

## Agata Ziobroń

MA, Maria Curie-Skłodowska University

orcid.org/ 0000-0003-0398-460X

26agata.z@wp.pl

# Political deepfake. Remarks *de lege lata* and postulates *de lege ferenda*

## Introduction

The essence of deepfake technology is the ability to create modified audio, video and audio-video recordings. It relies on substitution of the face of one person for a different one, with the result of creation of a convincing transformation. The first modifications of this type were created in 2017 having a pornographic nature.

Deepfake, according to Oxford Languages dictionary is “a video of a person in which their face or body has been digitally altered so that they appear to be someone else, typically used maliciously or to spread false information”<sup>1</sup>. Another definition coming from Cambridge Dictionary constitutes that “deepfake is a video or sound recording that replaces someone’s face or voice with that of someone else, in a way that appears real”<sup>2</sup>. These definitions refer to the effect of modification rather than to technique. Another definition coined by U.S. Government Accountability Office, indicates that deepfake

---

<sup>1</sup> *Deepfake* [term], oed.com/search/dictionary/?scope=Entries&q=deepfake [accessed: 7.06.2024].

<sup>2</sup> *Deepfake* [term], dictionary.cambridge.org/dictionary/english/deepfake [accessed: 30.05.2024].

“is a video, photo, or audio recording that seems real but has been manipulated with AI”<sup>3</sup>.

The definition that can express what deepfake can be, might be as follow: Deepfake is a technique or a result of the technique that assumes deep interference into sound or video basing on AI, that leads to realistic modifications showing actions and expressions that have never happened. Polish equivalent of deepfake could be: *falsz technologiczny* (eng. technological fake).

Due to the purpose, deepfake recordings can be classified as follows<sup>4</sup>:

1. Political deepfake – a political deepfake is used to ridicule or discredit a politician or a person holding a public office;
2. Deepfake porn – bears the features of pornography, depicting naked people/in the course of sexual activity or sexual intercourse. According to data for 2019, it is the most popular, as approximately 96% of deepfake productions are pornographic films with features of non-consensual pornography<sup>5</sup>;
3. Creative deepfake – it is not created for criminal purposes, it is used in areas such as education, art, film, etc.;
4. Satirical deepfake – the purpose is to ridicule the flaws and weaknesses of a group or individuals, may constitute statutory defamation;
5. Terrorist deepfake – created for the purpose of disseminating terrorist propaganda;
6. Evidence deepfake – created for the purpose of submitting false evidence to a court or other authority. Probably the most dangerous from the point of view of achieving the goals of legal proceedings;
7. Violent deepfake – used to intimidate or force a specific action or cause other harm or injury;

<sup>3</sup> *Deepfake* [term], [gao.gov/products/gao-20-379sp](https://gao.gov/products/gao-20-379sp) [accessed: 30.05.2024].

<sup>4</sup> A. Ziobroń, *Klonowanie głosu – wyzwaniem dla prawa karnego? O przestępczości z użyciem sztucznej inteligencji*, [in:] *Nowe technologie. Wyzwania i perspektywy dla prawa karnego*, red. J. Piskorski, M. Błaszczak, Łódź 2023, pp. 245–246.

<sup>5</sup> A. Sanocki, *Deepfakes, czyli postprawda objawiona*, [in:] *Zjawisko dezinformacji w dobie rewolucji cyfrowej. Państwo. Społeczeństwo. Polityka. Biznes*, red. M. Wrzosek, Warszawa 2019, p. 21.

8. Mixed deepfake – may contain features of two or more of the above types.

Currently, the most popular are pornographic deepfakes (data as of 2019), with features of non-consensual pornography, while non-pornographic ones are often satirical in nature, presenting the image of famous people, including politicians<sup>6</sup>. However, the literature argues that deepfakes can potentially be used for less innocent purposes: to unfairly gain and maintain power by discrediting the opponent, which is presented as a threat to the democratic legal order<sup>7</sup>. The literature indicates that ways to discredit a political opponent may include disseminating a processed recording in which he speaks in a socially unacceptable manner (e.g. he promotes pedophilia), is mentally disabled or under the influence of intoxicating substances<sup>8</sup>. Discrediting an opponent is not new in political games, as it happens in both authoritarian and democratic regimes<sup>9</sup>. The fairness of elections is a very important element of the legal characteristics of the electoral process and constitutes its principle<sup>10</sup>. Grzegorz Kryszewski formulates a number of electoral standards and conditions for fair elections, including: emphasizing the importance of effective procedures for submitting and examining

---

<sup>6</sup> M.B. Kugler, C. Pace, *Deepfake Privacy. Attitudes and Regulation*, "Northwestern University Law Review" 2021, vol. 116, iss. 3, p. 622.

