



Legal and Criminological Challenges of Deepfakes in Digital Media

Mostafa Pakniat

Assistant Prof., Department of Law, Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran. E-mail: m.pakniat@alzahra.ac.ir

Abstract

Objective

The rapid development of artificial intelligence has transformed digital media, introducing new opportunities alongside significant legal and criminological challenges. Among these developments, deepfake technology has attracted particular attention because of its ability to generate highly realistic synthetic audio, images, and videos. Although this technology has legitimate applications in education, entertainment, commerce, and healthcare, its misuse has created new opportunities for identity theft, defamation, extortion, privacy violations, and the dissemination of misleading information. The increasing accessibility of deepfake tools, together with the rapid circulation of digital content, has intensified concerns about individual rights, public trust, and the effectiveness of existing criminal justice systems. In Iran, these concerns are particularly relevant because existing criminal legislation does not expressly address every form of deepfake-related misconduct. The difficulty of distinguishing manipulated content from authentic material also raises important questions concerning criminal liability, digital evidence, and judicial jurisdiction. Furthermore, deepfakes have altered conventional patterns of offending and victimization, challenging traditional approaches to crime prevention and criminal policy.

Against this background, the present study examines the legal and criminological challenges associated with deepfakes in digital media, with particular emphasis on the Iranian legal system. It seeks to determine whether existing criminal laws adequately address deepfake-related offenses, identify the changes this technology has introduced into patterns of offending and victimization, and assess whether criminalization alone constitutes an effective response. The study also considers the need to balance the protection of individual rights and public security with freedom of expression and legitimate technological innovation.

Citation: Pakniat, Mostafa (2026). Legal and Criminological Challenges of Deepfakes in Digital Media. *Media Management Review*, 5(3), 405-425. (in Persian)

Research Methodology

This study adopts a descriptive-analytical research design supported by a comparative approach. Data was collected through library research involving domestic and international legislation, academic publications, international documents, and selected judicial decisions concerning deepfake technology and related offenses. The legal analysis examines the capacity of Iranian criminal legislation to address deepfake-related misconduct, particularly the provisions of the Computer Crimes Law and the Islamic Penal Code. It also considers the difficulties associated with criminal liability, the authentication of digital evidence, judicial jurisdiction, and the responsibilities of digital platforms. The comparative component examines regulatory developments and selects legal approaches in the United States and the European Union. These jurisdictions provide a basis for discussing different approaches to criminalization, transparency requirements, platform regulation, and the protection of fundamental rights.

The criminological analysis draws on contemporary criminological perspectives and relevant traditional theories to examine how deepfake technology influences offending, victimization, crime prevention, and criminal policy. Attention is given to the technological characteristics of deepfakes and the limitations they create for conventional legal and criminological responses.

Findings

The analysis identifies several interconnected legal and criminological challenges. From a legal perspective, the findings indicate that existing Iranian criminal legislation covers certain harmful uses of deepfakes but does not adequately address all forms of misconduct involving synthetic media. Article 16 of the Computer Crimes Law, for example, criminalizes certain forms of digital alteration and manipulation that damage an individual's reputation. However, difficulties arise when synthetic content is created entirely through artificial intelligence rather than by altering existing material, or when harmful content falls outside the specific conditions established by existing offenses. The absence of an explicit legal definition of deepfakes further complicates the application of criminal law. Determining criminal liability presents another challenge, particularly when deepfake content involves multiple individuals, organizations, or digital platforms. Establishing the responsibilities of content creators, distributors, and platform operators requires consideration of their respective actions, knowledge, and legal obligations.

The authentication of digital evidence constitutes an additional concern. Highly realistic synthetic content may be presented as genuine evidence, while authentic recordings may be challenged as fabricated. These circumstances complicate judicial proceedings and highlight the need for specialized digital forensic expertise. The transnational nature of deepfake-related offenses also creates difficulties concerning jurisdiction, international cooperation, and the enforcement of judicial decisions. The comparative analysis highlights the continuing development of regulatory approaches in the United States and the European Union, particularly regarding harmful synthetic content, transparency obligations, and platform responsibilities. It also demonstrates the difficulties involved in designing legal responses that protect individuals without unnecessarily restricting freedom of expression. From a criminological perspective, deepfakes have transformed conventional patterns of offending by reducing the technical skills, financial resources, and direct personal contact

traditionally required for certain offenses. They have also introduced new forms of victimization. Individuals may experience identity misuse, reputational damage, psychological harm, and economic consequences without any previous interaction with offenders. These harms can extend beyond direct victims to their families, professional environments, and wider communities. The analysis also identifies limitations in traditional crime prevention strategies and criminological theories when applied to technologically mediated offenses characterized by anonymity, automation, and rapid dissemination.

