RAIS RESEARCH ASSOCIATION for INTERDISCIPLINARY STUDIES

DOI:10.5281/zenodo.22801807

Deepfakes and the Right to Identity in the Age of Artificial Intelligence

Mariana-Alina Zisu

Legal Adviser, Member of the Methodological and Professional Guidance Commission of the "College of Legal Advisers Suceava" Association, Romania;
Member of the Methodological and Professional Guidance Commission of the "Order of Legal Advisers from Romania" Federation
stefanoaia.mariana@yahoo.com

Abstract: The accelerated development of generative artificial intelligence has made it possible to create increasingly realistic audio, video and synthetic images, capable of reproducing the image, voice and behavior of a person. Deepfake technology offers numerous possibilities, but at the same time generates important legal challenges when a person's identity is reproduced, modified or used without their consent. This article analyzes the relationship between deepfake technology and the legal protection of identity, with a focus on the unauthorized use of image and voice, violation of privacy, damage to reputation and the risks posed by digital impersonation. Particular attention is paid to the ability of existing legal mechanisms to protect individuals against the misuse of their identity in content generated by artificial intelligence. The analysis highlights the need to adapt legal protection to new forms of digital reproduction of identity. In this sense, consent, transparency, accountability and effective legal remedies must be essential elements of a protection framework capable of maintaining the balance between technological development, freedom of expression, and respect for individual rights.

Keywords: Deepfake, Artificial Intelligence, Right To Identity, Digital Identity

Introduction

The rapid development of generative artificial intelligence represents one of the most important technological transformations of contemporary society. The ability of artificial intelligence systems to generate and modify images, audio recordings and video materials has evolved at an accelerated pace, and the difference between authentic and artificially created content is becoming increasingly difficult to identify. In this context, deepfake technology is one of the most controversial applications of artificial intelligence, as it allows the realistic reproduction of a person's image, voice, expressions and behavior, creating the impression that they have performed actions or made statements that, in reality, did not take place.

Deepfake technology is not, by its nature, exclusively harmful. It can have legitimate applications in the film industry, entertainment, education, art, advertising or the preservation of cultural heritage. The possibility of digitally recreating characters or adapting audiovisual content demonstrates the creative potential of generative technologies. However, the same tools can be used to reproduce a person's identity without their consent, to make false statements, to manipulate their public image, to commit fraud or to create and distribute material that is likely to cause significant harm.

From a legal perspective, the essential characteristic of the deepfake phenomenon is the separation of a person's identity from the control they exercise over their own representation. Their image, voice and other distinctive elements can be digitally reproduced and used in contexts completely different from those in which the person initially expressed their consent. Thus, the development of artificial intelligence raises a fundamental legal question: to what extent should a person be able to control the artificial reproduction and use of their own identity?

This issue is all the more important since identity cannot be reduced to a person's name. The image, voice, physical features and other characteristics by which the individual can be recognized contribute to their social and digital representation. Generative technologies allow

the reproduction of these elements in a sufficiently realistic manner for the public to associate the artificially created material with the real person. Consequently, the harm does not only result from the use of an image, but also from the possibility of artificially attributing conduct, opinions or statements to an identifiable person.

The emergence of deepfakes therefore leads to a reconsideration of the relationship between personal identity and individual autonomy. Control over the way in which a person is represented in society constitutes an important element of the protection of personality. When technology allows a third party to create a realistic representation of the individual without their participation, the autonomy of the person over their own identity may be diminished. This situation is particularly problematic when the public cannot reasonably identify the artificial nature of the content.

The legal issue becomes even more complex due to the diversity of harms that can result from the unauthorized use of deepfakes. Depending on the circumstances, a synthetic representation may affect the privacy, reputation, dignity or economic interests of the person. Deepfakes can also be used for impersonation, fraud, manipulation or the creation of intimate material without consent. In such situations, the same technology can simultaneously affect multiple legal rights and interests, making it difficult to fit the phenomenon into a single traditional category of protection.

An additional difficulty is represented by the speed and scale of distribution of digital content. Fake material can be created in a short time and distributed to a very large number of people through online platforms. Even if the artificial nature of the content is subsequently demonstrated, the initial damage may be difficult to completely eliminate. Copies can continue to circulate, and the public perception of the affected person can be changed before he or she has an effective opportunity to react.

