broadcast in which he had endured prolonged humiliation and mistreatment by two co-presenters for purposes of viewer engagement and revenue. Investigations were swiftly launched by French authorities. An autopsy, conducted on August 21, revealed no signs of trauma or third-party intervention, suggesting that the cause of death was likely medical or toxicological in nature. The French Minister for Digital Affairs condemned the incident as an "absolute horror," highlighting critical failures in content moderation and platform responsibility. The streaming service Kick has pledged full cooperation, banned the involved streamers, and initiated a review of its content policies in France. This harrowing episode has ignited urgent debates around platform accountability, the ethics of extreme streaming, and the protection of vulnerable individuals in online environments. It underscores the necessity for robust regulatory frameworks to govern livestreaming practices, particularly when content veers into abusive or exploitative territory.[129]

Closely linked to risky behaviors is the frequent inattention to online safety and privacy precautions among children and young people. Children and young people often share vast amounts of personal information online, ranging from their favorite hobbies and daily routines to deeply sensitive and intimate details about their lives. This includes everything from photos and videos to private conversations and even location data. Unlike adults, minors frequently do not fully grasp the long-term consequences of putting such information into the digital world. Social media platforms, online games, and various apps have become integral to their social lives, encouraging constant sharing and interaction. Yet, this information is not only visible to their friends or followers; it is also collected, stored, and analyzed by the platforms them-

[127] Shared Hope International, 2014.

[128] Berkman Center, 2008; UNODC, 2015.

[129] https://www.srf.ch/news/gesellschaft/extreme-inhalte-streamer-jp-stirbt-vor-laufender-kamera-die-hintergruende.

selves, often without sufficient protective measures in place. Perpetrators often cast wide nets initially but successfully exploit children who indiscriminately share personal details, such as age, geographic location, school attendance, or real-time whereabouts via “check-in” functionalities.[130] This indiscriminate sharing escalates exposure to cyberenticement, cyberstalking, and the reception of harmful content. Publicly accessible social media profiles exacerbate these risks, as children with open profiles report greater instances of unsolicited contacts, cyberharassment, and bullying compared to peers with private settings.[131]

But even when the date is not shared with other users but with other third parties, such as the social media platform provider, when data breaches occur, whether through hacking, accidental leaks, or inadequate security protocols. This “treasure trove” of intimate data can fall into the wrong hands. The consequences can be severe, including cyberbullying, stalking, identity theft, and emotional distress. A breach involving a young person’s private information is not just a violation of privacy; it can fundamentally disrupt their sense of safety and trust in the digital world, potentially causing long-term psychological harm.

12. The threat to the ‘digital’ welfare of children

Despite the many new opportunities of the digital world, the dangers must not be overlooked. Unlike “traditional” child welfare endangerment, digital child welfare endangerment is difficult to predict, as the endless possibilities and ever-growing risks can hardly be assessed. However, a meaningful risk assessment requires at least a vague possibility of risk analysis. Depending on their maturity and age, children should and can be fully controlled neither by their legal guardians nor by third parties on the internet. This means that children must learn to identify problems themselves and be able to articulate them to adults. When it comes to adolescents, the challenge often lies in recognizing that, due to their natural process of detachment, they distance themselves from parental influence, and their socialization in digital experiences takes place partly without parental involvement. Consequently, child welfare endangerment often only comes to light when the resulting damage has already occurred. The challenge is to recognize risks in a timely and accurate manner. Identifying and preventing dangers while allowing an age-appropriate level of

[130] Berkman Center, 2008.

[131] Berkman Center, 2008.

autonomy is a balancing act that requires media literacy from parents. Protection requires knowledge – parenting and the responsible exercise of parental care have become more complex in the digital age due to the expansion of life into the digital realm.[132]

Vobbe and Kärgel describe digital child welfare endangerment as mediated child welfare endangerment. According to them, the following conditions must be met for such a case to exist:

a. The parents lack sufficient awareness of potential risks when a threat is suspected;
b. the parents are unable to avert the danger themselves; or
c. the parents themselves act in an abusive manner.[133]

Digital child welfare endangerment, therefore, occurs whenever harmful digital content or even assaults exceed the guardians' ability to manage the situation, preventing them from taking appropriate measures against the threat. Despite the difficulty in assessing risks in an ever-changing media landscape, dangers must be clearly identified. Many problem areas cannot even be anticipated.

A particularly concerning development for child welfare is that the porn industry hijacks almost every aspect of the digital landscape, and compared to other industries, operates on a high-investment and highly professional level. Given the profitability of the porn industry, this is hardly surprising. The revenues generated in the porn industry dwarf those of the traditional Hollywood film industry. Online platforms alone generate a daily turnover of over 12 million euros. Legislators must anticipate every new development in the digital world in a timely manner so that the inevitable regulatory lag does not become too vast. The constant and ubiquitous accessibility of pornography in the digital space poses challenges for adolescent development. However, it is not just the intentional consumption of pornographic material that is problematic, but also the unintentional exposure of minors to explicit content. For years, around 10% of surveyed minors have reported encountering inappropriate digital content. These materials primarily include erotic, sexualized, and pornographic depictions. Between 4% and 5% of surveyed children stated that the unintentional exposure to such content frightened them.[134]

[132] Oberlin et al., 2024.
[133] Vobbe & Kärgel, 2021.
[134] Oberlin et al., 2024

Violations of children's rights on social media have already risen sharply, with the phenomenon of Sharenting playing a central role. This trend not only raises concerns about media literacy but also challenges the ethical principles of modern society. Sharenting describes the practice of parents publicly sharing their childrens' lives online. Creating this content can severely impact children's safety and well-being. Moreover, due to the lasting nature of the internet, these harmful effects often persist into adulthood.[135] Authorities and tech companies can do little but observe as parents post images of their children—provided these do not breach criminal laws or platform regulations. However, this raises a critical question: What about the rights of the children themselves? Is it legally and ethically justifiable for parents to operate within these gray areas, potentially neglecting their child's best interests? Through the ever-growing digital landscape, society will face an unprecedented wave of personal rights violations, with criminal exploitation knowing no limits.[136]

[135] Simone, 2024.

[136] Oberlin et al., 2024

IV. International Guidelines and treaties on child online safety

In addition to existing legal frameworks, international guidelines and treaties on child online safety play a crucial complementary role in establishing coherent, cross-border standards aimed at protecting minors from digital threats in an increasingly interconnected world. These international instruments serve not only as legal benchmarks but also as policy blueprints to help governments, technology companies, and civil society actors align their efforts to ensure a safer digital environment for children.

1. OECD – Protecting Children Online

In 2011, the OECD addressed concerns regarding the increasing use of the internet by children and its associated risks, releasing a report and recommending measures for the protection of children online. This led to the OECD Recommendation on the Protection of Children Online, which was adopted in 2012, with a review scheduled for five years later. The recommendation aligns with the 1989 UN Convention on the Rights of the Child and calls for coordinated efforts among stakeholders to create a safer digital environment for children. A report was published to review the implementation of the recommendation, analyze emerging risks, and assess whether laws and policies have kept pace with technological advances. The findings were based on responses from 34 OECD countries to a 2017 survey on online child protection. The report highlights the significant changes in how children use digital platforms, with increased use of mobile devices and social media. It acknowledges that risks such as cyberbullying, sexting, and privacy issues have evolved, and new concerns have emerged, like sextortion. The report emphasizes that while the digital environment offers benefits such as connectivity and educational opportunities, it also presents new challenges. Children face increased privacy risks, exposure to harmful content, and targeted advertising, which they may be unprepared for. Policy makers are urged to find balanced solutions that address these risks while also promoting digital literacy and the benefits of the digital world. The report evaluates legal and policy responses, industry roles, and multi-stakeholder initiatives. It finds that current legislative responses are often fragmented, with risks being addressed in isolation by different ministries, potentially leading to gaps or duplication of efforts. Some countries have established oversight bodies, which have shown promise in coordinating responses. While digital and media literacy is recognized as essential, there is less focus on promoting positive digital content. There is a consensus on the importance of international and regional cooperation to address the global nature of the digital environment. However, challenges remain in regulating cross-border platforms and measuring the effectiveness of existing policies. The report concludes that balancing the promotion of digital media with the protection of children from its risks remains a complex issue for policy makers.[137]

The Council on Children in the Digital Environment recognizes the significant role the digital environment plays in the lives of children, offering both important opportunities and various risks. It emphasizes the need to create a digital space that both empowers and protects children, considering their different ages, maturity levels, backgrounds, and rights, particularly regarding privacy and data protection. The Council defines key groups involved in this environment—such as children (under 18), digital service providers, and other stakeholders—and calls on all actors to uphold five core principles to ensure a safe and beneficial digital environment for children. These principles include prioritizing the best interests of the child and safeguarding their rights online; supporting children and their caregivers by raising awareness of risks, rights, and available remedies; ensuring that protective measures are balanced, evidence-based, and respect fundamental freedoms without being unduly punitive or stifling innovation; addressing the diverse needs of children to avoid exclusion or discrimination based on social or economic circumstances; and fostering cooperation and dialogue among governments, businesses, families, educators, and children themselves. The Council recommends that governments demonstrate leadership by adopting clear, coordinated policies that integrate child rights into the digital domain. Legal frameworks should be fit for purpose, ensuring protection without limiting children's rights, promoting responsible business conduct, and providing effective remedies for harm. It also highlights the importance of promoting digital literacy tailored to children's developmental stages and circumstances, as well as using evidence-based policies supported by ongoing research. Furthermore, governments and stakeholders should encourage the design and adoption of age-appropriate safety features in digital products and services. International cooperation is strongly encouraged to share knowledge, harmonize approaches, build capacity, and ensure coordinated efforts globally. Digital service providers are called upon to follow guidelines and best practices that respect and protect chil-

[137] OECD, 2020.

dren's rights, considering their roles, services, and the legal contexts in which they operate.[138]

2. Council of Europe: Guidelines and advocating for respect, the protection and realization of the rights of the child in the digital environment

The Committee of Ministers of the Council of Europe, in accordance with Article 15(b) of its Statute, emphasizes the importance of ensuring that children enjoy their human rights as outlined in various international conventions, including the UN Convention on the Rights of the Child. The digital environment poses both opportunities and risks to children's well-being, necessitating appropriate protection and empowerment. The strategy acknowledges the role of information and communication technologies (ICT) in children's education and social integration, while also recognizing the risks of exploitation and abuse online. It reaffirms the responsibility of states to respect, protect, and fulfill children's rights, alongside the roles of parents, guardians, and businesses in safeguarding these rights in the digital space. The strategy highlights the need for a coordinated approach involving both public and private sectors and for children to be meaningfully involved in the process. The Committee recommends that governments review their laws and policies to align with these guidelines, ensuring their implementation and effectiveness. It encourages the dissemination of these recommendations, including in child-friendly formats, and urges businesses to fulfill their responsibility in respecting children's rights. Additionally, it calls for collaboration with the Council of Europe in developing and monitoring strategies to protect children's rights in the digital realm, with regular evaluations of progress.[139]

The guidelines aim to promote children's rights in the digital environment by supporting states and relevant stakeholders in implementing legal and policy measures that protect and realize children's rights. It emphasizes that the best interests of the child should be a priority, and the rights of all children, regardless of their background, must be respected. Children should be involved in decisions that affect them and be informed about their rights. Additionally, there is a responsibility for states, companies, and other stakeholders to ensure that children's rights are upheld online. The need for targeted measures

[138] OECD, 2021.

[139] Council of Europe, 2018.

for vulnerable children and close collaboration among all parties is also highlighted.[140]

The principles of actions outline key principles and measures for ensuring the respect, protection, and realization of children's rights in the digital environment. Access to the digital world is seen as essential for children to exercise their rights and freedoms, as it enables participation, education, and the maintenance of family and social connections. Therefore, states should ensure that all children have appropriate, affordable, and safe access to devices, internet connectivity, services, and content specifically designed for children. This includes providing public spaces with free internet access where possible and making sure that educational and care institutions offer children access to digital resources. Special measures should be taken for vulnerable groups, such as children in care, children who have been deprived of freedom, children in international migration situations, children living on the streets, and those in rural areas. Furthermore, online service providers should ensure accessibility for children with disabilities. This access should be complemented by education to help children understand and navigate issues like gender stereotypes or social norms that may limit their use of technology.[141]

Children's right to freedom of expression and information is crucial in the digital age. The digital environment has significant potential to support children's right to express their views, receive information, and share ideas. States should foster an environment where children can express their opinions freely, whether or not those opinions are accepted by others, including the state. Children should also be educated on how to exercise their right to free expression while respecting the rights and dignity of others. This education should address the concept of free speech and the legitimate limitations, such as respecting intellectual property rights and prohibiting the incitement of hatred or violence. States should promote the provision of diverse, high-quality online content for children that supports their development and participation in society, ensuring that the content is accessible, age-appropriate, and in a language that children can understand. The content should also include important resources about children's rights, health, and sexuality. Digital platforms should actively involve children in the creation of content, promote user-generated contributions, and recognize the importance of children's representation in online media. At the same time, any restrictions on children's right to freedom of expression must align with international human rights stan-

[140] Council of Europe, 2018.
[141] Council of Europe, 2018.

dards. Children must be informed of any restrictions, such as content filtering, in a manner appropriate to their evolving capabilities and must be provided with guidance on how to appeal or report issues. The digital environment also opens unique opportunities for children to participate in activities, play, and assemble peacefully, particularly through online communication, games, networks, and entertainment. States should work with other stakeholders to provide access to activities that promote participation, integration, and digital citizenship, both online and offline. The state must ensure that the digital content and services available to children support their development and meet their social, educational, and recreational needs. This includes promoting interactive and playful tools that foster creativity, teamwork, and problem-solving while being mindful of the needs of children in vulnerable situations. Additionally, children should have access to information about their participation rights in both online and offline settings.[142]

Regarding privacy and data protection, children's rights to privacy in the digital space must be upheld, including the protection of personal data and communication confidentiality. States should take necessary steps to ensure that children and all relevant actors, such as peers, parents, caregivers, and educators, are aware of and respect children's privacy rights. States should also ensure that children are informed of how to exercise their privacy rights, including understanding data collection practices and technological developments. Consent for processing personal data should be voluntary, informed, and given by the child and/or their parents or legal guardians, with special safeguards for sensitive data. The principle of data minimization must be respected, ensuring that personal data is processed in a way that is appropriate and necessary for the purposes it is intended for. States should assess the potential impacts of data processing on children's rights and ensure that the risks of harm are minimized. The processing of sensitive data, such as biometric information or health data, should only be allowed under strict legal safeguards. The use of age-appropriate privacy tools and settings should be promoted, ensuring that children's privacy is protected across digital platforms, especially those targeting children. Profiling children through automated data processing to create personal profiles should be prohibited, unless it serves the child's best interest or is necessary for a legitimate public purpose, with adequate safeguards in place.[143]

[142] Council of Europe, 2018.
[143] Council of Europe, 2018.

States should actively invest in and promote digital opportunities to fulfill children's right to education, aiming to develop their personalities, talents, and both their physical and mental abilities. Education should prepare children for an independent life in a free society, and digital resources must be made accessible to all children in an inclusive way, considering their developmental abilities and the unique circumstances of vulnerable children. In addition to access to education, the development of digital competence should be encouraged, including media and information literacy and digital citizenship. This will ensure that children have the skills to engage responsibly with the digital world and the resilience to manage its risks. Digital education should be incorporated into the core curriculum from an early age, adapted to children's evolving abilities, and should cover both technical skills and critical understanding of the digital environment's opportunities and risks. States should also support parents and caregivers to create a safer digital environment for their children. Furthermore, states must ensure that children with limited access to technology or those living in institutional care are not disadvantaged by the digital divide. Special attention should be given to children who face geographical, socio-economic, or disability-related barriers to engaging with technology, as well as ensuring that gender equality is promoted in access to digital resources. States must provide sufficient quality resources, devices, and infrastructures to support children's digital activities, and these should be developed in collaboration with relevant stakeholders. These resources need to be regularly evaluated for best practices to ensure effective education in the digital environment. Educational programs should teach children, parents, caregivers, and educators about prevention, rights, and duties in the digital world and include guidance on how children can identify and deal with harmful content like violence, pornography, and hate speech.[144]

At the same time, children's right to protection and safety in the digital space must be a priority. Protection measures should consider the child's developmental needs and avoid restricting their ability to exercise other rights. Special concerns include risks like sexual exploitation, grooming, online recruitment for crimes, and harmful content, as well as the physical and mental health risks of excessive internet use. To mitigate these risks, states should implement preventive measures and regularly analyze potential health risks related to new technologies. They should incentivize companies to adopt safety, privacy, and data protection measures in their products aimed at children, while encouraging the development of parental control tools that align with children's

[144] Council of Europe, 2018.

evolving abilities. These tools should also respect children's privacy rights and not reinforce discriminatory practices. States should also take action to protect children from premature exposure to the digital world and implement effective age verification systems to prevent access to inappropriate content. Additionally, laws should ensure that children are shielded from commercial exploitation in the digital environment, including age-inappropriate advertising and marketing practices. Finally, in the fight against child sexual exploitation, states must prioritize the identification, protection, and rehabilitation of children depicted in such abuse materials. They should work with companies to assist law enforcement in identifying perpetrators and collecting evidence for legal actions. Digital platforms must take responsibility for ensuring their networks are not used for illegal activities, such as the distribution of child sexual abuse materials, by using technical tools to prevent such abuse. In sum, it is crucial that states take proactive measures to ensure a safe and supportive digital environment where children can engage with technology responsibly and securely.[145]

3. EU: A European strategy for a better internet for kids (BIK+)

The BIK+[146] strategy, launched on May 11, 2022, is designed to ensure that children are protected, respected, and empowered in the digital world. Aligned with the European Digital Principles, it builds upon the original BIK strategy from 2012, considering advances in technology and changes in EU legislation. This updated strategy includes a child-friendly version, and importantly, children themselves will be involved in both its implementation and monitoring. The BIK+ strategy focuses on three key pillars: The first is ensuring safe digital experiences for children by protecting them from harmful content and online risks. The aim is to create digital environments that are age-appropriate and that prioritize children's best interests. The second pillar is digital empowerment, which ensures that all children, including those in vulnerable situations, are equipped with the necessary skills to navigate the digital world safely and make informed choices. Finally, the strategy emphasizes active participation, giving children a voice in the digital space and supporting child-led initiatives that encourage creativity and innovation in safe online environments. The Better Internet for Kids portal continues to play a key role in providing resources for children, parents, and educators, collaborating with Safer Internet Cen-

[145] Council of Europe, 2018.
[146] European Commission, 2012.

ters across EU Member States. Ongoing initiatives include the development of standardized age assurance methods, quick responses to harmful content, and support services for victims of cyberbullying, such as the "116 111" helpline. Youth participation is a central element of the strategy, with regular evaluations led by children, increased peer-to-peer activities, and expectations for the digital industry to consult young users. The strategy will continue to shape policy development across EU Member States, and also contribute to global efforts to protect children's digital rights. Since 2021, BIK+ has been an integral part of the broader EU Child Rights Strategy, ensuring that children's voices are at the heart of digital policy.[147]

A. European Commission: A report from the consultation with children and young people

The European Commission aims to promote the participation of children in EU policies and initiatives and to ensure that children in particularly vulnerable situations are also heard. In March 2021, the '2030 Digital Compass' was published, which contains a vision for Europe's digital transformation by 2030. It emphasizes that the rights and values of the EU should also be upheld in the digital space. The Commission proposes to develop a digital framework that promotes access to high-quality connectivity, digital skills, and fair online services and ensures that the same rights that apply offline can also be exercised online. Another important component of EU policy is the 'EU Strategy on the Rights of the Child', which aims to better protect children and promote their rights by placing them at the center of EU policy. This strategy also includes the goal of continuing to offer children and young people protection and empowerment in the digital space. As part of the 'Better Internet for Kids' initiative, European Schoolnet conducted a consultation from May to October 2021 to gather the views of children, young people, parents, educators, and other stakeholders on priorities related to children's rights in the digital world. Over 750 children and young people from Europe took part in this consultation, which was organized with the support of the Insafe network and other European internet safety and children's rights organizations. This consultation emphasized the importance of children's right to be heard in decisions that affect them and advocated for children's and young people's interests to be better integrated into the EU's digital ambitions by 2030. The consultation emphasized the importance of equipping children and young people with digital skills to navigate safely and responsibly online, as well as protecting them from

[147] European Commission, 2022a.

harmful and illegal content, such as sexual abuse and exploitation. Respondents expressed strong opinions about the activities they favor in the digital space and the risks they face, particularly cyberbullying, harmful content and fake news. Data protection and privacy were also high on the list of concerns. Concerns were raised about the lack of inclusivity and accessibility of the internet for children and young people with disabilities and vulnerable groups such as marginalized children and young people.

Recommendations for policy makers and digital providers included the need to prioritize inappropriate content and tackle cyberbullying as it affects children, and young people's behavior and self-esteem. It was pointed out that many of the risks and challenges that exist online arise from a lack of awareness and media literacy and that anonymous behavior online exacerbates the situation. It was also suggested that the EU should commit to harmonized rules for the protection of minors, including uniform age limits and better monitoring. Another key concern was the improvement of media literacy and digital safety education in schools to better prepare children for the risks of the internet. Consultation participants emphasized that schools and parents need to take their responsibilities in this area more seriously. It was also suggested to improve the enforcement of existing rules and impose stricter penalties for those who engage in harmful behavior in the digital space, such as bullying or spreading hate speech. Finally, the participants recommended that EU politicians should do more to encourage the industry to develop safe, child-friendly platforms and regulate their content according to age groups and categories. This includes, among other things, the regulation of sexualized or violent content and better protection against gambling advertising or spam. Overall, the consultation shows the need for a comprehensive and coherent strategy to enable the protection, empowerment, and participation of children and young people in the digital space, with all relevant stakeholders – from policy makers and digital providers to parents and teachers – sharing responsibility.[148]

B. Inclusion, Accessibility, and Digital Skills as Cornerstones of Child Empowerment

A central focus of the BIK+ strategy is not only to ensure protection and participation in the digital environment but also to promote digital equality of opportunity. This includes ensuring inclusive, accessible, and age-appropri-

[148] European Commission: Directorate-General for Communications Networks, Content and Technology, 2022

ate digital content for all children, particularly those in vulnerable situations. The EU-wide consultation carried out as part of the "Better Internet for Kids" initiative revealed that children and young people with disabilities or from marginalized backgrounds often face systemic barriers in accessing online platforms, educational content, and safe digital services. Participants strongly emphasized the need to improve digital accessibility and to ensure that the unique needs of vulnerable groups are considered in policy development and platform design.[149]

Moreover, the strategy highlights the importance of digital literacy and media education. Equipping children with the skills needed to navigate digital environments critically, responsibly, and safely is viewed as a key pillar of both the BIK+ and the broader EU Strategy on the Rights of the Child. Lack of media literacy was identified by young respondents as a major contributor to online risks such as cyberbullying, exposure to fake news, and uncritical sharing of personal data. As a result, there is a strong call for media education to be systematically integrated into school curricula, as well as for greater involvement of parents and educators in fostering these skills. The goal is to empower children to make informed decisions and engage meaningfully with digital technologies.[150]

In addition, the consultation highlighted the need for harmonized legal standards across the EU for the protection of minors online. Participants advocated for uniform age verification rules, clearer definitions of harmful content, and better enforcement mechanisms. Importantly, responsibility is not placed solely on governments but also on digital service providers, who are expected to design and maintain platforms that are safe and appropriate for young users. This includes stricter regulation of sexualized or violent content, gambling ads, and hate speech, as well as improved content moderation systems. These demands are reflected in the BIK+ strategy, which outlines cross-sectoral efforts at national, EU, and international levels to safeguard and promote children's digital rights.[151]

4. EU's Guide to age assurance

Protecting the safety of children and young people in the digital space without compromising their rights is a key issue being discussed worldwide. Age safety

[149] European Commission, 2022a.

[150] European Commission, 2021.

[151] European Commission, 2022a.

tools are seen as a solution to create safer online spaces for children. In line with the European Union's Better Internet for Kids+ (BIK+) strategy[152], a new section on age safety will be introduced on the BIK platform. The aim is to make this topic more understandable for everyone, from the public to young people. The new 'Old Age Security Toolkit' collection offers family-friendly explanations without jargon, a comprehensive explanation of terms, and visual content that is accessible to all age groups. They are designed to help families, teachers, and digital service providers take action to protect children online. In addition, a report[153] by the European Commission outlines various safeguarding methods, challenges, and legal frameworks. This helps to develop a better understanding of how safeguarding tools work. A self-assessment questionnaire and a handbook are available for digital service providers. These resources help to promote safeguarding practices that meet regulatory requirements and incorporate best practices to protect children online.

5. UN: General comment No. 25 (2021) on children's rights in relation to the digital environment

On March 24, 2021, the UN Committee on the Rights of the Child officially released General Comment No. 25 on children's rights in the digital environment.[154] This marked a historic moment, as it was the first time that children's digital experiences were explicitly mentioned within the UN Convention on the Rights of the Child (CRC).[155] The Comment affirms that children's rights apply both offline and online, marking a significant step forward in the global protection of children's rights in the digital age. The General Comment No. 25 aims to guide states on how to implement children's rights in the digital realm. It provides detailed instructions for the development of laws and policies to ensure that state obligations under the CRC are fulfilled. The Comment addresses critical issues such as privacy, non-discrimination, protection from harm, education, and the right to play and leisure in the digital world. It underscores that the digital environment offers both opportunities and risks for children, which must be carefully considered.

The document results from an extensive review process that involved state reports, legal rulings, and consultations with children from 28 countries. A total

[152] European Commission, 2022a.

[153] European Commission: Directorate-General for Communications Networks, Content and Technology et al., 2024.

[154] UNCRC, 2021.

[155] United Nations, 1989.

of 709 children participated in these consultations, including those from vulnerable groups such as minorities, refugees, children with disabilities, and incarcerated children. These consultations were essential in ensuring that the voices and perspectives of affected children were incorporated into the final text of the Comment. General Comment No. 25 urges states to take action to protect children from digital risks, including online harassment, abuse, and exploitation. It places responsibility on both states and businesses to ensure that children's rights are respected in the digital environment. Key measures include prohibiting targeted advertising and profiling of children for commercial purposes, as well as banning practices like neuro-marketing and emotional analysis that exploit children's digital data. The Comment also calls for states to launch awareness campaigns and promote educational programs for children, parents, caregivers, and policymakers to increase understanding of the opportunities and risks associated with the digital world. Additionally, businesses operating in the digital space are called upon to ensure that they respect children's rights and prevent any abuse related to children's digital experiences. Furthermore, the Comment highlights the importance of collaborating with civil society to develop, implement, monitor, and evaluate laws, policies, and programs related to children's rights in the digital environment. Humanium, an NGO, played an active role in the drafting process and submitted recommendations to emphasize the importance of children's rights in the digital world. Humanium supports the fundamental principles of the Comment, particularly the right of children to protection from abuse, exploitation, and violence online.[156]

6. The 24/7 Network established under the Convention on Cybercrime (known as the Budapest Convention)

The Council of Europe's Convention on Cybercrime[157], signed in 2001, established a legal framework for the production and distribution of CSAM and provided instruments for international cooperation among member states. The convention introduced a "24/7 Network" in Article 35, which allows immediate assistance in criminal investigations and digital evidence collection. The network operates around the clock, offering technical and legal support, preserving data, and locating suspects, with communication primarily through e-mail. However, due to strict legal requirements in many countries, data preservation requests are often preferred over immediate data collection. These re-

[156] General comment No. 25 (2021) on children's rights in relation to the digital environment.

[157] Council of Europe, 2001.

quests help secure digital evidence while the legal procedures for Mutual Legal Assistance (MLA) are followed. Data preservation requests have become a critical tool in expediting cybercrime investigations, especially in online child sexual abuse cases, ensuring evidence is preserved before it's lost. The second protocol to the Budapest Convention may introduce new functionalities to the network, such as directly requesting subscriber information from ISPs during emergencies.[158]

7. AEPD: Decalogue of principles for age verification and protection of minors from inappropriate content

The Decalogue of principles discusses the principles and challenges related to the protection of minors from inappropriate content on the Internet. Protection systems, although aimed at safeguarding minors, must respect the fundamental rights of all Internet users, including adults. These systems must be legitimate, necessary, and proportionate, ensuring that the data processing is transparent, auditable, and does not compromise users' privacy. The following 10 key principles are outlined for designing these systems and aim to protect the privacy, security, and safety of minors while ensuring that adults can access them. These principles emphasize the critical need to protect minors from inappropriate content while ensuring their privacy and safeguarding fundamental rights:[159]

1. Systems designed to protect minors must ensure that the identification or location of minors cannot be traced over the internet. These systems should not collect personal data or verify the age of users in ways that could lead to the identification of minors. Protecting minors' identity is crucial, as such systems could be exploited by malicious entities, such as predators or those seeking to target minors with harmful content.
2. The purpose of age verification systems should be to confirm that users are of appropriate age to access certain content, not to verify whether someone is a minor. Age verification mechanisms should not be designed to identify minors or collect excessive data, as doing so could expose minors to risks and violate data protection regulations.
3. The verification process should preserve users' anonymity, especially in relation to internet service providers and third parties. The collection and

[158] Açar, 2023.

[159] Agencia Española de Protección de Datos, 2023.

processing of personal data should be kept separate from the age verification process, ensuring that the system does not inadvertently expose or link a person's identity to their access to adult content.

4. The obligation to verify the status of "authorized to access" content should only apply to restricted content, not across all internet use. People should be allowed to browse the internet freely and anonymously unless they are attempting to access content with specific age restrictions. This principle ensures that minors are not automatically flagged simply by using the internet.
5. Age verification must be accurate and categorical, focusing on confirming that someone is authorized to access certain content without revealing specific age details or birthdates. Age verification systems should avoid providing age estimates or exact age data, as such practices can be error-prone and discriminatory, especially when applied to minors. Instead, verification should confirm the status of "authorized to access" without exposing further personal information.
6. This principle stresses that systems designed for this purpose must avoid profiling users based on their browsing behavior. While it is necessary to label content appropriately (e.g., "violent," "explicit"), there is a risk of profiling when content access policies are enforced on servers or third-party platforms. To mitigate these risks, content filtering should be carried out locally on users' devices, where it can be more effective and protect privacy.
7. Highlights the importance of ensuring that users' activities are not linked across different services, as such linking can lead to intrusive profiling. Systems must avoid using unique identifiers that track users across platforms, as this can compromise personal privacy and lead to the creation of comprehensive user profiles.
8. The focus is on ensuring that parents retain authority over their children's internet use, allowing families, educational institutions, and experts to be involved in determining what content is suitable for minors. Commercial entities should not be allowed to dictate what content minors can access.
9. Emphasizes that systems designed to protect minors must respect fundamental rights, including privacy, freedom of expression, and non-discrimination. These systems must not unduly limit access to content or generate unnecessary data that could jeopardize privacy.
10. Calls for a clear governance framework to oversee the implementation of these protections. Such systems must be effective, transparent, and auditable to ensure compliance with privacy laws and prevent manipulation, bias, or security breaches.

In conclusion, protecting minors from inappropriate online content is a shared responsibility that involves families, educational institutions, governments, and the private sector. While technology can support these efforts, it must be used within a framework that prioritizes privacy, effectiveness, and the active participation of all stakeholders.

V. Regulatory Approaches to Child Online Safety

1. Regulatory approaches to protect minors on the internet from a Youth Protection Perspective

Youth protection laws aim to safeguard minors from harmful influences while ensuring their healthy development. They regulate access to media, alcohol, tobacco, gambling, and public spaces, aiming to balance protection with personal freedom. A key objective is to prevent harm by restricting minors' exposure to violent, explicit, or addictive content in media and entertainment. Additionally, these laws control the sale of substances such as alcohol and tobacco, ensuring that young people are not exposed to them prematurely. Age restrictions for films, video games, and online content further help mitigate potential risks.

The following section presents the youth protection laws of Germany and Switzerland. As recently enacted regulations, they provide a comprehensive overview of the objectives pursued by youth protection legislation. Consequently, they serve as exemplary cases for understanding the core principles and goals of such legal frameworks.

A. Switzerland

a. JSFVG

In recent years, the protection of minors in digital media has gained increasing public attention. The Swiss Federal Council's "Youth and Media" report from May 13, 2015, examined key challenges and regulatory needs in this area.[160] Even a decade ago, it was evident that youth media protection should go beyond standardized media content and focus on the potential risks associated with the growing use of various (digital) media by adolescents—whether as consumers, communicators, or active participants. Child and youth media protection operates at the intersection of personal rights, data protection, and

[160] Jugend und Medien, 2015.

consumer protection, highlighting the need for systematic collaboration and joint regulatory development.[161]

The new Swiss Federal Law on Youth Protection in Films and Video Games (JSFVG) and the corresponding ordinance (JSFVV) have been in effect since January 1st, 2025. This law establishes a nationwide standardized legal framework to better protect children and adolescents from inappropriate media content. Approved by the Swiss Parliament on September 30, 2022, the JSFVG aims to shield minors from media content in films and video games that could harm their development, particularly violent and sexually explicit material. A key element of the law is the implementation of uniform age ratings and access control across the country. This ensures that parents receive clear guidance on age-appropriate content while also holding film and video game providers accountable for youth protection. Under the new law, all providers of films and video games in Switzerland will be required to follow these key regulations: Define and display minimum age ratings for their content; Implement strict age verification measures before granting access; Streaming services must verify users' age before providing access to adult content; Parental control systems must be available; Reporting mechanisms must be in place to allow users to flag inappropriate content.[162]

b. Controversy

A significant but debated advancement in the Swiss Youth Protection Act is the requirement of online platforms to implement age verification systems (Article 20, Paragraph 2, Letter a, JSFVG). However, this measure has sparked controversy, with critics expressing concerns that users might have to provide sensitive personal information, such as their passport, to prove their age. The adoption of zero-knowledge proofs (ZKPs) for age verification in the future would be a welcome improvement from a data protection standpoint. This technology would allow age verification without requiring users to disclose personal data, ensuring greater privacy and security in the digital realm.[163]

c. Enforcement

The JSFVG will be gradually enforced, following a co-regulation approach that fosters collaboration between the government and private sector. The fed-

[161] Kunz et al., 2024.

[162] BSV, 2024, June 6.

[163] Kunz et al., 2024.

eral government has set the minimum legal requirements, which took effect on January 1, 2025. Industry organizations representing film and video game providers will then have two years to develop their own youth protection regulations, ensuring compliance with at least the federal standards. By standardizing age ratings and enforcing parental controls, this new law seeks to balance child protection with industry responsibility, providing better safeguards for young audiences in the evolving digital entertainment landscape.[164]

B. Germany

a. JuSchG

Since 1 May 2021, the new regulations of the Youth Protection Act (JuSchG) have been in force, which are intended to create modern framework conditions, especially for the use of digital media, in order to protect minors in the best possible way and minimize possible dangers.[165] The reformed law includes several key regulatory approaches in which the legislator focuses on protection, guidance, and enforcement. Since then, relevant internet services have been obliged to implement appropriate and effective protective measures and structural precautionary measures[166] for underage users, such as safe default settings, easily accessible reporting and help systems, and age verification systems. The specific design of the measures varies depending on the platform, which leaves the legislator a great deal of room for maneuver for future developments in the digital space. In future, online platforms will also have to provide transparent age labeling for their content, where the risk of interactive elements, such as chats, must also be included in the age rating.[167]

Since the amendment of the JuSchG, the tasks of the Federal Centre for the Protection of Children and Young Persons in the Media, or BzKJ for short, have been regulated and significantly expanded in Section 17a of the JuSchG. The BzKJ, which emerged from the former Federal Review Board for Media Harmful to Young Persons, is now responsible for monitoring compliance with the new requirements of the DSA and Section 24a of the Protection of Young Per-

[164] BSV, 2024, June 6.

[165] BLJA, 2021.

[166] Structural precautionary measures include reporting and remedial procedures that enable users to quickly report harassment or harmful content, as well as secure default settings that prevent children from being contacted by strangers.