Discussion & Conclusion

The study concludes that deepfake technology presents interconnected challenges that cannot be adequately addressed through criminalization alone. Although existing Iranian legislation provides mechanisms for prosecuting certain forms of deepfake-related misconduct, legal gaps and difficulties concerning criminal liability, digital evidence, and jurisdiction limit its effectiveness. The criminological analysis further suggests that deepfakes require reconsideration of traditional approaches to offending, victimization, and crime prevention. Their accessibility, capacity for widespread dissemination, and potential to generate harm without direct interaction between offenders and victims demand responses adapted to digital environments. Accordingly, the study emphasizes the need for a multidimensional approach combining appropriate legislative reform, advanced detection technologies, specialized digital forensic procedures, improved media literacy, international cooperation, and responsible platform regulation. Such an approach should protect individual identity, privacy, and public trust while preserving freedom of expression and the legitimate applications of synthetic media.

The study acknowledges limitations arising from the relatively recent development of deepfake technology, the limited availability of relevant Iranian judicial decisions, and the continuing evolution of international regulatory frameworks. Future research should further examine platform liability, digital evidentiary standards, victim protection, civil liability, and the relationship between deepfake regulation and fundamental rights.

Keywords: Deepfakes; Synthetic media; Post-truth; Legal framework; Criminological explanation.



چالش‌های حقوقی و جرم‌شناسی دیپ‌فیک‌ها در رسانه‌های دیجیتال

مصطفی پاک‌نیت

استادیار، گروه حقوق، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهراء، تهران، ایران. رایانامه: m.pakniat@alzahra.ac.ir

چکیده

هدف: این پژوهش با هدف تحلیل چالش‌های حقوقی و جرم‌شناختی ناشی از فناوری دیپ‌فیک در رسانه‌های دیجیتال، بررسی ظرفیت و خلأهای نظام کیفری ایران در مواجهه با جرایم مبتنی بر این فناوری و تبیین تأثیر آن بر الگوهای بزه‌کاری، بزه‌دیدگی و سیاست کیفری انجام شده است.

روش: پژوهش حاضر از نوع توصیفی - تحلیلی و با رویکرد تطبیقی است. داده‌ها به‌شیوه کتابخانه‌ای و با بررسی قوانین داخلی و خارجی، منابع حقوقی و علمی، اسناد بین‌المللی و برخی رویه‌های قضایی گردآوری و تحلیل شده‌اند. همچنین، آثار دیپ‌فیک بر بزه‌کاری و بزه‌دیدگی با استفاده از برخی نظریه‌های جرم‌شناختی معاصر بررسی شده است.

یافته‌ها: یافته‌ها نشان می‌دهد که مقررات کیفری ایران، به‌ویژه قانون جرایم رایانه‌ای، تنها بخشی از رفتارهای مجرمانه مرتبط با دیپ‌فیک را پوشش می‌دهد. فقدان تعریف و جرم‌انگاری اختصاصی، دشواری انتساب مسئولیت کیفری، احراز اصالت ادله دیجیتال، مسئولیت پلتفرم‌ها و ماهیت فرامرزی جرایم، از مهم‌ترین چالش‌های حقوقی هستند. از منظر جرم‌شناختی نیز کاهش هزینه ارتکاب، افزایش گمنامی، گسترش بزه‌دیدگی و دشواری پیشگیری، الگوهای سنتی بزه‌کاری و بزه‌دیدگی را دستخوش تحول کرده است.

نتیجه‌گیری: مقابله مؤثر با دیپ‌فیک، فقط از طریق جرم‌انگاری امکان‌پذیر نیست و مستلزم رویکردی چندبعدی شامل اصلاح مقررات، توسعه فناوری‌های تشخیص، ارتقای سواد رسانه‌ای، تخصصی‌سازی احراز اصالت ادله، تنظیم‌گری پلتفرم‌ها و تقویت همکاری‌های بین‌المللی است.

کلیدواژه‌ها: دیپ‌فیک؛ رسانه مصنوعی؛ پسا حقیقت؛ چارچوب حقوقی؛ تبیین جرم‌شناسی.

استناد: پاک‌نیت، مصطفی (۱۴۰۵). چالش‌های حقوقی و جرم‌شناسی دیپ‌فیک‌ها در رسانه‌های دیجیتال. *بررسی‌های مدیریت رسانه*، ۵(۳)، ۴۰۵-۴۲۵.

