MARcin NIEDBAŁA¹

The Problem of Criminal Liability for Generating Pornography Using Artificial Intelligence²

Submitted: 1.07.2023. Accepted: 22.09.2023

Abstract

The process of advancement of artificial intelligence (hereinafter: AI) seen in recent years results in breakthroughs in many areas of human life, such as medicine, agriculture, and science. However, despite its many advantages and benefits, it also inevitably creates room for abuse. This is particularly true of child pornography, where AI systems are increasingly being used to generate nude images of minors – both existing in reality and fictional. These cases, given much attention in the public debate, cause doubts as to whether existing normative solutions are suitable to combat this new phenomenon. The article aims to answer the question of whether the existing law makes it possible to successfully prosecute perpetrators who use AI systems to generate child pornography (including that depicting fictional characters). In addition, it also offers an analysis of the criminal liability of the developers of the AI systems used in this process and of the hosting providers managing the websites used to distribute such pornographic content.

Keywords: artificial intelligence, child pornography, criminal law, deepfake.

¹ Marcin Niedbala, PhD, e-mail: marcin.niedbala@radcy-prawni.eu; ORCID: 0000-0001-9580-9707.
² The research in this article has not been supported financially by any institution. Translation of that article into English was financed under Agreement Nr RCN/SN/0331/2021/11 with funds from the Ministry of Education and Science, allocated to the “Rozwój czasopism naukowych” programme.
Introduction

The advancement of artificial intelligence (hereafter: AI), which is increasingly used in almost every aspect of daily life, has been gaining momentum in recent years. AI systems support human activity by enabling more efficient decision-making or better customisation of services provided. The use of AI in medicine makes it possible to diagnose patients more accurately and support them in their treatment and subsequent recovery more effectively. It can also be applied to counteract the process of climate change and its negative effects by, for example, reducing emissions generated by transportation, agriculture and industry, or by forecasting extreme weather conditions. The role of AI in the progress of science should not be underestimated either. A good example is biology, where an AI system has predicted the structure of almost every protein known to science in just 18 months, a breakthrough that could accelerate the development of new drugs.

Still, the development of artificial intelligence, apart from its numerous benefits, also creates room for new cases of abuse and leads to previously unknown forms of committing crimes or violating the rights of others. Using AI to create pornographic content is one infamous example of the above. Cases of computer-generated nude images of minors – both existing in reality and fictional ones – are especially contemptible. On the one hand, perpetrators who engage in such activities do not go unpunished, as evidenced by the high-profile criminal prosecutions and convictions in such cases. On the other hand, there are more and more doubts as to the legal classification of these reprehensible acts. The biggest question here is whether the mere generation of a nude image of a minor who does not exist in reality constitutes a crime or not yet. In view of this, it is necessary to analyse the existing legal regulations in the Polish normative system and determine if they are ready to deal with this new problem resulting from the advancement of artificial intelligence. The need for the analysis in question stems from two high-profile cases concerning

---

3 Insider Intelligence, Use of AI in healthcare & medicine is booming – here’s how the medical field is benefiting from AI in 2023 and beyond, https://www.insiderintelligence.com/insights/artificial-intelligence-healthcare/ (access: 19.05.2023).
the computer generation of nude images of minors – including fictitious ones – as well as from the doubts raised in Poland (including in a parliamentary interpella-
tion) with regard to the possibility of effectively countering this problem on the
grounds of the law in force.

Use of artificial intelligence to generate child pornography

At the end of April 2023, a Canadian court sentenced Steven Larouche, accused of
creating and possessing child pornography, to eight years in prison. The Canadian
was found guilty of generating at least seven videos using artificial intelligence
(so-called deepfake technology), which was used to superimpose the faces of
minors on the bodies of other persons (also minors). Although Larouche’s lawyers
requested a lower sentence, arguing that no violence against children occurred
during the making of the videos, the judge stressed the fact that the sexual integrity
of all minors whose images were used by the defendant had been violated. Accord­
ing to the judge, the criminal proceedings in question were the first in Canadian
history involving the creation of child pornography with the use of deepfake

technology.6

Under current Canadian criminal law, anyone in possession of child porno­
graphy is guilty of an indictable offence and is liable to imprisonment for a term
of not more than 10 years and to a minimum punishment of imprisonment for
a term of one year,7 with child pornography being defined as e.g. a photographic,
film, video or other visual representation, whether or not it was made by electronic
or mechanical means, that shows a person who is or is depicted as being under the
age of eighteen years and is engaged in or is depicted as engaged in explicit sexual
activity.8 Thus, Steven Larouche’s use of deepfake technology to generate a record­
ing showing nude minors could unquestionably be considered an act that is prohi­
bited under Canadian law as it clearly fulfils the criteria indicated above.