[167] Bundesministerium für Familie, Senioren, Frauen und Jugend, 2021; Oberlin, 2025, February.

sons Act.[168] In accordance with Section 24a JuSchG, a central component of the new tasks is the enforcement of the law in the area of systemic precautionary obligations, which obliges platforms to implement safe default settings and help systems. If the platforms within the scope of application of the standard do not co-operate or do not co-operate sufficiently, fines of up to 50 million euros may be imposed.[169] The process of review by the Office for the Enforcement of Children's Rights in Digital Services (KidD)[170] takes place in two phases in accordance with Section 24b JuSchG, consisting of the risk assessment and the risk encounter. Firstly, the KidD[171] must evaluate whether the provider has taken appropriate precautionary measures. If this is not the case, the provider is requested to implement the necessary measures.[172]

b. The Interstate Treaty on the Protection of Minors in the Media (JMStV)

The JMStV[173] is a treaty whose purpose is to ensure standardized protection for children and young people from harmful content in electronic information and communication media. This applies in particular to content that could jeopardize or impair the development or education of minors, as well as content that violates human dignity or other rights protected by the Criminal Code. The scope of this Interstate Treaty includes both broadcasting and telemedia within the meaning of the Interstate Media Treaty. In addition, it also applies to providers who are not based in Germany but provide content for use in Germany. In this case, providers must comply with the requirements of European Directives 2010/13/EU and 2000/31/EC, which regulate the provision of audiovisual media services and electronic commerce. A service is deemed to be intended for use in Germany if it is aimed at German users through language, content, or marketing measures or generates a significant proportion of its refinancing from Germany. This treaty also applies to providers of video sharing services based in Germany.[174]

[175]The Interstate Treaty defines key terms such as 'offer' (broadcasts or telemedia content), 'provider' (broadcaster or telemedia provider), 'child' (under

[168] BzKJ, n. d.
[169] BzKJ, n. d.
[170] BzKJ, n. d.
[171] BzKJ, n. d.
[172] BzKJ, n. d.; Oberlin, 2025, February.
[173] JMStV, 2002/2024.
[174] JMStV, 2002/2024.
[175] JMStV, 2002/2024.

14 years of age), ‘young person’ (14 to 17 years of age) and inappropriate content. Inappropriate content under the JMStV is the following:

a. Contain propaganda material or unconstitutional symbols.
b. Incite hatred against population groups or glorify violence.
c. depict or trivialize violent acts.
d. Glorify war or depict the suffering of people in an unethical manner.
e. depict children or young people in offensive, sexualized postures or contain pornographic content.
f. have been added to the list of media harmful to minors.

In addition, content that is pornographic or endangers the development of children and young people is also prohibited if it is not labelled as suitable for adults. Even after significant changes to the content of an offer, it remains illegal until a decision has been made by the Federal Review Board for Media Harmful to Young Persons.[176]

Providers that distribute content potentially harmful to the development of children or adolescents must ensure that such content is not accessible to the relevant age groups. The regulations apply to four age groups: 6, 12, 16, and 18 years old. If content is not approved for a specific age group under the Youth Protection Act, it is presumed to have a potentially harmful effect. Providers can comply with these regulations by using technical measures, age ratings, or adjusting the distribution time. Certain content, such as news and political programs, is exempt from these rules. Video-sharing service providers must also take measures to protect children and adolescents from developmentally harmful content. These measures include age verification, parental controls, and allowing users to rate content. Additionally, providers must have a process in place for users to report complaints about unlawful or harmful content, ensuring these complaints are promptly addressed. For broadcasts outside of the designated time restrictions, the content must be appropriately marked to minimize its potential harmful effect on children and adolescents. Advertising for indexed content is only permitted under the same conditions that apply to the distribution of the content itself. This means that advertising for offers listed as harmful to youth, according to § 18 of the Youth Protection Act, may not be distributed or made accessible. Additionally, advertisements cannot indicate that a procedure to add content or its equivalent to this list is ongoing or has occurred. Furthermore, advertisements must not physically or mentally harm children and adolescents. They must not di-

[176] JMStV, 2002/2024.

rectly encourage children or adolescents to purchase or rent goods or services by exploiting their inexperience or gullibility. It is also prohibited to take advantage of the special trust children and adolescents have in their parents, teachers, or other trusted adults. Advertisements should not prompt children or adolescents to persuade their parents or others to buy the advertised goods or services. Also, depicting children or adolescents in dangerous situations without a justified reason is prohibited. Advertising whose content is likely to impair the development of children and adolescents into responsible, community-oriented individuals must be separated from offers directed at children and adolescents. Advertising that targets children or adolescents or uses children or adolescents as actors must not exploit their inexperience or harm their interests. Special rules apply to alcohol advertisements, which may not be directed at children or adolescents, nor may their style appeal to this group or show children and adolescents consuming alcohol. Additionally, teleshopping advertisements must not encourage children or adolescents to enter into purchase or rental contracts for goods or services. Providers must take appropriate measures to reduce the impact of advertising for unhealthy foods, especially in children's programming, which contains nutrients and substances like fats, trans fats, salt, sodium, and sugar, whose excessive intake is not recommended within a balanced diet.[177]

Regarding youth protection officers, it is stipulated that providers of cross-border television or telemedia containing developmentally harmful or youth-endangering content must appoint a youth protection officer. This officer is responsible for advising the provider on all matters of youth protection and must be involved in relevant decisions in a timely and comprehensive manner. The youth protection officer must possess the necessary expertise and must be independent in their duties. They cannot be disadvantaged due to their responsibilities and must be provided with the necessary resources. Providers of telemedia with fewer than 50 employees or fewer than 10 million monthly accesses can waive the appointment of a youth protection officer if they join an accredited self-regulatory body and delegate the responsibilities to it.[178]

For broadcasting, specific rules apply to scheduling broadcast times. The regional public broadcasters, the ZDF, or recognized self-regulatory bodies can impose time restrictions on films and other formats if their content is likely to impair the development of children and adolescents. If a self-regulatory body has created a guideline for a film or program, this guideline takes precedence.

[177] JMStV, 2002/2024.
[178] JMStV, 2002/2024.

In certain cases, exceptions to general rules may be granted, such as when a content's evaluation dates back more than ten years.[179]

For telemedia, providers are required to implement youth protection programs that can read age classifications and detect content that might impair the development of children and adolescents. These programs must be presented to an accredited self-regulatory body for evaluation. They must be user-friendly and capable of allowing access to media in an age-appropriate manner. Even programs that are only aimed at specific age groups or those that allow access within closed systems can develop such youth protection programs. Providers of telemedia offering content like films or games on media carriers must clearly indicate the appropriate age classification and, if necessary, follow the classification procedure outlined in the Youth Protection Act.[180]

The state media authority ensures compliance with the provisions of the State Treaty for media providers and makes decisions based on the treaty's rules. The Commission for Youth Media Protection (KJM) supports the media authority in this process. The KJM consists of twelve experts, including representatives from state media authorities, state youth protection authorities, and the federal youth protection authority. KJM members serve a five-year term with the possibility of reappointment, and at least four members must have qualifications equivalent to judges. The KJM is responsible for making final decisions on various aspects of youth media protection, including age ratings, broadcasting times, and monitoring compliance. It can form review committees, which must make decisions by consensus, and it operates independently, bound only by confidentiality rules. The KJM also collaborates with self-regulation bodies, ARD, ZDF, and other authorities to implement the treaty's provisions. "Jugendschutz.net," a joint institution of state youth protection authorities, works with the KJM to monitor telemedia content and provide consultation and training. Self-regulation bodies for media can be recognized by the KJM if they meet specific criteria, such as independence and expertise. These bodies must follow procedural rules, allow for feedback, and address complaints. If a self-regulation body fails to comply with the treaty, the KJM can intervene and require corrections. Within this framework, the KJM in the Media examines and evaluates possible violations of the JMStV on

[179] JMStV, 2002/2024.

[180] JMStV, 2002/2024.

the Internet in the sense of telemedia supervision. It decides on appropriate measures, which are then implemented by the state media authorities.[181]

C. Regulatory Gaps

Despite the 2021 reform of the German Youth Protection Act (Jugendschutzgesetz, JuSchG) and the existing provisions of the Interstate Treaty on the Protection of Minors in the Media (Jugendmedienschutz-Staatsvertrag, JMStV), significant regulatory gaps persist regarding the effective and holistic protection of children's welfare in the digital environment. While the amended JuSchG introduced new obligations to better protect minors from online risks—such as requirements for age-appropriate default settings, accessible support systems, and transparent age ratings—there is still a lack of uniform minimum standards for these protective measures. Their concrete implementation largely remains at the discretion of platform providers, resulting in substantial discrepancies in the level of protection and hampering legal enforceability.

Although the establishment of the Federal Agency for the Protection of Children and Young People in the Media (Bundeszentrale für Kinder- und Jugendmedienschutz, BzKJ) marked an important step toward centralized oversight, enforcement faces considerable challenges, especially in cross-border cases involving service providers based outside the EU.

Another significant regulatory shortcoming arises in the context of interactive features such as chats, comment sections, or user-generated content. While these elements must be considered in age rating processes, no binding criteria exist to assess their risks. In practice, this leaves the evaluation and classification of interactive risks to the subjective judgment of providers, which increases the risk of under-classification. Furthermore, neither the JuSchG nor the JMStV currently provides adequate responses to new technological developments such as AI-generated content, deepfakes, or algorithmic recommendation and targeting systems. The potential use of artificial intelligence to manipulate, profile, or exploit minors remains unregulated.

A central protection gap can also be identified in relation to the phenomenon of "sharenting"—that is, the practice of parents or legal guardians publishing images of their children on social media platforms. This form of data disclosure can severely impact a child's privacy and long-term well-being. However, it is scarcely addressed in the current regulatory framework, as existing provisions

[181] KJM, n. d.

primarily focus on service providers. The child's personal rights are thus insufficiently protected against parental overexposure, despite the very real risks of misuse—such as grooming, doxxing, or the use of children's images in training AI systems. Moreover, the current legal framework treats the child largely as a passive object of protection. There is no legal obligation to involve the child in decisions about the publication of their personal data or images, even though Articles 12 and 16 of the UN Convention on the Rights of the Child (UNCRC) enshrine the right to be heard and the right to privacy. The regulatory regime fails to recognize the growing digital autonomy and participatory capacity of children and adolescents.

Further regulatory deficits are evident in the commercial exploitation of children's images, especially in the context of so-called child influencers or family-run social media accounts. The line between private family life and professionalized, profit-driven content creation is increasingly blurred, yet specific legal protections are lacking. While the JMStV contains certain provisions regarding advertising and its impact on minors, there is no tailored regulation addressing scenarios in which children themselves become the subject of monetized online content.

Another concern relates to the heavy reliance on industry self-regulation. Although tasks may be delegated to recognized self-regulatory bodies, this can weaken governmental oversight—particularly when providers' economic interests dominate. While the Commission for the Protection of Minors in the Media (Kommission für Jugendmedienschutz, KJM) has oversight powers and can intervene in cases of misconduct, enforcement tends to be reactive rather than preventive.

The legal framework also provides insufficient protection against phenomena such as doxxing or the unauthorized use of children's images through automated systems. Publicly accessible photos of children—often uploaded by parents—can be harvested, shared, or repurposed without the knowledge or consent of the children or their legal representatives, and with little to no access to effective legal remedies.

Finally, the division of regulatory responsibilities between the federal government, the German "Länder", and European authorities has led to fragmented enforcement mechanisms. The overlapping mandates of the BzKJ, KJM, state media authorities, and European regulators such as the Commission and Digital Services Coordinators (DSCs) hinder a coherent and coordinated enforcement regime, especially given the global and rapidly evolving nature of digital platforms.

2. Regulatory approaches to protect minors on the internet from a Data Privacy Law Perspective

Regulatory approaches to protecting minors online increasingly acknowledge the specific risks children face in the digital space. From a data privacy perspective, regulations are evolving to ensure that minors' personal data is handled with particular care through stricter consent requirements, enhanced transparency, and clear limitations on data processing. For this publication, a selection of well-balanced European data protection laws has been examined to illustrate how these protections are implemented in practice. Additionally, the U.S. Children's Online Privacy Protection Act (COPPA) is briefly explored as a comparative example from a non-European regulatory context.

A. EU – General Data Protection Regulation (GDPR)

The European Union's General Data Protection Regulation governs data protection across all member states, ensuring a uniform[182] or stricter level of privacy protection[183] and the strengthening of the rights of data subjects.[184] This strict regulation affects social media platforms, as GDPR applies to data processing within the context of an establishment[185] in an EU member state, regardless of where the processing takes place. According to Art. 3 GDPR, most international Social Media companies fall under GDPR's jurisdiction, meaning they must adhere to its strict rules.[186]

Regarding children's privacy protection on social media, personal data is inevitably processed when users post or share content. This includes information such as names, locations, photos, videos, and all other information, which makes the data subject identifiable.[187] Even if privacy settings are not changed,

[182] Rücker, 2018.

[183] Council of the European Union, n. d.; Detmering & Splittgerber, 2018.

[184] Rücker, 2018.

[185] Localization of the establishment pursuant to Art. 52 TEU, concretized by Art. 355 TFEU: Establishment is deemed to exist according to Recital 22, establishment is deemed to exist if the effective and actual exercise of an activity takes place in a fixed establishment, see establishment, see ECJ of 1 October 2015 – C-230/14, Weltimmo, EuZW 2015, 3636, 3639, para. 41, on Art. 4 para. 1 lit. a Directive 95/46/EC. (Local establishment pursuant to Art. 52 TEU, substantiated by Art. 355 TFEU: Pursuant to recital 22, establishment is deemed to exist if the effective and actual exercise of an activity takes place in a fixed establishment, see ECJ of 1 October 2015 – C-230/14, Weltimmo, EuZW 2015, 3636, 3639, para. 41, on Art. 4 para. 1 lit. a Directive 95/46/EC.)

[186] Husi-Stämpfli et al., 2024.

[187] European Parliament & Council 2016, GDPR Art. 4(1).

personal information may be shared with a wider audience, especially when users add strangers to their friend list or data is shared indirectly via screenshots with third parties.[188] The GDPR recognizes the sensitivity of children's data and grants them special protection, as children are less likely to fully understand the risks of online data processing. Recital 38 of the GDPR highlights that children deserve extra protection, particularly in areas like data collection for advertising or profile creation. Article 8 of the GDPR outlines the special conditions for a child's consent,[189] stating that data processing is only lawful if a child is 16 or older. For younger children, consent must be given by the parent or guardian. Consent is vital in this context. Children under the age of 16 cannot lawfully consent to data processing unless approved by a parent[190]. However, the ability of minors to understand data processing risks is increasingly acknowledged, and there is a shift in recognizing that younger users may be more digitally competent. Additionally, consent must be revocable at any time, as per Article 7 (3) of the GDPR. If a child withdraws consent, the platform must cease processing the data and delete any related information. However, the "Right to be Forgotten" may not always fully apply, as content could still exist elsewhere on the internet.[191]

Children must be provided with the same level of transparency and clarity regarding data processing as adults. According to Articles 12, 13 GDPR requires that information about data processing activities be clear, easily understandable, and accessible, enabling the individual (or their guardian) to make informed decisions, such as consenting to cookies or adjusting privacy settings. This transparency is particularly important for children, who may not be fully aware of the risks involved in data processing. Thus, the information should be communicated in a way that children can understand based on their cognitive ability or, if needed, with the help of their parents. The process and impact of data processing, as well as the extent of consent given, must be explained in a child-friendly manner, allowing children and parents to make informed choices. To create a child-friendly privacy policy, it is important to understand the age groups of the target audience. It is recommended to avoid long, complex texts and instead use videos or images to communicate the necessary information effectively. Regarding consent, the GDPR specifically addresses the issue of consent for children and adolescents in Article 8. For individuals under the age of 16, the processing of personal data is only lawful if parental con-

[188] Fankhauser & Fischer, 2017.
[189] Rücker, 2018.
[190] Oberlin & von Hoyningen-Huene, 2022.
[191] Husi-Stämpfli et al., 2024.

sent is given or if the child has obtained the consent of the parent. However, it is important to note that children still retain the right to transparency even when parental consent is involved. Therefore, both children and their parents must be provided with clear and accessible information about data protection. In practice, this may mean creating versions of the information tailored to both children and parents or providing a child-friendly version that is understandable by both. For very young children, the information should be accessible once they have the necessary understanding, potentially even before reaching the age of 16.[192]

For example, the problematic case of parents sharing photos of their kids on social media without prior consent from the person depicted is a violation of both their general personality rights and data protection laws. This holds true even if the initial publication was made with parental consent and the person in the photo has since turned 18. A case related to this issue was decided by the Frankfurt Regional Court on August 29, 2019 (LG Frankfurt a.M. Judgment from August 29, 2019, 2-03 O 454/18).[193] According to § 22 of the German Copyright Act (KunstUrhG), parental consent for the use of a child's image is generally binding and cannot be revoked, unlike consent under data protection laws, which can be withdrawn. However, this consent does not extend to an adult child once they reach the age of majority.

Once the individual depicted has come of age, they are entitled to make their own decision regarding the use of their image. If the person has not given consent, they have the right to exercise their personal autonomy and control over their image. This principle also applies to children who were depicted in images before they reached adulthood; upon reaching full legal capacity, they can decide whether they want their image to be used. Parental consent is only valid as long as the child is not yet of legal age. Once the child reaches adulthood, their own consent is required. Therefore, the legality of using a child's image cannot simply rely on the earlier consent granted by the parents, as the child, once of age, has the right to decide for themselves whether or not to allow the use of their image.[194]

Data breaches represent a significant and often underestimated risk to the online safety and privacy of children and adolescents. Under the General Data

[192] More information and also a great example of a child friendly privacy policy under: https://haerting.ch/wissen/datenschutzerklaerung-fuer-kinder/.

[193] https://www.rv.hessenrecht.hessen.de/bshe/document/LARE190035863.

[194] Gessner Legal, n. d.

Protection Regulation (GDPR)[195], Articles 33 and 34 set out clear obligations for data controllers and processors to notify authorities and affected individuals in the event of a personal data breach. Under Article 33, data controllers must notify the relevant supervisory authority within 72 hours of becoming aware of a breach, unless the breach is unlikely to result in a risk to individuals' rights and freedoms. When it comes to children's data, the threshold for risk assessment must be carefully considered, given the potential for long-lasting psychological, social, or even physical harm.

Article 34 requires that if a breach is likely to result in a high risk to the rights and freedoms of natural persons, the data controller must also communicate the breach directly to the affected individuals, meaning children and their guardians, without undue delay. This notification should be clear, transparent, and provide guidance on how to mitigate possible adverse effects.

Despite these legal protections, enforcement and practical application remain challenging. Data breaches involving minors often go unnoticed or are inadequately addressed, partly because children may not be aware that their data has been compromised or may not have the means to respond effectively. Moreover, some platforms fail to implement sufficient technical and organizational measures to prevent breaches or to comply fully with GDPR requirements. To strengthen protections for children online, it is essential not only to ensure rigorous compliance with Articles 33 and 34 but also to promote awareness among children, parents, educators, and service providers about the risks related to data sharing and breaches. This involves developing child-friendly communication strategies and providing accessible tools for managing personal data.

B. Switzerland

Data protection law specifies the right to informational self-determination enshrined in Article 13, Paragraph 2 of the Federal Constitution. Its aim is to protect the privacy of individuals and define the obligations of those who process personal data, as well as the rights of the individuals whose data is being processed.[196]

[195] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[196] BJ, n. d.

The current Swiss Data Protection Act (DSG) faces a significant gap when it comes to the protection of children's data. While the law outlines general guidelines for processing personal data, it fails to make a clear distinction between data belonging to children versus adults. According to the DSG, data processing is permitted as long as certain key principles are followed, but these principles apply equally to data from children and adults. This one-size-fits-all approach is where the problem begins. For children who are not yet of legal age to make decisions on their own, the responsibility for managing their personal data falls squarely on their parents or legal guardians. These guardians, however, are often not adequately equipped or informed to navigate the complex landscape of data privacy, making it difficult for them to protect their child's information effectively. The legal guardians need to be capable of making informed decisions about the privacy and security of their children's data. In reality, this expectation can be overwhelming for parents who may not have the expertise or time to fully understand the implications of data sharing in the digital age. While parents are crucial to the process, expecting them to shoulder this responsibility without sufficient support can leave significant gaps in the protection of children's personal data. This raises an important issue: the DSG does not sufficiently address the unique risks children face in the digital world. Unlike adults, children often do not fully comprehend the consequences of sharing their personal information online, nor do they understand the long-term impact that data processing might have on their lives. The act's reliance on parents to manage data protection without providing clear, child-specific guidelines places a heavy burden on families, potentially compromising the very privacy rights it seeks to safeguard.[197]

In addition, there is no specific regulation in Switzerland for the phenomenon of kidfluencers, as has already been established in France.[198] Kidfluencers are children who actively act as brand ambassadors or influencers on social media, and this has now increased in a way that goes far beyond so-called 'sharenting' where parents share their children's content. While sharenting mainly involves sharing private moments with Kidfluencers, the child is literally marketed and turned into a brand. This trend affects all areas of children's lives. Not only is what they do shown, but also what they consume, such as the toys they use, the places where they spend their holidays, and even their eating habits. This

[197] More about this under: https://datenschutz.law/ratgeber/datenschutz-bei-kindern-und-jugendlichen.

[198] The European Commission is already considering whether all countries in Europe should implement a law like this: https://www.europarl.europa.eu/doceo/document/E-9-2023-000496_EN.html.

extensive encroachment on children's personal rights is considerable and goes far beyond previous forms of private media use.[199]

C. Excursus: USA

a. COPPA[200]

The Children's Online Privacy Protection Act (COPPA) of 1998, which aims to prevent unfair or deceptive practices related to the collection, use, and disclosure of personal information from children online.[201] It applies to operators of websites or online services that collect personal information from children under 13 years old and mandates certain measures to protect children's privacy.[202]

16 CFR 312.2 defines terms used in the regulations. A "child" is someone under the age of 13. "Collection" of personal information includes not just requesting information directly but also enabling its publication or tracking online activities. Other terms, such as "parent" (including legal guardians), "operator" (the person running a website or online service), and "personal information" (including name, address, email, and other identifiers), are also defined to ensure operators understand the scope of these terms.

In section 312.3, the law outlines the general requirements for operators of websites or online services directed to children or those who knowingly collect personal information from children. These include:

- Operators must clearly state what information they collect from children, how they use it, and their disclosure practices.
- Parental consent must be obtained before collecting or disclosing personal information.
- Parents must have the ability to review their child's information and refuse its use.
- Operators cannot require children to disclose more information than necessary to participate in a game or activity.
- Measures must be in place to protect the security and confidentiality of collected information.

[199] Netzwerk Kinderrechte Schweiz, 2023.

[200] For a full overview see Husi-Stämpfli, 2025.

[201] Federal Trade Commission, 2023.

[202] Federal Trade Commission, 2023.

The regulations outlined in section 312.4 focus on the requirements for operators to provide notice and obtain verifiable parental consent before collecting, using, or disclosing personal information from children. Operators must ensure that the notice is clear, complete, and understandable, without any unrelated or confusing content. They must make reasonable efforts, based on available technology, to ensure that the parent receives direct notice of the collection, use, or disclosure of their child's personal information, including any significant changes to these practices. The direct notice must include details such as the child's or parent's contact information, the need for parental consent, the type of information collected, and how the parent can provide consent. Operators are also required to post a prominent link to an online notice on their websites or online services, which explains their practices regarding children's personal information. This notice must provide information about the operator's identity, the type of data collected, how the data is used, and the parent's ability to review, delete, or refuse further data collection.[203]

Regarding parental consent, operators must obtain verifiable consent from the parent before collecting, using, or disclosing children's personal information. They must give parents the option to consent only to certain aspects of data collection, such as allowing the collection without permitting disclosure to third parties. The regulations outline various methods for obtaining this consent, including forms, telephone calls, and video conferences. Safe harbor programs may approve alternative methods of obtaining consent.[204] There are exceptions to the requirement for prior consent, such as when collecting online contact information solely to provide notice or respond to a specific request from the child. In these cases, the information must not be used for any other purpose. Parents have the right to review their child's personal information upon request, and operators must make it easy for parents to refuse further data collection or have the information deleted.[205] The operator is also prohibited from conditioning a child's participation in an activity on the disclosure of more personal information than is necessary.[206]

Finally, operators must ensure the confidentiality, security, and integrity of children's personal information. They must take reasonable steps to protect the data and only share it with third parties who can ensure similar protections.[207]

[203] Federal Trade Commission, 2023, COPPA, Section 312.4

[204] Federal Trade Commission, 2023, COPPA, Section 312.5

[205] Federal Trade Commission, 2023, COPPA, Section 312.6

[206] Federal Trade Commission, 2023, COPPA, Section 312.7

[207] Federal Trade Commission, 2023, COPPA, Section 312.8

b. The proposed Kids Off Social Media Act

The Kids Off Social Media Act[208], co-sponsored by Senators Brian Schatz, Chris Murphy, Ted Cruz, and Katie Britt, represents a significant federal legislative effort to strengthen online protections for minors. Introduced as part of a broader initiative to address the mental health crisis among American youth, the bill seeks to establish stricter age-based limitations and content restrictions on social media platforms. The bill prohibits social media platforms from knowingly allowing children under 13 to create or maintain accounts. Existing accounts and any personal data collected from child users must be deleted. Platforms are also generally barred from using automated systems to recommend or promote content based on personal data from users under 17. Enforcement is assigned to the Federal Trade Commission (FTC), with states authorized to bring civil actions against platforms whose violations harm their residents. Additionally, schools receiving discounted telecommunications services under the E-Rate program must enforce policies that prevent access to social media via these supported services, networks, and devices, using blocking or filtering technologies. Schools failing to make good-faith efforts to comply and address violations are required to reimburse E-Rate funds. They must also submit their internet safety policies to the Federal Communications Commission (FCC) for public posting. The bill defines social media platforms as public-facing sites primarily serving as forums for user-generated content, explicitly excluding platforms focused mainly on videoconferencing, email, or educational services.[209]

On February 5, 2025, the bill, formally known as S. 278, was advanced by the Senate Committee on Commerce, Science, and Transportation, marking a critical step toward full Senate consideration. Despite gaining bipartisan support and public momentum, the bill faces vocal opposition from digital rights advocates, including organizations such as the Electronic Frontier Foundation (EFF) and Center for Democracy & Technology (CDT). Critics argue that certain provisions may pose constitutional challenges, particularly in relation to the First Amendment, privacy rights, and the feasibility of large-scale age verification. Legal challenges are anticipated should the legislation pass, especially concerning the potential chilling effect on lawful expression and access to information for teens, as well as broader implications for online anonymity and data collection practices.[210]

[208] US Congress, 2025.

[209] US Congress, 2025.

[210] Crespo et al., 2025.

c. Proposed and failed U.S. Kids Online Safety Act (KOSA)

The Kids Online Safety Act (KOSA) is a bipartisan bill introduced in the United States Congress in 2022, aimed at enhancing the protection of minors on online platforms. The legislation seeks to establish a duty of care for social media companies to prevent and mitigate harms to users under the age of 17. Key provisions of KOSA include the requirement for platforms to act in the best interests of minors by taking steps to limit exposure to harmful content such as material related to self-harm, eating disorders, substance abuse, and sexual exploitation. Additionally, the bill mandates that platforms provide parents and guardians with tools to supervise and manage their children's online activities, including privacy controls and screen time settings. Another critical aspect of KOSA is its focus on algorithmic transparency. Social media companies would be required to disclose how their recommendation algorithms operate and allow users to opt out of algorithmically generated content to reduce risks of addictive behavior and exposure to harmful material.[211]

The bill passed the Senate with overwhelming bipartisan support in July 2024 but subsequently stalled in the House of Representatives due to concerns about enforcement challenges and potential implications for free speech. Critics, including civil liberties organizations such as the Electronic Frontier Foundation, have raised concerns that the bill's provisions could result in over-censorship and restrict access to important information, particularly for marginalized communities.[212] Proponents argue that KOSA represents a necessary step to hold technology companies accountable and protect children from documented harms linked to social media use, including increased rates of anxiety, depression,[213] and exposure to dangerous content.[214]

Although the KOSA secured broad bipartisan approval in the Senate in July 2024, its progress stalled in the House of Representatives amid debates over the scope of regulatory authority and significant resistance from the tech industry. As the legislative session concluded without a final vote, the bill ultimately failed to pass in 2024. Nevertheless, Senator Richard Blumenthal has publicly affirmed his intention to reintroduce the measure in the 119th Congress which he has already done.[215] Given the arrival of a new congressional

[211] U.S. Congress, 2022.

[212] Electronic Frontier Foundation, 2024

[213] Common Sense 2024 and Crespo et al., 2025.

[214] Oberlin et al., 2024.

[215] U.S. congress, 2025.

cohort and a change in presidential administration, the bill's future developments merit close attention.[216]

d. Proposed and failed Protecting Kids on Social Media Act

Introduced in April 2023, the Protecting Kids on Social Media Act, S. 1291 shared similarities with the above-mentioned KOSA but featured some vague language and notable differences. The bill would have required social media platforms to take "reasonable steps" to verify user ages and prevent children under 13 from accessing these services. For minors aged 13 and above, platforms would need to obtain explicit parental or guardian consent before account creation and would be prohibited from using algorithmic recommendation systems for users under 18. The bill also proposed a voluntary pilot program through the Department of Commerce to develop secure digital ID credentials for age verification. Enforcement powers would lie with the Federal Trade Commission and state attorneys general. Despite these measures, S. 1291 did not pass in the 118th Congress, though its core provisions may resurface in future legislative initiatives.[217]

D. Regulatory Gaps

The COPPA establishes important safeguards to protect the personal information of children under 13 years old on websites and online services. It requires operators to obtain verifiable parental consent before collecting, using, or disclosing children's data and mandates transparency regarding data collection practices. However, COPPA has several regulatory gaps. One major limitation is the lack of robust, standardized age verification methods, which allows many platforms to rely on easily falsified self-reported ages. Although newer legislative proposals like the Protecting Kids on Social Media Act and the Kids Off Social Media Act aim to introduce stricter age verification requirements, including voluntary digital ID pilot programs, there is still no comprehensive federal standard to reliably verify users' ages, leaving enforcement difficult. Additionally, COPPA's scope is limited because it only applies to websites and services specifically directed at children or those who knowingly collect data from children. Many platforms that attract underage users without targeting them directly fall outside COPPA's reach. Furthermore, some proposed laws exclude categories such as videoconferencing, email, or educational platforms, creating regulatory blind spots for significant areas of children's online interactions.

[216] Crespo et al., 2025.

[217] Crespo et al., 2025.

Another critical gap concerns algorithmic content recommendations. While COPPA regulates data collection, it does not address how social media platforms use algorithms to suggest content, which can expose minors to harmful or addictive material. Recent bills like KOSA and the Kids Off Social Media Act seek to restrict the use of such algorithms for users under 17 or 18 and require greater transparency and opt-out options. However, there is currently no effective legal framework governing the opaque functioning of these algorithms.

Parental consent and control also present challenges. Although COPPA requires verifiable parental consent, the consent mechanisms are often cumbersome and can be circumvented or ignored in practice. Proposed legislation expands parental tools for supervision and control but faces enforcement difficulties, especially in balancing parents' oversight with children's privacy and autonomy. Enforcement itself remains a significant issue. COPPA enforcement is primarily handled by the Federal Trade Commission, which has limited resources and faces technological complexities in monitoring compliance. Newer bills empower state attorneys general and allow civil actions, but the fragmented jurisdictional landscape and the technical demands of enforcement create barriers to consistent oversight. Provisions that require schools to block social media access through federally subsidized networks add responsibilities to educational institutions, many of which may lack the capacity to effectively enforce these rules.

Balancing child safety with free expression is another complex challenge. Critics of proposed legislation warn that broad content restrictions could lead to over-censorship, limiting lawful speech and access to important information, especially for marginalized youth who rely on online platforms for support and community. Laws have struggled to find a clear balance between protecting children from harmful content and upholding First Amendment rights.

Finally, while COPPA requires data minimization enforcement around limiting secondary uses of data, such as targeted advertising, remains weak. Proposed legislation strengthens these protections but largely depends on platforms' good faith and faces challenges in auditing complex data ecosystems.

3. Regulatory approaches to protect minors on the internet from a Digital Law Perspective

In an increasingly digital world, children's personal data is constantly at risk—especially on social media platforms where privacy boundaries are easily blurred. As minors engage with online content, their data is often collected, analyzed, and shared in ways that they (and sometimes even their guardians) may not fully understand. Recognizing the vulnerability of children in the digital sphere, various legal frameworks such as the European Union's General Data Protection Regulation (GDPR), Switzerland's Data Protection Act (DSG), and the U.S. Children's Online Privacy Protection Act (COPPA) aim to protect minors' rights and privacy. However, the level and specificity of protection vary significantly across jurisdictions. We explore the challenges and gaps in regulating children's data protection across the EU, Switzerland, and the United States, with a particular focus on social media use, the role of parental consent, and the emerging phenomenon of "kidfluencers." It highlights how existing legal instruments address or fail to adequately address the complex realities of children's digital lives.

A. EU

a. Digital Service Act (DSA)

The Digital Service Act[218], which came into force in the EU on February 17, 2024, aims to create a safe, predictable and trustworthy online environment[219] that promotes innovation[220] while effectively protecting fundamental rights, such as consumer protection, as enshrined in the Charter. The legislation has three main objectives: enhancing consumer protection online, establishing a transparent and efficient framework for online platforms, and fostering innovation and competitiveness within the internal market.[221] A particular focus is placed on regulating large online services[222], especially platforms[223] like Instagram and search engines like Google, which have a high number of users within the EU.[224] The reason is because very large online platforms and very large online search engines can be used in a way that strongly influences

[218] European Parliament & Council, 2022.

[219] European Parliament & Council, 2022, Recital 3.

[220] European Parliament & Council, 2022, Recital 4.

[221] See European Commission, n. d.-a

[222] European Parliament & Council, 2022, Recital75.

[223] Given the importance of very large online platforms, due to their reach, in particular as expressed in their number of recipients of the service, in facilitating public debate, economic transactions and the dissemination to the public of information, opinions and ideas and in influencing how recipients obtain and communicate information online, it is necessary to impose specific obligations on the providers of those platforms, in addition to the obligations to all online platforms (European Parliament & Council, 2022, Recital 75).

[224] See: European Commission, 2023.

safety online and the shaping of public opinion and discourse, as well as online trade.[225]

The DSA acknowledges the special vulnerability of children and adolescents in the digital space.[226] Platforms primarily used by minors are required to design their terms and conditions and usage restrictions in a way that is easily understandable for this age group. Additionally, the right to information must be made accessible and comprehensible for young people. Furthermore, the DSA obliges platforms to implement appropriate technical and organizational measures to protect the privacy and safety of minors. Furthermore, platforms are not allowed to display targeted advertising to minors based on personal data. This requirement ensures that their data is not used for commercial purposes.[227]

Since large online platforms are primarily designed to optimize or benefit their often advertising-driven business models, which can raise societal concerns, they must assess the systemic risks arising from the design, functioning, and use of their services. Additionally, they should consider potential misuse by service recipients and take appropriate mitigating measures, ensuring respect for fundamental rights.[228] Providers of very large online platforms and search engines should conduct a thorough assessment of four categories of systemic risks. The first category addresses the risks related to the spread of illegal content, such as child sexual abuse material, illegal hate speech, or other forms of misuse for criminal activities. Such activities may pose significant systemic risks when access to illegal content can rapidly and widely spread through accounts with extensive reach or other amplification methods. Providers of these platforms and search engines must evaluate the risk of disseminating illegal content, regardless of whether it violates their terms and conditions. This assessment does not affect the personal responsibility of the service recipients of very large online platforms or the owners of websites indexed by these search engines, who remain accountable under applicable law for the legality of their activities.[229] The second, for this publication-relevant category concerns the actual or foreseeable impact of the service on the exercise of fundamental rights, as protected by the Charter, including but not limited to human dignity, freedom of expression and of information, including media freedom and pluralism, the right to private life, data protection, the right to non-dis-

[225] European Parliament & Council, 2022, Recital 79.

[226] See European Parliament & Council, 2022, Recital 81, 83, 89.

[227] Husi-Stämpfli et al., 2024.

[228] European Parliament & Council, 2022, Recital 79.