تاریخ دریافت: ۱۴۰۴/۰۵/۲۵

تاریخ ویرایش: ۱۴۰۵/۰۳/۱۸

تاریخ پذیرش: ۱۴۰۵/۰۵/۰۲

تاریخ انتشار: ۱۴۰۵/۰۶/۳۰

doi: <https://doi.org/10.22059/mmr.2026.400793.1198>

بررسی‌های مدیریت رسانه، ۱۴۰۵، دوره ۵، شماره ۲، صص. ۴۰۵-۴۲۵

ناشر: دانشکدگان مدیریت دانشگاه تهران

نوع مقاله: علمی - پژوهشی

© نویسندگان

مقدمه

فناوری‌های مبتنی بر هوش مصنوعی در سال‌های اخیر تحولات گسترده‌ای در حوزه رسانه‌های دیجیتال ایجاد کرده‌اند. یکی از مهم‌ترین این فناوری‌ها، «دیپ‌فیک»^۱ است که با استفاده از الگوریتم‌های یادگیری عمیق، امکان تولید محتوای صوتی و تصویری جعلی، اما بسیار واقعی را فراهم می‌کند. دیپ‌فیک‌ها قادرند چهره، صدا یا حرکات اشخاص را به گونه‌ای بازسازی کنند که تشخیص جعلی بودن آن‌ها برای کاربران عادی و حتی متخصصان دشوار باشد.

اصطلاح «دیپ‌فیک» را به کاربری در ردیت^۲ نسبت می‌دهد. ردیت یکی از پرطرفدارترین شبکه‌های اجتماعی دنیاست و می‌توان جدیدترین اخبار و ترندهای^۳ روز را در آن پیدا کرد. این شبکه اجتماعی در اواخر ۲۰۱۷، ویدئوهایی منتشر کرد که در آن‌ها، چهره بازیگران مشهور روی صحنه‌های پورن قرار گرفته بود (وسترلوند^۴، ۲۰۱۹).

در ۲۹ آگوست ۲۰۱۹ برنامه زائوآپ^۵ در چین معرفی شد. زائوآپ، یک برنامه جابه‌جایی چهره است که با استفاده از کلیپ‌های متنوع فیلم‌ها و برنامه‌های تلویزیونی، چهره بازیگران را با سلفی‌های کاربران جایگزین می‌کند. نتایج آن‌قدر خیره‌کننده بود که این برنامه تنها در سه روز به فروبارترین برنامه چین تبدیل شد و موفق‌ترین نمونه کاربری فناوری «دیپ‌فیک» تا به امروز لقب گرفت (دستا^۶، ۲۰۲۱). تولید ویدئو با استفاده از یک عکس، نماهنگ یا فایل گیف^۷ در کمتر از هشت ثانیه امکان‌پذیر است. اندکی پس از انتشار این نرم‌افزار، آلی‌پی^۸، بزرگ‌ترین اپراتور پرداخت آنلاین چین با بیش از ۱ میلیارد کاربر فعال، هشدار داد که کاربرانش باید [با] اطمینان خاطر بدانند که هرچقدر هم فناوری جابه‌جایی چهره پیشرفته باشد، قادر به فریب برنامه‌های پرداخت ما نیست^۹.

فناوری پشت دیپ‌فیک‌ها، به سرعت در حال پیشرفت است. همگانی‌سازی (دموکراتیزه‌شدن) دیپ‌فیک‌ها از طریق اپلیکیشن‌های عمومی، پرسش‌های متعددی را در حوزه‌های حقوقی، اجتماعی و اخلاقی از جمله ضرورت مداخله قانونی یا ممنوعیت کامل دیپ‌فیک‌ها و نیز شیوه کارآمد رهنمود جوامع آنلاین و نیز تکالیف اپراتورها و... مطرح کرده است (راملاکان^{۱۰}، ۲۰۲۴: ۲۸۲).

جعل‌های عمیق به دلیل پتانسیل محو کردن مرز میان واقعیت و جعلیت، نگرانی‌های جدی ایجاد می‌کنند (یاماوکا انکرلین^{۱۱}، ۲۰۲۴). در حالی که این فناوری، کارکردهای مشروع تجاری، هنری و آموزشی دارد، استفاده مجرمانه از آن در قالب جعل هویت، پورنوگرافی انتقامی، اخاذی، انتشار اخبار جعلی و دست‌کاری افکار عمومی، نظام‌های حقوقی را با چالش‌های جدی مواجه ساخته است.

1. Deep Fake

2. Reddit

3. Trends

4. Westerlund

5. ZaoApp

6. De Seta

7. GIF

8. Alipay

9. <https://www.bbc.com/news/technology-49570418>

10. Ramluckan

11. Yamaoka-Enkerlin

سؤال‌ها و فرضیه‌های پژوهش

این پژوهش در صدد پاسخ به پرسش بنیادین زیر است:

- مهم‌ترین چالش‌های حقوقی ناشی از دیپ‌فیک در نظام حقوقی ایران چیست؟
- آیا قوانین کیفری موجود ایران، توانایی پاسخ‌گویی به جرایم مبتنی بر دیپ‌فیک را دارند؟
- دیپ‌فیک چه تحولاتی در الگوهای بزه‌کاری و بزه‌دیدگی ایجاد کرده است؟
- آیا صرف جرم‌انگاری می‌تواند راه‌کار مؤثری برای کنترل دیپ‌فیک باشد؟

فرضیه اصلی پژوهش آن است که مقررات کیفری موجود ایران برای مواجهه با همه اشکال دیپ‌فیک کافی نیستند. دیپ‌فیک موجب تحول در الگوهای سنتی بزه‌کاری و بزه‌دیدگی شده است. سیاست کیفری صرف، بدون توسعه ابزارهای فنی و تنظیم‌گری پلتفرم‌ها، کارآمد نخواهد بود.