In this context, the law faces the difficulty of applying legal institutions developed in a period when realistic and automated reproduction of a person’s identity was not possible at the current level. Privacy, image rights, personality rights, reputation protection and civil liability rules may offer some remedies. However, the complex nature of deepfakes raises the question of whether these mechanisms are sufficient when a person's image and voice can be artificially replicated without the use of an authentic recording in its original form.

At the same time, any legal intervention must take into account freedom of expression and the existence of legitimate uses of synthetic content. Satire, parody, artistic creation and certain forms of public commentary may use the modification of a person’s image without the aim of fraudulently misleading the public. For this reason, a general ban on the digital reproduction of identity could in turn create problems for the exercise of other fundamental rights. The challenge is thus to differentiate legitimate uses from deceptive, abusive or harmful ones.

Consent becomes central to this analysis. The authorized use of a person’s image or voice in a particular context should not automatically be interpreted as an unlimited authorization for the digital reproduction of their identity. Artificial intelligence allows for the re-use and transformation of individual characteristics in ways that could not have been anticipated at the time of the initial consent. It is therefore necessary to analyze the limits of consent and the conditions under which a person can control the subsequent uses of their digital representation.

This article aims to analyze the impact of deepfake technology on the legal protection of individual identity and to determine the extent to which existing legal mechanisms can respond to the risks generated by the artificial reproduction of image and voice. The analysis focuses on the relationship between digital identity, consent, privacy, reputation and personality rights, as well as on the liability issues that arise when synthetic content causes harm.

The research hypothesis starts from the idea that the development of artificial intelligence not only creates new ways of using personal information, but also changes the very ability of third parties to reproduce a person's identity. For this reason, legal protection must go beyond the simple reaction to fake content and take into account the person's control over the use of characteristics by which they can be identified.

Consequently, the issue of deepfakes must be analyzed by striking a balance between technological innovation, freedom of expression and the protection of individual identity. The aim should not be to unduly limit the development of artificial intelligence, but to establish legal guarantees capable of preventing the use of technology as a tool for the substitution, manipulation or unauthorized exploitation of human identity. In a digital environment where images and voices can be reproduced with an ever-increasing degree of realism, protecting the individual's ability to control their own representation becomes an essential component of the legal protection of the person.

Deepfakes and Digital Identity: Conceptual and Legal Delimitations

The development of generative artificial intelligence has profoundly changed the way in which a person’s image, voice and other elements of identity can be reproduced in the digital environment. While the manipulation of photographs and audiovisual recordings existed long before the development of current artificial intelligence systems, contemporary technologies have changed both the quality of the result and the speed and accessibility of the creation process. Currently, the realistic reproduction of a person’s characteristics no longer necessarily requires exceptional technical resources, and this democratization of technology gives rise to legal issues that go beyond the simple falsification of an image.

The term “deepfake” is used, in a general sense, to describe synthetic or manipulated materials using artificial intelligence techniques, so that the image, voice or behavior of a person is reproduced or modified in a realistic manner. The result can create the impression that a person has said certain words, performed certain actions, or participated in a situation that did not exist in reality. The problematic legal nature of the phenomenon derives precisely from the ability of technology to use distinctive elements of a person to construct an apparently authentic representation. Recent research on deepfakes, digital replicas, and “digital twins” shows that unauthorized reproduction of identity can no longer be analyzed exclusively through traditional mechanisms for protecting intellectual property, since the damage can concern the individual’s very personality and control over the way in which he or she is represented (Bosher, 2026, pp.1-28).

In this context, the notion of digital identity is gaining increasingly important legal significance. A person’s identity is not represented exclusively by their name, but includes a set of characteristics through which the individual can be recognized by others. Facial image, voice, expressions, gestures, and other individual characteristics can function as identification elements. Artificial intelligence allows these characteristics to be detached from the natural person and reproduced in an environment where the individual no longer necessarily exercises control over the context in which they are used.

This separation between the real person and their digital representation is one of the fundamental legal characteristics of deepfakes. In the case of a traditional photograph, the image retains, in principle, a link to a moment in which the person was in front of a camera. In the case of a deepfake, this relationship can disappear almost completely. The person represented may never have performed the action depicted, may not have spoken the reproduced words, and may not even have participated in the process of creating the material. However, their identity is used to give the artificial representation the appearance of authenticity. For this reason, deepfake raises a different problem than simple falsification of information. When synthetic content reproduces an identifiable person, technology can lead to a form of appropriation of their personal characteristics. The image and voice become

resources that can be used to produce new representations, independent of the person's will. Recent legal literature on artificial intelligence and personality rights highlights precisely the difficulty of traditional law in responding to situations in which the image or voice of a person is artificially reproduced, without the problem being able to be fully resolved by the classical rules on copyright (Abdusattorov, 2025).