[229] European Parliament & Council, 2022, Recital 80.

crimination, the rights of the child, and consumer protection. Such risks may arise, for example, in relation to the design of the algorithmic systems used by the very large online platform or by the very large online search engine or the misuse of their service through the submission of abusive notices or other methods for silencing speech or hampering competition. When assessing risks to the rights of the child, providers of very large online platforms and of very large online search engines should consider, for example, how easy it is for minors to understand the design and functioning of the service, as well as how minors can be exposed through their service to content that may impair minors' health, physical, mental, and moral development. Such risks may arise, for example, in relation to the design of online interfaces that intentionally or unintentionally exploit the weaknesses and inexperience of minors or that may cause addictive behavior.[230] The fourth risk relevant for this publication arises from concerns related to the design, functioning, or use of very large online platforms and search engines, including through manipulation, which can have a tangible or foreseeable negative impact on public health, the protection of minors, or serious consequences for an individual's physical and mental well-being, including gender-based violence. This risk may also include coordinated disinformation campaigns related to public health or online interface designs that encourage behavioral addictions among users.[231]

The European Commission has just in 2024 launched a call for evidence to gather feedback for upcoming guidelines aimed at protecting minors online, as required by the DSA. These guidelines will advise online platforms on how to implement strong privacy, safety, and security measures for minors. The Commission is seeking input on the scope, approach, and best practices related to mitigating risks minors face online, encouraging stakeholders to contribute scientific reports and research. Platforms are expected to prioritize the rights and best interests of children by adopting a risk-based approach, conducting impact assessments, and implementing appropriate mitigation measures.[232]

Another important topic that puts children in danger in the digital space is the potential for generative AI (GenAI) to create harmful content, including AI-generated child pornography.[233] While the Digital Services Act (DSA) sets regulations for intermediary services, its applicability to GenAI is unclear. The DSA primarily covers conduit, caching, and hosting services, but GenAI doesn't fit

[230] European Parliament & Council, 2022, Recital 81.

[231] European Parliament & Council, 2022, Recital 83.

[232] European Commission, 2024.

[233] Oberlin/von Hoyningen-Huene, 2025, August; Oberlin/von Hoyningen-Huene 2025, November.

neatly into this model, as it produces new content rather than simply storing and sharing user input. Furthermore, the DSA's limited application to closed groups on messaging services and its lack of coverage for hybrid platforms that integrate GenAI leave significant gaps in regulation. For platforms that use GenAI tools, such as social media, gaming, and search engines, the resulting content may be modified by AI and shared with users, which raises questions about the platform's liability. While the DSA does include obligations for content moderation, such as notice and action procedures and an appeals system, these provisions may not fully address the complex nature of AI-generated content. This gap in regulation creates an additional risk of AI being used to generate harmful and illegal content, including child sexual abuse material, and calls for further clarification and guidance on how the DSA should apply to such technologies.[234]

b. Artificial Intelligence Act (AI-Act)

The impacts of AI-generated imagery and content are still not fully understood, and it is evident that legal regulation has struggled to keep pace with the rapid advancements in this field.[235] However, the EU has taken significant steps toward addressing this issue through various legislative initiatives, with one of the most important being the Artificial Intelligence Act (AI Act).[236] This regulation is a key part of the EU's digital strategy and aims to establish a unified legal framework for AI, ensuring the promotion of innovation[237] while maintaining a high level of protection for health, safety, and fundamental rights, as enshrined in the EU Charter of Fundamental Rights.[238] In April 2021, the European Commission introduced the first regulatory framework for AI, built on a risk-based approach.[239] This approach classifies AI systems based on the potential risks they pose to individuals, with different regulatory requirements[240] and obligations[241] based on the assessed risk level. The overarching goal of the European Parliament is to ensure that AI systems in the

[234] See Hogan Lovells, 2024.
[235] Kunz et al., 2024.
[236] European Parliament & Council, 2024, AI Act.
[237] European Parliament & Council, 2024, Recital 25 AI Act and Chapter VI of the AI Act on measures to promote innovation.
[238] European Parliament & Council, 2024, Recital 1, AI Act.
[239] European Parliament & Council, 2024, Recital 26, AI Act.
[240] European Parliament & Council, 2024, Section 2, AI Act.
[241] European Parliament & Council, 2024, Art. 26, AI Act.

EU are responsible[242], transparent[243], accountable, and non-discriminatory.[244] Additionally, AI should be subject to human oversight, rather than fully automated, to avoid harmful consequences or content. The AI Act was approved by the European Parliament in March 2024, with the Council giving its approval in May 2024. The regulation will be fully applicable 24 months after its entry into force, with some provisions coming into effect earlier.[245]

The AI Act establishes distinct obligations for providers and users of AI systems depending on the risk level of the technology. For instance, AI systems deemed to pose an unacceptable or excessively high risk to individuals will be banned, although exceptions exist for certain law enforcement applications.[246] Specifically, real-time identification systems may only be used in serious cases, and post-identification systems, where identification occurs with a delay, are permissible in the investigation of serious crimes but require prior judicial approval.[247] This could be relevant for the prosecution of suspected cases of sexual offenses with children. This framework also has specific provisions for the regulation of generative AI, which is not classified as high-risk but must still meet transparency requirements[248] and comply with EU copyright laws.[249] This includes disclosing that content was generated by AI[250] and preventing the creation of illegal material, such as child sexual abuse material.[251]

The regulation sets out different risk categories for AI systems, aiming to classify them according to the severity of the potential risks they pose. Some AI systems, which might pose a threat to the safety or fundamental rights of individuals, are categorized as high-risk.[252] These include AI systems used in sectors such as aviation, automotive, medical devices, and toys, which fall under EU safety regulations. Other high-risk AI systems are in fields like critical infrastructure, education, employment, law enforcement and border control, and are required to be registered in a dedicated EU database.[253] These high-risk systems must undergo rigorous evaluation before their market introduc-

[242] European Parliament & Council, 2024, Recital 34, AI Act.

[243] European Parliament & Council, 2024, Art. 50 and Recital 9, AI Act.

[244] European Parliament & Council, 2024, Recital 7, AI Act.

[245] Oberlin & von Hoyningen-Huene, 2025, August.

[246] European Parliament & Council, 2024, Art. 5 (1) h and Art. 5 (2), AI Act.

[247] European Parliament & Council, 2024, Art. 5 (3) [1], AI Act.

[248] European Parliament & Council, 2024, Art. 53 (1) a, b and d, AI Act.

[249] European Parliament & Council, 2024, Art. 53 (1) c, AI Act.

[250] European Parliament & Council, 2024, Art. 50 (2), AI Act.

[251] Oberlin & von Hoyningen-Huene, 2025, August.

[252] European Parliament & Council, 2024, Chapter III, AI Act.

[253] European Parliament & Council, 2024, Art. 6 and Annex III, AI Act.

tion and throughout their lifecycle.[254] Users of these systems have the right to file complaints with national authorities. Despite its focus on high-risk systems, the AI Act also addresses generative AI, which has become a major concern due to its potential for generating harmful content.[255] While not categorized as high-risk, only some specific generative AI must still comply with transparency[256] and EU copyright regulations.[257] For example, generative models must clearly disclose that content has been created by AI, and they must prevent the generation of illegal content.[258] This includes the risk of generative AI creating illicit material like deepfakes or CSAM. One significant regulatory challenge in relation to AI-generated CSAM is the categorization of such technologies under the AI Act. AI systems capable of generating child pornography are not classified as high-risk, even though they pose significant societal risks. The regulation specifically addresses high-risk systems in Article 6 and Annex III, but generative AI, which can produce illegal material, does not fall under this category, despite its potential to cause harm. This is a clear gap in the regulation, particularly when considering the fact that the creation and distribution of CSAM is a violation of the fundamental rights of every individual in the EU.[259] Recital 48 of the AI Act underscores the importance of protecting fundamental rights, particularly when AI systems are classified as high-risk. These rights include human dignity, privacy, freedom of expression, consumer protection, and the right to education. Special protection is also afforded to children, whose rights are enshrined in both the EU Charter of Fundamental Rights and the UN Convention on the Rights of the Child.[260] The AI Act mandates that high-risk AI systems be assessed for their potential impact on these rights and be subject to oversight to ensure compliance with EU law.[261]

Generative AI models, such as large general-purpose AI (GPAI) models, are of particular concern due to their widespread applicability and their ability to create harmful content. These models can perform a variety of tasks, from text generation to image and video creation. The AI Act aims to regulate these models by providing specific provisions[262] for GPAI models, which are defined as those that can be used across multiple applications and trained on vast

[254] European Parliament & Council, 2024, Art. 9 (2) and Art. 27, AI Act.

[255] European Parliament & Council, 2024 Chapter V, AI Act.

[256] European Parliament & Council, 2024, Art. 50, AI Act; Feiler & Forgó, 2024, p. 17.

[257] European Parliament & Council, 2024, Art. 53 (1) c, AI Act.

[258] European Parliament & Council, 2024, Art. 50 (2), AI Act.

[259] Oberlin & von Hoyningen-Huene, 2025, August.

[260] European Parliament & Council, 2024, Recital 48, AI Act.

[261] Oberlin & von Hoyningen-Huene, 2025, August.

[262] European Parliament & Council, 2024, Art. 50, AI Act.

amounts of data.[263] With this being said, the AI-Act acknowledges that GPAI models can pose systemic risks due to their power and versatility.[264] For example, models capable of generating text, images, audio, or videos could potentially be used to create harmful or illegal content, including deepfakes[265] or CSAM. The AI Act requires that providers of these models assess and mitigate the risks associated with their use, ensuring that their systems comply with the regulation's transparency and safety requirements. The lack of categorization for generative AI applications that can produce illegal content, such as CSAM, is a serious concern. Although these systems are not classified as high-risk, they still need to be evaluated thoroughly and adequately documented. In practice, the law enforcement community has already faced challenges, with such applications being widely available on the dark web. Once these technologies are circulating on decentralized platforms, they become exceedingly difficult to track and control.

B. Germany; The Netzwerkdurchsetzungsgesetz (NetzDG)

The Netzwerkdurchsetzungsgesetz (NetzDG)[266], effective since October 1, 2017, establishes binding compliance requirements for social network providers aimed at combating hate speech and other unlawful content on their platforms. The law targets telemedia service providers operating internet platforms where users can share content publicly or with other users. Exemptions apply to email and messenger services, professional networks, specialized portals, sales platforms, and social networks with fewer than two million registered users. Under NetzDG, social networks must implement a complaint procedure enabling users to report unlawful content easily and at any time. Providers must be promptly notified of complaints and are required to review them without delay. Clearly unlawful content must be removed or blocked within 24 hours, while other unlawful content must be addressed within seven days. Deleted content must be preserved for ten weeks and documented in accordance with the scope of Directives 2000/31/EC and 2010/13/EU. Providers must inform both complainants and content providers of decisions made regarding the complaints. Non-compliance with NetzDG provisions, including failure to timely produce and publish required reports, inadequacies in complaint handling procedures, insufficient oversight, or failure to

[263] European Parliament & Council, 2024, Art. 51, AI-Act.

[264] European Parliament & Council, 2024, Art. 51 (1) a and Art. 51 (2), AI-Act.

[265] Legal Definition Deepfakes in European Parliament & Council, 2024, Art. 3 No. 60, AI-Act.

[266] Netzwerkdurchsetzungsgesetz vom 1. September 2017 (BGBl. I S. 3352), das zuletzt durch Artikel 29 des Gesetzes vom 6. Mai 2024 (BGBl. 2024 I Nr. 149) geändert worden ist.

respond to information requests, constitutes an administrative offense punishable by fines up to five million euros. This applies even when violations occur outside Germany. Providers receiving over 1,000 complaints annually must publish biannual reports in German, both in the Federal Gazette and on their websites. These reports must detail measures taken to prevent unlawful acts, the complaint submission process, criteria for content removal or blocking, and average processing times. The law, however, leaves unresolved the crucial question of the criteria social networks should apply to distinguish punishable content from protected free speech, raising constitutional concerns about potential over-censorship. To address this, a new regulatory body staffed by approximately 50 employees from the Federal Office of Justice (BfJ) has been established to review difficult cases referred by private providers. Facebook has expressed reservations about NetzDG, citing uncertainty in interpreting legal obligations and criticizing the short three-month implementation period for complaint handling measures. To comply, Facebook created a dedicated 500-person team.[267]

The Netzwerkdurchsetzungsgesetz (NetzDG), despite its good intentions, has faced significant criticism and challenges. A major issue was that platform operators had to decide whether reported content was unlawful under strict time constraints, increasing the risk of errors. This pressure led to concerns about "overblocking," where platforms remove more content than legally necessary to avoid fines, potentially infringing on freedom of expression. Although evidence of systematic overblocking is inconclusive, the risk remains. Additionally, some platforms complicated the reporting process through "dark patterns," making it difficult for users to submit complaints, which reduced the number of reports and hindered enforcement. On the positive side, NetzDG increased platform accountability and encouraged proactive removal of illegal content. It mandated regular transparency reports, providing valuable data for oversight and research, and helped raise public awareness about online hate crimes. These developments influenced the European Union's Digital Services Act (DSA), which builds upon and improves many NetzDG principles. While NetzDG was flawed, it was a pioneering step in addressing digital hate crime. The lessons learned helped shape the DSA, which still requires further refinement to balance effective content moderation with protecting user freedoms.[268]

[267] Forum Verlag Herkert GmbH, 2017.

[268] Intersoft Consulting Services, 2025.

C. Switzerland – Preliminary Regulation Framework for Communication Platforms

On April 5, 2023, the Swiss Federal Council announced its intention to introduce stricter regulations for major communication platforms such as Google, Facebook, YouTube, and Twitter. The aim is to strengthen user rights in Switzerland and increase transparency in how these platforms operate. These platforms have become central to how people inform themselves and form opinions, yet they currently remain largely unregulated. The algorithms that determine which content users see are opaque, and users have limited means to contest the deletion of their content or the suspension of their accounts. To address these issues, the Federal Council has tasked the Federal Department of the Environment, Transport, Energy and Communications (UVEK), in cooperation with the Federal Office of Justice (BJ), with drafting a consultation proposal by the end of March 2024. The planned legislation will partly draw on the European Union's Digital Services Act for guidance. The proposal includes several key measures: major platforms will be required to designate a legal representative and a point of contact within Switzerland. Users whose content has been removed or whose accounts have been suspended will have the right to request a review directly from the platform. Additionally, an independent arbitration body is to be established in Switzerland, funded by the platforms, to handle such disputes. Transparency in advertising is another focal point. Platforms will be obliged to clearly label advertisements and disclose the main parameters used to deliver targeted ads. Furthermore, users should be able to easily report calls for hate, violent content, or threats. The platforms must investigate such reports and inform users of the outcome. However, the proposal does not foresee any extended state powers to intervene in content beyond what is currently allowed in the offline world, thus preserving freedom of expression.[269]

As of April 2025, the Swiss regulation of large communication platforms has once again been delayed. Although a draft proposal had already been postponed previously, it was most recently expected for release in autumn 2025, but that timeline was not met either. The Federal Office of Communications (BAKOM) confirmed that the consultation process ("Vernehmlassung") will no longer take place in 2025. According to BAKOM:

[269] Der Bundesrat, 2022.

> "The opening of the consultation will no longer take place this year. As this is a completely new law involving new legal questions, the revision process is taking some time."
> No updated timeline has been announced.

This delay comes shortly after it was revealed that Switzerland's AI regulation has also been postponed until at least later in the year of 2025.[270]

On 29 October 2025, the Federal Council opened the consultation process on a new federal law that aims to strengthen the rights of users in the digital space and oblige very large communication platforms and search engines to be more fair and transparent. Comments on the draft can be submitted until 16 February 2026. The law aims to create, for the first time, central regulations for digital services such as Facebook, X, TikTok, and Google, which, as communication infrastructures, have a significant influence on public opinion. The Federal Council is thus responding to the growing importance of a small number of international corporations whose privately set rules on content, access, and deletion are increasingly shaping democratic communication in Switzerland. The aim of the bill is to legally limit this influence, counteract abuse, and strengthen the rights of users. The law obliges very large communication platforms to provide procedures that enable users to easily report suspected illegal content, such as defamation, insults, or discrimination. When content is removed or accounts are blocked, the persons concerned must be informed, and the reasons for the decisions must be given. In addition, an internal complaints procedure and participation in out-of-court dispute resolution are provided for. The bill also contains far-reaching transparency requirements: advertising must be clearly labelled and addressed, recommendation systems must be disclosed, and the services concerned must maintain a publicly accessible advertising archive. Research institutions and authorities should also be given access to selected data in order to better investigate the functioning and social impact of these platforms. To ensure enforcement, the draft stipulates that providers not based in Switzerland must appoint a legal representative in Switzerland. The law is aimed exclusively at very large communication platforms and search engines, as these have a particularly strong influence on public debate and opinion-forming due to their reach. Services are considered very large if they are used by at least ten percent of the permanent resident population– currently around 900,000 people –on average each month. As part of the consultation process, participants are invited in partic-

[270] Steiger, 2024.

ular to comment on issues relating to the protection of minors and the design of the proposed reporting procedure.[271]

The current draft of the Federal Act on Communication Platforms and Search Engines (VE-KomPG) does not explicitly address the protection of minors. The draft refers exclusively to ‘users’ and thus does not address the specific needs and protection interests of minors, either in terms of terminology or substance. It would therefore be appropriate to include a separate section entitled ‘Minors’ accounts’ in the law, which regulates the obligations of platform operators when dealing with minors’ user accounts. Such a section could initially set general age restrictions and establish age verification mechanisms, thereby legally securing the age limits already provided for in the terms of use of many providers. A legal basis would enable the Confederation to adjust these age limits as necessary, for example, by raising the minimum age to 14. Furthermore, the legislator should specify the extent to which personalized advertising may be displayed to minors and the minimum age at which this is permissible. It would also be necessary to regulate how recommendation systems may be designed for underage users and, in particular, to prohibit profiling and algorithmic manipulation in connection with children and young people.

Furthermore, online platforms used by children should be required to implement appropriate and proportionate protective measures. This includes, in particular, providing user accounts with the highest privacy and security settings by default, which may be modified only by parents. Such safeguards are necessary to effectively ensure the protection of minors in the digital environment.The proposed reporting procedure under Article 4(1) of the draft law also has gaps in its content. According to the draft, users should be able to report content such as depictions of violence, defamation, slander, insults, threats, coercion, sexual harassment, public incitement to crime or violence, discrimination, and incitement to hatred. However, it is striking that other problematic content is not covered. Specifically, prohibited pornography, content that violates personal rights (in particular violations of the right to one’s own image), age-inappropriate content, unwanted or harmful content, and disinformation in the form of fake news or deepfakes are missing. These gaps significantly weaken the effectiveness of the reporting procedure and should be closed in the further legislative process.

Finally, Art. 20 VE-KomPG on risk assessment also needs to be supplemented. The draft does not contain an explicit obligation to assess the potential nega-

[271] Bundesamt für Kommunikation, 2025.

tive effects of platforms on the physical and mental health of underage users. In view of the empirically proven risks of digital platform use for children and young people, this omission is remarkable and should be urgently corrected in the further drafting of the law.

D. France – Kidfluencer Law & a new legislative proposal

On February 19, 2024, France passed the Children's Image Rights Law[272] to enhance the protection of children's privacy, particularly regarding the rights to their own image. This law targets the issue of "sharenting". Unlike other countries' regulations, this law specifically applies to parents or guardians and not to the controller. It complements previous laws, such as the "Child Influencer Law"[273], and builds on efforts to combat online risks like cyber harassment and access to inappropriate content. The Children's Image Rights Law underscores the right of children to privacy and control over their personal data. This includes rights to access, correct, erase, and object to the processing of their image. Parents or guardians can act on behalf of their children to request the deletion of any photos or videos shared without consent. The law also grants additional powers to the French data protection authority (CNIL), enabling it to take swift legal action when erasure requests are not honored or when children's rights are seriously infringed. The CNIL can now initiate summary proceedings to ensure faster resolution of such issues, including seeking interim injunctions to halt the publication of harmful content.[274] This French law represents a significant step forward in the protection of children's rights by explicitly including the right of minors to their own image in the civil code. Considering the growing phenomenon of 'sharenting' (the sharing of children's images on social media), this law aims to encourage parents to recognize the dangers of sharing children's images online. Previously, the protection of children's right to their own image was only indirectly regulated, but now the law obliges parents to consult with each other when publishing images and to take the child's consent into account. A judge can intervene in the event of disputes. This reform aims to protect children's privacy and protect them from potential dangers such as cyberbullying and the misuse of images.[275]

Moreover, on 26 January 2026, the French National Assembly adopted a legislative proposal that would generally prohibit children and adolescents under

[272] Assemblée nationale, 2024.

[273] Assemblée nationale, 2020.

[274] Dansac Le Clerc & Leportois, 2024.

[275] Quashie, 2024.

the age of 15 from accessing social networks. However, the law has not yet been finally adopted, as it must still be approved by the Senate, the second chamber of parliament. The draft provides that social networks would be inaccessible to users under 15, while online encyclopedias, educational and scientific services, and private messaging services would be exempt. Unlike earlier versions, the current text no longer allows use with parental consent. President Emmanuel Macron supports the initiative and aims for a swift implementation, ideally already by the next school year. It remains unclear, however, whether the measure would be compatible with EU law; a previous French attempt to introduce an age limit could not be applied for that reason. With such a measure, France would become one of the most restrictive countries in Europe, while several other states are also currently discussing or introducing similar age limits for social media.[276]

E. The Netherlands

On June 14, 2024, the Dutch government issued a decision regarding the implementation of the Wet bestuursrechtelijke aanpak online kinderpornografisch materiaal (Administrative Approach to Online Child Pornography Law). The law came into effect on July 1, 2024, except for certain provisions (Articles 9, 12, and 13, parts 4 and 5), which will take effect at a later date. Article 9 grants the Authority for Online Terrorist and Child Pornographic Material (ATKM)[277] the power to publicly disclose administrative fines imposed. Article 12 concerns the retention of child pornography material and associated personal data by the ATKM for proper execution of its duties, with rules on how this material can be used in criminal proceedings or administrative procedures. For the implementation of Articles 9 and 12, a general administrative regulation (amvb) is still being developed in consultation with the ATKM, after which it will be subject to public consultation and advice from the Advisory Division of the Council of State. Once finalized, these provisions will come into force. Additionally, a provision prohibiting child sex dolls (Article 253a of the Dutch Penal Code[278]) has been added through an amendment.[279] This part of the law, which will be notified before coming into force, will be activated once the notification process is complete.[280]

[276] Steinvorth, 2026.

[277] ATKM, n. d.

[278] The Netherlands, 2025.

[279] For more information see https://zoek.officielebekendmakingen.nl/kst-36377-14.html.

[280] Staatsblad van het Koninkrijk der Nederlanden, 2024.

F. Spain

a. The General Audiovisual Communication Law

The General Audiovisual Communication Law[281] transposes the European Audiovisual Media Services Directive (AVMSD)[282] and imposes obligations on video-sharing platforms like YouTube to protect minors. Platforms must implement age verification systems to prevent access to harmful content, including violence and pornography. They are also required to provide parental controls, classify content by age, and allow users to report inappropriate content. Influencers must adhere to these protections by classifying their content appropriately. The National Commission on Markets and Competition (CNMC)[283] will oversee compliance with these obligations.

b. Draft Organic Law for the Protection of Minors in Digital Environments

The Draft Organic Law for the Protection of Minors in Digital Environments[284] aims to create safer digital spaces for minors by amending the General Audiovisual Communication Law (LGCA).[285] The Draft requires platform service providers to include a visible link to the competent audiovisual authority's website and ensures that age verification systems are properly enforced to prevent minors from accessing harmful content. It also mandates default parental controls, aligns influencer regulations with those for audiovisual services, and imposes additional obligations on influencers regarding content restrictions for minors. The CNMC will oversee compliance and can sanction platforms that fail to implement effective age verification systems.[286]

G. Ireland – Online Safety Act

Coimisiún na Meán[287] Ireland's new media regulator for online safety has developed the Online Safety Code to ensure video-sharing platforms (VSP) pro-

[281] Spain, 2022.

[282] European Commission, n.d.-b.

[283] More information on the CNMC under https://www.cnmc.es/en.

[284] Draft accessed under https://technical-regulation-information-system.ec.europa.eu/en/notification/26278/text/I/EN.

[285] Spain, 2010.

[286] Villasante, 2025.

[287] More information about the Coimisiún na Meán under https://www.cnam.ie/.

tect users, especially children, from online harm.[288] Since Ireland is the home of the big Tech-Providers such as Google, Meta and Co, this law seems to be very important. The Online Safety Code aims to protect children and the general public from harmful and illegal content on video-sharing platforms (VSP).[289] It requires VSPS providers to implement protective measures, including content ratings, age verification, parental controls, and reporting systems. The Code covers various types of content, such as videos, user-generated content, and audiovisual commercial communications, with specific provisions to safeguard children from harmful content. Harmful content in this means is cyberbullying, self-harm, eating disorders, terrorism, and child sexual abuse material.

The Code was developed based on evidence from public consultations, research, and expert opinions, including input from Ireland's Youth Advisory Committee. Coimisiún na Meán will enforce the Code, investigating violations and taking action when necessary. Users can report harmful content, and the regulator provides resources to help parents protect children online. It also offers advice for online safety for adults, covering scams and privacy protection. The guidelines also mandate content moderation and compliance measures to ensure a safer online environment for all users.[290] Platforms must comply by November 2024 for general rules (Part A) and by July 2025 for detailed rules (Part B). Penalties for non-compliance can include fines up to €20 million.[291]

H. Greece

In March 2025, the Greek Ministry of Digital Governance published the *National Strategy for the Protection of Minors from Internet Addiction*, a policy and regulatory framework aimed at strengthening the protection of minors in

[288] Citizens Information, n. d.

[289] Article 1(1)(aa): 'video-sharing platform service' means a service as defined by Articles 56 and 57 of the Treaty on the Functioning of the European Union, where the principal purpose of the service or of a dissociable section thereof or an essential functionality of the service is devoted to providing programmes, user-generated videos, or both, to the general public, for which the video-sharing platform provider does not have editorial responsibility, in order to inform, entertain or educate, by means of electronic communications networks within the meaning of point (a) of Article 2 of Directive 2002/21/EC and the organisation of which is determined by the video-sharing platform provider, including by automatic means or algorithms in particular by displaying, tagging and sequencing, accessed under https://rm.coe.int/avmsdigest-2024-safe-screens-protecting-minors-online/1680b26ccf.

[290] Coimisiún na Meán, n. d.

[291] Citizens Information, n. d.

the digital environment, particularly in relation to social media use and excessive screen exposure. The strategy proceeds from the premise that digital transformation offers significant opportunities for education, access to information, and social participation, while simultaneously exposing children and adolescents to new risks, including problematic or addictive patterns of use, mental-health harms, cyberbullying, and the commercial exploitation of personal data. Particular emphasis is placed on algorithmic recommendation systems and personalized content used by large online platforms, which rely on profiling, automated content curation, and attention-capturing design features such as autoplay and infinite scrolling. These mechanisms are regarded as capable of influencing minors' behaviour and potentially fostering addictive usage patterns. Minors are therefore conceptualized as a particularly vulnerable user group whose rights and interests require enhanced legal and regulatory protection beyond parental supervision and existing forms of platform self-regulation. The strategy adopts a multi-layered approach combining regulatory intervention, technical safeguards, and educational measures. A central component is the establishment of a "digital age of consent" at 15 years for the use of social media and similar online services. Under the proposed framework, minors below this threshold would be permitted to use such services only with explicit parental consent, while service providers would be required to implement reliable age-verification mechanisms that go beyond mere self-declaration. This national measure is also intended to inform developments at the European Union level, with Greece advocating for the introduction of a harmonized digital age of consent across the EU and supporting amendments to the Digital Services Act as well as the development of a future "Digital Fairness Act." These initiatives aim to introduce clearer obligations for platforms to mitigate addictive design practices, restrict profiling of minors, and ensure that online services used by children are designed in a manner that prioritizes safety, privacy, and well-being by default. In addition, the strategy envisages a range of national measures intended to operationalize the protection of minors in practice. These include the introduction of mandatory parental-control functionalities on internet-enabled devices, the development of a state-supported digital identity and parental-control application for minors ("Kids Wallet"), public awareness campaigns directed at parents and young users, and the systematic collection of data on minors' internet use in Greece. Overall, the strategy constitutes a comprehensive regulatory and preventive framework designed to assign primary responsibility to digital service providers while simultaneously equipping parents and minors with information and tools. It thus serves as a policy blueprint intended to guide both domestic legislation and

European regulatory developments with the aim of creating a digital environment that effectively safeguards the rights, health, and well-being of minors.[292]

I. Austria

As Austrian media outlets reported in late January and early February 2026, the government is planning a “Social Media Order Act” (SOG) to be presented by the summer. Media Minister and Vice Chancellor Andreas Babler aims to use the law to better protect children and adolescents from harmful content on social media platforms and to place greater responsibility on platform operators. The core element of the proposal is an age restriction for social media use; the exact age limit has not yet been determined, though a possible ban for users under 14 has recently been discussed. The draft law is also expected to include sanctions for platforms that fail to comply, drawing on the EU Digital Services Act as a benchmark, which allows for substantial fines. Additional measures under consideration include greater transparency of algorithms and initiatives to strengthen young people’s media literacy. The draft is to be presented by the summer and then undergo consultation; if political agreement is reached quickly, it could be adopted by parliament in autumn 2026. The government is also preparing for the possibility that no rapid agreement on age limits for social media will be reached at the EU level.[293]

Already in late December 2025, Austrian government representatives had called on the European Commission to present a concrete legislative proposal introducing a minimum age for the use of algorithm-driven platforms across the EU. They warned of risks such as harmful content, radicalisation and mental-health impacts on minors, and indicated that Austria would pursue national legislation if a swift European solution failed to materialise.[294]

J. UK

a. Online Safety Act

The Online Safety Act 2023 introduced new laws to protect both children and adults online, placing significant responsibilities on social media platforms and

[292] Hellenic Republic, 2025.
[293] Weber, 2026.
[294] Bundesministerium Wohnen, Kunst, Kultur, Medien und Sport, 2025.

search services that are in scope of the act.[295] The Act requires these platforms to take measures to reduce the risks of illegal activities, remove illegal content, and protect users from harmful material. The strongest protections are aimed at children, including preventing access to harmful content and providing clear reporting mechanisms. For adult users, platforms must be more transparent and offer tools to control the content they encounter.[296] Key changes include senior accountability for safety, better moderation, easier reporting, and improved algorithm testing to reduce illegal content. For children, specific measures address online grooming, such as hiding profiles and locations from non-connected users and limiting direct messages. These steps aim to better protect users, especially children, from online harm.[297]

Ofcom[298], the independent regulator, oversees the enforcement of these duties, with the Act applying to various online services, including social media, video-sharing platforms, and search engines, regardless of their location if they impact UK users. The Act mandates platforms to assess risks, use age-assurance technologies, and ensure age-appropriate experiences for children. Platforms must also tackle illegal content like child abuse and hate speech while prioritizing the removal of suicide and self-harm content. New offenses, such as cyberflashing and intimate image abuse, have been introduced, and companies failing to comply face significant penalties. The Act also addresses the harmful effects of algorithms and misinformation, requiring platforms to assess their impact on user safety. The Act includes provisions to protect women and girls from online abuse, tackle mis- and disinformation, and ensures that platforms protect users from harmful online content effectively. Ofcom will implement the Act in phases, publishing guidance and codes of practice, and can impose substantial fines or criminal sanctions on non-compliant companies. Additionally, an independent review on pornography regulation is being conducted to adapt laws to the changing online landscape.[299]

The UK Parliament has given Ofcom 18 months from the passing of the Online Safety Act on 26 October 2023 to finalize its codes of practice and guidance on illegal harms and children's safety.[300] Ofcom can fine companies up to £18 mil-

[295] Information on which types of platforms are in scope of the Act can be found here: https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/guide-for-services/#who.

[296] Department for Science, Innovation and Technology, 2024.

[297] Ofcom, 2024.

[298] https://www.ofcom.org.uk/.

[299] Department for Science, Innovation and Technology, 2024.

[300] Ofcom, 2024.

lion or 10% of their qualifying worldwide revenue –whichever is greater. In very serious cases, Ofcom can also apply for a court order to block a site in the UK.[301]

Since the law's adoption, Ofcom has made significant progress. At the time of publication of this book, the first set of binding codes and guidance under the Online Safety Act is already in force, establishing clear obligations for tackling illegal content and safeguarding children online. Since July 2025, online services likely to be accessed by minors have been required to implement "highly effective" age-assurance systems and to conduct comprehensive children's risk assessments. These measures mark a major step toward operationalizing the UK's new digital safety framework and strengthening Ofcom's supervisory powers. Ofcom has already initiated enforcement actions, including investigations into adult-content providers suspected of failing to comply with age-verification duties. Further consultations are underway on additional regulatory elements, such as transparency requirements, the categorization of services, and the future fee regime for regulated entities. Although the Online Safety Act is now partially operational, its full implementation—including the finalization of all remaining codes and the expansion of Ofcom's supervisory framework—is expected to continue throughout 2026. In its current phase, the Act marks a decisive shift toward a proactive, risk-based model of online regulation aimed at making digital platforms safer for children and users across the UK.[302]

b. UK Age-Appropriate Design Code

The Children's Code, also known as the Age-Appropriate Design Code, sets out 15 standards that online services must follow to protect children's data and comply with data protection laws. It applies to a wide range of online services, including apps, games, connected toys and devices, and news services. Importantly, the code covers any "information society services" (ISS) likely to be accessed by children under 18, even if children are not the primary target audience. ISS are broadly defined as for-profit online services provided electronically and on individual request, which include social media platforms, search engines, messaging services, online marketplaces, streaming services, and websites offering goods or services. The code applies to both UK-based and non-UK companies processing personal data of UK children. To comply with the code, service providers may need to map the personal data they col-

[301] Ofcom, 2024.
[302] Ofcom, 2024.

lect from children, implement age verification, disable geolocation tracking, avoid using nudging techniques to encourage excessive data sharing, and provide high privacy settings by default. While the code does not apply directly to schools or educational institutions, which are not considered ISS, its principles—especially the requirement to prioritize the best interests of children—align with schools' legal obligations under UK GDPR and the Data Protection Act 2018. Providers of educational technology (edtech) used in schools may fall under the code if their services meet the ISS criteria. The code does not apply to government or local authorities providing public services not offered for remuneration, including educational services. However, when authorities are involved in procuring edtech services for schools, they may assume the role of data controllers and must fulfill their data protection responsibilities, including conducting due diligence on service providers to ensure compliance with data protection laws.[303]

K. Regulatory Gaps

Beginning with an examination of the EU regulations, the DSA and the AI-Act represent significant legislative efforts aimed at creating a safer and more transparent digital environment, with particular attention to the protection of children and the regulation of online content. However, both frameworks exhibit notable regulatory gaps, especially concerning generative AI and the safeguarding of minors.

The DSA primarily targets intermediary services such as hosting, caching, and transmission providers. Generative AI, which creates entirely new content independently, does not clearly fall within the scope of the DSA. This results in a regulatory gap when it comes to monitoring and managing potentially harmful or illegal AI-generated content, such as AI-CSAM. Furthermore, the DSA provides limited coverage for closed messaging groups and hybrid platforms integrating AI-generated content, complicating oversight and control of problematic content in more private or less visible areas. The issue of platform liability for AI-driven content also remains ambiguous, as existing content moderation mechanisms may not be sufficient to address the complexities posed by AI-generated materials. While the European Commission is currently developing guidelines to enhance the protection of minors, these are still in progress and lack binding enforcement.

[303] Information Commissioner̀s Office, n.d. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/introduction-to-the-childrens-code/.

The AI Act adopts a risk-based approach, classifying AI systems according to the potential risks they pose. Yet, generative AI models, despite their considerable capacity to produce harmful outputs such as deepfakes or AI-generated CSAM, are not designated as high-risk systems. As a result, they remain outside the scope of the strictest regulatory requirements. While the Act introduces certain transparency obligations, such as the disclosure of AI-generated content, it does not establish explicit prohibitions or binding safeguards aimed at preventing the creation of illegal material. The proliferation of generative AI through decentralized platforms and the dark web further exposes critical regulatory gaps that the AI Act leaves largely unaddressed. Moreover, given that the Act will only become fully applicable two years after its entry into force, existing risks remain insufficiently regulated in the interim. Although the legislation explicitly references the protection of fundamental rights, including those of children, it fails to translate this recognition into concrete provisions designed to shield minors from AI-related harms. This reveals a structural regulatory gap that undermines the effectiveness of the Act's protective framework.[304]

Furthermore, selected additional European countries with advanced regulations in this area, as well as Switzerland, have been examined. The findings reveal a variety of regulatory frameworks addressing child protection and digital platform governance but also highlight significant gaps and challenges in fully addressing the risks posed by AI-generated content and online harms.