روش‌شناسی پژوهش

این پژوهش از نوع توصیفی - تحلیلی و با رویکرد تطبیقی انجام شده است. داده‌های پژوهش به‌شیوه کتابخانه‌ای و از طریق مطالعه منابع حقوقی، مقاله‌های علمی، اسناد بین‌المللی، قوانین داخلی و خارجی و برخی رویه‌های قضایی مرتبط گردآوری شده‌اند. همچنین مقررات و رویه‌های ایالات متحده آمریکا و اتحادیه اروپا درباره دیپ‌فیک بررسی و وضعیت حقوق ایران تحلیل شده است. همچنین در بخش جرم‌شناختی، آثار دیپ‌فیک بر الگوهای بزه‌کاری، بزه‌دیدگی و سیاست کیفری با استفاده از نظریه‌های جرم‌شناسی معاصر مطالعه شده است.

پیشینه پژوهش

در سال‌های اخیر پژوهش‌های متعددی درباره دیپ‌فیک انجام شده است. در ادبیات فارسی، اکبری، آقاپور و آقاپور (۱۴۰۱)، به تحلیل سیاست کیفری ایران در قبال دیپ‌فیک و چالش‌های حقوق بشری آن پرداخته‌اند. همچنین شیرینی (۱۴۰۱)، دیپ‌فیک را از منظر حقوق کیفری ایران بررسی و به خلأهای قانونی موجود اشاره کرده است.

در ادبیات خارجی، دلفینو^۱ (۲۰۲۳) چالش‌های دیپ‌فیک در نظام دادرسی آمریکا و ضرورت توسعه نقش نظارتی قضات را بررسی کرده است. سیترون^۲ (۲۰۱۹) نیز بر تأثیر دیپ‌فیک بر حریم خصوصی جنسی و کرامت انسانی تأکید کرده است. همچنین پژوهش‌های لایوز^۳ (۲۰۲۳) و سایر نویسندگان این حوزه، ابعاد تنظیم‌گری و چالش‌های کلی این پدیده را مورد مطالعه قرار داده‌اند.

با وجود این، اغلب پژوهش‌های موجود یا صرفاً بر جنبه‌های فنی تمرکز داشته‌اند یا تنها به جرم‌انگاری دیپ‌فیک پرداخته‌اند. تحلیل هم‌زمان ابعاد حقوقی، جرم‌شناختی و سیاست کیفری این پدیده در بستر حقوق ایران کمتر مورد توجه قرار گرفته است.

1. Delfino
2. Citron
3. Labuz

چالش‌های حقوقی

فقدان چارچوب قانونی مشخص

یکی از اصول بنیادین حقوق کیفری آن است که هیچ رفتاری را نمی‌توان جرم دانست، مگر آنکه قانون‌گذار پیش‌تر آن را جرم‌انگاری کرده باشد؛ بنابراین نظام حقوقی برای مقابله با پدیده‌های نوظهور نیازمند مقررات روشن، شفاف و متناسب است.

فناوری دیپ‌فیک به دلیل نوظهور بودن، هنوز در بسیاری از نظام‌های حقوقی فاقد تعریف قانونی دقیق است. در ایران نیز هیچ قانون اختصاصی درباره دیپ‌فیک وجود ندارد و مقررات موجود تنها بخشی از رفتارهای مرتبط با آن را پوشش می‌دهند. این وضعیت موجب شده است تا بسیاری از رفتارهای زیان‌بار مبتنی بر دیپ‌فیک، خارج از قلمرو صریح قوانین کیفری قرار گیرند.

در حالی که تنها معدودی از دولت‌ها مقررات هوش مصنوعی که بستر تولید دیپ‌فیک‌هاست را تصویب کرده‌اند، اکثریت به دلیل نگرانی‌ها درباره آزادی بیان از این کار خودداری کرده‌اند (دلفینو، ۲۰۲۳). همچنین اگرچه اکثر پلتفرم‌های آنلاین، مانند یوتیوب، مکانیسم‌های قانونی متعددی برای کنترل محتوای منتشرشده اجرا کرده‌اند، این فرایند همچنان پرهزینه و زمان‌بر باقی مانده است (راملاکان، ۲۰۲۴: ۲۸۵).