The second case involved the arrest of a Spaniard who, like Steven Larouche,
was found liable for generating child pornography using artificial intelligence. The
perpetrator used images of minors commonly available on the Internet, which the

---

AI system then processed to create explicit paedophilic content. While the Spaniard’s methods are comparable to the case of Steven Larouche, the extent of the brutality of the pornography generated was so horrifying that it led to a heated debate on the use of images of not only real people, but also of fictional characters, which is becoming an increasing problem in the modern world. A good example is the Stable Diffusion algorithm, which, originally devoid of any restrictions censoring the content generated, made it possible to create pornographic content depicting both existing and fictional persons, including minors. The problem became so serious that in late 2022, the developers of the algorithm were forced to come up with an update to prevent the generation of child pornography. This issue also reached Poland, resulting in an interpellation submitted to the Minister of Justice on 24 January 2023 by MP Paweł Szramka, who drew attention to the threats and dangers of the uncontrolled use of artificial intelligence to create nude images of minors. It should therefore be considered whether the existing criminal-legal solutions would make it possible to hold the perpetrator liable if such an event occurred in Poland.

Generating nude images of minors – criminal liability of the creator

Child pornography is addressed by the provisions of Article 202 § 3-4c of the Act of 6 June 1997 – Penal Code. According to § 4b of these provisions, anyone who produces, distributes, presents, stores or possesses pornographic content presenting a produced or processed image of a minor involved in a sexual act shall be punished by imprisonment of up to three years. This is exactly this provision that will apply to cases of generation of pornographic content featuring the image of a minor with the use of deepfake technology, as further discussion will show. To

---


13 Act of 6 June 1997 – Penal Code, uniform text in the Journal of Laws of the Republic of Poland of 2022, item 1138 as amended, hereinafter referred to as the PC.
start with, it seems reasonable, for the purpose of clarity, to quote the definition of pornographic content established in the judicial practice by the Supreme Court in its judgement of 11 January 2017. The judgement in question states that the content must be characterised by all of the following features:

1. representation of sexuality and human sex life;
2. focusing exclusively on the technical aspects of sexuality and sex life, disregarding any intellectual and personal aspects;
3. showing human sex organs involved in their sexual activity;
4. the message of the content makes it clear that the main intention of the creator of the content is to cause sexual arousal in the recipient.\textsuperscript{14}

According to the views of legal academics and commentators, it is not possible \textit{a priori} to determine what constitutes pornographic content and what does not, as the judgement in this regard will have to be made each time by the body applying the law. However, in addition to the criteria listed by the Supreme Court, an auxiliary criterion may be the aesthetic value of the content considered, i.e. the lower this value is, the more it will suggest that the content in question is pornographic.\textsuperscript{15}

Within the meaning of Article 202 § 4b of the Penal Code, a minor is any person under the age of 18.\textsuperscript{16} Producing, in turn, means in this case all the activities necessary to produce what did not exist before. And it is irrelevant whether the actions taken by the perpetrator are professional or non-professional, or for what reasons they are taken – e.g. for personal use or for commercial purposes.\textsuperscript{17} In addition to production, said provisions criminalise the acts of distributing, displaying, storing or possessing a manufactured or processed image of a minor engaged in sexual activity. Taking into account the past experience resulting from e.g. the way the Stable Diffusion algorithm was used and the fact that its users shared content generated online, it should be considered that these users, too, may be held liable for committing a criminal act. On the one hand, publishing generated pornographic content featuring a minor on a particular website, as a rule, will not be considered an act of presentation since this concept means allowing others to view it without them having to take any initiative aimed at familiarising themselves with this content.\textsuperscript{18} An exception could be a situation where such content is posted on social

\textsuperscript{14} Supreme Court’s judgement of 11 January 2017, III KK 188/16, LEX no. 2321853.