In Switzerland, the preliminary regulatory framework for major communication platforms aims to enhance user rights and transparency, partly inspired by the EU's Digital Services Act. However, despite these positive intentions, a key regulatory gap remains in the significant delay of the legislation, leaving Switzerland without concrete rules at a time when platform-related risks are escalating. Furthermore, the proposed framework does not extend state powers to intervene beyond offline content regulation, which may limit its effectiveness against emerging online harms specific to AI-driven content moderation and misinformation. The absence of a clear timeline for implementation and coordination with AI regulation also weakens the framework's potential impact.

France's Children's Image Rights Law represents a strong legal step toward protecting children's privacy, particularly regarding parental sharing of children's images, the so called "sharenting". Nonetheless, this law focuses nar-

[304] Oberlin & von Hoyningen-Huene, 2025, August.

rowly on image rights and parental responsibility, leaving broader issues such as AI-generated content affecting children or algorithmic harms largely unaddressed. While the CNIL has enhanced powers to enforce erasure of unauthorized content, the law lacks comprehensive provisions on platform accountability for content moderation or the prevention of exposure to harmful AI-driven content, which could undermine protection in digital environments.

The Netherlands' law tackling online child sexual abuse material empowers enforcement agencies and introduces administrative fines, but gaps remain in the delayed activation of certain provisions and the slow regulatory process surrounding the administrative regulations. Moreover, the law primarily targets CSAM and does not comprehensively regulate the wider scope of AI-generated harmful content, digital manipulation, or platform transparency concerning algorithmic decision-making. This narrow focus limits the law's ability to adapt to the growing complexity of AI and digital threats to children.

Spain's regulatory approach, including the General Audiovisual Communication Law and the Draft Organic Law for the Protection of Minors in Digital Environments, sets clear requirements for age verification, content classification, and parental controls. However, significant challenges remain regarding enforcement and platform compliance, especially as the draft law is not yet finalized. Additionally, the regulation primarily targets audiovisual content and does not yet fully address AI-specific risks such as deepfakes, AI-driven content personalization, or algorithmic bias. The effectiveness of these laws depends heavily on the capabilities and resources of regulatory bodies like the CNMC.

Ireland's Online Safety Act introduces a comprehensive Online Safety Code with mandatory protective measures for video-sharing platforms. While this is a robust framework, there are concerns about the law's scope and the speed of enforcement, as the implementation deadlines could delay effective protection. Furthermore, while the Code covers many types of harmful content, it does not explicitly tackle AI-generated content risks or the transparency of AI systems used by platforms, which are increasingly critical for safeguarding children online.

The UK's Online Safety Act 2023 is one of the most advanced frameworks, with detailed provisions on age assurance, content moderation, and platform accountability enforced by Ofcom with significant penalties. Yet, despite these strengths, there are still gaps related to the regulation of emerging AI technologies, such as generative AI's role in producing harmful content or misinformation. The Act's phased implementation means full protections will only come into effect over several years, and certain enforcement mechanisms, like blocking websites, face legal and technical challenges. Moreover, while the Age-Appropriate Design Code sets important standards for protecting children's data, it does not cover all online environments where children interact, such as informal or decentralized platforms.

4. Regulatory approaches to protect minors on the internet from a Criminal Law Perspective

The protection of minors in the digital environment has emerged as a critical legal and policy concern across Europe. The increasing prevalence of online risks, ranging from exposure to harmful content to criminal exploitation such as grooming, cyberbullying, and the dissemination of child sexual abuse material (CSAM), has prompted national and international legislators to adopt a multifaceted criminal law framework aimed at prevention, deterrence, and prosecution.

Several criminal law provisions at the national and international levels protect children in the digital sphere. These legal frameworks aim to safeguard minors from sexual exploitation, cyber grooming, hate speech, cyberbullying, and other digital threats.

A. International Regulation

a. Lanzarote Convention

The Lanzarote Convention, officially known as the Council of Europe Convention on the Protection of Children against Sexual Exploitation and Sexual Abuse (CETS No. 201), was adopted in 2007 in Lanzarote, Spain, and came into force in 2010. It is the most comprehensive international legal instrument dedicated to protecting children from all forms of sexual exploitation and abuse. To date, all 46 member states of the Council of Europe, as well as two non-member states, the Russian Federation and Tunisia, have ratified the Convention, with other countries like Morocco expressing interest in accession. The Convention covers a wide range of sexual offenses against children, including abuse within the "circle of trust", such as by family members or people in positions of authority, as well as commercial exploitation. It builds upon existing United Nations and Council of Europe standards and obliges states to criminalize such behaviors, implement preventive measures, protect child victims, and ensure the prosecution of offenders. A key element of the Convention is its emphasis on the best interests of the child, guiding all measures from prevention to prosecution. Prevention measures under the Convention include raising children's awareness of sexual exploitation risks and empowering them to protect themselves, screening and training adults who work with children, and monitoring intervention programs for convicted or potential sexual offenders. Protection measures encourage reporting suspicions of abuse, the establishment of telephone and online helplines, the provision of support programs and therapeutic care for victims and their families, and the creation of child-friendly judicial procedures that respect the safety, privacy, and dignity of child victims.[305]

Regarding prosecution, the Convention defines specific criminal offenses such as sexual abuse, child prostitution, child pornography, involvement of children in pornographic performances, corruption of children, and solicitation of children for sexual purposes (online grooming). It also requires countries to extend statutes of limitations for sexual offenses so that legal proceedings can begin even after victims reach adulthood. Moreover, the Convention includes the principle of extraterritoriality, allowing countries to prosecute their nationals for sexual crimes committed abroad. International cooperation is a cornerstone of the Lanzarote Convention. It facilitates the fight against child exploitation by linking demand and supply in child prostitution, criminalizing all aspects of child abuse material, and addressing emerging challenges such as online grooming. Cooperation enables countries to share data, expertise, and best practices, improving prevention, protection, and prosecution efforts. The Convention encourages both European and non-European countries to become parties to strengthen this cooperation.[306]

The implementation of the Convention is monitored by the Lanzarote Committee, which conducts thematic monitoring rounds across all parties simultaneously. The first monitoring round focused on protecting children against sexual abuse within the circle of trust, producing reports on legislation, judicial procedures, prevention strategies, and victim protection. Subsequent rounds addressed the risks of sexual exploitation during refugee crises and emergencies, as well as the protection of children from abuse facilitated by information and communication technologies, such as the criminal exploitation of self-generated sexual content. In addition to monitoring, the Lanzarote Committee issues opinions, declarations, and recommendations on key issues including online grooming, the removal of websites promoting child sexual

[305] Council of Europe, 2024.

[306] Council of Europe, 2024.

abuse material, the handling of sexting among children, and the protection of children in care institutions. It has also issued statements responding to emerging challenges, such as the increased risks during the COVID-19 pandemic and military conflicts. To promote best practices, the Committee organizes regular capacity-building activities such as study visits, conferences, and seminars involving law enforcement, policymakers, and practitioners. These events facilitate the exchange of knowledge and foster cooperation to enhance child protection measures across member states.[307]

b. Budapest Convention on Cybercrime

At the Council of Europe level, the Convention on Cybercrime (Budapest Convention) and its First Additional Protocol serve as the principal international treaty frameworks. They provide for the criminalization of offenses committed via computer systems, facilitate cross-border cooperation, and establish standards for procedural powers.[308]

The Convention on Cybercrime of the Council of Europe, commonly known as the Budapest Convention, constitutes the cornerstone of international efforts to combat crimes committed via computer systems and networks. Adopted in 2001 and entering into force in 2004, it remains the first binding international treaty aimed specifically at harmonizing substantive criminal law, procedural powers, and international cooperation in the field of cybercrime. Although not conceived exclusively as a child protection instrument, the Convention plays a critical role in safeguarding minors from online threats, particularly in relation to the dissemination of child sexual abuse material (CSAM) and other forms of technology-facilitated exploitation. From a criminal law perspective, Article 9 of the Convention obliges State Parties to criminalize a broad spectrum of conduct related to child pornography when committed by means of computer systems. This includes the production, offering, distribution, procurement, and possession of such material. Importantly, the Convention adopts a technologically neutral approach, thereby encompassing not only traditional forms of CSAM but also emerging modalities such as digitally generated or altered images that may not involve real children but are nonetheless used for exploitative purposes. While the Convention does not address offenses such as online grooming or cyberbullying explicitly, its broad framing allows domes-

[307] Council of Europe, 2024.
[308] Council of Europe, 2001.

tic legislators to extend criminal liability to such acts, especially when they are closely linked to the digital distribution of illegal content.[309]

One of the Convention's most significant contributions lies in its procedural architecture. Recognizing that digital evidence is volatile and often stored across borders, the Convention provides law enforcement authorities with tailored tools for investigating and prosecuting cyber-enabled offenses. These include the expedited preservation of stored data and traffic information, production orders directed at service providers, the search and seizure of computer systems, and the interception of content data in real time. These instruments are particularly relevant for the prosecution of offenses involving minors, where timely identification of perpetrators and victims is essential to preventing ongoing harm.

Given the global reach of internet-based offenses, the Convention places particular emphasis on international cooperation. Chapter III establishes mechanisms for mutual assistance and requires each State Party to designate a 24/7 point of contact to facilitate urgent cross-border requests. This structure enables authorities to act swiftly in cases involving CSAM dissemination or live-streamed abuse, where delays can have immediate consequences for the safety of a child. The recently adopted Second Additional Protocol to the Convention further strengthens these provisions by introducing direct cooperation channels between law enforcement and private service providers, as well as streamlined procedures for data requests, thereby addressing long-standing obstacles in international legal assistance.

Nevertheless, the Budapest Convention is not without limitations. While it provides a robust procedural framework, it lacks a victim-centered approach and does not incorporate specific safeguards tailored to the needs of child victims. It also does not impose obligations on digital platforms regarding preventive or protective measures. As such, it must be viewed as part of a broader regulatory ecosystem, one that includes instruments such as the Lanzarote Convention, which focuses explicitly on sexual offenses against children, and EU Directive 2011/93/EU, which mandates both preventive and punitive responses to child exploitation.

Moreover, disparities in implementation and enforcement remain a persistent challenge. Not all States Parties have fully transposed Article 9 into their domestic legislation, and investigative capacities—especially in relation to encrypted communications or anonymized services—vary significantly. Conflicts

[309] Council of Europe, 2001.

between national legal frameworks, divergent data protection regimes, and limitations on extraterritorial jurisdiction further complicate effective prosecution.

In conclusion, the Budapest Convention provides a critical legal infrastructure for addressing cybercrime, including offenses that target children. Its value lies in its harmonization of legal definitions, its provision of specialised specialized investigative powers, and its facilitation of rapid international cooperation. However, its effectiveness in protecting minors online ultimately depends on its integration with complementary legal instruments and on the political and technical willingness of states to ensure its robust and dynamic implementation in light of ever-evolving technological threats.

c. EU Directive 2011/93/EU

The Directive 2011/93/EU[310], issued by the European Parliament and the Council on December 13, 2011, focuses on combating the sexual abuse and sexual exploitation of children, as well as child pornography. In 2015, it was incorporated as the 49th amendment to the criminal code[311], implementing the European guidelines related to sexual criminal law. This Directive harmonized the definitions of child sexual abuse offenses, including those committed online, across the EU. Additionally, it requires EU member states to adopt preventive measures and ensure the protection of child victims.[312]

B. Regulating Crimes Against Children: A Comparative Overview of National Legal Approaches

At the national level, Member States have transposed the aforementioned instruments into their domestic legal frameworks and, in many cases, complemented them with specific provisions in their criminal codes to address emerging forms of online harm.[313] Despite this progress, significant challenges remain. Fragmented implementation, limited prosecutorial resources, jurisdictional complexities in cross-border investigations, and the rapid pace of technological advancement continue to hinder effective enforcement. More-

[310] Richtlinie 2011/93/EU des Europäischen Parlaments und des Rates vom 13. Dezember 2011 zur Bekämpfung des sexuellen Missbrauchs und der sexuellen Ausbeutung von Kindern sowie der Kinderpornografie sowie zur Ersetzung des Rahmenbeschlusses 2004/68/JI des Rates

[311] Bundesgesetzblatt, 2015.

[312] Independent Federal Commissioner against Child Sexual Abuse (n.d.).

[313] European Parliamentary Research Service for the European Parliament, 2024.

over, the widespread use of end-to-end encryption and the emergence of AI-generated abusive content, such as deepfakes, require continual legal updates to ensure robust child protection while safeguarding fundamental rights like privacy and freedom of expression. Protecting children from harm remains a core goal of both national and international criminal law. Across jurisdictions, targeted legal provisions seek to combat a broad spectrum of offenses that jeopardize the safety, dignity, and well-being of minors. This overview highlights key categories of child-related crimes and exemplifies them through relevant legislation from Germany and Switzerland.

a. Sexual Abuse of Children

One of the most severe violations of children's rights involves sexual offenses. In Germany, § 176 of the Strafgesetzbuch (StGB) criminalizes all forms of sexual acts involving or committed against children. The provision encompasses both physical contact and acts conducted via digital means. Switzerland has a parallel provision in Article 187 of the Swiss Criminal Code (StGB) which criminalizes sexual acts with children under the age of sixteen, thereby establishing a clear statutory protection threshold. The provision applies irrespective of the child's apparent consent, as children below this age are deemed incapable of validly consenting to sexual activity. The law is particularly strict with children under twelve, where any sexual act constitutes an offense regardless of the circumstances. Between the ages of twelve and sixteen, the provision covers not only exploitative situations but also consensual conduct if it involves abuse of dependency, lack of maturity, or significant age differences. Violations of Article 187 StGB are punishable by imprisonment or a monetary penalty, with more severe sanctions in aggravated cases. Importantly, conduct falling within the scope of Article 187 StGB may simultaneously fulfill the elements of other sexual offenses under Swiss law, such as sexual coercion (Art. 189 StGB), rape (Art. 190 StGB), or sexual acts with dependent persons (Art. 188 StGB), depending on the circumstances of the case.

b. Child Sexual Abuse Material (CSAM)

The production, possession, and distribution of child sexual abuse material are addressed under § 184b StGB in Germany. This provision reflects a zero-tolerance approach toward CSAM, recognizing its role in perpetuating abuse. Similarly, Art. 197 StGB in Swiss law criminalizes the handling of pornographic content involving minors, including digital formats and media storage. Under Swiss criminal law, AI-generated child sexual abuse material (AI-CSAM) falls within the scope of the existing child pornography provisions. Article 197 of the Swiss Criminal Code (StGB) criminalizes the production, dissemination, possession, and consumption of pornographic material that depicts sexual acts with children, defined as persons under the age of eighteen. The provision explicitly covers both real and fictitious depictions, meaning that even if no actual child is involved in the production process, the material is nevertheless prohibited. Accordingly, AI-generated or synthetically created CSAM is treated no differently than conventional child pornography, as the decisive factor under Swiss law is the representation of minors in a sexual context rather than the method of production. This approach ensures that emerging technological developments, such as generative AI, do not create regulatory loopholes but remain subject to the established criminal framework for the protection of children.

c. Cyber Grooming and Digital Exploitation

With the rise of online communication, children face new threats in digital environments. §176b StGB in Germany explicitly targets cyber grooming, criminalizing the deliberate attempt to initiate sexual contact with a child via digital platforms. In Switzerland, Cybergrooming does not yet constitute a distinct criminal offense under the Swiss Criminal Code (StGB). Unlike Germany, Swiss law addresses such conduct indirectly through existing provisions, including sexual acts with children (Art. 187 StGB), sexual acts with dependent persons (Art. 188 StGB), coercion (Art. 181 StGB), or offenses involving pornography (Art. 197 StGB). However, these provisions typically require either the performance of a sexual act or at least a concrete preparatory step that goes beyond mere communication. As a result, situations in which an adult attempts to establish contact with a child online for the purpose of initiating sexual abuse may fall into a regulatory gap, as the mere attempt to initiate such interaction is not per se punishable unless it advances to a stage covered by other offences. The absence of an explicit grooming offense in Swiss law has been the subject of legal and political debate, with reform proposals highlighting the need to align national legislation more closely with international child protection standards.[314]

[314] The ongoing political debate: https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20180434.

d. Violation of Privacy and Deepfake Content

Digital technology also facilitates new forms of abuse, such as the creation or dissemination of manipulated private images. In Germany, § 201a StGB criminalizes the unauthorized capture or sharing of images that infringe upon a person's privacy —including deepfakes involving minors. This legal tool is particularly relevant in combating digital sexual violence and reputational harm.

Unlike Germany, Switzerland has not introduced a specific criminal offense targeting deepfakes. Instead, such cases are subsumed under general provisions of criminal law, which may create interpretative challenges and raise questions about the adequacy of protection considering rapidly evolving technologies. In this way, Swiss law provides a framework to address digital sexual violence and reputational harm, but without an explicit legal category dedicated to deepfakes.[315]

e. Hate Speech and Discrimination Affecting Children

Children can also become targets of hate speech, particularly in marginalized or minority groups. In the online environment, such as websites, social networks, comment sections, forums, or chat rooms, the threshold for racist expressions is often lower than in the "real" world, even though such expressions are equally illegal.

In Germany, Hate speech is not a clearly defined legal term. Under the German legal system, freedom of expression enjoys constitutional protection as a fundamental right under Article 5(1) of the Grundgesetz, GG, which grants it the highest level of protection. This right covers not only value judgments but also true factual statements as they contribute to opinion formation. However, false statements of fact are excluded from protection under Article 5(1) GG. The right to free expression is not unlimited; it is restricted where human dignity is affected, personality rights are violated, or where derogatory or defamatory criticism is expressed. Moreover, freedom of expression yields to criminal law provisions when these are infringed, including statutes protecting youth. Article 5 GG provides that everyone shall have the right to freely express and disseminate opinions in speech, writing, and images, as well as to access information from generally accessible sources without hindrance. Press freedom and freedom of reporting by broadcasting and film are guaranteed, and censorship is prohibited. These rights, however, find their limits in

[315] For a detailed analysis: Tag & Wyss, 2024.

the provisions of general laws, youth protection statutes, and the right to personal honor.[316] Offenses such as defamation, insult, and incitement to hatred (Volksverhetzung) fall outside the scope of freedom of expression, irrespective of whether the statements occur online or offline. Hate speech can satisfy several criminal offenses under the German Criminal Code (Strafgesetzbuch, StGB), including Section 111 (public incitement to commit crimes), Section 130 (incitement to hatred), Section 185 (insult), Section 186 (malicious gossip), and Section 187 (defamation). For example, insults under Section 185 StGB may involve extreme abusive language, while Section 130 StGB applies when a person incites hatred or violence against individuals or groups based on their ethnicity or religion in a manner capable of disturbing public peace.[317]

In Swiss criminal law, hate speech is not defined as a standalone offense but is primarily addressed through Article 261bis of the Swiss Criminal Code (StGB), the so-called anti-racism provision. This article criminalizes public incitement to hatred or discrimination against persons or groups on the basis of race, ethnicity, religion, or sexual orientation, as well as the dissemination of ideologies that seek to systematically denigrate such groups. It also prohibits the denial or justification of genocide and discriminatory denial of services intended for the general public. If a racist statement online is considered public, it may be punishable under Article 261bis StGB. This frequently involves calls to hatred ("hate speech") or the dissemination of racist ideologies, as covered under Article 261bis paragraphs 1 and 2 of the Swiss Criminal Code (StGB). "Public" in this context includes closed forums or groups with a larger circle of people not connected by personal relationships. Even private Facebook profiles are regarded as public if they are generally accessible, regardless of whether the profile is anonymous or uses a pseudonym. When a person is directly attacked, this can also constitute a violation of personal rights under Article 28 of the Swiss Civil Code (ZGB) and, if applicable, a criminally relevant offense of defamation under Article 177 StGB. Not only individuals making racist statements online may be criminally liable, but also network operators who fail to fulfill their responsibilities. However, the liability of network providers has not yet been conclusively determined.[318]

[316] Klicksafe, 2023.

[317] Klicksafe, 2023.

[318] Eidgenössische Kommission gegen Rassismus, n.d.

f. The Criminalization of Cybergrooming

In 2003, the German legislator first recognized the threat of cyber grooming in the digital space. Through the "Law Amending Provisions on Sexual Offenses" (SexualdelÄndG[319]) of December 27, 2003, cyber grooming was formally introduced into the German Criminal Code via § 176 (4) No. 3 (old version). The legislative rationale cited media reports of online grooming cases in the United States, some of which allegedly ended in rape. Additionally, the opinion of the European Economic and Social Committee was referenced, calling for legal adjustments to cover crimes in which children are lured into meetings through deception or manipulation. Until then, such conduct—especially the use of written communication in internet chatrooms aimed at initiating child sexual abuse—was considered a non-punishable preparatory act.[320] Now cyber grooming is classified as a form of child sexual abuse and is prohibited under § 176 of the German Criminal Code (StGB). Individuals who approach children or adolescents online with sexual intent may face prison sentences of up to five years. Prohibited actions under this provision include showing pornographic material to a child, attempting to induce a child to perform sexual acts on themselves, on the perpetrator, or a third party, or to be subjected to such acts, as well as intending to produce or obtain child sexual abuse material pursuant to § 184b (1) No. 3 or § 184b (3) StGB. Importantly, the communication does not have to be explicitly sexual in nature; even the attempt to initiate such contact with the intention of engaging the child in sexual acts falls under the offense of cyber grooming.[321] A completed sexual act is not required for criminal liability; mere intent is sufficient. Furthermore, it is not necessary for the child to respond to the messages; it is enough that the child has become aware of the communication for the conduct to be punishable.[322]

There is currently no specific criminal provision explicitly addressing cybergrooming. However, existing legal provisions, such as sexual harassment (Art. 198 StGB) or sexual acts with children (Art. 187 StGB), can be applied. The Legal Affairs Committee of the National Council has proposed incorporating a specific law against Cybergrooming into the Swiss Criminal Code.[323]

[319] Gesetz zur Änderung der Vorschriften über die Straftaten gegen die sexuelle Selbstbestimmung und zur Änderung anderer Vorschriften vom 27. Dezember 2003.

[320] Reschke, 2020.

[321] Bundestag, 2019.

[322] Bundeskriminalamt, n.d.

[323] See Chapter 4.B.c.

C. Regulatory Gaps

Despite significant progress in developing international and national criminal law instruments to protect minors online, important regulatory gaps persist. Instruments such as the Lanzarote Convention, the Budapest Convention on Cybercrime, and EU Directive 2011/93/EU establish a strong foundation by criminalizing child sexual abuse material (CSAM), sexual exploitation, and related offenses. Yet, their effectiveness is limited by uneven implementation, fragmented enforcement, and the rapid evolution of digital technologies.

At the national level, divergent approaches create further inconsistencies. While Germany explicitly criminalizes cybergrooming under § 176 StGB, Swiss law lacks a specific provision, relying instead on general offenses such as sexual acts with children (Art. 187 StGB) or coercion (Art. 181 StGB). This gap leaves situations in which adults attempt to establish online contact with children for abusive purposes insufficiently addressed until conduct escalates to a punishable act. Similarly, although Swiss law explicitly prohibits AI-generated CSAM under Art. 197 StGB by covering fictitious depictions, deepfakes involving minors are not yet codified as a standalone offense. Instead, they are prosecuted indirectly under general pornography or privacy provisions, which risks interpretative uncertainty and inconsistent protection.

The regulation of hate speech also reveals gaps: while Switzerland criminalizes racist or discriminatory incitement under Art. 261bis StGB, broader forms of online harassment targeting minors, particularly in the context of cyberbullying, remain under-regulated compared to more comprehensive frameworks in other jurisdictions. Austria holds a pioneering position in Europe with its explicit criminalization of cyberbullying. Since the amendment of the Austrian Criminal Code in 2016 (§ 107c StGB – *Fortgesetzte Belästigung im Wege einer Telekommunikation oder eines Computersystems*), cyberbullying has been recognized as a distinct criminal offense. The provision specifically targets the continuous harassment, humiliation, or public exposure of individuals through digital means, such as the dissemination of intimate images or persistent online attacks, and allows for prosecution even without direct physical threats or extortion. This makes Austria one of the first European countries to have introduced a clear, technology-specific legal basis to address online harassment and protect victims—particularly minors—from psychological harm in the digital sphere. In contrast, neither Switzerland nor Germany has an equivalent, stand-alone offence. In both countries, acts of cyberbullying are prosecuted under general criminal provisions such as defamation, coercion, insult, or unlawful use of personal data, which often fail to capture the specific dynamics and long-term impact of online harassment. Consequently, Austria's explicit cyberbullying provision is widely regarded as a progressive legal model, closing a significant normative gap that still exists in neighboring jurisdictions.

In addition, both Swiss and German law face challenges in addressing the use of encryption, anonymized networks, and decentralized platforms, which complicate investigations and allow offenders to evade detection.

Moreover, existing frameworks remain reactive rather than preventive. They emphasize punishment after offences occur but do not sufficiently mandate platform responsibility or proactive safeguards by digital service providers. This omission is particularly evident in the absence of binding obligations to detect, prevent, or remove harmful content beyond CSAM, leaving minors exposed to a wide range of digital risks including online grooming, harassment, and manipulative algorithmic practices.

Taken together, these gaps demonstrate that while the legal architecture to protect minors online is comparatively advanced, it remains fragmented, under-inclusive, and technologically outpaced. A more coherent approach is needed, one that closes substantive gaps such as grooming and mobbing offences in Switzerland, ensures harmonized enforcement across jurisdictions, and strengthens preventive duties for online platforms to safeguard children in the digital environment.

VI. Institutional Responses to the Challenges of the Digital World for Children

Several institutions have addressed the issue of dangers faced by children in the digitalized world. Their recommendations and research significantly shape the political landscape. The listed institutions represent only a fraction of the efforts but provide an overview of the ongoing initiatives.

1. The Watchdog Organizations

In the ongoing global fight against CSAM, several prominent watchdog organizations play a pivotal role in documenting, studying, and raising awareness about the extent of CSAM on the internet. These institutions conduct comprehensive studies to understand and publicize the trends and challenges surrounding CSAM, fostering a more coordinated and informed response worldwide.

The Internet Watch Foundation (IWF)[324], a leading organization in this area, works tirelessly to monitor, report, and remove child sexual abuse content online. It conducts annual reports that track the volume of CSAM across various online platforms, providing essential data on where and how this content is being shared and offering key insights into the emerging patterns of abuse. The UK-based is charity dedicated to reducing the availability of child sexual abuse material (CSAM) online. Established in 1996, the IWF works closely with law enforcement agencies, government bodies, and the internet industry to identify and remove illegal content, contributing to a safer digital environment. One of the foundation's core activities is the identification and removal of CSAM. The IWF operates a public hotline where individuals can report suspected illegal content. Trained analysts assess these reports and proactively search the internet to locate and facilitate the removal of such material. In addition, the IWF develops technological tools, such as IntelliGrade, which supports law enforcement and tech companies in accurately categorizing and addressing CSAM. Although headquartered in the UK, the IWF has a global reach, partnering with organizations worldwide to combat online child sexual abuse. In recent years, the foundation has reported a significant rise in AI-generated

[324] Internet Watch Foundation, n. d.

CSAM, with a 2024 study revealing a 17% increase in such content on dark web forums compared to the previous year. The IWF is advocating for legal reforms to criminalize the creation and distribution of AI-generated CSAM. Another growing concern highlighted by the IWF is sextortion, where minors—some as young as 11—are coerced into sharing explicit images. In 2024, the foundation handled 1,142 reports related to sextortion, marking a 44% increase from the previous year. Through its extensive efforts in content removal, technological innovation, and policy advocacy, the Internet Watch Foundation remains a key player in the global fight against online child sexual abuse, continuously working to protect vulnerable individuals in the digital world.[325]

The National Center for Missing and Exploited Children (NCMEC)[326], through its CyberTipline, receives and processes reports of online child exploitation. NCMEC's annual studies are invaluable in providing statistics on the prevalence and nature of CSAM in the U.S. and globally. Their research helps shape public policy and drive legislative changes, as well as providing law enforcement with crucial intelligence to combat child exploitation. The Canadian Centre for Child Protection (C3P), with its specialized expertise in child safety, works closely with international partners to address online child exploitation. Through its own research and partnerships, C3P contributes significantly to understanding the global scope of CSAM, publishing reports that outline the trends and the steps needed to prevent the distribution of this harmful material. The International Association of Internet Hotlines (INHOPE)[327] brings together national hotlines from around the world to create a global network that allows for the reporting and removal of CSAM. By collecting data from these national hotlines, INHOPE compiles yearly reports that offer a global perspective on the volume and nature of CSAM, contributing to the development of best practices and international collaboration. Finally, Interpol's International Child Sexual Exploitation Database (ICSE)[328] acts as a critical resource for law enforcement worldwide, enabling the tracking and identification of perpetrators and victims through a centralized system.[329] The ICSE offers an online platform where law enforcement officers and selected civilian experts from member states collaborate to identify victims, suspects, and crime scenes by thoroughly analyzing CSAM.[330] ICSE's detailed analysis and reporting on CSAM

[325] Internet Watch Foundation, n. d.

[326] National Center for Missing and Exploited Children, n. d.

[327] Safe Online, 2024.

[328] INTERPOL, n. d.

[329] Fry, 2024.

[330] Açar, 2023.

cases help inform policing strategies and improve international cooperation to dismantle child exploitation networks. Together, these organizations provide a wealth of valuable insights through their studies, documenting the ongoing challenge of CSAM on the internet. Their research not only serves as a tool for law enforcement but also educates the public and policymakers, driving global efforts to combat this heinous crime. Their work highlights the need for continued vigilance, increased cooperation, and comprehensive strategies to eradicate CSAM from the online world.

2. UNICEF: Children’s rights in the digital sphere – “promotion of media literacy instead of bans”

UNICEF Switzerland and Liechtenstein advocates for a digital world that respects and upholds children’s rights,[331] as defined in the UN Convention on the Rights of the Child (CRC). While the primary responsibility for implementing these rights rests with states and institutions, the cooperation of businesses and civil society is also crucial. To create a safe and child-friendly digital environment, several key measures are necessary: First, children’s rights must be fully implemented in the digital space, ensuring their protection, privacy, freedom of expression, and access to information. Upholding these rights requires a global effort and collaboration at all levels, including local, national, and international. Second, companies must be more responsible in creating digital platforms that consider children’s safety and rights. The internet and many digital technologies were not initially designed with children in mind, and this needs to change. Platforms should incorporate age-appropriate settings, content filters, and privacy protections. Additionally, companies must take effective action to combat issues like hate speech and cyberbullying. Third, fostering digital literacy from an early age is essential. Children, as well as parents and caregivers, should be equipped with the skills to navigate the digital world safely and critically. This includes offering high-quality, age-appropriate educational content and supporting initiatives that promote media literacy. Companies should also be held accountable for ensuring a secure online environment for children.[332]

[331] UNICEF Switzerland and Liechtenstein, n. d.

[332] UNICEF Switzerland and Liechtenstein, 2024.

A. UNICEF: Policy guidance on AI for children

UNICEF`s recommendations for AI systems impacting children emphasize the need for a rights-based approach, ensuring that AI development and deployment uphold children’s rights to protection, provision, and participation. It is essential to integrate these considerations, whether or not the AI system is specifically designed for children. A multi-stakeholder approach involving government, business, and other key actors is critical, with a focus on adapting solutions to local contexts to meet diverse needs. First, AI systems must prioritize children’s development and well-being. AI policies should be crafted with children’s interests at the forefront, ensuring that these systems not only support children but also contribute to broader goals like environmental sustainability. A child-rights approach should guide the design and application of these systems. Inclusion is another core principle, striving for diversity in the development and implementation of AI. This includes ensuring a range of voices in the design process and actively involving children in AI policy discussions and development stages. Meaningful participation and representation are essential for creating AI systems that truly benefit children. Fairness and non-discrimination should be prioritized, with particular focus on marginalized children. AI systems must be designed to be inclusive, with datasets that reflect the diversity of children’s experiences and needs. Any biases that lead to discrimination must be addressed to ensure equal access and opportunity for all children. Data privacy and protection are paramount. A responsible data approach is necessary when handling children’s data, promoting children’s agency over their information. Privacy-by-design principles must be incorporated, alongside group-level protections, to safeguard children’s personal information. Ensuring children’s safety in the digital space is critical. AI systems must be continually assessed for their impact on children’s safety and security, both in the development phase and throughout their life cycle. Rigorous testing for safety and robustness should be required to protect children from harm while leveraging AI systems to enhance their safety. Transparency, explainability, and accountability are also essential for AI systems that impact children. AI systems should be presented in age-appropriate language, making them understandable for children and caregivers. Additionally, AI systems must be designed to protect and empower children, aligning with legal and policy frameworks that prioritize child protection. Regulatory frameworks should be updated to integrate child rights, and oversight bodies should be established for accountability. Governments and businesses must be empowered with knowledge of both AI and children’s rights. Policymakers, management teams, and developers need ongoing capacity-building to navigate the intersection of these two domains. Consumer demand for trusted, transparent AI solutions can drive the development of child-centered technologies. To prepare children for the future, educational programs must evolve to include both technical and soft skills essential in an AI-driven world. Teachers should be equipped with the tools to enhance their own AI awareness, fostering an environment where students can thrive. Collaboration between businesses and educational institutions will be key, alongside campaigns to raise awareness among parents and society at large. Finally, creating an enabling environment for child-centered AI requires addressing the digital divide. Efforts must focus on equitable access to AI benefits, ensuring infrastructure development and funding for policies that prioritize children. Research into AI for and with children must be supported to continue advancing this important field.[333]

B. UNICEF: Protecting Children from Violence and Exploitation in relation to the digital environment – Policy Brief

This brief highlights both the vast opportunities and significant risks that the digital world presents for children. The UN Committee on the Rights of the Child emphasizes that access to digital technologies can help children realize their rights by enabling learning, connection, and self-expression. However, most digital environments were not designed with children's rights in mind, leading to serious risks that need to be addressed. These risks will only grow as technology advances and connectivity increases. UNICEF identifies four main areas of concern for protecting children online: sexual abuse and exploitation, bullying and harassment, economic exploitation and misuse of personal data, and harmful content. These risks also affect children's mental health, which must be an integral part of any protective response. The brief offers five key recommendations per priority area aimed at governments involved in children's welfare, justice, and digital policy. It also stresses the influential role of businesses—from major tech companies to smaller firms—in shaping children's digital experiences. All businesses must respect children's rights and be held accountable for any harm they cause, including conducting child rights impact assessments, as outlined by the UN Guiding Principles on Business and Human Rights. Children's rights in the digital sphere are interconnected and must be balanced carefully. Protective measures should not unnecessarily restrict other rights such as access to information, freedom of expression, or privacy. Governments should involve children in shaping policies, recognizing

[333] UNICEF Switzerland and Liechtenstein, 2023.

their unique vulnerabilities and capacities. The digital and physical worlds are inseparable when it comes to violence and exploitation; thus, solutions must reflect this reality. Children are diverse, with different needs depending on age, development, and background. Protective strategies must be tailored accordingly and guided by ongoing research and the child's best interests. Ultimately, any business engaging with children or affecting their digital lives carries the responsibility to uphold their rights in all digital interactions.[334]

3. CNIL's 8 recommendations

The CNIL (National Commission on Informatics and Liberty) has published 8 recommendations aimed at protecting children's privacy and rights in the digital space. These recommendations address the risks children face online, such as cyberbullying, exposure to inappropriate content, and the mass collection of personal data. Given that children are a significant portion of internet users and a target demographic in the data market, they are particularly vulnerable to online risks. The CNIL emphasizes the need to balance children's autonomy with their need for protection. The recommendations focus on three main areas: safeguarding children's online rights, strengthening the role of parents and educators in supporting children's digital experiences, and holding online service providers accountable for respecting children's privacy. These recommendations are meant as a starting point for further collaboration, with the CNIL planning to continue working on issues related to children's digital lives, such as education, healthcare, and finance.[335]

The 8 Key Recommendations are the following:

Recommendation 1: Children should be able to enter contracts online (e.g., social media or online games) based on their maturity, as long as the service complies with data protection regulations. Parental consent is still necessary, and the framework should protect children's autonomy.[336]

Recommendation 2: Children should be able to exercise their digital rights (access, rectification, deletion, and opposition) independently. The CNIL suggests that children should be informed about their rights and be able to exercise them, particularly regarding data they've uploaded (e.g., photos on social media).[337]

[334] UNICEF, 2024.