قالب‌های حقوقی موجود، از مواجهه با فناوری جدید ناتوان نیستند؛ اما قوانین می‌بایست همگام با تحولات فناوری تکامل یابد. اذعان به این امر که سرعت تحولات فناورانه، به تنظیمگری هم‌زمان مجالی نمی‌دهد و همواره یک گام جلوتر از آن است، خالی از وجه نیست.

قانون‌گذاران اتحادیه اروپا، در ژوئن ۲۰۲۴، قانون هوش مصنوعی (AI) را امضا کردند. این قانون هوش مصنوعی، اولین قانون الزام‌آور افقی جهانی در خصوص هوش مصنوعی، چارچوب مشترکی را برای استفاده و عرضه سیستم‌های هوش مصنوعی در اتحادیه اروپا تعیین می‌کند. این قانون جدید، طبقه‌بندی سیستم‌های هوش مصنوعی را با الزامات و تعهدات مختلف متناسب با یک رویکرد «مبتنی بر ریسک» ارائه می‌دهد. در این قانون مقرر شد محتوایی که با کمک هوش مصنوعی تولید یا اصلاح می‌شود، مانند تصاویر، فایل‌های صوتی یا ویدئویی (به عنوان مثال، دیپ‌فیک) باید به‌وضوح به‌عنوان تولید شده توسط هوش مصنوعی برچسب‌گذاری شوند تا کاربران هنگام مواجهه با چنین محتوایی آگاه شوند.

کالیفرنیا، تگزاس و ویرجینیا، اولین ایالت‌هایی بودند که در سال ۲۰۱۹ قوانینی را برای تنظیم دیپ‌فیک تصویب کردند. در سال ۲۰۲۵، اما قانون فدرال حذف محتوا^۱ در ایالات متحده نیز که برای مبارزه با پدیده «پورن انتقامی»^۲ طراحی شده بود، در کنگره تصویب و به قانون تبدیل شد. این قانون در مورد محتوای مستهجن ایجاد شده توسط دیپ‌فیک اعمال می‌شود و آن را به اولین قانون فدرال تبدیل می‌کند که به‌صراحت به خطرهای دیپ‌فیک ناشی از هوش مصنوعی می‌پردازد. قانون حذف محتوا، انتشار پورن انتقامی را جرم‌انگاری می‌کند و پلتفرم‌های رسانه‌های اجتماعی را

1. Take It Down
2. Revenge Porn

ملزم می‌کند که ظرف یک سال فرایندهای اطلاع‌رسانی و حذف را ایجاد کنند. پیش از این، قربانیان راه‌های قانونی کمی برای درخواست حذف چنین محتوایی داشتند. دیپ‌فیک‌ها به‌طور گسترده در پورن انتقامی استفاده شده‌اند، جایی که محتوای صریح جنسی با حضور یک فرد بدون رضایت او در وبسایت‌های مستهجن منتشر می‌شود.

در بسیاری از نظام‌های حقوقی، پرونده‌های کیفری مرتبط با دیپ‌فیک، هنوز نسبتاً تازه و در مراحل اولیه هستند. در آمریکا تا قبل از تصویب قانون حذف محتوا، یک قانون فدرال در خصوص دیپ‌فیک وجود نداشت؛ بنابراین در پرونده‌هایی که تا پیش از این طرح می‌شد حتی بدون وجود «تعریف خاص کیفری برای دیپ‌فیک»، از ساختارهای قانونی موجود در نظام قضایی آمریکا، علیه تولید و توزیع محتوای جنسی سوءاستفاده از کودکان بهره‌برده‌اند. این رویکرد عملی، در عمل قضایی می‌تواند سرآغازی برای توسعه مقررات جدید باشد؛ زیرا دادستان‌ها و قضات با موارد پیچیده‌ای مواجه می‌شوند که در آن فناوری هوش مصنوعی به‌گونه‌ای به‌کار رفته که آسیب واقعی ایجاد کرده است، بنابراین توان لازم برای استفاده از قوانین سنتی علیه بهره‌برداری و تعرض به حقوق کودکان را فراهم می‌آورد.^۱

رویه قضایی ایران تاکنون واکنش مناسبی در برابر فناوری دیپ‌فیک، ماهیت، آثار و گونه‌های مجرمانه آن نشان نداده است. این انفعال عمدتاً به دلیل نوظهور و ناشناخته بودن این فناوری در حوزه حقوق است. فقدان تعریف مشخص برای دیپ‌فیک در قوانین ایران و نبود مجازات‌های متناسب با شدت جرایم مرتبط با آن یک چالش اساسی است (شیری، ۱۴۰۱).