<sup>7</sup> L. Wilkerson, *Still Waters Run Deep(fakes). The Rising Concerns of 'Deepfake' Technology and Its Influence on Democracy and the First Amendment*, "Missouri Law Review" 2021, vol. 86, iss. 1, p. 410–412.

<sup>8</sup> M. Feeney, *Deepfake Laws Risk Creating More Problems Than They Solve*, 2021, [rtp.fedsoc.org/wp-content/uploads/Paper-Deepfake-Laws-Risk-Creating-More-Problems-Than-They-Solve.pdf](https://www.fedsoc.org/wp-content/uploads/Paper-Deepfake-Laws-Risk-Creating-More-Problems-Than-They-Solve.pdf) [accessed: 30.05.2024].

<sup>9</sup> For example, we can mention dissemination of a fake photo presenting Milard Tydings in the company of Earl Browder the leader of Communist Party of the USA by the political opponents during political campaign in 1950. D. Kaiser, *Ted Cruz is Not the First Politician to Cause Controversy With a Doctored Photo*, 19.02.2016, [time.com/4231131/ted-cruz-tydings-browder-photo](https://time.com/4231131/ted-cruz-tydings-browder-photo) [accessed: 30.05.2024].

<sup>10</sup> Wyrok Trybunału Konstytucyjnego z dnia 3 listopada 2006 r. sygn. akt K 31/06 [Judgement (Constitutional Tribunal) from 3 November 2006] (Dz.U. [Journal of Laws] 2006, nr 202, poz. [item] 1493); wyrok Trybunału Konstytucyjnego z dnia 20 lipca 2011 r. sygn. akt K 9/11 [Judgement (Constitutional Tribunal) 20 July 2011] (Dz.U. 2011, nr 149, poz. 889).

complaints about irregularities occurring in the electoral process, as well as examining electoral disputes and criminal law protection of elections<sup>11</sup>.

The aim of this study is to analyze political deepfake in terms of fulfilling the characteristics of the types of prohibited acts and crimes specified in the Polish Penal Code<sup>12</sup>. Is it justified to introduce an *expressis verbis* ban on the use of deepfakes in political games, following the example of California?<sup>13</sup>

### Legal good potentially threatened by political deepfake

Criminalization imposes the need to identify the legal good that is threatened or violated. According to Władysław Wolter, legal goods are “real or ideal objects that are associated with certain social values, and therefore are under criminal protection”<sup>14</sup>. Intuitively, we can assume that a political deepfake will threaten only collective legal goods, and not specific – individual, but this is not an accurate view. A deepfake may violate both the common good that an individual is unable to individually dispose of (e.g. freedom, democracy) as well as those that he or she can dispose of independently. The issue of precisely determining whether in the case of a political deepfake we are dealing with a collective good or a specific-individual good is crucial, because it refers to the possibility of using a non-statutory countertype of consent, which is excluded in the event of violating (threatening) the collective good. Therefore, it would be unacceptable to defend the creator of a political deepfake who explains that he obtained the presumed consent of the

<sup>11</sup> G. Kryszewski, *Uczciwość wyborów jako zasada prawa wyborczego*, “Studia Wyborcze” 2016, t. 21, p. 26.

<sup>12</sup> Ustawa z dnia 6 czerwca 1997 r. – Kodeks karny [Act of June 6, 1997 – Polish Penal Code] (Dz.U. 1997, nr 88, poz. 553).

<sup>13</sup> In 2020 in Canada the provision AB 730 was passed that criminalises dissemination of records audio or video which creates false harmful impression regarding words or acts of a politician. C. Lecher, *California has banned political deepfakes during election season*, 7.10.2019, [theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020](https://theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020) [accessed: 30.05.2024].