The distinction between the protection of the work and the protection of the person is essential. Copyright seeks, in principle, to protect an original creation and the interests of the holder of the rights to it. However, the natural image, voice or features of a person do not, by their mere existence, constitute a creative work belonging to that person in the traditional sense of copyright. Consequently, the synthetic reproduction of a voice or a face can cause significant harm without the mechanisms of copyright always offering a direct and complete remedy. This gap explains the growing interest in personality rights, the right to privacy, the right to image and the "right of publicity" mechanisms.

The problem is all the more relevant as deepfake can affect both public and private individuals. In the case of celebrities, artists or other people whose image has significant commercial value, unauthorized reproduction can cause obvious economic damage. A cloned voice can be used to perform a song, an actor's image can be introduced into a production, and the representation of a well-known person can be used to promote a product without the existence of a contract. In such situations, the use of deepfake can affect the person's ability to control the economic exploitation of their identity.

In the case of people who are not public figures, the damage can be of a different nature, but no less serious. A person's identity can be used for fraud, harassment, impersonation or the creation of false materials likely to affect their personal and professional life. The evolution of deepfakes has demonstrated that the legal issue cannot be limited to protecting celebrities, since the possibility of artificially reproducing the image and voice becomes accessible to any person for whom sufficient digital materials exist.

This reality leads to the need to differentiate between "identity" and "reputation". A deepfake can damage a person's reputation by attributing a false statement or conduct, but the damage to the identity can exist even in the absence of demonstrable damage to the reputation. A person may legitimately feel that they should have some level of control over the use of their own face or voice even if the artificial representation does not portray them in a negative light. Recent legal and philosophical research even goes so far as to suggest that the algorithmic use of a person's biometric characteristics can affect their authority over their own image and how their identity is projected in public space (Kirkpatrick, 2026).

This approach is important because it shifts the focus of the analysis from the question of whether the deepfake causes demonstrable material harm to the question of whether the person should have some legal authority over the artificial reproduction of their identity. In this sense, the issue is no longer exclusively one of repairing the harm after the material has been distributed, but also one of preventing unauthorized use.

Consent is a central element in delimiting legitimate uses from abusive ones. The use of the person's image or voice with their consent may be necessary in film productions, advertising, video games, or other creative activities. The problem arises when consent is absent, too general or has been given for a different purpose. The ability of artificial intelligence to reuse a person's characteristics requires a careful interpretation of the scope of consent. Consent to use a recording in a particular production should not automatically be considered consent to the unlimited generation of new artificial versions of the person.

The notion of "digital replica" becomes relevant precisely in this context. It aims to capture the situation where technology generates a realistic representation of a person's voice or image, such that the result can be perceived as coming from the real individual. The current legal debate on digital replicas shows that existing copyright and intellectual property mechanisms do not always cover the harm caused by the unauthorized reproduction of the

image and voice. Institutional analyzes have therefore highlighted the need to examine distinct forms of protection of the individual against unauthorized digital replicas (U.S. Copyright Office, 2024). However, the recognition of broad legal protection of identity is not without its difficulties. The image of individuals is legitimately used in journalism, documentary, satire, parody, art and social or political commentary. An absolute right of the individual to prohibit any artificial representation could lead to excessive restrictions on freedom of expression. For this reason, regulation must distinguish between content that seeks to mislead, exploit identity or cause harm and content that uses artificial modification in a clearly artistic, satirical or informative context.

This tension is particularly evident in the case of public figures. Overly broad protection of image could be used to limit criticism or satire. At the same time, the status of a public figure cannot lead to the conclusion that the voice or image of the individual becomes a resource available without any limit for the generation of false materials. The law must therefore identify criteria that allow the differentiation between legitimate representation and deceptive substitution of identity.