[335] CNIL, 2021a.

[336] CNIL, 2021b

[337] CNIL, 2021c.

Recommendation 3: Parents play a key role in guiding their children in the digital space. The CNIL calls for more support for parents, as many are unaware of how to protect their children's online rights or how to deal with issues like cyberbullying.[338]

Recommendation 4: For children under 15, parental consent is required for data processing, including functions like social media profile settings or geolocation. However, the maturity of the child should be considered when obtaining consent.[339]

Recommendation 5: Child protection tools, such as parental controls, are essential but should avoid invasive features that make children feel constantly monitored. The CNIL recommends evaluating and improving these tools to ensure they respect privacy and are appropriate for the child's age.[340]

Recommendation 6: Children must be adequately informed about how their data is used. Information should be age-appropriate and easy to understand, avoiding manipulative design techniques. Children should also be able to understand and exercise their rights regarding data usage.[341]

Recommendation 7: Age and parental consent verification mechanisms should be proportional, minimizing data collection and avoiding unnecessary intrusiveness. These systems should be robust and simple to use, ensuring children's safety online without violating their privacy.[342]

Recommendation 8: Specific safeguards should be put in place to protect children's interests on websites, services, and apps they use. These include stricter privacy settings by default, preventing profiling of children, and ensuring that children's data is not used for commercial or advertising purposes without clear justification in the child's best interest.[343]

These recommendations aim to create a safer and more respectful digital environment for children while maintaining a balance with their autonomy and privacy rights.

[338] CNIL, 2021d.

[339] CNIL, 2021e.

[340] CNIL, 2021f.

[341] CNIL, 2021g.

[342] CNIL, 2021h.

[343] CNIL, 2021i.

4. German Safer Internet Centre (SID) / Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM)

The FSM is a non-profit organization focused on protecting young people from harmful online content and promoting media literacy. Through its various initiatives, FSM educates and empowers children, parents, and educators to navigate the challenges of the digital world safely. For the Safer Internet Day 2025[344], FSM organized a series of events aimed at raising awareness about misinformation, extremist content, and cyberbullying. One of the key events is a nationwide interactive digital school lesson for students in grades 8 to 10, designed to help them recognize extremist narratives and manipulative deepfakes on social media. FSM also supports parents through its Elternguide.online project. They organize events for parents, focusing on how to educate children about cyberbullying and protect them from online harm. These events provide practical tips and resources for parents on how to discuss cyberbullying within the family. Additionally, FSM takes part in an event organized by the "Wake Up!" initiative to raise awareness of the societal consequences of disinformation. As part of its ongoing efforts, FSM is also promoting the FSM hotline, where users can report harmful or illegal online content, and highlighting its media education projects aimed at families and educators through a social media campaign on Safer Internet Day. FSM plays a crucial role in strengthening youth media protection, collaborating with its member organizations to monitor and regulate harmful online content. It has developed various educational resources, including an internet guide for parents and a platform with teaching materials for educators. FSM is also involved in the INHOPE network, which focuses on combating illegal content online, particularly child abuse imagery.[345]

5. Swiss Foundation for the Protection of Children (Kinderschutz Schweiz)

The Swiss Foundation for the Protection of Children emphasizes that children's rights must be upheld equally in both the offline and online worlds. The Foundation asserts in its comments on the draft of the General Comment on children's rights in relation to the digital environment on 12 November 2020,[346]

[344] European Union, 2025a.

[345] European Union, 2025b.

[346] Swiss Foundation for the Protection of Children, 2020.

that the internet was not initially designed with children in mind, but today a significant proportion of its users are minors. The Foundation states that approximately one-third of global internet users are under 18, and children are often the first to adopt new digital technologies. In Switzerland, nearly all young people between the ages of 12 and 19 own smartphones, and a large percentage of children between 6 and 13 use the internet regularly for various activities such as watching films, socializing, playing games, and seeking help with homework. Despite the positive aspects of digital media, the Foundation highlights that it also presents significant risks, particularly the increasing occurrence of online sexual abuse and exploitation. According to the Foundation, online grooming and the distribution of child sexual abuse material (CSAM) have escalated in recent years. In Switzerland, reports of CSAM cases have surged, yet convictions remain relatively low. This paradox is concerning to the Foundation, especially as a growing number of offenses, particularly since the COVID-19 pandemic, have been documented. However, Switzerland lacks a unified, nationwide strategy to combat cyber-pedocrime. The Foundation points out that, instead, each canton is responsible for managing investigations, despite the limited resources available. This fragmented approach and the absence of cohesive data collection make it difficult to assess the full extent of the issue and limit the effectiveness of efforts to tackle cyber-pedocrime.

The Foundation criticizes the term “child pornography” as imprecise and trivializing, stressing that any image involving children in sexual acts is a form of child sexual abuse, not pornography in the traditional sense. In Switzerland, individuals involved in producing or distributing CSAM face up to five years in prison. Despite this, the Foundation observes that the number of reported cases remains alarmingly high, yet convictions are insufficient. Moreover, the Foundation emphasizes the lack of reliable data on the material hosted or consumed within Switzerland and points out that service providers are under limited obligations to report CSAM unless they stumble upon it by accident or are informed by third parties.

The Foundation also highlights a disturbing rise in reports from international agencies, yet Swiss authorities collect little data on CSAM hosted or consumed within the country. The responsibility for cyber-pedocrime investigations, handed over to individual cantons in 2021, remains underfunded and underresourced. Additionally, preventive measures such as counseling for individuals with sexual interests in children are lacking, making it difficult to address the root causes of this issue.

The Foundation also observes a similar increase in online child sexual abuse material in the European Union, particularly during the COVID-19 lockdown. In response, the EU Commission launched a comprehensive strategy to combat child sexual abuse both online and offline between 2020 and 2025. This strategy aims to improve detection and removal of CSAM, enhance cooperation among EU member states, and collaborate with the private sector to tackle the problem. The EU has introduced measures requiring service providers to actively search their platforms for CSAM and is exploring ways to address encryption challenges that hinder the detection of online abuse. Considering these alarming trends, the Swiss Foundation for the Protection of Children calls for immediate action. The Foundation demands that the production of CSAM be punished as severely as real child abuse, with penalties for sexual child abuse offenses increased. It also insists that cantons prioritize the prosecution of cyber-pedocrime and apply consistent punishment. Preventive measures should be systematically implemented to prevent children from becoming perpetrators themselves, and nationwide data on the hosting and consumption of CSAM should be collected.

Furthermore, the Foundation calls for a nationwide obligation for service providers to actively monitor their platforms for CSAM and report it to relevant authorities. The Foundation urges the Federal Council to create regulations that protect children from digital dangers and require service providers to educate customers about child protection. The Foundation stresses the need for a federal police force dedicated to cyber-pedocrime and a national strategy to coordinate inter-cantonal investigations and address issues like encryption. Lastly, the Foundation emphasizes the importance of expanding prevention programs and educating parents, teachers, and children on online safety and media literacy. The Foundation asserts that these measures must be implemented quickly and comprehensively to protect children from the growing dangers of online sexual abuse.

VII. Platforms and their actions/activities

1. Initiatives by the platforms

Alongside legal frameworks and international guidelines, initiatives by platform operators have emerged as a critical component in advancing child online safety, reflecting growing industry responsibility to prevent harm and promote safer digital environments for minors.

A. Instagram

a. Introducing Instagram Teen Accounts: Built-In Protections for Teens

Instagram introduced new measures on September 17, 2024, to increase the safety of teenagers who have reached the age of 13. These measures involve both restrictions on account usage and enhanced control options for parents. Accounts identified as "Teen Accounts" are automatically assigned to a special protection category, which includes these new safety features. The goal is to regulate the online activities of teenagers and strengthen parents' trust, as users under 16 will only be able to make changes to their accounts with parental consent. These new features are initially available in the US, UK, Canada, and Australia, with plans for deployment in the EU by the end of 2024, followed by global implementation in 2025. Teen accounts will also be made available across other Meta platforms, such as Facebook, WhatsApp, and Threads. Key features include the ability to set accounts to private, meaning teenagers can manually approve followers before they can view posted content. A new sleep mode feature has been introduced, automatically muting notifications at night and sending automated responses to direct messages to promote better sleep. Parents can also monitor usage times and set daily limits for their children, ensuring that the app becomes inaccessible after a set amount of time. Additionally, harmful content such as violent depictions or cosmetic procedures is filtered to ensure a safer environment for communication. Parents are also provided with an overview of their children's chat partners from the past seven days, though without access to the content of the messages. Another innovative feature is the use of AI to identify underage users who might be pretending to be adults. If such users are detected, their accounts are automatically switched to Teen Accounts. If the AI incorrectly determines the user's age, the restrictions can be lifted.[347]

b. Policies

Meta enforces strict policies on all its products, including Instagram,[348] to prevent child sexual abuse and exploitation across its platforms. The company does not allow any content or activities that involve or endanger children through sexual abuse. If Meta identifies potential child exploitation, it reports the content to the National Center for Missing and Exploited Children (NCMEC) in compliance with applicable laws. While Meta acknowledges that some users may share nude photos of their own children with good intentions, it removes these images to prevent misuse by third parties. Meta also collaborates with external experts, including the Meta Safety Advisory Board, to continuously improve its online safety policies, particularly in relation to children. The company strictly prohibits several types of content. This includes the sexual abuse of children, such as any depiction or promotion of child sexual abuse, explicit or implied sexual acts involving minors, and the sexualization of children—even in non-realistic depictions like art or AI-generated content. Additionally, Meta prohibits the request or sharing of sexual material involving children, including explicit or sexualized images. It also bans content that arranges sexual encounters with children, entices them into sexual acts, or involves the exchange of sexual material with them. Meta also addresses abusive intimate images and sexual extortion, removing content that involves the exploitation or abuse of children through threats to share intimate images or private sexual conversations. Content that sexualizes children in any form, including images, videos, or verbal representations, is also strictly prohibited. This includes content that depicts nudity of minors, such as close-ups of genitalia or visible genitals in naked images of minors, whether real or fictional. Additionally, Meta removes content showing non-sexual child abuse, unless the content is from legitimate sources such as news media or art with appropriate context. To enforce these policies, Meta uses both automated systems and human reviewers to assess content based on a wide range of signals. If adults engage with inappropriate content or belong to communities that violate Meta's guidelines, their access to certain features may be restricted. For content that may disturb viewers, such as non-sexual child abuse or sensitive images of children, Meta adds warning labels to ensure users are aware of the content's

[347] Husi-Stämpfli et al., 2024.

[348] Meta, n. d.-a

potentially disturbing nature. Meta is also committed to protecting the privacy and safety of children. It may remove content that identifies potential victims of child abuse or content that could endanger the safety of a child, whether it's reported by law enforcement, trusted partners, or the child's own family. Through these efforts, Meta continuously strives to create a safer online environment, improving its detection, reporting, and removal processes to protect children from exploitation and harm.[349]

Meta has also updated its child safety policies, which are also applicable for Instagram, to address subtler forms of content that might sexualize children without being explicit. While Meta has always removed content that explicitly sexualizes children, the new policy clarifies that accounts sharing innocent images of children with inappropriate captions, hashtags, or comments will also be removed. This adjustment is intended to better capture content that may not show child nudity but could still be harmful by sexualizing the children depicted.

c. Technical measures to combat CSAM

Meta is committed to preventing child exploitation and ensuring the safety of children across its platforms[350], with an emphasis on both prevention and detection. The company has been working on improving its systems and tools to combat this issue by focusing on several key areas, including preventing abuse, detecting harmful content, reporting violations, and collaborating with experts and authorities to safeguard children. To detect such content, Meta uses not only photo-matching technology but also leverages artificial intelligence and machine learning to proactively detect child nudity and previously unseen exploitative content as it is uploaded. These technologies enable Meta to quickly identify such content, report it to NCMEC, and identify accounts that may engage in inappropriate interactions with children.[351]

One of the major efforts Meta has undertaken is the development of a research-backed taxonomy to understand how and why users share child exploitative content. In collaboration with the National Center for Missing and Exploited Children (NCMEC) and other experts,[352] Meta conducted an in-depth analysis of reported child exploitative content from October and November 2020. This analysis revealed that over 90% of the reported content

[349] Meta, n. d.-b

[350] Meta, n. d.-a

[351] Meta, 2018.

[352] Meta, 2021a.

was either identical to or visually similar to previously reported content.[353] Additionally, just six videos were responsible for more than half of the reported child exploitative material. This finding suggested that the same content, often slightly altered, was being shared repeatedly, causing further harm to victims. To better understand the intent behind the sharing of such content, Meta worked with experts to create a taxonomy to categorize users' apparent intent. The research found that over 75% of users who uploaded exploitative content did not exhibit malicious intent but shared it for other reasons, such as outrage or poor humor. This insight has driven Meta's focus on targeted solutions to prevent the sharing of this content, including new tools and policies to reduce its spread.[354]

Meta is currently testing two new tools to address this issue. The first tool is a pop-up notification shown to users who search for terms associated with child exploitation on Facebook and Instagram. This pop-up provides information about the consequences of viewing illegal content and offers support from diversion organizations. The second tool is a safety alert that informs users who have shared viral child exploitative content about the harm it causes, warns them that it violates Meta's policies, and outlines the legal consequences. This alert, which is shown alongside content removal, aims to educate users and help identify those at risk of sharing such material. In addition to these tools, Meta has enhanced its detection capabilities. For years, the company has used technology to identify child exploitative content and potential grooming behaviors. It has expanded these efforts to include the detection and removal of networks that violate child exploitation policies, similar to its work against coordinated inauthentic behavior and dangerous organizations.

Finally, Meta has improved its reporting tools to make it easier for users to report content that violates child exploitation policies. The option to select "involves a child" under the "Nudity & Sexual Activity" category when reporting content has been added to more places on Facebook and Instagram. These reports are prioritized for review to ensure quick action. Additionally, Meta has integrated Google's Content Safety API to better prioritize content that may contain child exploitation, making it easier for content reviewers to assess and address the issue efficiently.[355]

[353] Meta, 2021b.
[354] Meta, 2021a.
[355] Meta, 2021b.

B. TIKTOK

a. Privacy and Security Settings for Teenagers

TikTok offers a variety of privacy settings to help users create a safer online presence, particularly for younger users. The platform is available to people aged 13 and older, with privacy options tailored to the user's age. For teens aged 13 to 15, the account is set to "Private" by default, allowing only approved followers to access content. Additionally, features like creating Duets, Stitches, and Stickers, as well as video downloads, are disabled. For teens aged 16 and 17, the account remains private but can be switched to public, allowing all users on the platform to view and share content. In addition to controlling visibility and interaction on the platform, teens can adjust their direct messages, comments, and the ability to create Duets or Stitches. For private accounts, these interactions are limited to approved followers. TikTok also allows users to manage access to video downloads and stickers, with these settings automatically restricted for younger users. If the account is public, users can opt to enable or disable these features. For parents and guardians, TikTok provides resources to monitor and control their teens' online activities. It is recommended that parents have regular discussions about online presence and TikTok usage. TikTok also offers a "Family Pairing" mode that allows parents to adjust their teens' privacy settings. Additionally, TikTok provides support for managing accounts of children under 13, who can either have their accounts deleted or switched to a child-friendly experience.[356] TikTok describes it as followed in their description in regard to the Teen privacy and safety settings:[357]

TikTok is available for people 13 years and older (or other ages as indicated in our Privacy Policy[358] and Terms of Service[359]). To safeguard your experience, we have default privacy settings for features based on your age, and some features may not be available to you until you turn 16 or 18. Depending on your age, you can adjust the privacy settings at any time, including when you sign up for TikTok or when a feature becomes available to you.

[356] TikTok, n. d.-a.

[357] TikTok, n. d.-b.

[358] TikTok, n. d.-c.

[359] TikTok, 2025.

➜ **Ages 13 to 15**

- Your account is set to private by default. Only people you approve can follow you and view your videos, bio, likes, as well as your following and followers' lists.
- Others can't duet, stitch, create stickers with your videos, download your posts, or add your posts to their Stories.

➜ **Ages 16 to 17**

When you sign up for TikTok, your account is set to private by default. You can adjust it at any time in your privacy settings.
If you choose a private account:

- Only people you approve can follow you and view your videos, bio, likes, as well as your following and followers' lists.
- Others won't be able to duet, stitch, download your posts, make stickers from, or add your posts to their Stories.

You can manage who can send you direct messages from your privacy settings:[360]

➜ **Ages 13 to 15**

- Whether your account is private or public, direct messaging isn't available.

➜ **Ages 16 to 17**

Whether your account is private or public:

- This control is set to No one by default.
- You can change this control to suggested friends (followers you follow back and people you may know) or Friends (followers you follow back) in your privacy settings.

TikTok also offers special resources and guides in its Safety Center for parents who want to learn more about using the platform safely. In the U.S., parents can contact TikTok directly for specific concerns, such as deleting accounts or

[360] The same for Duets and Stitches, also for 13- to 15-year-old users whether the account is private or public, others can't create stickers from any of the posted videos and future videos. This control is set to "Only you" in the privacy settings, see: https://support.tiktok.com/en/account-and-privacy/account-privacy-settings/privacy-and-safety-settings-for-users-under-age-18.

reporting privacy violations. Teens can also deactivate or delete their accounts independently, with parents being involved when necessary.[361]

TikTok has also introduced a variety of privacy settings and safety tools to help manage the content access, ensuring it is appropriate for a family, including minors. One such tool is Restricted Mode, which limits access to content that may not be suitable for all audiences, such as adult content or complex topics. Certain features, such as the "Follow Me" feed, going LIVE, and receiving LIVE gifts, are disabled when this mode is active. Restricted Mode can be turned on or off at any time, and parents or guardians can manage it for their teens through the Family Pairing feature. Another useful tool is the Comment Management Mode, which allows the user to filter out inappropriate or offensive comments on the content. The user is also able to customize the feeds even further with Keyword Filters. These filters allow the user to block specific words or hashtags from appearing in their feeds, such as the "For You" or "Following" feeds. For parents and guardians, Family Pairing Mode provides a way to adjust safety settings for their teens based on individual needs. This feature offers an overview of tools designed to help parents and guardians ensure their teens explore TikTok in a safe manner. To further support online safety, TikTok provides various Safety Resources. The Safety Center in the app offers access to tools, guides, and helpful information for both users and parents. The Safety Center also includes links to global and local safety organizations, such as the Family Online Safety Institute, Internet Watch Foundation, WePROTECT Global Alliance, and the National Center for Missing and Exploited Children. The user is also able to find important resources on preventing child sexual abuse, bullying prevention, mental health support, and digital well-being, ensuring that TikTok is a safer place for everyone.[362]

b. Technical measures to combat CSAM

TikTok employs advanced detection technologies to combat Child Sexual Abuse Material (CSAM) and other violations on its platform. When content is uploaded, it undergoes an automated moderation process to detect and remove violations of community guidelines, including CSAM. Approximately 98.5% of violations are proactively detected and removed before being reported by users, with 93.9% having zero views. TikTok uses technology from other platforms like Google's Content Safety API, YouTube's CSAI Match, and Microsoft's PhotoDNA to enhance its internal detection systems and partners

[361] TikTok, n. d.-b.

[362] TikTok, n. d.-d.

with organizations like NCMEC and the IWF to detect and remove known CSAM. TikTok also utilizes text-based prevention strategies, such as redirecting users searching for harmful content to educational resources. It develops and deploys natural language processing models to detect harmful behaviors like grooming and other predatory actions. Audio models are also used to identify speech violations when no visual or text-based violations are present. Additionally, TikTok reviews accounts showing predatory behaviors and escalates them for further investigation. TikTok is committed to supporting law enforcement agencies (LEAs) by responding to valid legal requests for user information and cooperating with authorities to protect children. In emergency situations, TikTok will disclose user information quickly to prevent imminent harm. The platform regularly engages with LEAs, sharing insights and data to improve safety measures. This collaboration has led to an increase in law enforcement requests for data and proactive engagement in child safety forums.[363]

C. Youtube

YouTube recognized the need for strong protection measures to safeguard young users online. To ensure their child's safety, YouTube has implemented various child and youth protection policies. These include clear guidelines on children's data, content protection, and the appropriate use of the platform. YouTube Kids, a dedicated app for younger users, offers a safer environment, with curated, age-appropriate content. Additionally, YouTube works to eliminate harmful content, such as videos promoting eating disorders or dangerous challenges. The platform also allows for parental oversight, empowering parents to manage what their children watch. YouTube's advisory board focuses on issues related to children, adolescents, and families, helping shape these protective measures. The overarching goal is to create a secure, enriching space for young users, with efforts to continuously improve and adapt to emerging challenges.[364]

YouTube is committed to protecting children and teenagers on the platform, guided by several key principles. First, the focus is on safeguarding privacy, safety, and well-being, ensuring that young users have safer and more enriching online experiences. Mental health is also a priority, with YouTube developing safeguards to address young people's mental health needs and advocating for their well-being. Another important principle is the role of parents and

[363] TikTok, n. d.-e.

[364] YouTube, n. d-a.

caregivers. YouTube offers various features and settings that allow families to tailor online experiences to suit their needs. Parents can balance supervision and independence according to the developmental stage of their children or teenagers. Furthermore, YouTube ensures that all children and teenagers have access to high-quality, age-appropriate content. The platform does not show personalized ads to kids, and its recommendations system is designed to provide safer, inspiring content that encourages curiosity and creativity. YouTube also acknowledges that the developmental needs of children and teenagers vary greatly. As such, the platform adapts content and features to different age groups, working with experts in child development, digital learning, and media studies to meet these needs. Finally, the platform provider recognizes the potential of innovative technologies and ensures that new tools are assessed for risks from the outset. With the help of industry experts and parents, YouTube ensures that families understand how new technologies can help children learn, create, and grow, while also being aware of their limitations.[365]

a. Account with parental supervision

In 2021 YouTube has introduced a new feature designed to give parents more control over their older children's and teenagers' use of the platform. The new feature is a step forward in this direction, catering specifically to children who have outgrown YouTube Kids but are not yet ready for the full, unrestricted YouTube experience. In response to feedback from both parents and older children, YouTube has developed a managed Google account option with parental supervision. This account allows parents to set content restrictions while still enabling their children to explore YouTube in a more independent way. The feature offers parents three distinct content settings, allowing them to tailor the YouTube experience to their child's age and maturity level. The first option, "Explore," is for children around the age of 9 and provides a broad range of videos, including tutorials, vlogs, music videos, and educational content. The second option, "More Explore," is intended for children around 13 years old and includes an even wider selection of videos, such as livestreams. Finally, "Most YouTube Content" allows access to almost all videos on the platform, with some exceptions, such as age-restricted content. This option is meant for older teenagers and provides a near-complete YouTube experience. The feature also gives parents tools to manage their children's YouTube usage. They can monitor their child's watch and search history, use

[365] YouTube, n. d-b.

Family Link to set screen time limits, and customize the account's content restrictions. Additionally, certain features will be disabled, such as personalized ads and the ability to make in-app purchases or leave comments, ensuring a safer, more age-appropriate environment. Alongside this new functionality, YouTube Kids will continue to be the recommended option for younger children, with updates to allow parents more control over specific content and channels. The Freiwillige Selbstkontrolle Multimedia-Diensteanbieter (FSM)[366], a German non-profit organization focused on youth media protection, has praised these updates, especially the options that help parents strike a healthy balance between screen time and offline activities. The FSM emphasized the importance of resources that encourage family discussions about media use and help parents make informed decisions about their children's media consumption.[367]

b. Youtube's "Made for Kids"

YouTube launched "made for kids" on January 6, 2020, and its designation is part of its effort to comply with the Children's Online Privacy Protection Act (COPPA) and other relevant laws, as agreed upon with the U.S. Federal Trade Commission (FTC). All creators, regardless of their location, are required to specify whether their videos are "made for kids" to ensure compliance with COPPA and to avoid potential legal consequences. Failing to properly classify content could result in penalties on YouTube or legal repercussions under COPPA and other regulations. YouTube provides some general guidelines for determining whether content is considered "made for kids," but creators are advised to seek legal counsel if they are unsure about how to categorize their videos. In general, content may be considered "made for kids" if children are the primary audience or if the video features elements like actors, characters, games, songs, or stories that clearly target children, even if the video is not specifically intended for them. On the other hand, content that includes mature themes such as sexual content, violence, or other inappropriate material not suitable for young audiences, or content restricted to viewers over 18, should be classified as "not made for kids." For videos classified as "made for kids," YouTube imposes certain restrictions, such as disabling personalized ads, comments, and features like autoplay. These measures help ensure a safer environment for young users while complying with COPPA regulations. Creators are encouraged to refer to the FTC's blog[368] for further clarification on

[366] For more information see: https://www.fsm.de/.

[367] YouTube, 2021.

[368] Federal Trade Commission, 2019a.

how to determine if their content fits the "made for kids" criteria.[369] In addition, YouTube employs machine learning algorithms to automatically label unclassified videos or override creators' classifications.[370]

c. Child Safety Policy

YouTube is committed to protecting the emotional and physical well-being of minors, defined as individuals under 18 years old. The platform prohibits content that endangers minors in any way. If users encounter such content, they are encouraged to report it. In cases where a child is in danger, local law enforcement should be contacted immediately. Content that violates YouTube's guidelines includes:[371]

1. Sexualization of minors: Explicit or exploitative content involving minors, including nudity intended for comedic purposes, is strictly prohibited. YouTube reports such content to the National Center for Missing and Exploited Children for further investigation.
2. Harmful or dangerous acts: Content showcasing minors engaging in dangerous activities or encouraging others to do so, like consuming alcohol or using firearms unsupervised, is not allowed. It also includes content that inflicts or advocates for the infliction of physical, sexual, or emotional abuse.
3. Misleading family content: Content aimed at young audiences but containing inappropriate themes such as violence, sexual content, self-harm, or medical procedures is banned.
4. Cyberbullying and harassment: Content that targets minors with insults, personal information, or sexualization is prohibited, as well as content that encourages bullying.

Content violating these guidelines, including links to external sites that lead to harmful material, will be removed. If a user violates this policy, they may receive a warning and can take a policy training course. Repeated violations or severe cases may result in channel termination. YouTube also has a zero-tolerance policy for predatory behavior and will cooperate with law enforcement if a child is at risk based on reported content.[372]

[369] YouTube, n. d.-c.
[370] Ma et al., 2024.
[371] YouTube, n. d.-d.
[372] YouTube, n. d.-d.

d. Harassment & cyberbullying policies

YouTube prohibits content that targets individuals with prolonged insults, slurs, or harmful behaviors based on their physical traits, protected group status, or other intrinsic attributes. This includes threats, doxxing, and any form of harassment, especially when it involves minors. YouTube takes a stricter approach to content targeting minors and has a policy in place for reporting harmful content. There are exceptions for educational, documentary, scientific, or artistic content, but even in those cases, harassment is not allowed. Examples of prohibited content include making derogatory comments about someone's physical appearance, using slurs, wishing harm on someone, and sharing private information to encourage abuse. Content creators must avoid repeatedly targeting or insulting individuals based on intrinsic attributes. Violations of this policy can lead to content removal, warnings, strikes, or channel termination, especially in cases of repeated or severe abuse. The platform may also act against creators who incite hostility for personal financial gain or expose individuals to physical harm. YouTube provides resources for reporting harmful content and offers safety guidelines for users.[373]

e. Nudity and Sexual Content Policy

YouTube enforces strict policies regarding explicit content, particularly when it comes to sexual material. Content intended for sexual gratification, including pornography, fetishes, and explicit depictions of minors, is not allowed on the platform. Such content may be removed, and in severe cases, the channel may be terminated. This includes videos, images, and other content types that depict sexual acts, fetishes, or sexualized content, whether in real-life, animated, or dramatized forms, including video games and music. The platform also prohibits sexually explicit content involving minors, including any material that sexually exploits or abuses children. YouTube reports such content to the National Center for Missing and Exploited Children, which works with global law enforcement agencies. Violations of this policy can lead to content removal, as well as penalties such as warnings, strikes, and channel termination, especially in the case of repeated or severe violations. If users encounter content that violates these policies, they are encouraged to report it. YouTube provides clear instructions on how to report violations through various features such as videos, comments, and live streams. In the case of a first violation, YouTube typically issues a warning with no penalty, but repeated viola-

[373] YouTube, n. d.-e.

tions can lead to more serious consequences, such as strikes and, eventually, account termination if three strikes are accumulated within 90 days. Additionally, YouTube may prevent repeat offenders from taking policy training in the future. Overall, YouTube aims to maintain a safe and respectful environment by prohibiting explicit content, particularly that which involves minors or is intended to be sexually gratifying, and enforces this through a combination of policy enforcement and collaboration with law enforcement.[374]

D. Google

a. Technical tools to combat CSAM

Since its inception, Google has worked actively to prevent the spread of illegal child sexual abuse material (CSAM) across its platforms. While CSAM represents a small portion of the content shared on these platforms, Google takes the matter seriously and removes such content when discovered, reporting it and often suspending accounts associated with it. Google is transparent about its efforts and aims to minimize the risk of incorrect suspensions while ensuring child safety. To detect CSAM, Google relies on two main technologies: hash matching and artificial intelligence (AI). Hash matching compares digital signatures of images and videos with a database of known CSAM, provided by trusted sources such as NCMEC and the Internet Watch Foundation. Approximately 90% of reported content matches previously identified CSAM. AI helps identify new CSAM by recognizing patterns of known abusive content while ensuring benign imagery is not flagged. Google also provides a Child Safety Toolkit to other companies and NGOs, offering access to its detection technologies like the Content Safety API, which uses AI to identify and triage CSAM for review by organizations, and the 'CSAI Match' technology to detect CSAM videos[375], and CSAI Match, which have helped process billions of images. Additionally, human content reviewers with specialized expertise confirm the findings from technology. Once CSAM is identified, it is reported to NCMEC, which evaluates the content and may involve law enforcement. Google's Transparency Report provides regular updates on its CSAM detection efforts, including over 1 million reports to NCMEC, the removal of millions of pieces of content, and the disabling of hundreds of thousands of accounts. Google is committed to continuing improvements in its processes, ensuring

[374] YouTube, n. d.-f.

[375] Edwards et al., 2021.

transparency for users, refining its appeals process, and exploring additional ways to combat CSAM while improving user experience.[376]

b. Transparency Reports: Awareness and Education

As mentioned above, Google employs a combination of industry-leading automated detection tools and specially trained reviewers who work around the clock to identify, remove, and report CSAM. This effort is complemented by reports from users and third parties. Between July and December 2024, Google reported a total of 2,528,409 pieces of CSAM content to the National Center for Missing and Exploited Children (NCMEC), which acts as the primary reporting hub in the United States and forwards reports to law enforcement agencies globally. These pieces of content include images, videos, URLs, and texts soliciting CSAM. Since content can appear in multiple accounts or be reported multiple times, the number may include duplicates. During the same period, Google submitted 576,980 CyberTipline reports to NCMEC. These reports contain information about the illegal content itself, users responsible, victims, and other contextual details. Google also escalates supplemental reports for severe cases involving direct or ongoing abuse and the production of CSAM, supporting law enforcement investigations. Google took enforcement actions on 282,584 accounts linked to CSAM violations between July and December 2024. These actions include disabling or restricting access to services, and affected users are notified and given the chance to appeal. The top ten countries with the most enforced accounts are Indonesia, India, Brazil, the United States, Russia, Mexico, the Philippines, Vietnam, Thailand, and Bangladesh. Additionally, Google identified and removed 882,941 URLs containing CSAM from its Search index during this period. While Google cannot remove content hosted on third-party websites, it de-indexes these URLs, making them inaccessible through Google Search. These removals are carried out both automatically and manually. A critical part of Google's fight against CSAM involves the use of hashing technology, which creates unique digital fingerprints of known illegal content. Google has contributed a total of 2,807,013 CSAM hashes to the NCMEC database. Sharing these hashes enables other providers to detect and block previously identified CSAM, strengthening industry-wide efforts.[377]

[376] Google, n. d.

[377] Google, 2024.

2. On the Sufficiency of Provider Measures in Child Online Safety

A. Instagram

Instagram's recent safety measures for young users aim to address significant online risks mentioned above, such as cyber grooming, sextortion, cyberbullying, exposure to harmful content, and addiction. The effectiveness and proportionality of these measures, particularly in terms of protecting minors' rights to privacy and autonomy, are central concerns. Measures like account privatization and chat partner controls help protect against cyber grooming by limiting who can view content and communicate with the user. While these protections are not foolproof, they significantly reduce the risk by preventing unknown individuals from accessing content or initiating contact. However, issues may arise when minors accept friend requests from people they don't personally know. The chat partner control acts as an additional safety net, enabling parents to monitor potentially harmful interactions, though it doesn't fully eliminate the risk. For sextortion, similar measures to cyber grooming, such as chat partner controls and content moderation, serve as preventive tools. The content control system helps limit the spread of explicit material, while media literacy programs encourage minors to handle these risks responsibly. In addressing cyberbullying and hate speech, Instagram uses AI to detect harmful content. While AI improves content moderation efficiency, challenges remain in distinguishing between harmful content and legitimate expressions of opinion, especially across different legal contexts. Further, limiting exposure to harmful content, such as unrealistic beauty standards, could help protect minors' mental health, especially by implementing restrictions on usage time. Labeling edited images, such as those altered with filters, might help mitigate unrealistic body image pressures. Instagram's increased efforts to filter content related to suicide and self-harm are crucial for protecting minors from psychological distress. The effectiveness of these measures depends on the AI's ability to detect and remove harmful material. Regarding addiction, Instagram introduces features like a sleep mode to reduce screen time and promote healthy sleep habits, aiming to break the cycle of phone addiction by limiting daily usage.[378]

Overall, Instagram's measures provide a solid framework for reducing online risks for minors. The effectiveness of the measures will depend on the AI's

[378] Husi-Stämpfli et al., 2024.

progress and consistent enforcement of safety protocols. In terms of proportionality, there is a balancing act between safeguarding minors' physical and mental well-being and protecting their privacy and freedom of expression. In the U.S., concerns about the constitutional and proportional nature of child protection measures often arise, particularly regarding parental oversight of children's privacy. The authors believe that, given the significant online dangers, these safety measures are justified, especially for younger teenagers who may lack awareness of such risks. Mild interventions, like monitoring chat partners for a week, are seen as acceptable for initiating dialogue between parents and children and identifying dangerous contacts early. However, reading entire conversations could be viewed as disproportionate and even a form of abuse from a psychological standpoint. For older teens (16-17), continuous surveillance is deemed unnecessary. It's believed that less intensive measures, such as content moderation to protect against harmful media, are sufficient. Ultimately, Instagram's new measures strike a balance between protecting minors and respecting their rights to privacy and autonomy, with the need for a proportional approach that takes age and risk into account.[379]

B. TIK TOK

TikTok's safety measures do offer substantial protection for minors, though like any platform, the effectiveness depends on how they are used. The platform's default privacy settings for younger users, particularly for those aged 13-15, are a strong preventative measure. The automatic setting of accounts to private, combined with restrictions on certain features like Duets, Stickers, and video downloads, significantly reduces the chances of minors encountering unwanted interactions. These automatic privacy settings are a notable strength –also compared to other providers. TikTok's proactive approach in ensuring that the default for younger users is a private and controlled experience puts it ahead in protecting minors by default. TikTok also provides parents with the "Family Pairing" feature, which is another layer of protection, allowing them to manage privacy settings, screen time, and monitor content. This is comparable to Instagram's parental controls, but TikTok goes a step further in offering a comprehensive and integrated system for parental oversight. Instagram does offer tools for reporting and setting privacy, but TikTok's Family Pairing mode allows for more specific, real-time adjustments to privacy and interaction settings, making it more user-friendly and effective in ensuring minors' safety. Regarding content moderation, TikTok's "Restricted Mode"

[379] Husi-Stämpfli et al., 2024.

and “Comment Management Mode” are strong features for controlling what minors are exposed to. The option to filter inappropriate comments and block certain keywords adds another layer of safety. TikTok’s combination of these features, along with its restricted content settings, presents a more complete safety system than Instagram’s, which focuses primarily on reporting and blocking users rather than preventing exposure to harmful content in the first place.