<sup>14</sup> W. Wolter, *Zarys systemu prawa karnego. Część ogólna*, Kraków 1933, pp. 81–82.

recipients of the deepfake who read his work with knowledge of its compromising nature, or a political opponent who, by omitting legal action or not paying attention to its content, 'expressed consent' to its dissemination. Objective (external) freedom to vote is a legal good protected by, among others: art. 249 Penal Code. It may be considered whether the opposite subjective (internal) freedom of voting should be subject to criminal law protection, which can be understood as a conscious decision to vote, free from threats, violence or deceit. It seems that expressing even defamatory content 'under one's own name' is one thing and 'attributing' it to a political opponent is another thing. It is something other than objective sincerity of voting, it is a situation in which: the entitled person takes part in voting and votes according to his or her freely held conviction, free from any unauthorized pressure. Therefore, a special state of free exercise of electoral rights, both voting and being elected, is protected here<sup>15</sup>. In relation to subjective freedom, we would talk about the conformity of this belief with the truth, in terms of truth or falsehood. The social harm of political deepfakes lies in the limitation of this freedom or, to put it more precisely, in the disruption of its shaping.

The above problem is multi-threaded. A political deepfake may threaten not only the public good, such as the state system and its functioning, morality or public order, but also violate individual good, e.g. the mental state of the recipient, including freedom from sexual shame (e.g. in the situation of processing the statements of a politician, resulting in the creation of a sexual message). Legal goods potentially threatened by political deepfakes are divided into collective and individual goods, and they themselves can be further differentiated (e.g. protection against sexual shame, protection against disruptions to the proper psychosexual development of a minor), therefore one of the premises for criminalizing the phenomenon is the existence of a threatened legal good – is fulfilled. Following Jan Kulesza, it should be pointed out that: "determining the social reality of behavior perceived by society as dangerous

---

<sup>15</sup> *Kodeks karny. Komentarz*, red. A. Grzeńkowiak, K. Wiak, wyd. 7, Warszawa 2021.

to legal goods”, and then the fulfillment of three conditions should be identified: proportionality, usefulness and necessity<sup>16</sup>. In the context of the principle of proportionality, it is extremely important to distinguish between a satirical deepfake, which is a joke, and a deepfake leading to the discrediting of a political opponent. In the case of the first situation, introducing criminalization of the dissemination of political deepfakes would constitute an undesirable example of ‘shooting a sparrow with a cannon’ (as Lech Gardocki described such a situation)<sup>17</sup>.

An extremely different situation is attributing socially unacceptable, illegal content to opponents or presenting them in a way that could undermine trust in them. It is true that similar situations enjoy criminal law protection in the form of, for example, art. 212, art. 216 or art. 200b of the Penal Code, however, the specific nature of the subjective side in such a situation (the desire to gain and maintain power in the state) should be taken into account. This particular coloring of the subjective side is reflected in art. 256 § 1a of the Penal Code in the form of the intention to influence political or social life by promoting a totalitarian system or inciting hatred. Therefore, if one of the politicians created and spread a deepfake in which his opponent ‘appeared’ expressing e.g. anti-Semitic content or praising communism, his behavior could meet the statutory criteria of the new article. Art. 256 § 1a of the Penal Code if it referred to the allegedly existing negative characteristics of a given person (e.g. suggested that he is addicted to drugs), such behavior could be met with a criminal law response under art. 212 or art. 216 of the Penal Code, however, in the case of prosecution by private prosecution. As you can see, criminal law protection for politicians against the use of deepfakes to unfairly influence election results is fragmentary and inconsistent.

---

<sup>16</sup> J. Kulesza, *Zarys teorii kryminalizacji*, “Prokuratura i Prawo” 2014, nr 11–12, p. 92.

<sup>17</sup> L. Gardocki, *Zagadnienia teorii kryminalizacji*, Warszawa 1990, p. 116.

## Political deepfake and electoral proceedings

Based on art. 111 § 1 of the Polish Electoral Code<sup>18</sup>: if disseminated, including in the press within the meaning of the Act of January 26, 1984 – Polish Press Law<sup>19</sup>, election materials, in particular posters, leaflets and slogans, as well as statements or other forms of electioneering contain false information, the candidate or the election representative of the interested election committee has the right to submit to the district court an application for a ruling:

1. prohibition of dissemination of such information;
2. forfeiture of election materials containing such information;
3. order the correction of such information;
4. ordering the publication of responses to statements violating personal rights;
5. ordering an apology to the person whose personal rights have been violated;
6. ordering a participant in the proceedings to pay an amount of up to PLN 100,000 to a public benefit organization.

