POLICY BRIEF

April 2022

DEEPFAKE TECHNOLOGY AND CHILDREN'S INFORMATIONAL RIGHTS IN EUROPE

Loreto Corredoira

Key messages

- Every child should have the right to receive truthful information.
- Deepfake technologies pose risks to children's rights and therefore should be controlled and regulated in the European Union.
- Media education for youth is of critical importance to ensure children's understanding of online algorithms and threats.

Introduction

The topic of protecting children's informational rights cannot be more relevant In the current context of the debate on mis- and disinformation. In terms of the universal informational rights under Article 19 of the Universal Declaration of Human Rights (UDHR) or Article 10 of the European Convention, children's rights are not only limited to two rights - imparting and seeking information - but also include the right to receive factual information. As the youth of today is suffering the consequences of the post-truth era, the latter is highly relevant.

This policy brief unfolds the issues of the right to receive truthful information and the ways in which this process is threatened by current "cyberthreats". One of the biggest cyberthreats of today and of the future, among others, is deepfakes (DF). This policy brief focuses on deepfake videos and how they should be regulated and controlled in the European Union.

Deepfakes: Future Threats and Challenges

Deepfakes, an emergent type of threat falling under the greater umbrella of synthetic media, utilize a form of artificial intelligence/machine learning (AI/ML) to create believable, realistic videos, pictures, audio, and text of events, which never happened [1]. DFs makes it possible to fabricate media — swap faces, lip-syncing, and puppeteer — mostly, without consent and bring threat to psychology security, political stability, and society in general.

Advances in DF technology offer many opportunities for a variety of aims both with positive (in art, education, autonomy, medicine, film industry, e-commerce, etc.) or negative consequences. Because of their rapid spread, DFs can be used for malicious and criminal aims, for example: disinformation; distortion and manipulating public opinion; political hostility; manipulation of election campaigns results; violation of intellectual property rights; defamation; blackmail, hating, bullying; harm to childhood: child pornography and other prohibited sexual practices and many more [2].

The real threat of deepfakes and synthetic media comes not only from the technology used to create it, but also from people's natural inclination to believe what they see, and as a result deepfakes and synthetic media do not need to be particularly advanced or believable in order to be effective in spreading mis/disinformation. Deepfake videos are extraordinarily effective on adults, let alone on fragile children's minds.

The impact of deepfake on children's informational safety

DFs can pose a threat to the safety and well-being of children as they intensify cyberbullying and online violence against children on the Internet. Realistic depiction of scenes of aggression, horror, violence, or explicit sexual acts can be harmful to a child's mental health. DF production tools are available to everyone and are quite easy to use. Applications for creating DF are available in smartphones. With the help of such applications, it is possible to carry out destructive actions, both directly involving images of children, and aimed at a children's audience.

For instance, one of the most popular activities for children is to create videos of themselves and post them on various social media. In addition, children are active users of applications that can create a realistic image of a child at an older and mature age from a photo of a child. By uploading their photographs for processing, children provide them to third parties, unconsciously contributing to the creation of vast libraries of content that can be analyzed and used for the production of DFs. There have been cases where images of children have been collected and used to create sexualized DFs.

European Union policy-making: current stage

There are three main points that are currently being discussed in draft EU policy recommendations or regulations.

1. One group of regulations aims to mitigate malicious DF videos by ensuring that they are properly labelled as non-authentic.

---

[1], [2] *Increasing Threat of DeepFake Identities, 2021.* Homeland Secuirty, USA.
https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

This refers to the European Parliament's call for mandatory labelling of deepfake videos; in particular, article 76 of the Report on Artificial intelligence [3], adopted in January 2021 states that the parliament "is deeply concerned about deepfake technologies that allow increasingly realistic photo, audio and video forgeries to be produced [4]".

This is the reason why the EU Parliament *"calls for an obligation for all that kind of materials to be labelled as 'not original' by the creator, with strict limits on their use for electoral purposes and robust enforcement"*

2. The draft (at the time of writing) Digital Services Act (DSA) sets out rules for flagging and removing illegal content, which could help to interrupt its circulation and amplification.

There is a concern however, among scholars, about the compatibility of the measures with communication rights. In terms of protecting minors, the initial DSA draft relied as heavily on self-regulation as the eCommerce Directive does [5]. However, a recent European Parliament briefing [6] shows that things are changing. Several views expressed in the parliament's briefing recommend that the Council amend "the draft DSA text to, inter alia,
·   ensure a high level of privacy
·   provide guidelines on 'dark patterns',
·   ensure that minors under 18-years-old are not subject to targeted advertising"

The Council has indicated that it is willing to amend the DSA in line with these recommendations.

3. The final point is suggested in the document "Key social media risks to democracy" [7], which urges the EU to empower citizens and enforce "several measures available to protect elections and democratic process from the risks posed by social media platforms. These include general measures, such as on data protection and general transparency, and specific measures to regulate targeted political advertising".