\textsuperscript{18} K. Lipiński, op. cit., p. 669.
media, although in this case it would most likely be promptly removed or at least censored. On the other hand, publishing pornographic content even on a closed online message board will be an act of distribution,\textsuperscript{19} which is also penalised by Article 202 § 4b of the Polish Penal Code. Any Internet user who saves such content on their own storage device, computer, server, etc., and thus is in possession of such content, will also be liable for prosecution on said grounds.\textsuperscript{20}

Pornographic content the production, distribution, presentation, storage or possession of which is penalised under Article 202 § 4b of the Penal Code may depict both a generated/produced and a processed image of a minor. ‘Generated’ and ‘produced’ shall mean ‘crafted’, depicting a person that does not exist in reality, while ‘processed’ means content in which the image of a real person has been so altered that it resembles a completely different person.\textsuperscript{21} The above means that according to the provision of Article 202 § 4b of the Penal Code, criminal liability will be imposed both on the person who uses artificial intelligence to create child pornography depicting persons that do not exist in reality and on the perpetrator who creates such content with the use of deepfake technology. On the grounds of the provision considered, therefore, any action involving the artificial generation of pornographic content showing minors is punishable, regardless of whether the depicted persons exist or are fictitious. This conclusion is supported by the \textit{ratio legis} behind § 4b of Article 202 of the Penal Code, which, according to the intention of the drafters of the 2008 amendment to the Penal Code incorporating it,\textsuperscript{22} aims to penalise criminal actions involving the widespread distribution of child pornography – especially computer-generated images of minors.\textsuperscript{23} This implemented Article 1a (iii) in conjunction with Article 3 of the Council Framework Decision of 22 December 2003 on combating the sexual exploitation of children and child pornography,\textsuperscript{24} which was replaced by the current Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the

\textsuperscript{19} Ibidem, p. 673.
\textsuperscript{20} Ibidem, p. 672.
\textsuperscript{21} V. Konarska-Wrzosek, op. cit., p. 1023.
\textsuperscript{22} § 4b was added to Article 202 of the Polish Penal Code on the grounds of Article 1 item 3 of the Act of 24 October 2008 on the amendment of the Penal Code act and certain other acts, Journal of Laws of the Republic of Poland of 2008, no. 214, item 1344.
sexual abuse and sexual exploitation of children and child pornography. Article 2(c)(iv) of that directive defines child pornography as e.g. realistic images of a child engaged in sexually explicit conduct or realistic images of the sexual organs of a child, for primarily sexual purposes. Thus, it should be assumed that it does not matter whether the image depicts a real or fictional (computer-generated) child, as long as the image is realistic. To conclude, it needs to be emphasised that the Polish normative system may impose criminal liability on the perpetrator of any act of use of artificial intelligence to produce child pornography, regardless of whether the content produced features existing or fictional minors.

**Generating nude images of minors – criminal liability of the hosting service provider and the creator of the AI system used**

Simply answering the question of the possibility of convicting those who create child pornography using AI in the affirmative does not resolve all the doubts concerning this issue. It needs to be pointed out that the problem under analysis has recently become even more serious because of two factors. First, because of the increasing availability of the AI algorithm that makes it possible to generate the content in question, and second, because of a kind of consent on the part of hosting providers managing the websites through which such content has been and continues to be distributed and disseminated.

Referring first to the liability of hosting service providers, one should consider in particular Article 14 of the act of 18 July 2002 on providing services by electronic means, which excludes the hosting service provider’s liability for stored data in the following conditions: if the service provider is unaware of the unlawful nature of the data stored or activities performed in relation to such data (1), and, in the event of receiving official notification or obtaining reliable knowledge of the unlawful nature of the data stored or activities performed in relation to such data, if the service provider immediately prevents access to the data in question (2). In the second case, it will be relatively easy to make the judgement because it will be necessary to determine whether the service provider has taken the action required under the regulation, i.e. in the context of the considerations undertaken, whether


26 Act of 18 July 2002 on providing services by electronic means, uniform text in the Journal of Laws of the Republic of Poland of 2020, item 344, hereinafter referred to as the APSEM.
they blocked access to child pornography content. However, the issue of their lack of prior knowledge of the unlawfulness of the content calls for a more in-depth analysis. The views of legal academics and commentators stress the importance of the tendency to interpret the aspect of service providers’ awareness in an expansive manner.\textsuperscript{27} For instance, the Supreme Court, in its judgement of 10 January 2014, recognised that the establishment of an automatic content moderation system by the administrator of a hosting service may prove their awareness of the fact that the users of the provided services post illegal content (in the case considered by the Supreme Court – explicit content) in this system.\textsuperscript{28} In a judgement of 30 September 2016, in turn, the Supreme Court expressed the view that the service provider’s awareness of the unlawful nature of the data posted can be justified by the fact that, due to the service provider’s experience in the field of hosting services, they are aware of the actual possibility of Internet users posting content that violates the personal rights of specific individuals and know that the moderators employed and the automatic filtering system are ineffective, but at the same time they accept this state of affairs and do not take appropriate action.\textsuperscript{29} Therefore, hosting providers whose websites are used to host child pornography content should be judged harshly as well. It should be emphasised that, taking into account the huge scale of the phenomenon in question and the simultaneous media coverage thereof, administrators of hosting sites, as professional hosting entities, should be aware of the problem under analysis, so it is more than reasonable to require them to exercise the necessary caution by applying appropriate content filters and preventing the spreading of such content. Thus, it is necessary to consider that it would be unlikely to exclude their liability for stored data on the grounds of Article 14(1) of the APSEM. The European Commission’s recommendation on measures to effectively combat illegal content online on the Internet deserves particular attention here as an expression of recognition of the ineffectiveness of existing legal solutions. According to its provisions, hosting providers should take proactive measures to combat illegal content – including content that features child sexual abuse. Said entities are therefore required to identify and prevent the spread of such content, but are left free to choose specific methods of action.\textsuperscript{30} In view of the above, the inaction of hosting providers could be perceived as tacit approval of the use of their ICT systems to share child pornography.