A clear distinction must be made here between message and information. At this point it is worth quoting the view of Mariusz Grabowski and Agnieszka Zajęc:

It is indicated that information is data contained in the message, interpreted by the recipient, having significance for him and bringing an element of novelty to his awareness, i.e. reducing his ignorance. For data to become information, the recipient must decide, firstly, whether he wants to interpret the data, and secondly, whether they are understandable to him and to what extent. Then the data becomes a message for the recipient. Then the recipient determines whether the message is a repetition of something he already knows or whether it is something new for him, and if so, the message becomes information. Because information depends on the interpretive abilities of the recipient, it is subjective. The above process of knowledge formation indicates that the element of novelty distinguishes

<sup>18</sup> Ustawa z dnia 5 stycznia 2011 r. – Kodeks wyborczy [Act of January 5, 1994 – Polish Electoral Code] (Dz.U. 2011, nr 21, poz. 112).

<sup>19</sup> Ustawa z dnia 26 stycznia 1984 r. – Prawo prasowe [Act of January 26, 1984 – Press Law] (Dz.U. 1984, nr 5, poz. 24).

information from news. If the message is new to the recipient, it becomes information<sup>20</sup>.

Furthermore, potential civil liability is possible on the ground of Polish Civil Code<sup>21</sup>, having regard to violation of personal rights. However, these solutions may be not sufficient. It is commonly known, that dissemination of fake information may cause national or even international chaos or disinformation and result in serious threat to stability or safety of state or causing damage to privacy or property.

'Untrue' according to linguistic interpretation means: 'inconsistent with the truth' or 'replacing something'. Therefore, it can be said that false information is information with untrue content, replacing true information, while containing an element of novelty. A political deepfake may convey various information: that a right-wing politician expresses extreme left-wing views, that he is an alcoholic, etc. Therefore, it is a message that influences the recipients' knowledge, so the content of a political deepfake can be considered information.

As for the measures that can be imposed on a person spreading false information, not all of them seem to be equally effective, and some are impossible to actually implement. If a deepfake is widely and quickly spread by many people, it may be impossible to effectively remove false information, and the correction may escape the attention of those interested. Similarly, an order to pay or apologize to a person whose personal rights have been violated may not receive the same attention from recipients as a controversial alteration. It can therefore be said that the current provisions of the Electoral Code are insufficient, as are the provisions regarding the protection of personal rights or liability for damages in tort. In this situation, civil law seems to protect only the interests of individuals, while in the case of a specific deepfake, such as

<sup>20</sup> M. Grabowski, A. Zając, *Dane, informacje, wiedza – próba definicji*, "Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie" 2009, nr 798, p. 116.

<sup>21</sup> Ustawa z dnia 23 kwietnia 1964 r. – Kodeks cywilny [Act of April 23, 1964 – Polish Civil Code] (Dz.U. 1964, nr 16, poz. 93).



a political deepfake, collective legal goods, such as fair elections or the democratic system, are primarily at risk.

The right to free elections means, first of all, that the election result is not the result of pressure exerted on the voter<sup>22</sup>. It is connected with the principle of fairness of elections, which is most often combined with the principle of free elections (especially the terminological version of free and fair election). However, it sometimes happens that it is proposed to separate it and treat it as an independent principle of electoral law. It is intended to emphasize not only the requirement of legality and reliability of all electoral activities, but also the use of specific, honest practices in the election campaign, the opposite of which are, for example, the so-called negative campaigns aimed at discrediting the opponent, not promoting the candidate<sup>23</sup>.

It can therefore be concluded that current measures from other branches of law may prove insufficient to deal with the phenomenon of political deepfake, so it may be justified, at least to consider it, to introduce criminal law measures, in accordance with the principle of *ultima ratio*.