Contemporary legal analyzes of deepfakes show that the phenomenon is located at the intersection of several areas of law, including the protection of privacy, personality rights, intellectual property, defamation and legal liability. Comparative studies highlight the fact that the abusive uses of deepfakes can simultaneously affect the identity, reputation and patrimonial interests of the person, while the existing legal mechanisms are often fragmented between several institutions and regulatory regimes (Kukreti et al., 2025). This fragmentation can lead to practical difficulties for the victim, who must identify the appropriate legal basis depending on the concrete nature of the damage.

The cross-border nature of the digital environment further complicates the situation. A deepfake can be created in one state, published through a platform administered in another jurisdiction and viewed in many other countries. Personality rights and image protection are not regulated uniformly, and differences between legal systems can affect both the existence of the right and the possibility of obtaining a remedy. In this context, the speed of content distribution can exceed the speed of traditional legal mechanisms. Moreover, the harm caused by deepfakes can be of very different natures. In some situations, the aim is to obtain a financial advantage through impersonation. In others, the deepfake can be used to manipulate public opinion, defame or produce intimate content without consent. Recent legal literature on the global impact of technology highlights precisely this diversity of illicit uses and the need to adapt liability mechanisms to the complex nature of synthetic content (Kukreti et al., 2026).

The problem of identifying the author is, in turn, a major obstacle to legal protection. Content can be distributed through anonymous accounts, multiplied by other users and redistributed on multiple platforms in a very short time. Even if the victim theoretically has a legal remedy, exercising it can become difficult when the original author cannot be identified or is located in another jurisdiction. For this reason, the legal debate inevitably extends to the liability of platforms and the mechanisms through which harmful content can be quickly removed. Effective protection of digital identity must therefore pursue several objectives simultaneously. The person must benefit from mechanisms against the unauthorized reproduction and exploitation of his or her distinctive characteristics, but the legal framework must preserve the necessary space for freedom of expression and legitimate creation. Clear criteria must also be established regarding consent, the creator's liability, the distribution of the material and the obligations of digital intermediaries.

Deepfakes thus demonstrate that human identity has acquired a new dimension in the digital environment. The image and voice are no longer just ways in which a person is recognized, but can become data and resources used by generative systems to build completely new representations. This transformation requires the law to establish to what

extent the autonomy of the individual over his or her own representation extends and in which situations the interests of other people or freedom of expression justify its use.

Consequently, the protection of identity in the era of artificial intelligence cannot be reduced to verifying the true or false nature of a material. The fundamental legal issue concerns the control over the use of the elements by which the person is recognized and the consequences of their detachment from the real individual. As deepfakes become more realistic and easier to produce, the need to build a legal framework capable of protecting identity without hindering legitimate uses of technology becomes increasingly evident.

Consent and unauthorized use of image and voice

Consent is one of the central elements of identity protection in the context of deepfake technologies. The ability of artificial intelligence to realistically reproduce the image and voice of a person makes it possible to create digital representations without their participation and even without the person knowing the existence of the material. From a legal perspective, the problem concerns not only the false nature of the content, but also the use of distinctive elements of identity without the authorization of their owner. Recent legal literature highlights that deepfakes can cause harm to identity, reputation and dignity precisely through this non-consensual appropriation of the representation of a person (Tandy, 2024, art.6).

Image and voice are of particular importance because they allow the identification of the individual even in the absence of his or her name. Generative technologies can create a replica that is realistic enough for the public to believe that the person represented actually participated in the creation of the material. In such situations, harm can occur even if the original image or recording was not copied in the traditional sense. Artificial intelligence can generate new material that reproduces individual characteristics and creates the appearance of authenticity. This ability has led legal literature to pay increasing attention to the right of publicity, through which an individual can benefit from protection against certain unauthorized uses of their name, image, voice, or other elements of their identity (Marlan, 2025, art.4).

The problem is obvious in the case of commercial use. An artist's voice can be cloned to produce audio material, and the image of an actor or a well-known person can be artificially inserted into an advertisement without consent. Deepfakes thus put traditional models of identity protection under pressure, because digital reproduction can be achieved quickly and at relatively low cost. The legal literature on the right of publicity in the age of artificial intelligence highlights that the replication of the face and voice can affect a person's ability to control and economically exploit their own identity (Koski, 2024).

Consent, however, should not be analyzed solely in terms of the existence or absence of an agreement. Its scope is also important. A person may authorize the use of their image or voice in a particular production, without this authorization allowing the unlimited generation of new materials through artificial intelligence. Consent to use an existing recording should not automatically be interpreted as a general waiver of control over all digital replicas that may be created later. For this reason, contracts and authorization mechanisms need to more clearly specify the purpose, duration and ways in which the image or voice may be reproduced through generative technologies.