شناسایی جرایم مرتبط با دیپ‌فیک‌ها

یکی از قواعد مهم دادرسی کیفری آن است که قاضی باید رفتار ارتكابی را با یکی از عناوین مجرمانه مصرح در قانون تطبیق دهد. هرچه رفتارهای نوظهور پیچیده‌تر باشند، امکان انطباق آن‌ها با جرایم سنتی دشوارتر می‌شود. در نظام حقوقی ایران، به‌صورت سنتی، جرم جعل که شاید نزدیک‌ترین عنوان مجرمانه مرتبط با دیپ‌فیک‌ها باشد، در نوشته‌ها یا امضا و مهر و نظیر آن محقق می‌شود و حال آن که در دیپ‌فیک‌ها با صوت و تصویری غیرحقیقی مواجهیم. همچنین ماده ۱۶ قانون جرایم رایانه‌ای تغییر و تحریف فیلم یا صوت یا تصویر را به‌وسیله سامانه‌های رایانه‌ای یا مخابراتی، به نحوی که عرفاً موجب هتک حیثیت او شود، جرم‌انگاری کرده است؛ اما این جرم شامل همه انواع دیپ‌فیک نمی‌شود؛ چراکه برخی دیپ‌فیک‌ها ممکن است بدون قصد هتک حیثیت یا حتی بدون انتشار عمومی تولید شوند. در چنین مواردی،

۱. برای مثال در پرونده کین جیمز کلهم، مرد ۳۹ ساله‌ای از شهر Baytown در نزدیکی هوستون، به اتهام چهار فقره جرم فدرال از جمله بهره‌برداری جنسی از کودکان و تولید تصاویر منافی با حقوق کودک تحت تعقیب قرار گرفت. او با استفاده از ابزارهای هوش مصنوعی، ویدئوهایی ساخته که در آن چهره دو کودک (یک کودک ۳ ساله و یک نوجوان ۱۵ ساله) روی بدن‌های دیگران در صحنه‌های جنسی قرار گرفته‌اند. این پرونده از نظر قضایی قابل توجه است؛ زیرا یکی از نخستین موارد در ایالات متحده است که در آن تولید محتوای مصنوعی توسط هوش مصنوعی با استفاده از چهره‌های کودکان به‌عنوان بخشی از ادله فساد جنسی مورد پیگرد فدرال قرار می‌گیرد. مقامات FBI گفته‌اند که تعداد گزارش‌های مربوط به سوءاستفاده و محتوای کودک‌آزاری که شامل عناصر AI می‌شود، بسیار افزایش یافته است، اگرچه هنوز قانون فدرال در آمریکا به‌طور مشخص دیپ‌فیک را تعریف نکرده است و دادستان‌ها از قوانین موجود علیه بهره‌برداری جنسی از کودکان برای پیگیری استفاده می‌کنند. -<https://www.houstonchronicle.com/news/houston-texas/crime/article/baytown-man-charged-ai-child-abuse-material-21324118.php>

مقررات موجود قابلیت پاسخ‌گویی کامل ندارند. همچنین در بسیاری از موارد، دیپ‌فیک با استفاده از هوش مصنوعی ساخته می‌شود (لابوز، ۲۰۲۳) و بخش زیادی از دیپ‌فیک‌ها از ابتدا خلق می‌شوند (نه تغییر و تحریف یک محتوای واقعی، بلکه خلق محتوای جعلی بر پایه داده‌های چهره یا صدا)؛ بنابراین همسان دانستن ساخت با تغییر یا تحریف، مستلزم دست کشیدن از تفسیر مضیق قوانین کیفری است. با این وصف، خلق کامل محتوای جعلی منتسب به دیگری (نه صرفاً تغییر یا تحریف) در این ماده جرم‌انگاری نشده است.

در برخی موارد، فرد اعم از اینکه چهره سیاسی یا هنری شناخته شده باشد یا نباشد، فیلم یا تصویری از خود می‌سازد. همچنین در بسیاری از موارد نیز دیپ‌فیک به هتک حرمت فرد نمی‌انجامد، بلکه ممکن است به ناکامی در رقابت‌های سیاسی بین احزاب، اخلال در نظم و امنیت کشور و... منجر شود. تولید دیپ‌فیک بدون انتشار (مثلاً تولید در گروه‌های بسته یا برای تهدید خصوصی) تحت این ماده قابل پیگیری نیست، مگر همراه با انتشار؛ بنابراین برای دیپ‌فیک‌های تولید شده بدون قصد هتک حیثیت یا با کاربردهای غیرموهن و نیز برای صرف تولید بدون انتشار، این ماده کفایت نمی‌کند.

چنانچه دیپ‌فیک به عنوان یک وسیله متقلبان در جهت بردن مال دیگری به کار گرفته شود، منوط به حصول سایر شرایط، تحقق کلاهبرداری انکارپذیر نخواهد بود؛ اما در بیشتر موارد، دیپ‌فیک‌ها برای بردن مال دیگران به کار گرفته نمی‌شوند یا حتی با وجود بردن مال دیگران نیز ممکن است شرایط کلاهبرداری محقق نشده باشد. به هر روی امکان انطباق رفتار مرتکب با یکی از جرایم مقرر در نظام حقوقی همراه فراهم نیست.