### Promoting an undemocratic system or hate speech

When focusing on the aspects of criminal law liability for promoting an undemocratic system and defamation, attention should be paid to the fact already mentioned above that political deepfakes are often satirical in nature. It is a literary or journalistic genre including lyric poetry and epic, the aim of which is to stigmatize and ridicule phenomena, customs, politics and social relations, through a caricatured presentation of characters through the

---

<sup>22</sup> I.C. Kamiński, *Prawo do wolnych wyborów w Europejskiej Konwencji Praw Człowieka*, "Problemy Współczesnego Prawa Międzynarodowego, Europejskiego i Porównawczego" 2015, nr 13, p. 14.

<sup>23</sup> J. Szymanek, *Nowe zasady prawa wyborczego*, "Infos. Zagadnienia Społeczno-Gospodarcze" 2021, nr 9, p. 4.

prism of the author's critical attitude<sup>24</sup>. Satire often uses caricature, characterized by exaggeration of people and phenomena, and both of these forms may take visual or textual form. Satire is legally protected under press law (art. 41 of the Press Law). This is a very strong argument against recognizing it as unlawful under criminal law. However, it is pointed out that the limit of satire is human dignity<sup>25</sup>. According to Marek Mozgawa, communication can take place not only orally, but also in writing and using, among others, printing, drawings, as well as using technical means of transmitting information<sup>26</sup>.

Moreover, the legislator formulated a code circumstance excluding the unlawfulness of promoting hatred or an authoritarian regime (art. 256 § 3 of Penal Code). It states that the perpetrator does not commit a crime if he committed the act as part of artistic, educational, collecting or scientific activities. There is no doubt that caricature is a form of artistic expression<sup>27</sup>, so in the situation of a mocking, processed recording, a countertype of artistic activity may be used. According to Gardocki, even the strictest artistic or literary criticism cannot be considered defamation; this statement will be fully justified in relation to most literary and artistic works. However, in the context of political deepfakes, the situation is different. Artistic activity cannot be equated with a deepfake, even if it is an innocent, satirical joke, because in the situation mentioned by Gardocki and other authors, the recipient is aware that he is coming into contact with art (e.g. he is reading the satire *To the King [Do króla]* by Ignacy Krasicki), but it is one thing to create the impression that it is a real and serious piece of news (a speech

<sup>24</sup> P. Grabarz, *Ochrona czci człowieka a satyra i karykatura jako element wolności sztuki – rozważania na gruncie polskiego i niemieckiego prawa karnego*, "Zeszyt Studencki Kół Naukowych Wydziału Prawa i Administracji UAM" 2016, nr 6, p. 121.

<sup>25</sup> L. Gardocki, *Prawo karne*, Warszawa 2010, p. 277.

<sup>26</sup> M. Mozgawa, *Komentarz do Rozdziału XXVII kodeksu karnego – Przestępstwa przeciwko czci i nietykalności cielesnej (art. 212–217a k.k.j.)*, [in:] *Kodeks karny. Komentarz*, red. idem, Warszawa 2007, Lex/el. 2024.

<sup>27</sup> K. Kakareko, *Karykatura jako narzędzie polityki państwa*, "Czasopismo Prawno-Historyczne" 2017, t. 69, z. 2, p. 295.

by a politician allegedly touching on important social and moral issues). A situation in which a political opponent transforms the image of a right-wing politician who, after this transformation, begins to 'state' demands such as abortion on demand, reimbursed access to euthanasia, or the legalization of homosexual marriage, or vice versa, may prove problematic in a criminal law assessment. Extremely conservative views are attributed to the leftist orientation. Can we then speak of slandering about properties that may humiliate such an individual in public opinion?

The answer to this question requires possible empirical research, but it can be assumed that this could potentially lead to the loss of part of a given person's electorate.

The element that distinguishes a harmless political deepfake of a satirical nature from an unlawful deepfake seems to be a specific element of the perpetrator's subjective side, which is the goal of gaining power, but currently only the transmission of specific content (e.g. hateful, fascist, authoritarian) is criminalized. Further doubts arise as to whether – apart from the use of the countertype of artistic activity – it can be said that all the statutory features of the act specified in art. 256 § 1a of the Penal Code, namely the verbal mark of 'propagation'. What would the criminal law assessment of a situation look like in which an opponent creates a deepfake in which his political opponent says, for example, that homosexual people should be exterminated, while a moment later the image of the creator who clearly condemns such views appears on the screen? Would it then be possible to consider that the creator of deepfake promotes hatred? It seems that this question should be answered in the negative.