In conclusion, TikTok’s safety features are generally good at protecting minors online, with robust privacy controls, content moderation tools, and resources for parents. It is particularly positive to emphasize that users aged 13 to 15 cannot access certain services, which can be seen as an exemplary regulation. However, consideration should be given to whether the age limit should be raised. Even 16-year-olds could possibly still have difficulties correctly assessing the dangers online, especially when interacting with third parties. A sensible adjustment of the limit could be made up to the minimum age of majority. Should a relaxation of the age limit be considered, it would be advisable to combine this with mandatory training developed by the platform itself and integrated into the verification process. This training should be assessed with a ‘pass/fail’ system. Only users who successfully complete the training would then be allowed to access the content of the next age group up. Users who do not pass the training would continue to be excluded from accessing certain services. The authors are critical of the early unblocking of various TikTok features, as there is still a risk that underage users will fall victim to abuse, for example, through the trend of dance videos, which often contain physical elements and portray the dancers as attractive. Critics warn that such content could be misused by pedophiles or criminals if children and young people share their own videos. There is also a risk of Cybergrooming, where offenders make contact with the adolescents.[380]

An alarming newspaper report sheds light on a worrying issue that, in the opinion of the authors, has not yet been brought under control yet. It describes the dangers to which children and young people are exposed through the consumption of extremely violent and sexual content on the internet. The example of a 13-year-old schoolgirl who came across a video of a girl performing sexual acts on a small child while surfing TikTok illustrates how quickly such content can end up on the smartphones of minors. In addition, another horrifying video showing a gruesome offense was shared on Twitter. Social

[380] see also https://www.schau-hin.info/sicherheit-risiken/tiktok-musically-mehr-sicherheit-fuer-kinder.

networks, particularly platforms such as TikTok and Twitter, are increasingly a space for the dissemination of disturbing content such as torture, murder, and pornography. This content, which often comes from anonymous sources abroad, regularly ends up in the hands of children and young people, which can lead to considerable psychological damage. In one case, a pupil reported sleep disorders and nausea after receiving a murder video from a war zone.[381]

C. YouTube

In 2019, the Federal Trade Commission (FTC) imposed its largest fine under the COPPA against YouTube for non-compliance. Along with paying the fine, YouTube agreed to implement changes to improve its adherence to the law.[382] The main complaint against YouTube was that, despite being a child-directed platform under COPPA, it insisted to advertisers and channel owners that the site was intended for a general audience and didn't require COPPA compliance. The FTC argued that YouTube had actual knowledge that children were using the site, as evidenced by its actions with advertisers and channel owners. The FTC revealed that YouTube earned nearly $50 million from child-directed content, and a study showed that 85% of children's videos contained ads. While the $170 million fine[383] against YouTube was significant, it did little financial damage, and the FTC's penalties appeared limited in their impact. The question remains whether the consent decree will effectively prevent future violations of COPPA or if stronger measures are needed.[384]

Since then, especially to comply with the court order, YouTube implemented several policy changes affecting creators and channel owners. These include requiring channel owners to designate their videos as either "made for kids" or not[385], which impacts both new and existing content. YouTube will also use machine learning to assist in identifying videos for kids, though creators are warned not to solely rely on these systems. The "made for kids" designation limits key features such as personalized ads, commenting, and autoplay, potentially reducing creators' revenue. In response, YouTube committed over $100 million to support the creation of family-friendly content. While privacy

[381] NDR, accessed under: https://www.ndr.de/nachrichten/niedersachsen/oldenburg_ostfriesland/Lehrerin-alarmiert-ueber-Gewalt-und-Pornografie-auf-Tiktok-und-Co,klassenchat100.html.

[382] Federal Trade Commission, 2019b.

[383] https://www.ftc.gov/news-events/news/press-releases/2019/09/google-youtube-will-pay-record-170-million-alleged-violations-childrens-privacy-law.

[384] Cobb, 2021.

[385] See YouTube, n. d.-c.

advocates praised the changes, content creators expressed dissatisfaction. A study shows that creators and consumers viewed the “made for kids” classification as inconsistent with their actual practices.[386] Creators faced unexpected consequences when implementing the labeling, and both creators and consumers recognized how the “made for kids” classification intersected with other platform features.[387]

Many creators and users have reported issues with the system, leading to false positives and false negatives in video classifications.[388] A major concern among parents and creators is the lack of autonomy in managing children’s content consumption. Some parents feel that the MFK classification undermines their control over what their children watch, while others point out that it doesn’t provide enough customizable options for selecting appropriate content for their kids. This frustration has led some users to develop alternative methods for managing the content children are exposed to, which in turn can lead to false negatives. The system has also caused creators to struggle with content creation. The ambiguity of the MFK criteria has left many unsure whether their content qualifies as “Made for Kids,” leading to additional labor and uncertainty in content labeling.[389]

While YouTube has taken steps to protect minors through policy enforcement, parental controls, and reporting features, the effectiveness of these measures is still a subject of debate. Parents can monitor their children’s usage and set restrictions, but many minors are technically savvy enough to bypass parental controls, potentially exposing them to unsafe content. Moreover, while YouTube’s moderation teams do remove inappropriate content, there are still challenges in addressing harmful material, such as hidden predatory behavior in comments, livestreams, and user-generated content that may not be easily categorized as violating the platform’s rules.

[386] Ma et al., 2024.

[387] Ma et al., 2024.

[388] False Positives occur when videos are mistakenly labeled as “Made for Kids” even though they don’t meet the criteria. This can happen when age-restricted or inappropriate content is flagged as suitable for children. Some creators exploit the system by labeling adult content as “Made for Kids” without facing consequences. The main issue is that YouTube’s algorithms fail to account for the complexity of content, leading to misclassifications, especially for videos aimed at older audiences but featuring child-friendly elements like video games. False Negatives happen when content that should be classified as MFK is not recognized as such. Users often bypass restrictions, adding MFK videos to playlists or using parent accounts to access non-MFK content, undermining the purpose of the classification system.

[389] Ma et al., 2024.

The platform's safety measures also rely on user reports to identify harmful content, which may not always capture issues that slip under the radar. Additionally, some older minors may not fully understand the risks associated with online interactions, leading to situations where they unknowingly expose themselves to harmful or exploitative content.

D. Google

Google has invested heavily in technological and procedural safeguards to prevent the circulation of child sexual abuse material (CSAM) across its platforms. Its reliance on industry-leading tools such as hash matching and AI-driven detection demonstrates a proactive approach, reinforced by the scale of its operations—millions of reports to the National Center for Missing and Exploited Children (NCMEC), the removal of nearly a million CSAM-related URLs, and the disabling of hundreds of thousands of accounts within a six-month reporting period. Moreover, Google's publication of transparency reports and its decision to make detection tools available to external stakeholders reflect a commendable commitment to accountability and knowledge-sharing within the industry.

Yet, the sufficiency of these measures must be viewed critically. First, the company's approach is inherently reactive: most identified content corresponds to previously known material (around 90% of reported content), which underscores the persistence of circulation rather than the prevention of creation. While AI tools are designed to identify new CSAM, they face significant challenges in balancing sensitivity with the avoidance of false positives, and it remains unclear how effective these systems are in identifying emerging forms of abuse such as live streamed or deepfake CSAM.

Second, jurisdictional and enforcement gaps remain significant. Google reports content to NCMEC, which then disseminates information to law enforcement worldwide. However, the scale of reports—millions of pieces of content within months—risks overwhelming investigative capacity. Without parallel investments in law enforcement infrastructure, the sheer volume of reports risks diluting their effectiveness. Moreover, while Google removes or de-indexes illegal content from its own services, it cannot directly eliminate CSAM hosted on third-party platforms or the dark web. This highlights a structural limitation of provider measures: their capacity is restricted to the scope of their own ecosystems.

Third, the company's enforcement actions raise questions of due process and proportionality. While hundreds of thousands of accounts are disabled for CSAM violations, Google itself acknowledges the risk of incorrect suspensions. Appeal mechanisms exist, but there is limited independent oversight of these processes, leaving affected users reliant on the same provider that imposed the sanction. Transparency reports enumerate numbers, but they do not fully clarify the criteria, thresholds, or error rates underlying enforcement.

Finally, despite repeated references to the protection of children's rights, preventive and child-centered measures remain underdeveloped. The focus lies overwhelmingly on technological detection and takedown, rather than on structural strategies that could reduce children's exposure to grooming, sextortion, or coerced self-produced imagery. In this respect, Google's measures, while technologically advanced, may be insufficiently aligned with a broader rights-based framework that requires collaboration with educators, parents, civil society, and policymakers to address the root causes of online child sexual exploitation.

In sum, Google's efforts represent an important contribution to child online safety, but they cannot be considered sufficient in isolation. The measures are technologically sophisticated but largely reactive, limited to the boundaries of Google's services, and not fully integrated with broader systemic safeguards. Addressing CSAM effectively demands not only provider-level interventions but also regulatory clarity, law enforcement capacity, and cross-sectoral prevention strategies that extend beyond the private sector's technological remit.

3. Enhancing Online Safety for Minors: A Provider Perspective

In today's digital age, social media platforms play a pivotal role in shaping the online experience for minors. As the internet continues to be an integral part of their lives, the responsibility to create a safe, age-appropriate environment falls squarely on platform providers. With children and young people increasingly engaging with digital content, it is essential for these platforms to implement robust measures to protect their users from harm. From age verification systems to personalized content filters, social media providers must take proactive steps to ensure that minors are shielded from inappropriate or harmful material. This responsibility extends beyond the technical aspects of content moderation and touches upon the ethical and legal challenges of balancing user safety with privacy rights. In this context, platforms must evolve their strategies to not only protect young users from explicit content but also from emotional harm caused by bullying, body shaming, or social humiliation. Through a combination of innovative technologies, collaboration with law enforcement, and enhanced media literacy, platform providers can foster a safer online environment for minors while also respecting their rights to digital expression. This article explores the various strategies and challenges faced by social media providers in their mission to enhance online safety for minors.

A. Platform Providers ensuring an age-appropriate digital environment

Social media providers play a key role in making their platforms a safe and age-appropriate digital space for children and young people. An important step here is the implementation of effective age verification systems that ensure that only users of the appropriate age group have access to certain content or functions. These systems should not only take effect at registration but also during the use of the platforms to ensure that young people do not access content that is unsuitable for their age group. By utilizing advanced technologies, these systems can protect users' privacy while ensuring that younger users do not see inappropriate content.

In addition, social media platforms should offer personalized filter mechanisms that adapt content based on the age and interests of users. These filters can block violent, sexualized, or otherwise harmful content so that children and young people do not interact in inappropriate or dangerous online environments. In addition, it is important that the communication functions on the platforms are designed in such a way that they are safe and only allow age-appropriate interactions. Algorithms could be used here that recognize problematic messages or harassment and report them immediately, thereby ensuring a safe environment for young users.

B. Expanding the definition of inappropriate content

In the context of social media platforms, the protection of children and minors has become a growing concern due to the widespread exposure to potentially harmful content. Social media giants have taken various steps to ensure the safety of younger users. However, a significant issue that has been raised is the definition of "inappropriate content." Currently, most platforms primarily focus on explicit material, hate speech, cyberbullying, and other forms of harmful communication. However, there is a broader range of content that can be damaging to minors, such as images that may humiliate, embarrass, or negatively affect the social reputation of individuals, especially young users.

The authors suggest expanding the definition of inappropriate content to include images or videos that could potentially expose minors to social humiliation or negatively impact their social environment. This includes images that may not be overtly harmful but could cause emotional distress, such as those that make someone a target for ridicule or bullying or portray unrealistic expectations of beauty or behavior. For example, images that have been altered through filters, showcasing unattainable beauty standards, could contribute to issues like low self-esteem and body dysmorphia, particularly for young users still developing their sense of identity. To tackle this challenge, one possible solution is the use of advanced artificial intelligence (AI) to detect and flag such content. AI has already been deployed in social media platforms for content moderation, such as identifying explicit material or hate speech. However, its capabilities could be expanded to identify more nuanced forms of inappropriate content, such as images that might cause social embarrassment, emotional distress, or damage to one's social standing. This more advanced AI could analyze not only the visual content but also the context surrounding the image. For example, AI could assess whether the image has been widely shared or commented on in a way that could lead to negative social consequences for the individual in the picture. It could also recognize subtle forms of cyberbullying or harassment, such as certain types of edited images, memes, or context in which someone's image is being used maliciously to embarrass them. AI would need to be trained to detect patterns in social media interactions that might suggest bullying or shaming.[390]

Furthermore, the idea of expanding content moderation to protect minors from the social implications of certain images could be linked with the broader challenge of ensuring privacy and autonomy for young users. Social media platforms must tread carefully when implementing measures that protect users from harm, as they need to balance the potential overreach of content moderation with the rights of minors to express themselves freely. One potential solution is the implementation of "soft" moderation, where AI flags potentially harmful content, and the user or a trusted adult (if the user is a minor) is alerted. This allows for a more contextual and individualized response, rather than simply removing content. Additionally, these AI tools could help identify instances of content that might lead to negative mental health outcomes, such as excessive body-shaming or exclusionary behavior.

[390] Husi-Stämpfli et al., 2024.

The need for such measures is particularly urgent given the increasing prevalence of mental health challenges among young users of social media, exacerbated by online interactions. Issues such as body image concerns, cyberbullying, and social comparison are disproportionately affecting children and teens, and social media platforms have a responsibility to minimize the potential harm caused by such exposure. As the technology behind AI evolves, platforms can continue refining the detection of more complex forms of inappropriate content, ensuring the protection of vulnerable users while maintaining a balance with their right to expression.

C. Presence of police on the social media platforms

In the context of social media platforms, the shift of criminal activity, especially in areas like fraud, to the digital realm presents significant challenges. Social media platforms are increasingly being used for various illicit activities such as cybercrime, scams, harassment, and even more severe crimes like online grooming or exploitation. However, despite the increasing prevalence of online crime, law enforcement agencies have struggled to establish a significant presence in the digital world. Although some efforts are being made to improve the reporting of online crimes—such as the development of digital crime reporting tools or online patrols—the actual presence of law enforcement on social media platforms remains limited. The idea of creating a "police authority" within social media platforms raises concerns. Such an initiative could potentially challenge the state's monopoly on force and law enforcement. While social media companies may have their own internal content moderation systems, the responsibility of enforcing the law and maintaining public order is generally a state function. Therefore, handing over such powers to private companies could lead to potential legal and ethical issues, especially in terms of accountability and transparency. Despite these concerns, a cooperation model between social media companies and law enforcement agencies could be more practical and effective in combating digital crime. By partnering with law enforcement, social media platforms could help identify and flag illegal activities or suspicious behavior more quickly, without necessarily taking on the role of law enforcement themselves. These collaborations could ensure that platforms do not overstep legal boundaries while still providing a safer online environment for users, especially minors.[391]

In addition to collaboration with law enforcement, the concept of private security services or moderation within social media platforms, like security

[391] Husi-Stämpfli et al., 2024.

guards or entry controls in physical spaces, could be a viable option. Many social media platforms already have automated systems or human moderators in place to identify and flag inappropriate content. However, adding a layer of security by employing private security services (either in the form of dedicated security personnel or technology-driven tools) could further bolster the safety of users. For example, some platforms already have moderators or administrators who manage chatrooms or online communities, ensuring that conversations and interactions stay within acceptable bounds and adhere to the platform's terms of service.

This concept of security measures could extend beyond just monitoring content and include measures that help prevent fraud, identity theft, and other forms of online exploitation. By improving the internal security mechanisms within social media platforms, it is possible to prevent minors from falling victim to various types of online crime, including scams or fraudulent schemes that target vulnerable users.

Moreover, for social media platforms, implementing effective security measures is not only a matter of ethical responsibility but also a business necessity. Platforms that fail to protect users, particularly minors, from crime and harmful content risk damaging their reputations, facing legal action, or losing users. As public awareness about the dangers of social media grows, platforms could face mounting pressure from governments, users, and advocacy groups to improve safety measures. Considering these concerns, social media platforms may increasingly need to embrace safety measures to ensure that they maintain user trust, minimize liability, and contribute to the overall security of the digital ecosystem.

D. Promoting Media Literacy and Parental Awareness for Safer Social Media Use Among Minors

In today's digital landscape, developing strong media literacy is crucial, particularly for minors, to protect them from various online dangers. Media literacy can help young people critically assess the credibility of online content, especially political posts, and understand the potential risks of digital media. A key solution is comprehensive sexual education that includes digital media literacy, helping minors navigate online content like pornography, body image issues, and sexual identity. This education should focus on teaching young people about consent, gender equality, and how to cope with negative emotions triggered by unrealistic portrayals of sexuality. The authors argue that media literacy should not only address sexual content but also equip minors with the tools to deal with other online threats, such as violent content, cyberbullying, and unrealistic beauty standards. Understanding the consequences of cyberbullying is vital, as its impact can be severe, even leading to suicide. Minors should also learn about available resources for seeking help if they experience bullying. In addition to educating minors, it is essential to raise awareness among parents. Parents need to understand the risks their children face online to effectively support them and provide a good example in their own social media use. Educated parents can foster healthy digital habits in their children and be credible guides when challenges arise. Social media companies and the entire internet industry also have a responsibility to provide clear and accessible risk education. One approach could be a user-friendly “warning label” that explains risks and safety guidelines, tailored for both adults and children. Additionally, integrating mandatory educational videos or informational sessions on the risks of social media use for minors could become part of future regulations. These combined efforts can help create a safer digital space for young users.

E. Digital Safety Risk Assessment Framework & Risk Classifications

In the context of child safety on social media platforms, Digital Safety Risk Assessment Frameworks could play in the future a crucial role in identifying and mitigating potential risks. These frameworks should be designed to assess the specific dangers minors face online, such as cyberbullying, exposure to inappropriate content, grooming, and privacy breaches. By systematically evaluating the threats associated with online platforms, these frameworks help ensure that the safety measures implemented are effective and comprehensive.

A well-designed Digital Safety Risk Assessment Framework would function by continuously analyzing various risk factors like content moderation efficiency, the platform’s age-verification mechanisms, data protection protocols, and the potential for harmful interactions between users. It could involve automated tools, such as AI-powered systems, that flag harmful content, alongside regular manual reviews. Importantly, these frameworks should also incorporate regular feedback from users, parents, and experts in child psychology, cybersecurity, and digital ethics.

Vinh-Thong Ta describes that traditionally security assessments have focused mainly on protecting businesses, leaving a gap in addressing children’s unique online challenges. He proposes a safety risk assessment framework specifically for children’s online safety, using automated mathematical reasoning to evaluate weaknesses, particularly in age verification and parental control measures. However, it does not address data breaches or privacy violations. The framework has limitations, including reliance on human expertise, challenges in automating safety measure extraction, and a lack of dynamic risk assessment based on children's online behavior. Future research could involve AI and Machine Learning, but challenges with explainability, especially in Deep Learning, should be considered.[392]

Given the growing concerns over children's exposure to online dangers, it is essential that governments and regulatory bodies make these risk assessments a legal requirement for social media platforms. Legislation should mandate that platforms provide annual assessments of their safety measures, disclose risk factors, and outline how they plan to address emerging threats. Such legal requirements would ensure that platforms continuously improve their protection strategies and are held accountable for any negligence regarding user safety.

Social media companies themselves must take the initiative to implement these frameworks proactively. This includes creating dedicated teams to monitor and update safety protocols, investing in research and development of better content moderation tools, and providing transparency reports to the public. Additionally, platforms should work closely with child protection organizations, educators, and governments to align their safety measures with best practices. By making digital safety a core aspect of their service, platforms can ensure a safer online experience for minors, empowering parents and guardians while fostering a responsible and secure digital environment for children.

a. World Economic Forum – Digital Safety Risk Assessment Framework

In an increasingly interconnected world, measuring digital safety has become essential for understanding risks, allocating resources effectively, and ensuring compliance with regulations. However, the task of establishing clear and reliable metrics is far from simple. The rapid pace of technological evolution, the need for flexible yet consistent measurement standards, and the challenge of balancing privacy concerns with transparency all create significant obstacles. Moreover, the nature and context of digital harms differ widely across platforms and services, further complicating the creation of universally applic-

[392] Ta, 2024.

able metrics. As a result, there is currently a lack of standardized and agreed-upon measures that can be used to assess digital safety in a consistent manner. To address this gap, the Global Coalition for Digital Safety has produced a comprehensive paper outlining the most effective approaches to measuring digital safety. Drawing on the expertise of a diverse group of stakeholders—including platforms, regulators, safety providers, NGOs, academics, and international organizations—the paper aims to create a shared understanding of what constitutes effective digital safety metrics. By promoting this common framework, the coalition seeks to ensure that risks are properly identified, mitigated, and continuously monitored over time. The paper introduces a structured approach to evaluating digital safety, categorizing metrics into three broad groups: impact, risk, and process. Impact metrics focus on the consequences of digital engagement on individuals, providing insights into the lived experiences and challenges faced by users. These metrics help us understand the real-world effects of digital interactions on people's well-being. Risk metrics, on the other hand, are designed to detect potential harms and allow for proactive mitigation strategies. By identifying areas of vulnerability, these metrics help to prevent or reduce the likelihood of negative outcomes. Lastly, process metrics evaluate the systems, strategies, and outcomes related to digital safety initiatives. These include the methods used to implement safety measures and the effectiveness of these measures over time. Practical application of these metrics is vital to ensure that digital safety efforts are effective and responsive to emerging threats. They guide the ongoing improvement of safety measures, help evaluate interventions in real time, and provide a basis for holding service providers accountable. When aligned with the goals and challenges of the ever-evolving digital landscape, these metrics serve as a valuable tool for both regulators and digital service providers. They allow for the tracking of progress, the identification of gaps, and the enhancement of user trust in digital platforms, if privacy concerns are carefully considered.[393]

Furthermore, the paper emphasizes the importance of collaboration between various stakeholders to improve access to data while addressing privacy and security concerns. The risk framework highlights the need for responsible data handling and encourages partnerships between researchers and data custodians to improve the availability of data for safety assessments. By fostering such collaborations, stakeholders can better monitor the safety of digital platforms and ensure that the risks to users are continuously addressed. In this context, digital safety metrics also play a critical role in supporting regulatory

[393] World Economic Forum, 2024.

frameworks. They help ensure that policies are based on accurate, evidence-driven insights, enabling regulators to harmonize safety benchmarks across industries. Rather than imposing multiple conflicting requirements, regulators can focus on creating cohesive standards that promote meaningful progress in measuring and improving digital safety. Finally, the paper acknowledges that while digital safety metrics are an important tool, they are not the only means of assessing progress in protecting users from online harms. Other complementary approaches should also be considered to provide a comprehensive understanding of digital safety. Metrics are a starting point for deeper, ongoing efforts to create safer digital environments and ensure that children and other vulnerable groups are adequately protected from online threats.[394]

b. UK: Ofcom’s Children’s Risk Assessment Framework

Under the Online Safety Act 2023, Ofcom, the UK’s independent communications regulator, has introduced a comprehensive Children’s Risk Assessment Framework to strengthen online child protection. From April 2025, online services likely to be accessed by children must carry out two key assessments: a children’s access assessment, to establish whether their platform is used by minors, and, if so, a children’s risk assessment, to identify and mitigate potential harms.

The framework is structured around a four-stage process:

1. Identify harmful content – platforms must consider risks from categories such as pornography, self-harm, bullying, hate speech, and other harmful material relevant to their service.
2. Assess risks – providers evaluate how likely children are to encounter such harms and how severe the impact could be, considering features like algorithms, recommendation systems, or private messaging.
3. Implement mitigations – services are required to introduce safety measures such as age verification, content moderation, or safer design of platform features. Providers may follow Ofcom’s Codes of Practice or propose equivalent alternatives, provided they are effective.
4. Report and review – the outcomes must be documented, monitored, and updated regularly, especially when significant changes are made to the service.

[394] World Economic Forum, 2024.

Providers are legally obliged to complete these assessments by July 2025, with enforcement powers enabling Ofcom to hold non-compliant companies accountable. The framework is designed to embed safety by design, ensuring that child protection is built into the architecture and operation of digital services rather than left to parental controls or user actions. This approach represents a major step forward in digital child safety in the UK, setting a potential model for international regulatory frameworks.[395]

c. Theoretical framework for children's internet use

Sonia Livingstone and Leslie Haddon developed one of the most influential theoretical frameworks for understanding children's internet use within the EU Kids Online project. Their approach situates children's online experiences at the intersection of opportunities and risks, emphasizing that digital technologies provide not only educational, social, and creative benefits but also expose children to threats such as cyberbullying, sexual exploitation, or harmful content. The framework highlights the mediating role of parents, peers, and schools in shaping how children engage with the internet, while also stressing the impact of broader socio-cultural, regulatory, and technological contexts. Central to their model is the notion of a "risk–opportunity balance," which underlines that effective child protection requires not only shielding children from harm but also enabling them to fully benefit from digital participation.[396]

d. EU: The 4Cs: Classifying online risk to children

The European framework of the "4Cs" provides a widely adopted model for classifying online risks to children. It distinguishes four categories of risk: content, contact, conduct, and contract. Content risks arise when children are exposed to harmful or age-inappropriate material, such as violent or sexual imagery. Contact risks involve harmful interactions with others, for instance grooming, cyberbullying, or coercion by adults or peers. Conduct risks refer to situations where children themselves may behave in harmful ways, either towards others or through self-exposure, such as sharing personal information or engaging in risky challenges. Finally, contract risks concern the commercial and legal dimensions of children's online engagement, including data exploitation, hidden costs, and manipulative design practices. By structuring risks along these four dimensions, the 4Cs framework helps policymakers, educators, and regulators to develop more holistic protection strategies, balanc-

[395] Ofcom, 2025.

[396] Livingstone, Haddon & Görzig, 2012.

ing the prevention of harm with the promotion of safe and empowering digital participation for children.[397]

F. Classifying Generative AI Models Capable of Producing Harmful Content as High-Risk Systems

It is crucial that generative AI models capable of producing potentially illegal content, such as CSAM, be classified as high-risk AI systems within various regulations. These systems should undergo deeper risk assessments and meet stringent requirements before they are deployed or marketed. Given the severe harm that can result from the misuse of generative AI for creating CSAM, it is imperative that these models be subject to thorough testing and documentation to assess their potential for producing such illegal content. Such assessments should not be limited to theoretical evaluations but must include practical, real-world testing to ensure these AI systems cannot generate harmful or illegal material. If any generative AI model is found to be capable of producing CSAM, immediate measures should be implemented to prevent this functionality. These measures could include disabling the model's ability to generate such content, implementing built-in safeguards, or imposing strict access controls to restrict the model's use. In addition, the developers of these AI systems should be required to document and report the results of these risk assessments, ensuring full transparency and accountability. Only by adopting these proactive measures can we help prevent the creation and distribution of illegal content while ensuring that the technology is used in a responsible and ethical manner. Regulatory frameworks should reflect the evolving capabilities of AI and include provisions to prevent its misuse, protecting both the public and vulnerable individuals, particularly children, from the dangers of harmful AI-generated content.[398]

G. Legal Requirement for Mandatory Social Media Training to Protect Minor Users

In the course of the research, it became clear that a key aspect of the debate on child protection in the digital space is the inadequate training of parents and children with regard to risks, safety measures, and a general understanding of the topic. One aspect that is still lacking in the legislation analyzed by the authors is mandatory training that is imposed on platform operators. Be-

[397] Livingstone & Stoilova, 2021.

[398] Oberlin & von Hoyningen-Huene, 2025, August.

fore the use of digital services by underage users, it should be mandatory for both the children and their legal guardians to take part in a corresponding training programme.

Therefore, Social media providers should be legally required to offer mandatory social media training for minor users to ensure their safety and well-being online. These training programs should cover essential topics such as data protection, where users learn how to safeguard their personal information, manage privacy settings, and understand the risks associated with sharing sensitive data online. Additionally, the training should highlight the potential dangers minors face on social media platforms, including cyber grooming, cyberbullying, and exposure to inappropriate content. It should educate users on how to recognize warning signs of harmful interactions and provide guidance on how to protect themselves. Furthermore, the program should outline clear behavioral rules, offering guidance on what to do when encountering illegal content or inappropriate conduct, such as how to report such incidents to platform moderators or appropriate authorities. Finally, the training should incorporate a basic cybersecurity module, teaching minors essential practices for staying secure online, such as recognizing phishing attempts, creating strong passwords, and avoiding malware.

To ensure the effectiveness of these programs, the training should be structured in a way that it is evaluated with a “pass” or “fail” grade. Only those users who successfully complete and pass the training test should be granted unrestricted access to the platform. This approach will ensure that minors are fully equipped with the knowledge they need to navigate the digital world responsibly and safely before they are allowed to use social media without limitations. By implementing these mandatory training programs with clear assessment criteria, social media providers can create a safer environment for young users and help prevent potential online dangers.

H. The EU law on chat control: The ultimate solution?

The European Union’s proposal for a law on chat control, officially known as the “Regulation on the Prevention of Child Sexual Abuse”[399], was first introduced in May 2022. The proposal for a European regulation on combating online child sexual abuse is rooted in the protection of children’s rights, as enshrined in the United Nations Convention on the Rights of the Child (UNCRC) and the Charter of Fundamental Rights of the European Union. The need

[399] European Commission, 2022b.

for such a proposal was highlighted by the United Nations Committee on the Rights of the Child in 2021, emphasizing the importance of safeguarding children's rights in the digital environment.[400] With increasing evidence of children falling victim to online sexual abuse, the EU has recognized the necessity of addressing this growing threat. A significant number of children face high risks of sexual abuse, both offline and online. While the sexual abuse and exploitation of children, including the creation and distribution of CSAM, is criminalized across the EU through the Child Sexual Abuse Directive[401] adopted in 2011, it has become evident that current efforts are insufficient, particularly in addressing the online aspect of the problem. The digital space has proven to be a major challenge, with offenders increasingly exploiting the internet to access and abuse children. In response to these alarming trends, the European Commission adopted the EU Strategy for a More Effective Fight Against Child Sexual Abuse on July 24, 2020.[402] This strategy sets out a comprehensive framework to combat child sexual abuse both offline and online, with a focus on improving prevention, enhancing investigation capabilities, and providing better support for victims. The strategy includes eight key initiatives aimed at strengthening the legal framework for child protection, improving law enforcement efforts at both the national and EU levels, promoting industry collaboration to protect children on digital platforms, and enhancing global cooperation to safeguard children worldwide.[403] The strategy is complemented by other EU efforts, including the adoption of the EU Strategy on the Rights of the Child in March 2021, which advocates for stronger measures to protect children from violence, including online abuse.[404] In line with the vision for a digitally transformed Europe by 2030, as outlined in the "Digital Compass 2030: A European Way Forward for the Digital Decade"[405] the Commission presented a European Declaration on Digital Rights and Principles for the Digital Decade on 26 January 2022. Furthermore, the proposed European Declaration on Digital Rights and Principles for the Digital Decade includes a commitment to protecting children from illegal content and exploitation in the digital space.[406] On 15 December 2022, the Presidents of the Council of the EU,

[400] UNCRC, 2021.
[401] European Parliament & Council, 2011.
[402] European Commission, 2020.
[403] European Parliament, 2025.
[404] European Union, 2021.
[405] Council of the European Union, 2021.
[406] European Commission, n. d.-c.

the European Parliament, and the Commission signed the European declaration on digital rights and principles for the digital decade.[407]

A critical component of this proposal is the role of providers of hosting or interpersonal communication services. These providers play a key part in ensuring a safe, predictable, and trustworthy online environment for users, particularly children. The proliferation of CSAM online perpetuates the harm experienced by victims, as it allows offenders to easily access and exploit children through digital platforms. Despite the voluntary efforts of some online service providers to detect, report, and remove CSAM, the measures taken have been inconsistent, and many providers have failed to take sufficient action. In fact, a large proportion of the reports received by EU law enforcement authorities come from just a handful of providers, while many others take no action at all. For example, the National Center for Missing and Exploited Children (NCMEC) in the U.S. received over 21 million reports in 2020, with more than 1 million of those related to EU member states, highlighting the scale of the problem. Despite the contributions of certain providers, voluntary actions have proven insufficient to address the misuse of online services for child sexual abuse. As a result, several EU member states have started developing national laws to combat online child sexual abuse, which has led to fragmentation in the Digital Single Market. The proposal seeks to address this issue by establishing uniform EU rules on the detection, reporting, and removal of CSAM to complement the Digital Services Act, remove barriers to the Digital Single Market, and prevent further fragmentation.[408]

The proposal aims to establish a clear, harmonized legal framework to combat online child sexual abuse, providing legal certainty for service providers about their obligations. These obligations include assessing and mitigating risks associated with their services, and, where necessary, detecting, reporting, and removing CSAM in a manner that is supposed to respects the fundamental rights of users, particularly their privacy. Balancing the protection of children with the privacy rights of other users is a core principle of the regulation, and the measures are designed to be proportionate to the risks associated with different services. To support the implementation of the regulation, the proposal calls for the establishment of a European Centre to prevent and counter child sexual abuse. This center will facilitate cooperation between national authorities and help providers meet their obligations under the regulation. The EU Centre will also operate and maintain databases of indicators of on-

[407] Pingen, 2023.

[408] European Commission, 2022b.

line child sexual abuse that service providers will be required to use to detect CSAM. The databases will be prepared before the regulation is applied, and the European Commission has already allocated funding to assist member states with these preparations. The EU Centre will also assist national authorities, provide victim support, and facilitate information exchange and cooperation across EU member states. This proposal is aligned with the broader EU legal framework, including the Child Sexual Abuse Directive and the interim Regulation (EU) 2021/1232, which will remain in place until 2024. By requiring providers to detect, report, block, and remove CSAM from their platforms, the proposal will enhance the detection and investigation of child sexual abuse and improve law enforcement efforts in line with EU and international law. The proposal also supports the objectives of the European Strategy for a Better Internet for Children, which seeks to create a safer digital environment for children while promoting their digital empowerment. The EU Centre will work closely with Europol, receiving and reviewing reports from providers before forwarding them to Europol and national authorities. Europol's involvement will ensure effective coordination and cooperation across the EU and international partners in the fight against online child sexual abuse.[409]

In conclusion, this proposal builds on existing EU and international legal instruments, such as the Lanzarote Convention on the Protection of Children against Sexual Exploitation and the Budapest Convention on Cybercrime, reinforcing the EU's commitment to preventing and combating child sexual abuse in all its forms. It marks an essential step towards ensuring that online platforms take responsibility for the safety of children and providing law enforcement with the necessary tools to protect children from exploitation and abuse in the digital world. This law, as proposed in the EU, could potentially play a significant role in combating child abuse and identifying CSAM. On the one hand, such a law could help in the early detection and reporting of CSAM, enabling quicker interventions by authorities. Automated technologies could potentially detect harmful material faster, reducing the suffering of victims and helping authorities apprehend offenders sooner. A unified legal framework across the EU would also prevent fragmented regulations in different member states, which could undermine the protection of children and complicate efforts to tackle child exploitation online. Furthermore, coordinated efforts at the European level would improve cross-border cooperation, making it easier to address online child abuse that often transcends national borders.

[409] European Commission, 2022b.

However, there are several challenges and concerns that must be addressed. One of the main issues is the potential threat to privacy. The monitoring and scanning of messages and communications in a fully automated real-time chat-control could be seen as the abolition of digital letter secrecy, as disproportionately invasive , and as potentially posing serious risks to individual privacy rights.[410] This becomes especially problematic when it comes to encrypted messaging services, where the challenge of implementing such measures could undermine the very concept of secure communications.[411] Moreover, automated detection technologies are not perfect, and false positives are a real concern. These systems could mistakenly flag harmless content as abusive,[412] which would not only lead to unnecessary disruptions but could also undermine trust in the technology. The risk of further misuse of surveillance systems is another concern of the proposed law. Especially the widespread surveillance of communication could foster a sense of distrust among users, potentially affecting how they engage with online platforms. People might become more reluctant to use digital services if they feel their communications are being constantly monitored, leading to a negative impact on the overall online environment. The perception of overreach could also damage the public's trust in digital platforms, which may ultimately harm the goal of ensuring online safety for children. In conclusion, while a law for chat control could certainly help in protecting children from abuse, it is crucial to balance this goal with strong protections for privacy and data security. Even though the proposal aims to respect the privacy and fundamental rights of the users such laws could pave the way for broader surveillance of online communications, extending beyond child protection and potentially leading to an erosion of digital freedoms. Without careful, transparent regulation and the development of secure, privacy-conscious technologies, such a law could end up causing more problems than it solves. It is essential to ensure that such laws are well-thought-out, with clear safeguards in place to prevent misuse and protect fundamental rights, including the right to privacy. Finding the right balance between protecting children and respecting online freedoms will be key to the success of such a law.

[410] Gesellschaft für Freiheitsrechte, n. d.; Gesellschaft für Freiheitsrechte, 2023; Gesellschaft für Informatik, 2024; BfDI, n. d.