The consequences of lack of consent become particularly serious in the case of synthetic intimate content. Deepfakes can artificially place an image of a person in a sexual context in which they never participated. The harm does not disappear because the material is artificial, as the public may believe the representation to be authentic, and the victim may suffer consequences for their dignity (Rotaru, 2016), reputation and privacy. Legal analysis of synthetic intimate images shows that traditional definitions of intimate images do not always provide uniform protection when the material does not represent a real event, but has been artificially generated or modified (Dunn, 2024).

Moreover, the legal problem does not end at the moment the material is created. Distribution through digital platforms can quickly multiply the harm and make it difficult to completely remove the content. Even if the original material is deleted, copies may continue to circulate. Research on non-consensual intimate deepfakes has highlighted the limitations of mechanisms based predominantly on the removal of content after publication and the need for more effective preventive measures (Kira, 2024). However, identity protection cannot lead to the establishment of an absolute right to any reproduction of an image or voice. Satire, parody, artistic creation, journalism and other legitimate forms of expression may involve the representation or transformation of a person's image. For this reason, regulation must strike a balance between the individual's control over their identity and freedom of expression. Recent legal analyzes of deepfakes have highlighted the need to delimit deceptive and exploitative uses from forms of expression that benefit from legal protection.

Consent must therefore constitute an essential element of legal protection against the misuse of deepfakes, but its effectiveness depends on clearly defining the limits of authorization. The individual must be able to know and control the purpose for which their image or voice is used, and consent for a specific use should not be transformed into an unlimited authorization to create digital replicas. At the same time, the legal framework must ensure effective remedies against non-consensual uses that affect identity, privacy or reputation, without unduly restricting legitimate uses of the technology.

Legal liability and victim protection in the case of deepfakes

The misuse of deepfakes raises the issue of identifying the person or entity that should be held liable for the harm caused. Unlike traditional forms of infringement of personality rights, synthetic content can involve multiple actors: the person who creates the material, the user who publishes it, the people who redistribute it, and the platform through which it becomes accessible to a wide audience. This complex structure makes it difficult to establish a unitary form of liability, especially when the creator of the deepfake is anonymous or located in another jurisdiction. The legal literature highlights that existing law offers certain tools against the misuse of deepfakes, but these are dispersed between the protection of privacy, defamation, identity rights, and other forms of liability (Chesney & Citron, 2019).

The liability of the creator must be analyzed primarily in relation to the nature and purpose of the material. The creation of an artificial representation does not automatically give rise to legal harm, as deepfakes can have legitimate artistic, educational or satirical uses. The situation is different when the technology is used to mislead the public, to attribute false statements to a person, for fraud or to deliberately damage reputation. In such cases, the artificial nature of the material does not eliminate the harm. On the contrary, the realism of the technology can amplify its effects, as the recipient may believe that the reproduced image or voice is authentic. Legal analyzes of the phenomenon show that deepfakes create difficulties for traditional rules precisely because they can simultaneously affect reputation, privacy and the autonomy of the person over their own identity (Citron & Chesney, 2019).

A significant issue concerns liability for defamation. When a deepfake attributes a false statement or conduct to a person that is likely to damage their reputation, defamation rules can constitute a protective mechanism. However, they do not solve all forms of misuse of identity. A deepfake can make unauthorized use of a person's image or voice without necessarily conveying a defamatory message. For this reason, American legal literature has emphasized the limits of the exclusive application of defamation and the need to correlate it with the protection of privacy and the right of publicity (Silbey & Hartzog, 2019). Protecting the victim also becomes particularly difficult because of the speed of distribution of content. Once published online, a material can be copied and redistributed on numerous platforms before the affected person can obtain its removal. In these circumstances, the damages awarded subsequently are not always sufficient to restore the previous situation. Reputational,

professional or emotional damage can continue even after the removal of the initial source, since copies of the material may remain accessible. This reality justifies the development of rapid remedies capable of limiting distribution before the damage becomes difficult to control (Delfino, 2019).