Such doubts, as well as the inconsistent criminal law protection shown above (private prosecution in the case of defamation vs. public prosecution in the context of hate speech) of the subject of deepfake satire may make it necessary to criminalize the dissemination of discrediting political deepfakes in order to gain power by influencing electoral decisions.

Types penalizing hate crimes are not sufficient because not all political deepfake must be connected with attacks on specific groups of people.

## Assessment of the validity of criminalizing political deepfakes

It was argued above that criminalization of political deepfakes could be justified. The conditions are met: the existence of a threatened legal good and the conditions of necessity and proportionality. Proportionality consists of three elements<sup>28</sup>:

1. Suitability – it is assessed whether the means used are suitable for achieving the goal;
2. Necessity – to use the least burdensome measure for the individual;
3. Proportionality in the strict sense – assessment of the degree of nuisance in relation to the value of the good subject to protection.

Assuming that collective goods have primacy over individual goods may carry the risk of implementing authoritarian solutions and excessive restrictions on freedom of speech and freedom of artistic expression and, as a consequence, contradiction of any regulation with art. 31 section 3 of the Constitution<sup>29</sup>. Therefore, in criminal law evaluation, it is necessary to clearly distinguish between a 'calculated' deepfake intended to manipulate voters by influencing their electoral decisions and a satire that has the nature of a joke. It is worth emphasizing that a satirical deepfake does not imitate reality and it is very easy for the recipient to detect its falseness. The features of a creative political deepfake of a satire nature, which should not constitute grounds for criminal liability, may be:

1. Humorous context;
2. Easily recognizable falsity of the alteration (involving the addition of an appropriate marking or other elements);
3. No racist, xenophobic or hateful content;
4. No attribution to a politician of views that are inconsistent with the arguments presented by him.

<sup>28</sup> J. Bojke, *Rola zasady proporcjonalności w wykładni przepisów prawa karnego materialnego*, "Internetowy Przegląd Prawniczy TBSP UJ" 2015, nr 1, p. 108.

<sup>29</sup> Ustawa z dnia 2 kwietnia 1997 r. – Konstytucja Rzeczypospolitej Polskiej [Act of April 2, 1997 – Constitution of Poland Republic] (Dz.U. 1997, nr 78, poz. 483).

Content that is 'serious' and contains the characteristics of hate speech or false views poses a risk of unfair influence on voters' decisions and thus threatens fair elections. Due to the specific context (influencing voters' decisions) and the entity 'expressing' the views, it is insufficient to apply provisions penalizing hate speech, due to the particular social harmfulness of the situation when a fraudulent deepfake has (or creates a real possibility of having) the effect desired by the creator, what is the influence on the opponent's electorate.

A potential regulation could be:

§ 1. Whoever, in order to influence the electoral decision of a recipient entitled to vote, creates and distributes processed visual or audio content shall be subject to the penalty of imprisonment for up to 2 years, a fine or the penalty of restriction of liberty.

§ 2. Prosecution takes place at the request of the injured party.

The regulation could be placed in chapter XXXI entitled: *Crimes against elections and plebiscites*. In the context of people who indirectly create deepfake (programmers, graphic designers the sufficient solutions are constructions currently existing in Penal Code (complicity).

The criterion distinguishing a manipulative political deepfake from an innocent joke or artistic expression would therefore be the subjective side marked by the goal (the desire to influence the electoral decision) and would be distinguished by the court in criminal proceedings taking into account the individual context of the alteration, if necessary with the help of an expert whose role would be to assess whether the processing has the characteristics of processing using artificial intelligence algorithms. For this reason, it is unnecessary to create a countertype of artistic activity. A clear indication that the content has been modified using the deepfake technique would make it impossible to hold the creator criminally liable. Formulating verb characteristics using conjunctions would eliminate the risk of excessive punitiveness, which would result in criminal liability for people who only share processed content on social media, for example, and who did not participate in the interference. Likewise, persons who process processed content, giving

it a humorous context and thus 'exposing' its falseness, would be protected against excessive criminal law interference.