دشواری شناسایی مسئولیت کیفری

اصل شخصی بودن مسئولیت کیفری یکی از پایه‌های اساسی حقوق کیفری معاصر است. در حقوق ایران نیز این اصل مورد پذیرش قانون‌گذار قرار گرفته و مطابق ماده ۱۴۱ قانون مجازات اسلامی مصوب ۱۳۹۲، «مسئولیت کیفری، شخصی است». همچنین به موجب ماده ۱۴۲ همین قانون، «مسئولیت کیفری، به علت رفتار دیگری، تنها در صورتی ثابت است که شخص به‌طور قانونی مسئول اعمال دیگری باشد یا در رابطه با نتیجه رفتار ارتكابی دیگری، مرتکب تقصیر شود». بر این اساس، انتساب جرم و مجازات اصولاً مستلزم احراز رفتار مجرمانه و عنصر روانی در شخص مرتکب است و نمی‌توان صرفاً به دلیل ارتباط سازمانی یا مدیریتی، مسئولیت کیفری را به اشخاص دیگر تسری داد.

با این حال، حقوق کیفری ایران همانند بسیاری از نظام‌های حقوقی نوین، امکان مسئولیت کیفری اشخاص حقوقی را نیز پذیرفته است. مطابق ماده ۱۴۳ قانون مجازات اسلامی، «در مسئولیت کیفری اصل بر مسئولیت شخص حقیقی است و شخص حقوقی در صورتی دارای مسئولیت کیفری است که نماینده قانونی شخص حقوقی به نام یا در راستای منافع آن مرتکب جرمی شود». بنابراین، هرچند عنصر مادی جرم توسط اشخاص حقیقی انجام می‌شود، اما در صورت تحقق شرایط قانونی، شخص حقوقی نیز به دلیل انتفاع از رفتار مجرمانه یا ارتکاب جرم در چارچوب فعالیت سازمانی خود، واجد مسئولیت کیفری خواهد بود. افزون‌براین، ماده ۲۰ قانون مجازات اسلامی نیز ضمانت اجرای کیفری قابل اعمال بر اشخاص حقوقی را پیش‌بینی کرده است؛ از جمله انحلال، مصادره اموال، ممنوعیت از فعالیت، ممنوعیت دعوت عمومی برای افزایش سرمایه و انتشار برخی از احکام محکومیت.

اما در حوزه مسئولیت کیفری اشخاص حقوقی در جرایم رایانه‌ای، مهم‌ترین مستند قانونی ماده ۱۹ قانون جرایم رایانه‌ای است که بر اساس آن، هرگاه جرایم رایانه‌ای به نام شخص حقوقی و در راستای منافع آن ارتکاب یابد، شخص حقوقی نیز دارای مسئولیت کیفری خواهد بود. مطابق این ماده، مسئولیت کیفری شخص حقوقی، در چهار فرض قابل تحقق است:

- نخست، زمانی که مدیر شخص حقوقی شخصاً مرتکب جرم رایانه‌ای شود؛
 - دوم، هنگامی که مدیر دستور ارتکاب جرم را صادر کرده و جرم واقع شود؛
 - سوم، در صورتی که یکی از کارکنان با اطلاع مدیر یا در نتیجه عدم نظارت وی مرتکب جرم شود؛
 - چهارم، هنگامی که تمام یا بخشی از فعالیت شخص حقوقی به ارتکاب جرایم رایانه‌ای اختصاص یافته باشد.
- تبصره نخست ماده مزبور، مدیر را شخصی می‌داند که اختیار نمایندگی، تصمیم‌گیری یا نظارت بر شخص حقوقی را دارد. همچنین مطابق تبصره دوم، مسئولیت کیفری شخص حقوقی مانع مجازات مرتکب حقیقی نخواهد بود و در صورت فقدان شرایط مقرر در ماده، تنها شخص حقیقی مسئول شناخته می‌شود.
- بر این اساس، در فرض انتشار یا ترویج دیپ‌فیک‌های مجرمانه از سوی پلتفرم‌های دیجیتال، چنانچه مدیران پلتفرم آگاهانه دستور انتشار چنین محتوایی را صادر کنند یا با علم به ماهیت مجرمانه آن، از حذف و جلوگیری از انتشار خودداری کنند یا در کل، بخشی از فعالیت اقتصادی پلتفرم بر تولید، توزیع یا بهره‌برداری از محتوای جعلی و فریب‌کارانه مبتنی باشد، امکان انتساب مسئولیت کیفری به شخص حقوقی وجود خواهد داشت. بنابراین، مسئولیت پلتفرم صرفاً به مدیر یا کارمند منتشرکننده محدود نمی‌شود و در صورت احراز شرایط قانونی، خود پلتفرم نیز به‌عنوان شخص حقوقی قابل تعقیب و مجازات خواهد بود. این امر به‌ویژه در مواردی اهمیت می‌یابد که انتشار گسترده دیپ‌فیک‌ها از طریق سازوکارهای الگوریتمی، تبلیغاتی یا اقتصادی پلتفرم صورت گرفته و منافع مستقیم یا غیرمستقیم برای آن ایجاد کرده باشد.