An important role also falls to digital platforms. They are not usually the authors of the material, but their infrastructure can facilitate its distribution to millions of users. Regulation of deepfake content must therefore strike a balance between the responsibility of the author, the obligations of intermediaries and the protection of freedom of expression. Imposing a general obligation to remove all manipulated content would be excessive, as it could affect satire, artistic creation or other legitimate uses. In contrast, synthetic content used for fraud, impersonation or causing serious harm justifies faster notification, assessment and intervention mechanisms. The literature on deepfake governance highlights precisely the need to combine legal solutions with platform responsibility and technological mechanisms to identify manipulated content (Diakopoulos & Johnson, 2021).

Legal liability must also be analyzed in the light of the cross-border nature of the internet. The author, the victim and the platform may be located in different jurisdictions, and the material may produce effects simultaneously in several states. Differences between national regulations on privacy, image, defamation and identity can make it difficult to determine the applicable law and to enforce a decision. This highlights the limits of exclusively national protection in the face of a technology that enables the almost instantaneous global distribution of content (Kietzmann et al., 2020). Effective victim protection therefore requires more than the possibility of obtaining compensation after the damage has occurred. The legal framework must combine preventive measures with the possibility of rapid removal of content, identification of authors, right to compensation and proportionate obligations for digital intermediaries. At the same time, liability must be established according to the concrete role of each actor and the nature of the use of the deepfake. Such a model allows the protection of the identity and reputation of the person without transforming any synthetic content into a prohibited legal act.

Deepfakes, freedom of expression and the need for a legal balance

The regulation of deepfakes cannot be constructed exclusively from the perspective of protecting individual identity, since the use of a person's image or voice can also take place in contexts protected by freedom of expression. Artificial intelligence technologies can be used for satire, parody, artistic creation, education or commentary on issues of public interest. For this reason, the mere fact that a material has been artificially modified or generated cannot, in itself, constitute a sufficient criterion for its prohibition. Legal literature has highlighted that excessively broad regulation of deepfakes can have negative effects on freedom of expression and can even eliminate socially valuable uses of the technology (Silbey & Hartzog, 2019).

The main difficulty lies in delimiting legitimate content from that designed to mislead or cause harm. A satirical material in which the artificiality is obvious presents a different legal situation from a video created with the aim of making the public believe that a real person made a statement that he or she never made. The intention of the creator, the context of distribution, the likelihood of misleading and the nature of the harm are therefore important elements for the legal assessment. Research on the legal implications of deepfakes shows that one of the most difficult challenges lies precisely in developing rules that are precise enough to combat harmful uses without disproportionately restricting legitimate expression (Chesney & Citron, 2019).

This issue is particularly sensitive when the deepfake reproduces public figures or is used in political debate. Freedom of expression implies the possibility of criticizing, caricaturing and satirizing people involved in public life. At the same time, technology can be used to fabricate seemingly authentic statements or actions and to influence public perception.

Deepfakes can thus affect not only the rights of the person depicted, but also the quality of the information space in which citizens form their opinions (Vaccari & Chadwick, 2020).

One possible solution is transparency obligations regarding the artificial nature of the material. Clearly labelling content generated or manipulated by artificial intelligence can reduce the risk that the public will consider it authentic, without necessarily requiring its removal. Such an approach is particularly relevant for content that is not inherently illegal, but which can become misleading if the artificial origin is concealed. Research on the perception of deepfakes shows, however, that the mere existence of warnings does not eliminate all the effects of disinformation, which means that transparency must be combined with other protective measures (Dobber et al., 2021).

At the same time, regulation must take into account the development of technical tools to detect synthetic content. Automatic identification, digital markings and mechanisms regarding the provenance of content can help to differentiate authentic from artificially generated material. However, detection technology cannot be the only solution, as generation and identification methods are evolving simultaneously. The literature describes this relationship as a continuous competition between the improvement of deepfakes and the development of techniques for their detection. From a legal perspective, an effective framework must focus in particular on uses that cause or create a serious risk of harm. The consent of the individual, the purpose of the use, the misleading nature of the material, the context of publication and the consequences for the victim can constitute criteria for differentiating legitimate uses from abusive ones. In this way, both the lack of protection of the individual and the establishment of a general ban on a technology that may also have legitimate uses can be avoided. The legal debate on deepfakes shows, moreover, that an effective response requires a combination of legal rules with technological measures, digital education and the responsibility of the actors involved in the distribution of content.