The act could only be committed intentionally, with direct and specific intention, so the perpetrator must want to influence the election results, and not just agree to it, because the author of a 'humorous' deepfake may also agree to such an effect. The crime of a political deepfake would therefore be of a formal nature, because it would be legally irrelevant whether the perpetrator actually caused the voter to change or develop a specific decision, which would be difficult, if not impossible, to prove.

Criminalizing political deepfakes may entail a number of threats and constitutional concerns. While artistic freedom is not in possible opposition to the penalization of deepfakes (as discussed above, the type of deepfakes subject to penalization would be clearly separated from deepfakes of a creative, satirical nature), the risk of violating the freedom to conduct an election campaign is questionable. However, it is impossible not to note that this freedom not only should not be, but is also not unlimited<sup>30</sup>.

However, a criminal law solution could become excessively paternalistic and may be a source of abuse. Therefore, in cases of political deepfake use, it would be reasonable to expand protection measures beyond criminal law (especially administrative law), and treat the above proposal as a contribution to further observation of the phenomenon

## Bibliography

### Legal acts

Ustawa z dnia 23 kwietnia 1964 r. – Kodeks cywilny [Act of April 23, 1964 – Polish Civil Code] (Dz.U. [Journal of Laws] 1964, nr 16, poz. [item] 93).

Ustawa z dnia 26 stycznia 1984 r. – Prawo prasowe [Act of January 26, 1984 – Polish Press Law] (Dz.U. 1984, nr 5, poz. 24).

<sup>30</sup> An example of limiting the freedom of running a political campaign is the institution of election silence, limitation regarding place of running such a campaign or the type of campaign, personal limitations (e.g. age qualification). R. Balicki, K. Piech, *Ograniczenia swobody prowadzenia kampanii wyborczej w świetle regulacji Kodeksu wyborczego*, "Polityka i Społeczeństwo" 2015, nr 3(13), pp. 36–43.

Ustawa z dnia 2 kwietnia 1997 r. – Konstytucja Rzeczypospolitej Polskiej [Act of April 2, 1997 – Constitution of Poland Republic] (Dz.U. 1997, nr 78, poz. 483).

Ustawa z dnia 6 czerwca 1997 r. – Kodeks karny [Act of June 6, 1997 – Polish Penal Code] (Dz.U. 1997, nr 88, poz. 553).

Ustawa z dnia 5 stycznia 2011 r. – Kodeks wyborczy [Act of January 5, 1994 – Polish Electoral Code] (Dz.U. 2011, nr 21, poz. 112).

### Judicature

Wyrok Trybunału Konstytucyjnego z dnia 3 listopada 2006 r. sygn. akt K 31/06 [Judgement (Constitutional Tribunal) from 3 November 2006] (Dz.U. 2006, nr 202, poz. 1493).

Wyrok Trybunału Konstytucyjnego z dnia 20 lipca 2011 r. sygn. akt K 9/11 [Judgement (Constitutional Tribunal) 20 July 2011] (Dz.U. 2011, nr 149, poz. 889).

### Literature

Balicki R., Piech K., *Ograniczenia swobody prowadzenia kampanii wyborczej w świetle regulacji Kodeksu wyborczego*, "Polityka i Społeczeństwo" 2015, nr 3(13), pp. 33–45.

Bojke J., *Rola zasady proporcjonalności w wykładni przepisów prawa karnego materialnego*, "Internetowy Przegląd Prawniczy TBSP UJ" 2015, nr 1, pp. 107–115.

*Deepfake* [term], [dictionary.cambridge.org/dictionary/english/deepfake](https://dictionary.cambridge.org/dictionary/english/deepfake) [accessed: 30.05.2024].

*Deepfake* [term], [gao.gov/products/gao-20-379sp](https://gao.gov/products/gao-20-379sp) [accessed: 30.05.2024].

*Deepfake* [term], [oed.com/search/dictionary/?scope=Entries&q=deepfake](https://oed.com/search/dictionary/?scope=Entries&q=deepfake) [accessed: 7.06.2024].