---

[4] "76. *Is deeply concerned about deepfake technologies that allow increasingly realistic photo, audio and video forgeries to be produced that could be used to blackmail, generate fake news reports, or erode public trust and influence public discourse; believes such practices have the potential to destabilise countries, spreading disinformation and influencing elections; calls, therefore, for an obligation for all deepfake material or any other realistically made synthetic videos to be labelled as 'not original' by the creator, with strict limits on their use for electoral purposes and robust enforcement; calls for adequate research in this field to ensure that technologies to counter these phenomena keep pace with the malicious use of AI;"*

[5] About the DSA, EU Legislation in Progress, Briefing, 04-04-2022 stated "*new societal challenges are emerging, such as the spread of counterfeit goods, hate speech and disinformation online". Against this backdrop, the European Commission tabled a new legislative proposal on a digital services act in December 2020. The proposal aims at amending the e-Commerce Directive and sets out a horizontal framework for content moderation of the EU online space. Interinstitutional negotiations between the Commission, the European Parliament and the Council seeking to reach an agreement on a compromise text started in January 2022. Controversial issues currently being discussed by the co legislators include the scope of the new regulation, the enforcement mechanisms proposed, and to what extent targeted advertising and techniques to influence users' behaviour such as 'dark patterns' should be curbed. Second edition. The 'EU Legislation in Progress' briefings are updated at key stages throughout the legislative procedure".*

[6] EU Legislation in Progress, Briefing, 04-04-2022
[7] By Costica Dumbrava, Members' Research Service, EPRS. This paper has been drawn up by the Members' Research Service, within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament. (13-12-2021)
https://www.europarl.europa.eu/thinktank/en/document/EPRS_IDA(2021)698845

Further Policy Implications and Recommendations

It cannot be stressed enough how important the regulations and policies on protecting children from synthetic technologies, such as DF, are. The combination of children's consumption of the growing volumes of information, development of new technologies and children's vulnerability calls for more attention to the threats emerging from it and consequently new regulations.

The modern post-millennial generations, the so-called Generation Z and Generation Alpha, who are being educated in the AI era, talk to Siri, Alexa and Google Assistant, follow suggestions on the Internet generated by algorithms, do not go to reliable sources of information or verify their sources of information in any way. Left unchecked, DFs are set to become the cybercriminals' next big weapon. The blind trust to provided information is the biggest threat that children and society as a whole face nowadays. In order to avoid distrust, the right to receive truthful information must be fostered.

In the present moment, in the European law, DF videos need to be "malicious" to be considered illegal. For example, they need to contain elements of fraud, extortion, bullying, non-consensual pornography or hate-speech. However, there are several proposals under consideration to regulate the dissemination of DFs that do not meet the threshold of illegality. It has been argued that Facebook was created to connect us, however, it has been proven on several occasions, social networks have also been an effective tool to polarise and divide citizens through using profiling and micro-targeted ads without any respect for privacy rights or pluralism.

The concern is that Facebook and other social media practices complicate the right to receive truthful information and the separation of it from other messages, such as political campaigns, manipulated information, synthetic pieces like DFs. These issues must be addressed, clearly defined and comprehensively regulated.

Educational systems should be adjusted to modern technological developments and demands and therefore include more subjects on media education, such as media literacy, Internet and social media safety and technology.

Along with implementation in educational systems, it is important to promote children's informational safety among parents through educational facilities and by implementing new laws. The goal is for parents and the older generation as a whole to acknowledge this problem and understand the importance of this issue. As a result, parents should be empowered to gain a thorough understanding in order to be able to discuss DF and the importance of image consent with children in their care. Parents should make sure their children know why they should ask someone before using an image of them to create a DF or manipulated picture. Parents must be thoroughly informed through respective regulations how to make sure that all the devices their children own or have access to have the best safety settings enabled.

Conclusion

Deepfakes, synthetic media, and disinformation in general pose challenges to our society. These are not only threats to children and adults, these are threats to democracy. They can impact individuals and institutions from small businesses to nation states. As discussed above, there are some approaches which may help mitigate these challenges, and there are undoubtedly other approaches we have yet to identify. the situation needs to be improved through committing to three types of action: policy measures, education (including algorithm literacy) and good practices from all the industry actors implied: the big tech, application developers, and toys manufacturers.

The Policy Brief is published in the framework of the FREuDe project. The project aims to intervene for positive future social change that derives from the commitment and intellectual input across disciplines, such as Sociology, Law, Education, Childhood and Youth studies, European studies and Politics, as well as Communication scholarship and Security studies. Moreover, the Centre addresses the question from the perspective of future autonomous citizens, today's children, and explore closely the ways in which information and Europe feature in their lives.

Jean Monnet Communication, Facts and Regulation for European Democracy (FREuDe) Centre of Excellence

– stimulates new forward thinking with regards the role of facts and place of regulation for securing a future democratic Europe
- generates new research and policy-oriented thinking about integration on the basis of informational rights and enabling informational environments across disciplines not traditionally involved in studying Europe:
- develops new agendas for research, policy and teaching across disciplines and across stakeholder communities
- provides an impetus for future oriented thinking, by researching the needs and perceptions of Europe's future autonomous citizens, young people and in particular children for factual information in and about Europe
- mobilises knowledges and competencies of a range of experts and especially aiming to "hear from" stakeholders which have historically been permitted least input to questions of right to accurate and comprehensive information as a civil and human right.

Loreto Corredoira is a Professor of Communication policy at the Complutense University in Madrid, Spain.