Consequently, the protection of the right to identity must be achieved by striking a balance between the autonomy of the person over his or her image and voice and freedom of expression. Not every synthetic representation should be considered illicit, just as freedom of expression cannot justify every unauthorized use of identity. The essential criterion must be the concrete context in which the technology is used and the capacity of the material to mislead, exploit identity or cause harm. Such an approach allows the protection of the person without transforming the regulation deepfakes in a disproportionate restriction on digital creation and communication.

Conclusion

The development of deepfake technologies demonstrates that artificial intelligence is not only changing the way digital content is created, but also the relationship between the individual and their own identity. The possibility of realistic reproduction of a person’s image, voice and behavior without their participation or consent generates significant risks for privacy, reputation, dignity and control over personal identity.

The analysis highlights the fact that existing legal mechanisms offer some forms of protection, but the complex and rapidly evolving nature of deepfakes requires strengthening guarantees regarding consent, use of image and voice, liability for damages and protection of victims. At the same time, regulation must differentiate abusive and misleading uses from legitimate ones, such as satire, parody, artistic creation or expression on issues of public interest.

Consequently, protecting the right to identity in the age of artificial intelligence requires a balance between technological innovation, freedom of expression and individual autonomy. An effective legal framework should not aim to ban deepfake technology, but to ensure that its development does not transform a person's image and voice into resources that can be reproduced and exploited without limits and without responsibility.

References

Abdusattorov, S. (2025). AI-generated deepfakes and personality rights: A new frontier for intellectual property law. *International Journal of Artificial Intelligence, 5*(6). https://doi.org/10.71337/inlibrary.uz.ijai.114452

Bosher, H. (2026). Do deepfakes, digital replicas and human digital twins justify personality rights? *The Journal of World Intellectual Property,* 1–28. https://doi.org/10.1111/jwip.70020

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820. https://doi.org/10.15779/Z38RV0D15J

Citron, D. K., & Chesney, R. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs, 98*(1), 147–155.

Delfino, R. A. (2019). Pornographic deepfakes: The case for federal criminalization of revenge porn's next tragic act. *Fordham Law Review, 88*(3), 887–938.

Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society, 23*(7), 2072–2098. https://doi.org/10.1177/1461444820925811

Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2021). Do (microtargeted) deepfakes have real effects on political attitudes? *The International Journal of Press/Politics, 26*(1), 69–91. https://doi.org/10.1177/1940161220944364

Dunn, S. (2024). Legal definitions of intimate images in the age of sexual deepfakes and generative AI. *McGill Law Journal,* 69(4). https://doi.org/10.26443/law.v69i4.1626

Kira, B. (2024). When non-consensual intimate deepfakes go viral: The insufficiency of the UK Online Safety Act. *Computer Law & Security Review*, Volume 54, 106024, ISSN 2212-473X, https://doi.org/10.1016/j.clsr.2024.106024

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons, 63*(2), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006

Kirkpatrick, J.R. (2026). Deepfakes at face value: Image and authority. *arXiv.* https://arxiv.org/abs/2604.12490

Koski, R. M. (2024). Warhol, Drake, and deepfakes: Monetizing the right of publicity in the generative AI era. *Georgia State University Law Review, 40*(4), 981–1020.

Kukreti, A., Kumar, V., & Raj, A. (2025). Deepfakes, copyright, and personality rights: A cross-jurisdictional analysis. *Journal of Informatics Education and Research, 5*(4). https://jier.org/index.php/journal/article/view/4056

Kukreti, A., Kumar, V., & Raj, A. (2026). The impact of deepfake technology on legal systems: A global perspective. *ShodhKosh: Journal of Visual and Performing Arts, 7*(1S). https://doi.org/10.29121/shodhkosh.v7.i1s.2026.7068

Marlan, D. (2025). Generative identity theft: Criminalizing deepfakes using the right of publicity. *Akron Law Review*, *58*(3), Article 4.

Rotaru, I.-G. (2016). Plea for Human Dignity. *Scientia Moralitas. Human Dignity - A Contemporary Perspective, 1*(1), 29-43.

Silbey, J., & Hartzog, W. (2019). The upside of deep fakes. *Maryland Law Review, 78*(4), 960–966.

Tandy, B. (2024). Deepfakes: Identity misappropriation in the digital age. *Belmont Law Review, 12*(1), Article 6.

U.S. Copyright Office. (2024). *Copyright and artificial intelligence, Part 1: Digital replicas*. https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society, 6*(1). https://doi.org/10.1177/2056305120903408