[411] Threema, 2023.

[412] Harding-Zentrum für Risikokompetenz, 2025.

I. Social Media Ban for minors

In recent years, concerns about the effects of social media on children and adolescents have led to discussions about implementing a minimum age requirement or even a full ban for users under 16. While such a policy might offer some protection, it also raises significant ethical, practical, and social questions. A blanket ban on social media for children under the age of 16 currently seems difficult to envision and practically almost impossible to implement, although calls for such a measure are periodically raised and discussed. In a previous publication, the authors, along with Dr. iur. Sandra Husi, already explored this option.[413] Internationally, various approaches and political initiatives are emerging that consider stricter regulations or even a ban, yet a complete ban has not been realized. This debate is complex and requires a nuanced evaluation of both the pros and cons:

The debate surrounding a potential ban on social media use for individuals under the age of sixteen is marked by both compelling advantages and significant disadvantages. Advocates argue that such a ban could serve as an important protective measure, particularly in relation to mental health. Numerous studies suggest a strong correlation between excessive social media use and heightened risks of anxiety, depression, and low self-esteem among adolescents. By restricting access during this vulnerable developmental stage, a ban could mitigate these psychological risks. Beyond mental health, young users are especially susceptible to various forms of online harm, including cyberbullying, hate speech, sexual exploitation, and the pervasive influence of unrealistic beauty standards. A ban might therefore act as a preventive mechanism, shielding minors from exposure to such content. Moreover, restricting access would reduce the extent to which social media companies exploit children's personal data, thereby limiting their digital footprint and protecting their privacy from the mechanisms of surveillance capitalism. Finally, by decreasing digital distractions, a ban could foster healthier patterns of real-life socialization, promoting face-to-face interactions, physical activity, and offline relationships that are essential for adolescent development.

Critics, however, caution that such restrictions come at a high cost. Social media platforms constitute key arenas for self-expression, civic participation, and activism, and a blanket ban could unduly infringe upon the rights of minors to freedom of expression and democratic engagement. Additionally, preventing adolescents from accessing online environments risks creating a digital liter-

[413] Husi-Stämpfli et al., 2024.

acy gap: rather than learning to navigate the digital sphere responsibly, young people may be left unprepared for the realities of an increasingly interconnected world. Questions of enforceability further complicate the proposal, as bans are notoriously difficult to implement effectively. Adolescents may circumvent age restrictions by providing false information, leading to hidden and unregulated usage that could be even more dangerous in the absence of adult oversight. Finally, since social media is an important medium through which young people build and maintain peer relationships, exclusion from these digital spaces could contribute to feelings of isolation, marginalization, or social exclusion, particularly if peers continue to interact online.

A ban on social media for minors under 16 offers certain protective benefits but also raises concerns regarding children's rights, digital literacy, and social inclusion. Rather than an outright ban, many experts advocate for a more balanced approach —one that includes stricter age verification, content moderation, digital education, and parental involvement.

VIII. Evaluation and Outlook

1. The Effectiveness of Existing Laws in Protecting Minors online

The legal landscape regarding online child protection in Europe has evolved considerably in recent years. It is becoming increasingly clear that lawmakers are increasingly concerned with advancing technological developments and have recognized the growing potential risks posed by the use of artificial intelligence.

To start with the GDPR it does specifically mention children's data and sets a minimum age for children to consent to the processing of their personal data, but it primarily focuses on privacy and data protection rather than addressing other significant risks minors face online, such as cyberbullying, sexual exploitation, or exposure to harmful content. While it provides a framework for safeguarding children's privacy in the digital space, it does not offer a comprehensive approach to tackling the broader range of dangers that children encounter online, such as online abuse, exploitation, and the spread of inappropriate material.

The Digital-Service-Act (DSA) aims to make the digital space safer by regulating the responsibility of online platforms in relation to illegal content, misinformation, and other harmful activities and makes the platform providers accountable according to Art. 28 DSA that they offer a high level of privacy, safety, and security to young users. The DSA requires platforms to assess every year and minimize risks to minors on their services, support parents with settings to protect their kids from online risks, implement a system to verify the age of users before they are able to access the service, and implement tools to help young people to signal abuse or get support.[414] However, the implementation and enforcement of these regulations is crucial to ensure that minors are adequately protected. Despite the advancing legal framework, enforcement and monitoring remain a significant challenge. Online platforms often operate across borders, and the speed at which new risks emerge makes it difficult for regulators, but also for the platforms, to keep pace and implement security measures before the risk occurs.

[414] European Commission: Directorate-General for Communications Networks, Content and Technology, 2023.

The AI-Act has been critiqued for not adequately addressing the potential dangers posed by generative AI, particularly in relation to AI-models that are able to “produce” CSAM. One significant issue is the categorization of generative AI systems. These AI models, capable of creating content such as images, videos, and text, including harmful or illegal material like CSAM, are not per se classified as high-risk systems under the AI Act. This raises concerns about whether the regulation goes far enough to protect vulnerable groups, especially children. While high-risk AI systems are subjected to rigorous assessments and oversight, generative AI systems that could be misused for creating illegal content are not subject to the same level of scrutiny. The Act’s provision that generative AI must comply with transparency requirements and prevent the generation of illegal content is a step in the right direction. However, this may not be sufficient, particularly when it comes to the creation of CSAM. Despite its potential to cause significant harm, generative AI is not automatically classified as a high-risk system and thus falls outside the most stringent regulatory measures. Furthermore, there are practical challenges in enforcing the regulation. Many generative AI applications are already in circulation and can be easily modified for malicious purposes, often spreading through decentralized platforms like the dark web. Once these systems are in the hands of criminals, it becomes difficult to track or control their misuse. This highlights a critical gap in the AI Act’s current framework: it does not account for the dynamic, evolving nature of AI systems and their potential to be exploited after their release. For the AI Act to better protect children, stronger preventive measures are necessary. This could involve classifying AI applications that can generate harmful content, including CSAM and deepfakes, as high-risk systems. Additionally, AI developers should be required to implement safeguards to prevent the generation of illegal content, such as automatic detection mechanisms. Lastly, a more proactive regulatory approach is needed, focusing on independent third-party assessments and ensuring that such AI models cannot be misused once they are deployed.[415]

The Youth Protection Laws in Switzerland and Germany aim to protect minors from harmful content in the digital space by establishing clear regulations for age verification, content classification, and parental control systems. They strengthen data protection and parental oversight while ensuring that content providers take responsibility for the protection of minors. The implementation and enforcement of these laws are overseen by the respective regulatory authorities, which can impose penalties for non-compliance. However, the effec-

[415] Oberlin & von Hoyningen-Huene, 2025, August.

tiveness of these protective measures depends on the specific implementation by the platforms and their cooperation with authorities.

France is far ahead in this respect by explicitly including the right to one's own image of minors in the civil code. This new law places greater obligations on data controllers, particularly social media platforms, to fulfill requests for deletion without delay. It also signals tighter control by the CNIL, which will investigate the handling of minors' data in 2024. Platforms must implement appropriate age verification, security measures, and data minimization practices to meet the requirements of the law.[416]

Another challenge is the global inconsistency in laws and regulations. While some countries have robust protections in place, others lack adequate legal frameworks, leaving minors in certain regions more vulnerable to online risks.

2. Outlook

To ensure comprehensive protection for minors in the future, laws must evolve in several key areas:

- First, stronger enforcement mechanisms on a global scale are needed to make platforms more proactive in identifying and removing harmful content.
- Age verification systems must be enhanced to effectively prevent minors from accessing inappropriate content. Legislation should enforce regular audits of these systems by an independent third-party auditor to ensure their reliability and effectiveness.
- Global cooperation and consistent regulation are also necessary to ensure that minors are protected regardless of where they are in the world.
- Furthermore, education and awareness programs for both parents and minors should be expanded, empowered, and enforced by the regulator.
- Child pornography should be treated like a cyber security incident. Users should be rewarded when finding such content, but there should also be a notification obligation to a government entity to keep track of harmful content online. The wording could be as follows:

  ...in the event that harmful content is discovered by the provider on the platform, the provider shall without undue delay and, where feasible, not later than 72 hours after having become aware of it, notify the incident to

[416] Le Clerc & Leportois, 2024.

the supervisory authority competent in accordance with Article… *and erase such content…*

Through this approach, the state would not only gain an overview of the extent of harmful content actively present on platforms but also be able to assess whether the security measures implemented by providers are adequate. Additionally, it would allow for the identification of which providers perform particularly well or poorly in ensuring online safety.[417]

- There is a pressing need for a clear legal definition of harmful content to ensure consistent protection of minors in the digital environment.
- A comprehensive harmonization of legislation must be pursued to ensure that, in all countries, providers are legally obliged to promptly report any illegal material—particularly child sexual abuse material (CSAM)—to law enforcement authorities. Such a framework should establish binding international standards to guarantee seamless cross-border cooperation and to strengthen the global fight against these grave offences. It is of paramount importance that every state enacts statutory obligations enabling the early detection and reporting of this criminal material, thereby enhancing the protection of children and adolescents worldwide.

[417] Günal Rütsche et al., 2025.

IX. Conclusion

Children today are increasingly exposed to a wide range of risks in the digital world, including sexualized content, violence, cyberbullying, and inappropriate interactions with strangers. These dangers are compounded by the excessive screen time many children spend online, which can negatively affect their physical and mental well-being, leading to issues such as sleep deprivation, anxiety, and diminished social skills.[418] As children's lives become more intertwined with digital spaces, it is essential to address these risks holistically. However, simply banning access to social media or restricting digital engagement does not resolve the deeper challenges that children face in the online world. While such restrictions may seem like an immediate solution, they tend to overlook the root causes of the issues. Bans on social media use for minors often fail to address fundamental problems such as a lack of media literacy, digital safety education, or the fact that many technologies are designed with little consideration for children's needs. Rather than offering true protection, such restrictions could inadvertently push children toward unregulated, potentially more dangerous platforms. Furthermore, these measures can hinder open dialogue between children and their parents, which is crucial for guiding children through online spaces safely. Without proper guidance and communication, children may find themselves increasingly isolated in navigating digital spaces, further exposing them to risks.

Instead of outright bans, a more nuanced approach is necessary—one that addresses both the immediate concerns and the long-term challenges children face in the digital age. This approach must involve ongoing legal reform, where laws keep pace with technological developments. For example, current laws like the GDPR are beneficial in that they are flexible enough to apply across various technological contexts, ensuring data protection. However, as new technologies, such as artificial intelligence and deep learning, rapidly emerge, these laws often fall short. Legislators must not only account for technological advancements but also consider the broader societal, ethical, and security implications of these changes. In a world where new technologies are being developed at an unprecedented rate, it is vital that laws remain adaptable and forward-thinking, rather than rigid and outdated. At the core of this approach to safeguarding minors online is a multi-layered framework that involves various stakeholders: lawmakers, platform providers, parents, and the

[418] Chen, 2023.

users themselves. One of the first steps is to introduce clear, robust legislative regulations that ensure social media platforms provide safe environments for minors. These regulations should mandate the implementation of effective age verification systems and enforce security measures, such as data protection, privacy safeguards, and security filters, to shield children from harm. Social media platforms should also be held accountable for actively monitoring and managing content that may pose a risk to young users.

In addition, parents must play a crucial role in guiding their children through the digital world. Parents need to be more involved, not only in monitoring their children's online activities but also in educating them about the potential risks and how to navigate them. Parents should foster open, honest conversations about what their children encounter online and encourage critical thinking about digital content. Empowering children with the knowledge to distinguish between credible information and misleading or harmful content is essential for their online well-being. Equally important is the need to promote digital literacy, particularly among teenagers. As young people spend more time on social media platforms, it becomes critical to help them develop the skills to engage with content in a more critical and informed manner. Media literacy should go beyond just understanding how to use technology; it should include teaching children how to recognize and respond to harmful content, how to protect their privacy, and how to engage in respectful and safe online communication. Schools, community organizations, and digital platforms themselves all have a role to play in fostering this critical skill set.

The developments to limit access to social media for minors illustrate a broad international trend toward redefining the digital rights of minors, with age-based access models gaining traction in response to a perceived regulatory gap. While some countries pursue formal age restrictions and parental consent mechanisms, others focus on platform accountability, education, and design-based interventions. The underlying legal and cultural frameworks vary, but the shared objective remains clear: to create safer and more developmentally appropriate digital spaces for children and adolescents.

Ultimately, any strategy to protect children online must not be reactive but proactive. The laws, technological tools, and educational frameworks need to be continually updated to reflect the evolving nature of the digital world. It is only through a holistic, collaborative approach that involves multiple stakeholders—governments, technology companies, parents, educators, and children themselves—that we can hope to create an online environment that is both safe and empowering for the next generation. The focus must not just be on protecting children from immediate dangers but also on preparing them to thrive in a digital world, equipping them with the skills and knowledge to use technology safely, responsibly, and with a critical eye.

References

Açar, K. V. (2023). *International cooperation mechanisms for online child sexual abuse investigations*. http://dx.doi.org/10.13140/RG.2.2.14630.80969

Adel, A., & Norouzifard, M. (2024). Weaponization of the growing cybercrimes inside the Dark Net: The question of detection and application. *Big Data and Cognitive Computing*, 8(8), 91. https://doi.org/10.3390/bdcc8080091

Agencia Española de Protección de Datos (AEPD). (2023). *Decalogue of principles – Age verification and protection of minors from inappropriate content*. https://www.aepd.es/guides/decalogue-principles-age-verification-minors-protection.pdf

Alexiou, E. (2018). Cyber-Grooming: Eine kriminologische und strafrechtsdogmatische Betrachtung, Berlin. https://doi.org/10.3726/b14449

Ali, K., Farrer, L., Gulliver, A., & Griffiths, K. M. (2015). Online peer-to-peer support for young people with mental health problems: A systematic review. JMIR *Mental Health*, 2(2), e19. https://doi.org/10.2196/mental.4418

American Psychological Association. (2013). *Violence in the media: Psychologists study potential harmful effects*. https://www.apa.org/topics/video-games/violence-harmful-effects

Amnesty International. (2023). *Driven into the darkness: How TikTok's 'For You' feed encourages self-harm and suicidal ideation*. https://www.amnesty.ch/de/themen/wirtschaft-und-menschenrechte/unternehmensverantwortung/dok/2023/tik-tok-setzt-jugendliche-schaedlichen-inhalten-aus/report-driven-into-darkness-how-tiktoks-for-you-feed-encourages-self-harm-and-suididal-ideation.pdf

Andreassen, C. S. (2015). Online social network site addiction: A comprehensive review. *Current Addiction Reports*, 2(2), 175–184. https://doi.org/10.1007/s40429-015-0056-9

Arendt, F., Scherr, S., & Romer, D. (2019). Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media & Society*, 21(11–12), 2422–2442. https://doi.org/10.1177/1461444819850106

Assemblée nationale. (2020, October 19). *LOI n° 2020-1266 du 19 octobre 2020 visant à encadrer l'exploitation commerciale de l'image d'enfants de moins de seize ans sur les plateformes en ligne*. Journal officiel de la République française, n° 0255, October 20 2020. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042439054

Assemblée nationale. (2024, Febuary 19). *LOI n° 2024-120 du 19 février 2024 visant à garantir le respect du droit à l'image des enfants*. Journal officiel de la République française, n° 0042, Febuary 19 2024. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049163317

Autoriteit online Terroristisch en Kinderpornografisch Materiaal (ATKM). (n.d.). Autoriteit online Terroristisch en Kinderpornografisch Materiaal. https://www.atkm.nl/

Bayerisches Landesjugendamt (BLJA). (2021). *Änderungen im Jugendschutzgesetz (JuSchG) seit dem 1. Mai 2021* (Mitteilungsblatt 2/2021, S. 9–12). https://www.blja.bayern.de/imperia/md/content/blja_2024/pdf/mitteilungsblatt_02_2021_barrierefrei.pdf

Benner, A. D., Wang, Y., Shen, Y., Boyle, A. E., Polk, R., & Cheng, Y.-P. (2018). Racial/ethnic discrimination and well-being during adolescence: A meta-analytic review. *American Psychologist*, 73(7), 855–883. https://doi.org/10.1037/amp0000299

Bennett, W. (2019, April 16). *Going offline with a body-positive queer Instagram sensation. them.*. https://www.them.us/story/mina-gerges-interview

Beran, T., & Li, Q. (2005). Cyber harassment: A study of a new method for an old behavior. *Journal of Educational Computing Research*, 32(3), 265–277. https://doi.org/10.2190/8YQM-B04H-PG4D-BLLH

Berger, M. N., Taba, M., Marino, J. L., Lim, M. S. C., & Skinner, S. R. (2022). Social media use and health and well-being of lesbian, gay, bisexual, transgender, and queer youth: Systematic review. *Journal of Medical Internet Research*, 24(9), e38449. https://doi.org/10.2196/38449

Bergmann, M. C. & Baier, D. (2016). Erfahrungen von Jugendlichen mit Cybergrooming: Schülerbefragung – Jugenddelinquenz. *Rechtspsychologie*, 2(2), 172–189.

Berkman Center (2008). The Berkman Center for Internet & Society at Harvard University, https://cyber.harvard.edu

Blackburn, M. (2024, December). *I've seen too much heartbreak — House must pass landmark Kids Online Safety Act* [Press release]. https://www.blackburn.senate.gov/2024/12/i-ve-seen-too-much-heartbreak-house-must-pass-landmark-kids-online-safety-act

Brand, M., Young, K. S., Laier, C., Wölfling, K., & Potenza, M. N. (2019, September). The interaction of person-affect-cognition-execution (I-PACE) model for addictive behaviors: Update, generalization to addictive behaviors beyond internet-use disorders, and specification of the process character of addictive behaviors. *Neuroscience & Biobehavioral Reviews*, 104, 1–10. https://doi.org/10.1016/j.neubiorev.2019.06.032

Bryant, M. (2024, November 30). Instagram actively helping spread of self-harm among teenagers, study finds. *The Guardian*. https://www.theguardian.com/technology/2024/nov/30/instagram-actively-helping-to-spread-of-self-harm-among-teenagers-study-suggests

Bundesamt für Justiz (BJ). (2024). Datenschutz. https://www.bj.admin.ch/bj/de/home/staat/datenschutz.html

Bundesamt für Kommunikation (BAKOM). (2025). Vorentwurf des Bundesgesetzes über Kommunikationsplattformen und Suchmaschinen (VE-EKomPG), Erläuternder

Bericht zur Eröffnung des Vernehmlassungsverfahrens https://cms.news.admin.ch/dam/de/der-schweizerische-bundesrat/atfTk-DOkgjEA/Erl%C3%A4uternder+Bericht_DE.pdf and https://www.ebg.admin.ch/de/newnsb/6TmEAde4htulaWG9CWYtK

Bundesamt für Sicherheit in der Informationstechnik (BSI). (2023). *Risiken und Schutzmaßnahmen für Kinder im Internet*. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Kinderschutz-im-Internet/Risiken-und-Schutzmassnahmen-fuer-Kinder-im-Internet/risiken-und-schutzmassnahmen-fuer-kinder-im-internet_node.html

Bundesamt für Sozialversicherungen (BSV). (2024, June 6). *Jugendschutz in den Bereichen Film und Videospiele*. https://www.bsv.admin.ch/bsv/de/home/sozialpolitische-themen/kinder-und-jugendfragen/jugendschutz/jugendschutz-bei-filmen-und-videospielen.html

Bundesgesetzblatt. (2015, January 26). Neunundvierzigstes Gesetz zur Änderung des Strafgesetzbuches – Umsetzung europäischer Vorgaben zum Sexualstrafrecht. https://www.bundesgerichtshof.de/SharedDocs/Downloads/DE/Bibliothek/Gesetzesmaterialien/18_wp/StrAendG_SexualstrafR/bgbl.pdf;jsessionid=C262D38D71F32430A8F64561A02C41C3.2_cid359?__blob=publicationFile

Bundeskriminalamt. (n.d.). Cybergrooming. https://www.bka.de/DE/UnsereAufgaben/Aufgabenbereiche/Zentralstellen/Kinderpornografie/Cybergrooming/Cybergrooming_node.html

Bundesministerium für Familie, Senioren, Frauen und Jugend. (2021). *Reform des Jugendschutzgesetzes tritt in Kraft*. https://www.bmfsfj.de/bmfsfj/aktuelles/alle-meldungen/reform-des-jugendschutzgesetzes-tritt-in-kraft-161184.

Bundesministerium Wohnen, Kunst, Kultur, Medien und Sport. (2025). Klare Regeln für Social-Media gefordert – EU-Kommission soll konkreten Gesetzesvorschlag für mehr Schutz für Jugendliche vorlegen. https://www.bmwkms.gv.at/themen/aktuell/social-media-eu.html

Bundestag. (2019, June 26). *Gesetz zur Änderung des Strafgesetzbuches – Versuchsstrafbarkeit des Cybergroomings* (BT-Drs. 19/13836). https://dip.bundestag.de/vorgang/gesetz-zur-%C3%A4nderung-des-strafgesetzbuches-versuchsstrafbarkeit-des-cybergroomings/251731

Bundeszentrale für Kinder- und Jugendmedienschutz (BzKJ). (n. d.). *Aufgaben der Bundeszentrale für Kinder- und Jugendmedienschutz*. https://www.bzkj.de/bzkj/ueberuns/aufgaben

Burdick, J., & Sandlin, J. A. (2017). Social media and youth activism: Networked publics and the fight for racial justice. *Journal of Adolescent Research*, 32(5), 507–530. https://doi.org/10.1177/0743558417712013

Center for Countering Digital Hate. (2022). *Deadly by design: TikTok pushes harmful content promoting eating disorders and self-harm into young users' feeds*. https://counterhate.com/research/deadly-by-design/

Chen, M. L. (2023, September 7). *How screen time is affecting teens' sleep and mental health*. World Economic Forum. https://www.weforum.org/stories/2023/09/screen-time-affecting-sleep-mental-health

Childlight Global Child Safety Institute. (2024). *Over 300 million children are victims of online sexual exploitation and abuse*. https://www.childlight.org/newsroom/over-300-million-children-a-year-are-victims-of-online-sexual-exploitation-and-abuse

Citizens Information. (2025). Online safety. https://www.citizensinformation.ie/en/consumer/buying-digital-content-and-services/online-safety

CNIL. (2021a). CNIL *publishes 8 recommendations to enhance the protection of children online*. https://www.cnil.fr/en/cnil-publishes-8-recommendations-enhance-protection-children-online

CNIL. (2021b). *Recommendation 1: Regulate the capacity of children to act online*. https://www.cnil.fr/en/recommendation-1-regulate-capacity-children-act-online

CNIL. (2021c). *Recommendation 2: Encourage children to exercise their rights*. https://www.cnil.fr/en/recommendation-2-encourage-children-exercise-their-rights

CNIL. (2021d). Recommendation 3: *Support parents in digital education*. https://www.cnil.fr/en/recommendation-3-support-parents-digital-education

CNIL. (2021e). *Recommendation 4: Seek parental consent for children under 15*. https://www.cnil.fr/en/recommendation-4-seek-parental-consent-children-under-15

CNIL. (2021f). *Recommendation 5: Promote parental controls that respect the child's privacy and best interests*. https://www.cnil.fr/en/recommendation-5-promote-parental-controls-respect-childs-privacy-and-best-interests

CNIL. (2021g). *Recommendation 6: Strengthen information and rights of children by design*. https://www.cnil.fr/en/recommendation-6-strengthen-information-and-rights-children-design

CNIL. (2021h). *Recommendation 7: Check the age of the child and parental consent while respecting the child's privacy*. https://www.cnil.fr/en/recommendation-7-check-age-child-and-parental-consent-while-respecting-childs-privacy

CNIL. (2021i). *Recommendation 8: Provide specific safeguards to protect the interests of the child*. https://www.cnil.fr/en/recommendation-8-provide-specific-safeguards-protect-interests-child

Cobb, S. (2021). It's COPPA-cated: Protecting Children's Privacy in the Age of YouTube. *Houston Law Review*, 58(4), 965–986. https://houstonlawreview.org/article/22277-it-s-coppa-cated-protecting-children-s-privacy-in-the-age-of-youtube

Coimisiún na Meán. (n. d.). Online Safety Framework. https://www.cnam.ie/general-public/online-safety/online-safety-framework

Common Sense Media. (2024). A Double-Edged Sword: How Diverse Communities of YoungPeople Think About the Multifaceted Relationship Between Social Media and Mental Health. https://www.commonsensemedia.org/sites/default/files/research/report/2024-double-edged-sword-hopelab-report_final-release-for-web-v2.pdf

Council of Europe. (2001). *Convention on Cybercrime* (*Budapest Convention*) (ETS No. 185). https://rm.coe.int/1680081561

Council of Europe. (2018). *Leitlinien zur Achtung, zum Schutz und zur Verwirklichung der Rechte des Kindes im digitalen Umfeld. Empfehlung CM/Rec(2018)7 des Ministerkomitees an die Mitgliedstaaten.* https://rm.coe.int/168092dd25

Council of Europe. (2024). *Information Note – The Council of Europe Convention on the Protection of Children agains Sexual Exploitation and Sexual Abuse.* https://rm.coe.int/information-note-the-council-of-europe-convention-on-the-protection-of/16807962a7%20#portlet_com_liferay_journal_content_web_portlet_JournalContentPortlet_INSTANCE_fYB7DYRDaV4M

Council of Senators: Blackburn & Blumenthal. (2025, May 14). *Blackburn, Blumenthal, Thune, and Schumer Introduce the Kids Online Safety Act* [Press release]. U.S. Senate. https://www.blackburn.senate.gov/2025/5/technology/blackburn-blumenthal-thune-and-schumer-introduce-the-kids-online-safety-act

Council of the European Union. (2021, September 15). *Interinstitutional file: 2021/0293(COD)* (proposal COM 2021/574 final). https://data.consilium.europa.eu/doc/document/ST-11900-2021-INIT/en/pdf

Council of the European Union. (2024). *The general data protection regulation* (*GDPR*)*: What is the GDPR and how does it apply?*. https://www.consilium.europa.eu/en/policies/data-protection-regulation/

Crespo M., Kallet K., Newmark J., Rubin, A. (2025). Just a Minor Threat: Online Safety Legislation takes off, *Socially Aware*, https://www.sociallyawareblog.com/topics/just-a-minor-threat-online-safety-legislation-takes-off

Cox Communications (2012). *Tween Internet safety survey*, http://ww2.cox.com/wcm/en/aboutus/datasheet/takecharge/tween-Internet-safety-survey.pdf

da Silva Gioia, R. (2023, December 28). *KI kann auch Bilder von sexuellem Kindsmissbrauch generieren: Ein Albtraum für den Kinderschutz – und die Polizei.* NZZ am Sonntag. https://www.nzzas.nzz.ch/magazin/kinderpornografie-im-dunklen-reich-der-digitalen-triebtaeter-ld.1587830

DAK Forschung. (2018). *WhatsApp, Instagram und Co.* – So süchtig macht Social Media. *DAK-Studie: Befragung von Kindern und Jugendlichen zwischen 12 und 17 Jahren.* https://www.schau-hin.info/fileadmin/content/Downloads/Sonstiges/dak-studie-sucht-nach-sozialen-medien.pdf

Dansac Le Clerc, M., & Leportois, J. (2024, March 19). *France introduces new law to enhance the protection of children's rights in France. Connect on Tech.* https://connectontech.bakermckenzie.com/france-introduces-new-law-to-enhance-the-protection-of-childrens-rights-in-france

Denham, E., & Wood, S. (2023). *Children's privacy laws and freedom of expression: Lessons from the UK Age-Appropriate Design Code*. International Association of Privacy Professionals (IAPP). https://iapp.org/news/a/childrens-privacy-laws-and-freedom-of-expression-lessons-from-the-uk-age-appropriate-design-code

Department for Science, Innovation and Technology. (2024, May 8). *Online Safety Act: explainer*. https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer

Der Bundesrat. (2022, September 30). *Neues Bundesgesetz über den Jugendschutz in den Bereichen Film und Videospiele*. https://www.news.admin.ch/de/nsb?id=90545

Detmering, F., & Splittgerber, A. (2018). DSGVO und nationale Umsetzungsgesetze. *digma*, 18, 172–173.

Deutsches Kinderhilfswerk. (n. d.). Children's rights in the digital world. https://www.dkhw.de/informieren/unsere-themen/kinder-und-medien/kinderrechte-in-der-digitalen-welt/

Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (BfDI). (n. d.). *Die geplante EU-Verordnung zur Prävention und Bekämpfung des sexuellen Missbrauchs von Kindern – die sogenannte „Chatkontrolle"*. https://www.bfdi.bund.de/DE/Fachthemen/Inhalte/Telemedien/CSA_Verordnung.html

Döring, N. (2005). Sexting – Fakten und Fiktionen über den Austausch erotischer Handyfotos unter Jugendlichen. Merz. *Medien + Erziehung*, 56(5), 47–52.

Edwards, G., Christensen, L. S., Rayment-McHugh, S., & Jones, C. (2021, September). *Cyber strategies used to combat child sexual abuse material* (Trends & issues in crime and criminal justice, No. 636). Australian Institute of Criminology. https://www.aic.gov.au/sites/default/files/2021-09/ti636_cyber_strategies_used_to_combat_csam.pdf

Eidgenössische Kommission Gegen Rassismus (EKR). (n.d.). Internet und Soziale Netzwerke. https://www.rechtsratgeber-rassismus.admin.ch/lebensbereiche/d206.html

Electronic Frontier Foundation. (2024). Concerns over KOSA's potential for over-censorship. https://www.eff.org/issues/kosa

European Commission, Directorate-General for Communications Networks, Content and Technology. (2023). *The Digital Services Act (DSA) explained – Measures to protect children and young people online*. Publications Office of the European Union. https://data.europa.eu/doi/10.2759/576008

European Commission: Directorate-General for Communications Networks, Content and Technology. (2022). *How to make Europe's Digital Decade fit for children and young people? – A report from the consultation with children and young people*. Publications Office of the European Union. https://data.europa.eu/doi/10.2759/096742

European Commission: Directorate-General for Communications Networks, Content and Technology, Center for Law and Digital Technologies (eLaw), Leiden University Leiden The Netherlands, LLM, Raiz Shaffique, M., & van der Hof, S. (2024). *Mapping age assurance typologies and requirements*. Publications Office of the European Union. https://data.europa.eu/doi/10.2759/23620

European Commission. (2012). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: European strategy for a better internet for children* (COM(2012) 196 final). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0196

European Commission. (2020, July 24). COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS: EU *strategy for a more effective fight against child sexual abuse* (COM 2020/607 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52020DC0607

European Commission. (2021). *Consultation on children's rights in the digital environment: Summary report. Conducted by European Schoolnet.* https://www.betterinternetforkids.eu/en/policy/childrens-consultation

European Commission. (2022a). *European strategy for a better internet for kids – BIK+.* https://digital-strategy.ec.europa.eu/en/policies/strategy-better-internet-kids

European Commission. (2022b, May 11). *Proposal for a* REGULATION *of the European Parliament and of the Council laying down rules to prevent and combat child sexual abuse* (COM 2022/209 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52022PC0209

European Commission. (2023, April 25). *Gesetz über digitale Dienste: EU-Kommission benennt erste Gruppe von sehr großen Online-Plattformen und Suchmaschinen.* https://germany.representation.ec.europa.eu/news/gesetz-uber-digitale-dienste-eu-kommission-benennt-erste-gruppe-von-sehr-grossen-online-plattformen-2023-04-25_de

European Commission. (2024, August 19). *Commission launches call for evidence for guidelines on protecting minors under the Digital Services Act.* https://digital-strategy.ec.europa.eu/en/news/commission-launches-call-evidence-guidelines-protection-minors-online-under-digital-services-act

European Commission. (n. d.-a). *Digital Services Act (A Europe fit for the Digital Age).* https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_de

European Commission. (n. d.-b). *Audiovisual and Media Services Directive* – AVMSD. https://digital-strategy.ec.europa.eu/en/policies/audiovisual-and-media-services

European Commission. (n. d.-c). *European Digital Rights and Principles.* https://digital-strategy.ec.europa.eu/en/policies/digital-principles

European Parliament & Council. (2011, December 13). *Directive 2011/92/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing*

Council Framework Decision 2004/68/JHA. https://eur-lex.europa.eu/eli/dir/2011/92/oj/eng

European Parliament & Council. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR)*, Art. 4(1). Official Journal of the European Union, L 119, 1–88. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679

European Parliament & Council. (2022, October 19). *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act). Official Journal of the European Union*, L 277, 1–102. https://eur-lex.europa.eu/eli/reg/2022/2065/oj

European Parliament & Council. (2024, June 13). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European* Union, L 1689, 1–. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

European Parliament. (2020). *Curbing the surge in online child abuse: The dual role of digital technology in fighting and facilitating its proliferation* (Briefing No. 659360). https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/659360/EPRS_BRI(2020)659360_EN.pdf

European Parliament. (2024, November). *Criminalisation of Hate Speech and Hate Crime in selected EU countries*. https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/766226/EPRS_BRI(2024)766226_EN.pdf#:~:text=Directive%20(EU)%202024/1385%20of%2014%20May%202024,by%20publicly%20disseminating%20material%20containing%20incitement%20online

European Parliament. (2025). *EU strategy for a more effective fight against child sexual abuse* (Legislative Train Schedule file spotlight-JD 23-24). European Parliament. https://www.europarl.europa.eu/legislative-train/spotlight-JD%2023-24/file-eu-strategy-to-fight-child-sexual-abuse

European Union. (2021). *EU strategy on the rights of the child and the European child guarantee*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=legissum:4540916

European Union. (2025a). *Safer Internet Day (SID) 2025. Better Internet for Kids*. https://better-internet-for-kids.europa.eu/en/events/safer-internet-day-sid-2025

European Union. (2025b). FSM. https://better-internet-for-kids.europa.eu/en/safer-internetday/supporter-listing/fsm

Eurostat. (n. d.). *Young people – digital world*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Young_people_-_digital_world

Eyal, N., & Hoover, R. (2014). *Hooked: How to build habit-forming products*. Portfolio/Penguin.

Fankhauser, R., & Fischer, N. (2017). Kinderfotos auf Facebook oder wenn die Eltern die Persönlichkeitsrechte ihrer Kinder verletzen. In M. A. Besson, A. Rumo-Jungo & V. Vetter-Schreiber (Eds.), *Brennpunkt Familienrecht. Festschrift für Thomas Geiser zum 65. Geburtstag* (S. 205). Schulthess.

Fatt, S. J., & Fardouly, J. (2023, December). Digital social evaluation: Relationships between receiving likes, comments, and follows on social media and adolescents' body image concerns. *Body Image*, 47, 101621. https://doi.org/10.1016/j.bodyim.2023.101621

Federal Trade Commission. (2019a, November 22). *YouTube channel owners: Is your content directed to children?* https://www.ftc.gov/business-guidance/blog/2019/11/youtube-channel-owners-your-content-directed-children

Federal Trade Commission. (2019b, September 4). *$170 million FTC-NY YouTube settlement offers COPPA compliance tips for platforms and providers* [Press release]. https://www.ftc.gov/news-events/blogs/business-blog/2019/09/170-million-ftc-ny-youtube-settlement-offers-coppa

Federal Trade Commission. (2023). *Children's Online Privacy Protection Rule* (COPPA), 16 C.F.R. § 312.1 *et seq.* https://www.ecfr.gov/current/title-16/chapter-I/subchapter-C/part-312

Feierabend, S., Rathgeb, T., Kheredmand, H., & Glöckler, S. (2023). *JIM-Studie 2023: Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. Medienpädagogischer Forschungsverbund Südwest. https://www.mpfs.de/studien/jim-studie/2023/

Feiler, L., & Forgó, N. (2024). *KI-VO, EU-Verordnung über künstliche Intelligenz*. Verlag Österreich.