\textsuperscript{28} Supreme Court’s judgement of 10 January 2014, I CSK 128/13, LEX no. 1526612.

\textsuperscript{29} Supreme Court’s judgement of 30 September 2016, I CSK 598/15, LEX no. 2151458.

The liability of the developer of the AI system used to generate illegal content – and to commit a crime as a result – requires a separate analysis. The question is whether they can be held liable as an accessory, i.e. whether they have facilitated the commission of a criminal act through their behaviour (the creation of an AI system without adequate safeguards). Legal academics and commentators argue that perpetration includes both direct intent and conceivable intent, so that the accessory should be aware of both the form and impact of the act they intend to facilitate and of the fact that their behaviour constitutes facilitation of the commission of a criminal act by another person. Applying these considerations to the case of the Stable Diffusion algorithm, it should be acknowledged that this may be an example of a form of conceivable intent since the creators of the algorithm, by providing an AI system lacking appropriate restrictions, created in fact the opportunity of using the algorithm not only for socially acceptable purposes – such as by artists, but also for the purpose of generating illegal and inappropriate content, especially child pornography. The efficient response of the developers of the algorithm in question solves the problem only partially, as the consequences resulting from the production and widespread sharing of pornographic content featuring minors online are often impossible to reverse.

Conclusion

The development of artificial intelligence brings with it tremendous opportunities for the growth and prosperity of modern societies, the emergence of new technologies, the advancement of medical science, or simply improved overall standards of living. Nevertheless, in addition to numerous benefits, the progress in the analysed area also creates new risks and threats which must be addressed by existing law. Artificial intelligence, like any revolutionary invention, is used frequently in socially unacceptable ways. In this context, its use to generate child pornography is particularly outrageous. At the same time, it should be stressed that the provisions of the criminal law in force do take into account the specifics of the problem under analysis and provide tools to prosecute those who create such content using artificial intelligence. However, this response is in part merely neutralising the effects of the phenomenon rather than attempting to nip it in the bud, so to speak, as its usual sources take the form of AI algorithms lacking incorporate adequate safeguards due to their developers’ inactivity and the inaction of administrators.

of hosting sites used for sharing child pornography. It is these two groups of actors that should be required to undertake specific actions correlated with the ongoing development of artificial intelligence. In this context, the imminent entry into force of the so-called Artificial Intelligence Act (the “AI Act”), which stipulates e.g. that AI developers will be obliged to design AI systems in such a way that they cannot be used to generate illegal content, looks very promising.\(^{32}\) It must be emphasised here that the advancement of technology, despite the myriad resulting benefits, cannot occur with the tacit approval of the emergence of new ways of committing crimes – including those involving the unacceptable objectification of children, a group that the law should protect in a special way.

**Bibliography**

**Normative acts**


Act of 18 July 2002 on providing services by electronic means, uniform text in the Journal of Laws of the Republic of Poland of 2020, item 344


Commission Recommendation (EU) 2018/334 of 1 March 2018 on measures to effectively tackle illegal content online, OJ EU L 63, 6.3.2018


**Judicial decisions**

Supreme Court’s judgement of 10 January 2014, I CSK 128/13, LEX no. 1526612.

Supreme Court’s judgement of 30 September 2016, I CSK 598/15, LEX no. 215458.

Supreme Court’s judgement of 11 January 2017, III KK 188/16, LEX no. 2321853.

---

Literature sources


Insider Intelligence, *Use of AI in healthcare & medicine is booming – here’s how the medical field is benefiting from AI in 2023 and beyond*, https://www.insiderintelligence.com/insights/artificial-intelligence-healthcare/ (access: 19.05.2023).