Feeney M., *Deepfake Laws Risk Creating More Problems Than They Solve*, 2021, [rtp.fedsoc.org/wp-content/uploads/Paper-Deepfake-Laws-Risk-Creating-More-Problems-Than-They-Solve.pdf](https://rtp.fedsoc.org/wp-content/uploads/Paper-Deepfake-Laws-Risk-Creating-More-Problems-Than-They-Solve.pdf) [accessed: 30.05.2024].

Gardocki L., *Prawo karne*, Warszawa 2010.

Gardocki L., *Zagadnienia teorii kryminalizacji*, Warszawa 1990.

Grabarz P., *Ochrona czci człowieka a satyra i karykatura jako element wolności sztuki – rozważania na gruncie polskiego i niemieckiego prawa karnego*, "Zeszyt Studencki Kół Naukowych Wydziału Prawa i Administracji UAM" 2016, nr 6, pp. 117–130.

Grabowski M., Zajac A., *Dane, informacja, wiedza – próba definicji*, "Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie" 2009, nr 798, pp. 99–116.

- Kaiser D., *Ted Cruz is Not the First Politician to Cause Controversy With a Doctored Photo*, 19.02.2016, [time.com/4231131/ted-cruz-tydings-browder-photo](https://time.com/4231131/ted-cruz-tydings-browder-photo) [accessed: 30.05.2024].
- Kakareko K., *Karykatura jako narzędzie polityki państwa*, "Czasopismo Prawno-Historyczne" 2017, t. 69, z. 2, pp. 295–311.
- Kamiński I.C., *Prawo do wolnych wyborów w Europejskiej konwencji praw człowieka*, "Problemy Współczesnego Prawa Międzynarodowego, Europejskiego i Porównawczego" 2015, nr 13, pp. 7–40.
- Kodeks karny. Komentarz*, red. A. Grześkowiak, K. Wiak, wyd. 7, Warszawa 2021.
- Kryszewski G., *Uczciwość wyborów jako zasada prawa wyborczego*, "Studia Wyborcze" 2016, t. 21, pp. 7–31.
- Kugler M.B., Pace C., *Deepfake Privacy. Attitudes and Regulation*, "Northwestern University Law Review" 2021, vol. 116, iss. 3, pp. 611–680.
- Kulesza J., *Zarys teorii kryminalizacji*, "Prokuratura i Prawo" 2014, nr 11–12, pp. 87–111.
- Lecher C., *California has banned political deepfakes during election season*, 7.10.2019, [theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020](https://theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020) [accessed: 30.05.2024].
- Mozgawa M., *Komentarz do Rozdziału XXVII kodeksu karnego – Przepisy przeciwko czci i nietykalności cielesnej (art. 212–217a k.k.)*, [in:] *Kodeks karny. Komentarz*, red. idem, Warszawa 2007, Lex/el. 2024.
- Sanocki A., *Deepfakes, czyli postprawda objawiona*, [in:] *Zjawisko dezinformacji w dobie rewolucji cyfrowej. Państwo. Społeczeństwo. Polityka. Biznes*, red. M. Wrzosek, Warszawa 2019, pp. 21–22.
- Szymanek J., *Nowe zasady prawa wyborczego*, "Infos. Zagadnienia Społeczno-Gospodarcze" 2021, nr 9, pp. 1–4.
- Wilkerson L., *Still Waters Run Deep(fakes). The Rising Concerns of 'Deepfake' Technology and Its Influence on Democracy and the First Amendment*, "Missouri Law Review" 2021, vol. 86, iss. 1, pp. 407–432.
- Wolter W., *Zarys systemu prawa karnego. Część ogólna*, Kraków 1933.
- Ziobroń A., *Klonowanie głosu – wyzwaniem dla prawa karnego? O przestępczości z użyciem sztucznej inteligencji*, [in:] *Nowe technologie. Wyzwania i perspektywy dla prawa karnego*, red. J. Piskorski, M. Błaszczak, Łódź 2023, pp. 241–256.



**Abstract****Political deepfake. Remarks *de lege lata* and postulates *de lege ferenda***

In the paper the phenomenon of political deepfake was discussed, constituting the type of political deepfake deriving from original deepfake classification regarding to the aim. It was concluded, that currently functioning measures may be insufficient. The proposition of new potential type of illicit act was formulated, differentiating socially harmful deepfake from artistic deepfake.

**Key words:** deepfake, election, politician, artistic deepfake, illicit act