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann Publishers

Forum Verlag Herkert GmbH. (2017). Das Netzwerkdurchsetzungsgesetz (NetzDG) ist in Kraft. https://www.forum-verlag.com/fachwissen/datenschutz-und-it-sicherheit/netzwerkdurchsetzungsgesetz/

Fry, D. (2024, June 4). *Hidden pandemic: study uncovers scale of online child harm*. University of Edinburgh – Edinburgh Impact. https://impact.ed.ac.uk/research/future-health-and-care/hidden-pandemic-study-uncovers-scale-of-online-child-harm

Gassó A.M., Mueller-Johnson K. & Montiel I. Sexting, Online Sexual Victimization, and Psychopathology Correlates by Sex: Depression, Anxiety, and Global Psychopathology. *Int J Environ Res Public Health*. 2020 Feb 6;17(3):1018. doi: 10.3390/ijerph17031018. PMID: 32041115; PMCID: PMC7036947

Gessner Legal. (n. d.). *Das Recht am eigenen Bild von Kindern – was man wissen sollte.* https://www.rechtsanwalt-gessner-berlin.de/recht-am-eigenen-bild-von-kindern/

Gesellschaft für Freiheitsrechte. (2023). *Chatkontrolle. Der Verordnungsentwurf der EU-Kommission und seine Unvereinbarkeit mit der Europäischen Grun-*

drechtecharta. https://www.bundestag.de/resource/blob/935250/7ecae89c214ef74dc6e40bd922c854e9/Stellungnahme-Reda.pdf

Gesellschaft für Freiheitsrechte. (n. d.). *Chatkontrolle: Mit Grundrechten unvereinbar*. https://freiheitsrechte.org/themen/freiheit-im-digitalen/chatkontrolle

Gesellschaft für Informatik. (2024). *GI-Arbeitskreis warnt vor EU-Plänen zur Chatkontrolle*. https://gi.de/meldung/gi-arbeitskreis-warnt-vor-eu-plaenen-zur-chatkontrolle

GFI Software. (2011). *2011 parent-teen Internet safety report*. https://www.gfi.com/pages/parent-teen-internet-safety-report?sc_lang=ru-ru

Google. (2023). *How we detect, remove and report child sexual abuse material*. Google Blog. https://blog.google/intl/en-au/company-news/outreach-initiatives/how-we-detect-remove-and-report-child-sexual-abuse-material/

Google. (2024). *Transparency Report*. https://transparencyreport.google.com/child-sexual-abuse-material/reporting?hl=en

Gorwa, R., & Emmert, M. (2021). The governance of online hate speech: A systematic literature review. *Social Media + Society*, 7(4).

Government of the Netherlands. (2020, June 9). *Speech by Minister of Justice and Security Ferdinand Grapperhaus on EU action to combat child sexual abuse*. https://www.government.nl/documents/speeches/2020/06/09/speech-by-minister-of-justice-and-security-ferdinand-grapperhaus-on-eu-action-to-combat-child-sexual-abuse-9-june-2020

Government of the Netherlands. (2024, June 5). *Authority may require hosting companies to remove online child sexual abuse material from servers*. https://www.government.nl/topics/child-abuse/news/2024/06/05/authority-may-require-hosting-companies-to-remove-online-child-sexual-abuse-material-from-servers

Griffiths, M. D. (2005). A 'components' model of addiction within a biopsychosocial framework. *Journal of Substance Use*, 10(4), 191–197. https://doi.org/10.1080/14659890500114359

Günal Rütsche, S., Oberlin, J. & von Hoyningen-Huene, S. (2025, June 16). Die neue Meldepflicht für Cyberangriffe – ein Vorbild für den Umgang mit pädokriminellen Inhalten. *Jusletter*. https://jusletter.weblaw.ch/juslissues/2025/1244/die-neue-meldepflich_b5e295cadf.html__ONCE&login=false

Harding-Zentrum für Risikokompetenz. (2025, June). *Unstatistik des Monats: Falsch positive Chatkontrolle*. https://www.hardingcenter.de/de/unstatistik/unstatistik-des-monats-falsch-positive-chatkontrolle

Hellenic Republic, Ministry of Digital Governance (2025), *National Strategy for the Protection of Minors from Internet Addiction*, March 2025, https://www.mindigital.gr/wp-content/uploads/2025/03/%CE%9DStrategy_ProtectM_long_ENG.pdf?

Hern, A., & Milmo, D. (2022, December 15). TikTok self-harm study results 'every parent's nightmare'. *The Guardian*. https://www.theguardian.com/technology/2022/dec/15/tiktok-self-harm-study-results-every-parents-nightmare

Hogan Lovells. (2024, January 23). *The sorcerer's apprentice conundrum: Generative AI content under the EU DSA and UK Online Safety Act.* https://www.lexology.com/library/detail.aspx?g=14cd1786-2587-49da-9a30-a784ef3c6d4e.

Husi-Stämpfli, S. (2025, April 7). 25 *Jahre Children's Online Privacy Protection Act*, Jusletter. https://jusletter.weblaw.ch/juslissues/2025/1236/25-jahre-children-s-_aaffb58b03.html__ONCE&login=false

Husi-Stämpfli, S., Oberlin, J. S., & von Hoyningen-Huene, S. (2024, November 4). *Instagram-Teen-Accounts: Hoffnung für den Kinderschutz*? Jusletter. https://jusletter.weblaw.ch/juslissues/2024/1217/instagram-teen-accou_bb60f3f347.html

Independent Federal Commissioner against Child Sexual Abuse. (n.d.). *International and European Law.* https://beauftragte-missbrauch.de/en/themen/recht/international-and-european-law

Information Commissioner's Office. (n. d.). *Introduction to the Children's code.* https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/introduction-to-the-childrens-code/

Insoll, T., Ovaska, A. K., & Vaaranen-Valkonen, N. (2021). *CSAM Users in the Dark Web: Protecting Children Through Prevention (ReDirection Survey Report).* Protect Children / Suojellaan Lapsia ry. https://www.suojellaanlapsia.fi/en/post/csam-users-in-the-dark-web-protecting-children-through-prevention

International Justice Mission (IJM). (2020). Online sexual exploitation of children in the Philippines: Analysis and recommendations for governments, industry, and civil society (Full report). https://ijmstoragelive.blob.core.windows.net/ijmna/documents/Final-Public-Full-Report-5_20_2020.pdf

Internet Watch Foundation. (2020, May 1). *Pixels from a Crime Scene Podcast: Episode 5 – The industry can't solve it alone* [Podcast]. https://www.youtube.com/watch?v=YBdmeeCEenI

Internet Watch Foundation. (2021). *IWF Annual Report 2020 – Face the Facts.* https://www.iwf.org.uk/about-us/who-we-are/annual-report-2020/

Internet Watch Foundation. (2022). *Europe remains 'global hub' for hosting of online child sexual abuse material.* https://www.iwf.org.uk/news-media/news/europe-remains-global-hub-for-hosting-of-online-child-sexual-abuse-material/

Internet Watch Foundation. (2023). *EU still hosts the most child sexual abuse material in the world.* https://www.iwf.org.uk/news-media/news/eu-still-hosts-the-most-child-sexual-abuse-material-in-the-world/

Internet Watch Foundation. (n. d.). *Internet Watch Foundation.* https://www.iwf.org.uk/

INTERPOL. (n. d.). *International Child Sexual Exploitation database.* https://www.interpol.int/Crimes/Crimes-against-children/International-Child-Sexual-Exploitation-database

Intersoft Consulting Service – Dr. Datenschutz. (2025, January 17). Das Netzwerkdurchsetzungsgesetz (NetzDG) im Rückblick. https://www.dr-datenschutz.de/das-netzwerkdurchsetzungsgesetz-netzdg-im-rueckblick/.

Jackson, S. J., Bailey, M., & Foucault Welles, B. (2020). *#HashtagActivism: Networks of race and gender justice*. MIT Press.

Joleby, M., Löfgren-Mårtenson, L., Månsson, S. A., & Plantin, L. (2020). "All of me is completely different": Experiences and consequences among victims of technology-assisted child sexual abuse. Frontiers in Psychology, 11, Article 606218. https://doi.org/10.3389/fpsyg.2020.606218

Jugend und Medien. (2015). *Zukünftige Ausgestaltung des Kinder- und Jugendmedienschutzes der Schweiz: Bericht des Bundesrates in Erfüllung der Motion Bischofberger* 10.3466 "*Effektivität und Effizienz im Bereich Jugendmedienschutz und Bekämpfung von Internetkriminalität*". https://www.jugendundmedien.ch/

Jugend und Medien. (n. d.). *Extremismus und Radikalisierung: Jugendliche im Internet zunehmend mit digitalen Gewaltformen, radikalen Inhalten und antidemokratischen Strategien konfrontiert*. https://www.jugendundmedien.ch/extremismus-radikalisierung

Kattenberg, T. (2024). Cybergrooming – eine Bestandsaufnahme und zwei Schlussfolgerungen. https://www.kriminologie.de/index.php/krimoj/article/download/349/206/1084

Kaye, R. (2010, October 7). *How a cell phone picture led to girl's suicide*. CNN Living. https://edition.cnn.com/2010/LIVING/10/07/hope.witsells.story/index.html

Klicksafe. (2024, June 25). Schockinhalte im Internet – Richtig reagieren auf verstörende Gewaltdarstellung im Netz. https://www.klicksafe.de/news/richtig-reagieren-auf-verstoerende-gewaltdarstellung-im-netz

Klicksafe. (2023, October 12). Hate Speech – Rechtslage. https://www.klicksafe.de/hate-speech/rechtslage

Knobloch-Westerwick, S., & Westerwick, A. (2023). Algorithmic personalization of source cues in the filter bubble: Self-esteem and self-construal impact information exposure. *Journal of Computer-Mediated Communication*, 28(2), 94–112. https://doi.org/10.1177/14614448211027963

Kommission für Jugendmedienschutz (KJM). (n. d.). *Über uns*. https://www.kjm-online.de/ueber-uns

Kriechbaumer, T. & Nath, C. (2015). The darknet and online anonymity (POSTNOTE No. 488). Parliamentary Office of Science and Technology, UK Parliament. https://researchbriefings.files.parliament.uk/documents/POST-PN-488/POST-PN-488.pdf

Külling-Knecht, C., Willemse, I., Deda-Bröchin, S., & Streule, P. (2024). JAMES – *Jugend, Aktivitäten, Medien: Erhebung Schweiz 2024 – Ergebnisbericht*. https://www.zhaw.ch/storage/psychologie/upload/forschung/medienpsychologie/james/2018/JAMES_2024_DE.pdf

Kunz, C., Oberlin, J., & von Hoyningen-Huene, S. (2024). Kinder und soziale Medien – der Datenschutz bleibt gefordert. In Datenschutzforum Schweiz (Eds.), *Persönlichkeitsschutz zwischen Mensch und Maschine: Festschrift zum 25-jährigen Bestehen des Datenschutzforums Schweiz* (pp. 43–66). Schulthess Verlag.

Livingstone, S. & Haddon, L. (2009). EU Kids Online. *Zeitschrift für Psychologie – Journal of Psychology*, 217, 236-239

Livingstone, S., Haddon L. & Görzig, A. (2012). Children, Risk and Safety on the Internet: Research and Policy Challenges in Comparative Perspective. *Policy Press*, 1-14.

Livingstone, S., & Stoilova, M. (2021). The 4Cs: Classifying Online Risk to Children. (CO:RE Short Report Series

on Key Topics). Hamburg: Leibniz-Institut für Medienforschung/Hans-Bredow-Institut (HBI); CO:RE – Children Online: Research and Evidence, abrufbar unter: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5df252f14&appId=PPGMS

Ma, R., Zhang, Z., Gui, X., & Kou, Y. (2024). Labeling in the dark: Exploring content creators' and consumers' experiences with content classification for child safety on YouTube. In A. Vallgårda, L. Jönsson, J. Fritsch, S. Fdili Alaoui, & C. A. Le Dantec (Eds.), *Proceedings of the Designing Interactive Systems Conference* (DIS '24) (pp. 1518–1532). Association for Computing Machinery. https://doi.org/10.1145/3643834.3661565

Madriaza, P., Hassan, G., Brouillette-Alarie, S., Mounchingam, A. N., Durocher-Corfa, L., Borokhovski, E., Pickup, D., & Paillé, S. (2025). Exposure to hate in online and traditional media: A systematic review and meta-analysis of the impact of this exposure on individuals and communities. *Campbell Systematic Reviews*, 21(1), e70018. https://doi.org/10.1002/cl2.70018

Malamud, O., Cueto, S., Cristia, J. P., & Beuermann, D. W. (2019). Do children benefit from internet access? Experimental evidence from Peru. *Journal of Development Economics*, 138, 41–56. https://doi.org/10.1016/j.jdeveco.2018.11.005

Maniglio, R. (2013). Child sexual abuse in the etiology of anxiety disorders: A systematic review of reviews. *Trauma, Violence, & Abuse*, 14(2), 96–112. https://doi.org/10.1177/1524838012470032

Manzoni, P., Baier, D., Kamenowski, M., Isenhardt, A., Haymoz, S., Jacot, C. (2019). *Einflussfaktoren extremistischer Einstellungen unter Jugendlichen in der Schweiz.* zhaw Soziale Arbeit. https://doi.org/10.21256/zhaw-18673

Martinez Sainz, G., & Hanna, A. (2023). Youth digital activism, social media and human rights education: The Fridays for Future movement. *Human Rights Education Review*, 6(1), 116–136. https://doi.org/10.7577/hrer.4958

May, M. (2018). Hate Speech analog – Eine situative Herausforderung in Schule und Unterricht. *GWP – Gesellschaft. Wirtschaft. Politik*, 67(3), 399–408. https://doi.org/10.3224/gwp.v67i3.12

Meshi, D., Morawetz, C., & Heekeren, H. R. (2013). Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Frontiers in Human Neuroscience*, 7(2013), Article 439. https://doi.org/10.3389/fnhum.2013.00439

Meta. (2018, October). *New technology to fight child exploitation.* https://about.fb.com/news/2018/10/fighting-child-exploitation/

Meta. (2021a, February). *Understanding the intentions of child sexual abuse material* (CSAM) *sharers.* https://research.facebook.com/blog/2021/02/understanding-the-intentions-of-child-sexual-abuse-material-csam-sharers

Meta. (2021b, February 23). *Preventing Child Exploitation on Our Apps.* https://about.fb.com/news/2021/02/preventing-child-exploitation-on-our-apps/

Meta. (n. d.-a). *Was sind die Meta-Produkte?.* https://www.facebook.com/legal/meta-products

Meta. (n. d.-b). *Child sexual exploitation, abuse and nudity. Policy details.* https://transparency.meta.com/en-gb/policies/community-standards/child-sexual-exploitation-abuse-nudity/

Mitchell at al., (2003). The Exposure Of Youth To Unwanted Sexual Material On The Internet: A National Survey of Risk, Impact, and Prevention, *Youth & Society* 34, 330-358 accessed under: https://www.researchgate.net/publication/237257967_The_Exposure_Of_Youth_To_Unwanted_Sexual_Material_On_The_Internet_A_National_Survey_of_Risk_Impact_and_Prevention/citations

Montag, C., Sindermann, C., & Baumeister, H. (2020). Digital phenotyping in psychological and medical sciences: A reflection about necessary prerequisites to reduce harm and increase benefits. *Current Opinion in Psychology*, 36, 19–24. https://doi.org/10.1016/j.copsyc.2020.03.013

Müller-Johnson, K. (2018). Cyber-Grooming. In J. Gysi & P. Rüegger (Hrsg.), Handbuch sexualisierte Gewalt: Therapie, Prävention und Strafverfolgung (1. Auflage, S. 63–72). Hogrefe.

National Center for Missing & Exploited Children. (2024). *2023 CyberTipline® report. National Center for Missing & Exploited Children.* https://www.missingkids.org/content/dam/missingkids/pdfs/2023-CyberTipline-Report.pdf

National Center for Missing and Exploited Children. (n. d.) *National Center for Missing and Exploited Children.* https://www.missingkids.org/home

Negreiro Achiaga, M. D. M. (2020). *Curbing the surge in online child abuse: The dual role of digital technology in fighting and facilitating its proliferation* (Briefing No. 659.360). European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/659360/EPRS_BRI(2020)659360_EN.pdf

Netzwerk Kinderrechte Schweiz. (2023, September 22). "Das Recht am eigenen Bild muss auch innerhalb der Familie gewahrt werden": Interview with Sandra Husi-Stämpfli. *Netzwerk Kinderrechte Schweiz.* https://www.netzwerk-kinderrechte.ch/aktuell/2023/sandra-husi-staempfli-das-recht-am-eigenen-bild-muss-auch-innerhalb-der-familie-gewahrt-werden

Ngo, V. M., Gajula, R., Thorpe, C., & McKeever, S. (2024). Discovering child sexual abuse material creators' behaviors and preferences on the dark web. *Child Abuse & Neglect*, 147(January 2024), 106558. https://doi.org/10.1016/j.chiabu.2023.106558

Oberlin, J. & von Hoyningen-Huene, S. (2022). Innocence in Danger. Wenn Likes auf Kosten der Kinder gehen. *Schweizerische Juristen-Zeitung*, 23, 1123–1140

Oberlin, J., & von Hoyningen-Huene, S. (2025, January). Missbrauchspotentiale von KI im Kontext von Kinderpornografie: Eine juristische Analyse. *IusNet Digitales Recht und Datenrecht*, (2025).

Oberlin, J., & von Hoyningen-Huene, S. (2025, August). KI und Kinderpornografie, Eine verhängnisvolle Entwicklung, *Zeitschrift für Strafrecht*, 3(2025), 337-378

Oberlin, J., & von Hoyningen-Huene, S. (2025, November). Kinderpornografie im Lichte der KI-VO, MMR *Beck*, 11(2025), 867.

Oberlin, J., von Hoyningen-Huene, S., & Fassbind, P. (2024). Kindeswohlgefährdungen und Kindesschutz im Metaverse. ZKE-RMA, 2(2024), 78.

OECD. (2020). *Protecting children online: An overview of recent developments in legal frameworks and policies* (*OECD Digital Economy Papers*, No. 295). OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/06/protecting-children-online_0c385619/9e0e49a9-en.pdf

OECD. (2021, May 31). Recommendation of the Council on Children in the Digital Environment. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0389?_ga=2.138548851.1597661473.1707343531-2141752145.1706178830

Ofcom. (2024, December 16). Time for tech firms to act: UK online safety regulation comes into force. https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/time-for-tech-firms-to-act-uk-online-safety-regulation-comes-into-force

Ofcom. (2025, April 24). Children's Risk Assessment Guidance and Children's Risk Profiles. https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/statement-protecting-children-from-harms-online/main-document/childrens-risk-assessment-guidance-and-childrens-risk-profiles.pdf?v=396653

Papalia, N., Mann, E., & Ogloff, J. R. P. (2021). Child sexual abuse and risk of revictimization: Impact of child demographics, sexual abuse characteristics, and psychiatric disorders. *Child Maltreatment*, 26(1), 74–86. https://doi.org/10.1177/1077559520932665

Parti, K., & Szabó, J. (2024). The legal challenges of realistic and AI-driven child sexual abuse material: Regulatory and enforcement perspectives in Europe. *Laws*, 13(6), 67. https://doi.org/10.3390/laws13060067

Paul, K. (2025, February 16). Parents are desperate to protect kids on social media. Why did the US let a safety bill die? *The Guardian*. https://www.theguardian.com/us-news/2025/feb/16/kids-social-media-online-safety-act

Pellegrino, A., Stasi, A., & Bhatirasevi, V. (2022, November 10). Research trends in social media addiction and problematic social media use: A bibliometric analysis. *Frontiers in Psychiatry*, 13, Article 1017506. https://doi.org/10.3389/fpsyt.2022.1017506

Pingen, A. (2023, February 8). *European Declaration on Digital Rights and Principles Signed*. eucrim. https://eucrim.eu/news/european-declaration-on-digital-rights-and-principles-signed/

Quashie, J.-M. (2024, September 17). *The protection of children's privacy in France: A reform of image rights law*. Humanium. https://www.humanium.org/de/der-schutz-der-privatsphaere-von-kindern-in-frankreich-eine-gesetzesreform-bezueglich-des-rechts-am-eigenen-bild/

Ramaswamy, S., & Seshadri, S. (2020). Children on the brink: Risks for child protection, sexual abuse, and related mental health problems in the COVID-19 pandemic. *Indian Journal of Psychiatry*, 62(7), 404–413. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_1032_20

Reschke, S. (2020). Strafbarkeit des Cybergroomings, Kripoz. https://kripoz.de/wp-content/uploads/2021/05/reschke-strafbarkeit-des-cyber-grooming.pdf

Rücker, D. (2018). Development and importance of the data protection reform. In D. Rücker & T. Kugler (Eds.), *New European General Data Protection Regulation*.

Rüdiger, T.-G. (2020). Die onlinebasierte Anbahnung des sexuellen Missbrauchs eines Kindes: Eine kriminologische und juristische Auseinandersetzung mit dem Phänomen Cybergrooming. Frankfurt.

Safe Online. (2024, April 9). International Association of Internet Hotlines (INHOPE). https://safeonline.global/international-association-of-internet-hotlines-inhope/

Saldías, B. (2024). *Designing child-centered content exposure and moderation*. arXiv. https://doi.org/10.48550/arXiv.2406.08420

Saleem, J., Islam, R., & Kabir, M. A. (2022). The anonymity of the dark web: A survey. IEEE Access, 10, 33628–33660. https://doi.org/10.1109/ACCESS.2022.3161547

Sayyed, H., & Paul, S. R. (2025, March 17). Exploring the role of encryption and the dark web in cyber terrorism: Legal challenges and countermeasures in India. *Cogent Social Sciences*, 11(1), Article 2479654. https://doi.org/10.1080/23311886.2025.2479654

Schweitzer, J. (2024, February 27). "Soziale Medien können ein Trigger für Essstörungen sein": Interview with Katrin Giel. ZEIT ONLINE. https://www.zeit.de/gesundheit/2024-02/social-media-essstoerung-jugendliche-koerperbild-studie

Schweizerische Kriminalprävention. (n. d.). Sextortion (Erpressung mit Nacktbildern). https://www.skppsc.ch/de/themen/internet/sextortion-erpressung/

Sempach, N. (2023). Strafbarkeit von Gewaltaufrufen in sozialen Medien. *medialex*, 09/2023. https://doi.org/10.52480/ml.23.13

Shared Hope International (2014). *FAQs about child sex trafficking*. http://sharedhope.org/learn/faqs/

Sherman, L. E., Greenfield, P. M., Hernandez, L. M., & Dapretto, M. (2016, August). The power of the like in adolescence: Effects of peer influence on neural and behavioral responses to social media. *Psychological Science*, 27(7), 1027–1035. https://doi.org/10.1177/0956797616645673

Simone, C. (2024). When parents decide that all the world's a stage: Expanding publicity rights to protect children in monetized social media content. *Columbia Journal of Law and Social Problems*, 57(4), 553–599. https://jlsp.law.columbia.edu/files/2024/10/Simone.pdf

Skinner, B. F. (1953). *Science and human behavior*. The Free Press.

Spain. (2010, April 1). *Ley 7/2010, de 31 de marzo, General de la Comunicación Audiovisual*. Boletín Oficial del Estado No. 79. https://www.boe.es/eli/es/l/2010/03/31/7

Spain. (2022, July 8). *Ley 13/2022, de 7 de julio, General de Comunicación Audiovisual*. Boletín Oficial del Estado, núm. 163. https://www.boe.es/buscar/act.php?id=BOE-A-2022-11311

Staatsblad van het Koninkrijk der Nederlanden. (2024, November 19). *Besluit van 14 juni 2024 tot vaststelling van het tijdstip van inwerkingtreding van artikel 13, onderdelen 4 en 5, van de Wet bestuursrechtelijke aanpak online kinderpornografisch materiaal* (Staatsblad 2024, Nr. 363). https://zoek.officielebekendmakingen.nl/stb-2024-363.html

Staatsvertrag über den Schutz der Menschenwürde und den Jugendschutz in Rundfunk und Telemedien (Jugendmedienschutz-Staatsvertrag – JMStV). (2002). Zuletzt geändert durch Artikel 2 des Fünften Medienänderungsstaatsvertrags (27. Februar – 7. März 2024); in Kraft seit 1. Oktober 2024. https://www.gesetze-bayern.de/Content/Document/JMStV

Statista. (2023). *Number of Internet and Social Media Users Worldwide as of January* 2023. https://www.statista.com/statistics/617136/digital-population-worldwide/

Steiger, M. (2024, November 25). *Plattform-Regulierung verzögert sich in der Schweiz für unbestimmte Zeit*. https://steigerlegal.ch/2024/11/25/plattform-regulierung-schweiz-sankt-nimmerleins-tag/

Steinbacher, M., Drechsler, J., & Schröder, L. (2024). KI-generierte Abbildungen von Kindesmissbrauch – technische Grundlagen und rechtliche Einordnung. In M. Klein, D. Krupka, C. Winter, M. Gergeleit & L. Martin (Eds.), INFORMATIK 2024 (pp. 247–261). Gesellschaft für Informatik. https://publikationen.bibliothek.kit.edu/1000176585/155728685

Steinvorth D., (2026). Social-Media-Verbot für Kinder und Jugendliche: In Frankreich gibt es künftig eine Altersgrenze für Tik Tok und Co. *Neue Zürcher Zeitung*, 2026 January 27

Stelzmann, D., Amelung, T. & Kuhle, L. F. (2020). Grooming-Umgebungen von pädophilen und hebephilen Männern in Deutschland: Erste Ergebnisse einer qualitativen Befragung. In T.-G. Rüdiger & P. S. Bayerl (Hrsg.), Cyberkriminologie: Kriminologie für das digitale Zeitalter (1st ed. 2020, S. 475–485). Wiesbaden

Sui, M., Hawkins, I., & Wang, R. (2023). When falsehood wins? Varied effects of sensational elements on users' engagement with real and fake posts. *Computers in Human Behavior*, 142, Article 107654. https://doi.org/10.1016/j.chb.2023.107654

Suter, L., Bernath, J., Willemse, I., Külling, C., Waller, G., Skirgaila, P., & Süss, D. (2023). *MIKE – Medien, Interaktion, Kinder, Eltern: Ergebnisbericht zur MIKE-Studie* 2021. Zürcher Hochschule für Angewandte Wissenschaften. https://www.zhaw.ch/storage/psychologie/upload/forschung/medienpsychologie/mike/Bericht_MIKE-Studie_2021.pdf

Swiss Foundation for the Protection of Children. (2020, November 14). *Sexual violence against children online – Input from the Swiss Foundation for the Protection of Children for the General Comment on children's rights in relation to the digital environment*. https://www.ohchr.org/sites/default/files/Documents/HRBodies/CRC/GCChildrensDigitalEnvironment/OtherStakeholders/Swiss%20Foundation%20for%20the%20Protection%20of%20Children.docx

Ta, V.-T. (2024, January). *A safety risk assessment framework for children's online safety based on a novel safety weakness assessment approach* (arXiv preprint arXiv:2401.14713). Department of Computer Science, Edge Hill University. https://arxiv.org/pdf/2401.14713

Tag, B. & Wyss, M. (2024), Die strafrechtliche Einordnung von pornografischen Deepfakes, *Jusletter*, 2024 April 29

Tait, A. (2019, June 6). *Greta Thunberg: How one teenager became the voice of the planet*. *Wired*. https://www.wired.com/story/greta-thunberg-climate-crisis/

The Netherlands. (2025, January 1). Criminal Code of the Netherlands (applicable from 01-01-2025 to the present day). https://wetten.overheid.nl/BWBR0001854/2025-01-01

The Verge. (2024). KOSA bill stalls in House amid free speech concerns. https://www.theverge.com/2024/07/

Threema. (2023). *Open Letter to EU Member States on the Proposed CSA Regulation*. https://threema.com/en/blog/open-letter-csa-regulation

TikTok. (2025, July 25). *Terms of Service*. https://www.tiktok.com/legal/page/global/terms-of-service-eea-archive/en/

TikTok. (n. d.-a). *For parents and guardians – Account and user safety*. https://support.tiktok.com/de/safety-hc/account-and-user-safety/for-parents-and-guardians/

TikTok. (n. d.-b). *Privacy and safety settings for users under age 18 – Account privacy settings*. https://support.tiktok.com/en/account-and-privacy/account-privacy-settings/privacy-and-safety-settings-for-users-under-age-18/

TikTok. (n. d.-c). *Privacy Policy*. https://www.tiktok.com/legal/page/us/privacy-policy/en

TikTok. (n. d.-d). *User Safety – Account and user safety*. https://support.tiktok.com/en/safety-hc/account-and-user-safety/user-safety/

TikTok. (n. d.-e). *TikTok Safety Features and Settings*. https://www.aph.gov.au/DocumentStore.ashx?hearingid=31015&submissions=true.

UK Safer Internet Centre. (2016). *Creating a Better Internet for All: Young people's experiences of online empowerment and online hate (Executive Summary)*. UK Safer Internet Centre / ResearchBods. https://childnetsic.s3.amazonaws.com/ufiles/SID2016/Executive%20Summary%20-%20Creating%20a%20Better%20Internet%20for%20All.pdf

UN Committee on The Rights oft he Child (UNCRC). (2021). *General comment* No. 25 (2021) *on children's rights in relation to the digital environment* (CRC/C/GC/25). https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CRC%2fC%2fGC%2f25&Lang=en

UNICEF Office of Research – Innocenti. (n. d.). *The Sale and Sexual Exploitation of Children: Digital Technology*. https://www.unicef.org/innocenti/documents/sale-and-sexual-exploitation-children-digital-technology

UNICEF Switzerland and Liechtenstein. (2023, June 5). *Policy guidance on AI for children*. https://www.unicef.ch/sites/default/files/2023-06/AI%20Statement%20EN_050623.pdf

UNICEF Switzerland and Liechtenstein. (2024, December 13). *Children's rights in the digital sphere – promotion of media literacy instead of bans*. https://www.unicef.ch/en/current/statements/2024-12-13/children-s-rights-digital-sphere-promotion-media-literacy-instead

UNICEF. (2024). *Protecting children from violence and exploitation in relation to the digital environment – Policy Brief*. https://www.unicef.org/media/164421/file/Policy%20brief_Protecting%20children%20from%20violence%20in%20the%20digital%20environment.pdf.pdf

UNICEF Switzerland and Liechtenstein. (n. d.). *Digitalization*. https://www.unicef.ch/en/what-we-do/national/digitalization

UNICEF. (2021). *Study on the children's rights situation*. https://www.unicef.ch/en/what-we-do/national/digitalization

UNICEF. (n. d.). *Soziale Medien und Gewalt*. https://www.unicef.de/_cae/resource/blob/340676/982900005bbf365febda20aac864dd94/soziale-medien-und-gewalt-fuer-homepage-data.pdf

United Nations Office on Drugs and Crime (UNODC). (2014). *Study on the effects of new information technologies on the abuse and exploitation of children*. https://www.unodc.org/documents/Cybercrime/Study_on_the_Effects.pdf

United Nations. (1989). *Übereinkommen über die Rechte des Kindes. Abgeschlossen in New York am 20. November 1989*. Entered into force for Switzerland on 26 March 1997. https://www.fedlex.admin.ch/eli/cc/1998/2055_2055_2055/de

United Nations. (n. d.). *The impact of digital technologies*. https://www.un.org/en/un75/impact-digital-technologies

UNODC. (2015). *Study on the Effects of New Information Technologies on the Abuse and Exploitation of Children*, https://www.unodc.org/documents/Cybercrime/Study_on_the_Effects.pdf

U.S. Congress. (2022). Kids Online Safety Act. https://www.congress.gov/bill/117th-congress/house-bill/3663

Villasante, C. (2025, February 3). *Children's online safety in Spain. In Online safety update*. Taylor Wessing. https://www.taylorwessing.com/en/interface/2025/online-safety-update/childrens-online-safety-in-spain

Vobbe, F., & Kärgel, K. (2021). *Sexualisierte Gewalt und digitale Medien: Reflexive Handlungsempfehlungen für die Fachpraxis*. Springer VS

Wachs, S. (2014). Cybergrooming – Erste Bestandsaufnahme einer neuen Form sexueller Onlineviktimisierung. https://doi.org/10.3262/EEO18140331

Wachs, S. (2017). Gewalt im Netz: Studien über Risikofaktoren von Cyberbullying, Cybergrooming und Poly-Cyberviktimisierung unter Jugendlichen aus vier Ländern. Schriftenreihe Studien zur Kindheits- und Jugendforschung. Hamburg.

Wachs, S., Ballaschk, C., & Krause, N. (2023). Hatespeech im Netz: Eine Herausforderung für den Kinder- und Jugendschutz. In K. Biesel, P. Burkhard, R. Heeg & O. Steiner (Eds.), *Digitale Kindeswohlgefährdung – Herausforderungen und Antworten für die soziale Arbeit* (S. 152–168). Verlag Barbara Budrich.

Wachs, S., Gámez-Guadix, M., & Wright, M. F. (2022). Online hate speech victimization and depressive symptoms among adolescents: The protective role of resilience. *Cyberpsychology, Behavior, and Social Networking*, 25(7), 416–423. https://doi.org/10.1089/cyber.2022.0009

Weber, C., Anonymität im Internet – Schau mir in die Augen, Troll, *Süddeutsche Zeitung 22. Dezember 2012*, abrufbar unter: https://www.sueddeutsche.de/wirtschaft/anonymitaet-im-internet-schau-mir-in-die-augen-troll-1.1557187.

Werner, M., (2026). Babler: „Social-Media-Ordnungs-Gesetz" soll bis zum Sommer am Tisch liegen, Der Standard 2. Februar 2026, abrufbar unter: https://www.derstandard.at/story/3000000306694/babler-social-media-ordnungs-gesetz-soll-bis-zum-sommer-am-tisch-liegen.

Wolak J., Finkelhor, D., Mitchell, K.J. & Ybarra, M.L. (2008). "Online 'Predators' and Their Victims: Myths, Realities, and Implications for Prevention and Treatment", *American Psychologist* 63:118.

World Economic Forum. (2023). *Toolkit for digital safety, design interventions and innovations: Typology of online harms*. https://www3.weforum.org/docs/WEF_Typology_of_Online_Harms_2023.pdf

World Economic Forum. (2024, June 6). *Making a Difference: How to Measure Digital Safety Effectively to Reduce Risks Online* [White paper]. https://www3.weforum.org/docs/WEF_Making_a_Difference_2024.pdf

YouTube. (2021, February 24). A *new choice for parents of tweens and teens on YouTube.* https://blog.youtube/intl/de-de/news-and-events/neue-moglichkeiten-fur-eltern-von-alteren-kindern-und-jugendlichen-auf-youtube/

YouTube. (n. d.-a). *Fostering Child Safety*. https://www.youtube.com/intl/de_be/howyoutubeworks/our-commitments/fostering-child-safety/

YouTube. (n. d.-b). *Our Youth Principles*. https://www.youtube.com/howyoutubeworks/kids-and-teens/youth-principles/

YouTube. (n. d.-c). *How to add the made for kids designation.* https://support.google.com/youtube/answer/9528076?hl=en

YouTube. (n. d.-d). *Describing regulated goods – YouTube help*. https://support.google.com/youtube/answer/2801999?hl=en&ref_topic=9282679

YouTube. (n. d.-e). *Harassment & Cyberbullying Policies*. https://support.google.com/youtube/answer/2802268?hl=en&ref_topic=9282436

YouTube. (n. d.-f). *Nudity & Sexual Content Policy*. https://support.google.com/youtube/answer/2802002?hl=en&ref_topic=9282679

In an era where digital technologies shape nearly every aspect of daily life, children and young people are growing up more connected than any previous generation. Across the European Union, most youth use the internet daily and encounter digital environments from an early age. This connectivity opens pathways for learning, creativity, and social engagement, yet also exposes young people to increasingly complex and global risks. While many thrive online, nearly one in three reports feeling unsafe. Violations of privacy, exposure to disturbing content, unwanted sexual approaches, and cyberbullying are becoming more common. At the same time, Europe has emerged as a major hub for hosting child sexual abuse material (CSAM), including newer forms such as deepfake abuse content and AI-generated "DeepNudes." Without effective safeguards, the digital world can shift from a space of opportunity to one of harm.

This book explores how law, policy, and institutional practice can address these urgent challenges. It provides a comprehensive and accessible overview of the legal frameworks governing online safety for minors across Europe, enriched with insights from the United States, and evaluates how effectively current regulations protect children in an evolving digital landscape. Through comparative legal analysis, it highlights best practices, persistent gaps, and the growing need for internationally coordinated approaches.

Beyond legislation, the book examines the shared responsibilities of technology companies, service providers, and civil society. How effective are current measures? Where do they fall short? And what innovations are needed to better safeguard young people's rights, privacy, and well-being?

Ultimately, the book offers essential guidance for policymakers, researchers, and educators. It underscores a central message: only with forward-looking, robust regulation can the digital world remain a place where young people can explore, grow, and participate safely—with their rights fully protected.

EIZ Publishing