البته بحث میزان مسئولیت پلتفرم‌های دیجیتال مانند شبکه‌های اجتماعی در میزبانی، انتشار یا ترویج دیپ‌فیک‌ها، یکی از چالش‌برانگیزترین مباحث حقوق فناوری اطلاعات محسوب می‌شود؛ زیرا باید میان آزادی بیان و گردش آزاد اطلاعات از یک سو و حمایت از حریم خصوصی، امنیت عمومی و جلوگیری از انتشار محتوای زیان‌بار یا مجرمانه از سوی دیگر تعادل برقرار شود (سیترون، ۲۰۱۹). در بسیاری از نظام‌های حقوقی، پلتفرم‌ها تا زمانی که از ماهیت غیرقانونی محتوا آگاهی نداشته باشند و پس از اطلاع نیز اقدامات متعارف و سریع برای حذف یا محدودسازی آن انجام دهند، از مسئولیت مستقیم معاف شناخته می‌شوند. اما هرگاه پلتفرم در تولید، هدایت، ترویج یا تقویت انتشار محتوای مجرمانه نقش فعال ایفا کند، برای مثال از طریق الگوریتم‌های پیشنهاد محتوا، برجسته‌سازی در فهرست ترندها یا بهره‌برداری اقتصادی از انتشار دیپ‌فیک‌های زیان‌بار، امکان انتساب مسئولیت مستقیم به آن تقویت می‌شود.

از سوی دیگر، شناسایی دیپ‌فیک‌ها مستلزم بهره‌گیری از فناوری‌های پیچیده و پرهزینه است و همه پلتفرم‌ها از توانایی فنی یکسان برخوردار نیستند (دهقانی و صابری^۱، ۲۰۲۵). از این رو، این پرسش مطرح می‌شود که آیا پلتفرم‌ها

باید به شناسایی پیشگیرانه تمامی دیپفیک‌ها مکلف باشند یا فقط پس از اطلاع از وجود محتوای مجرمانه مسئولیت پیدا می‌کنند؟ همچنین آیا سطح مسئولیت، باید متناسب با ظرفیت فنی، منابع مالی، گستره فعالیت و تعداد کاربران هر پلتفرم تعیین شود؟ به نظر می‌رسد پاسخ مثبت به این پرسش منطقی‌تر باشد؛ زیرا نمی‌توان مسئولیت یک پلتفرم محلی و نوپا را با یک پلتفرم جهانی دارای زیرساخت‌های گسترده و فناوری‌های پیشرفته یکسان دانست. با وجود این، تردیدی وجود ندارد که هرگاه یک پلتفرم یا مدیران آن با علم به ماهیت مجرمانه محتوا، در انتشار، تقویت یا بهره‌برداری اقتصادی و سیاسی از دیپفیک‌های غیرقانونی مشارکت کنند، زمینه انتساب مسئولیت کیفری به اشخاص حقیقی و حقوقی دخیل فراهم خواهد شد.

دشواری اثبات جرم

یکی از اصول اساسی دادرسی کیفری آن است که قاضی باید بر پایه ادله معتبر و قابل اعتماد تصمیم‌گیری کند. ماده ۳۷۴ قانون آیین دادرسی کیفری تصریح می‌کند که رأی دادگاه باید مستدل، موجه و مستند به مواد قانون و اصولی باشد که براساس آن صادر شده است.

هرچه اصالت یک دلیل بیشتر قابل تردید باشد، احتمال خطای قضایی افزایش می‌یابد. دیپفیک‌ها این اصل بنیادین را با چالش جدی مواجه ساخته‌اند. امروزه ویدئوها، تصاویر و فایل‌های صوتی جعلی به گونه‌ای تولید می‌شوند که تشخیص مصنوعی بودن آن‌ها حتی برای متخصصان نیز دشوار است. در چنین شرایطی، امکان ارائه ادله جعلی در فرایند دادرسی افزایش می‌یابد.

طرف دعوا ممکن است ویدئویی جعل کند تا در پرونده به پیروزی برسد یا ممکن است با ویدئوی ساخته‌شده توسط شخص ثالث مواجه شود و در حالی که از جعلی بودن آن بی‌اطلاع است، بخواهد آن را به عنوان شاهد ارائه دهد. همچنین ویدئوهای جعلی ممکن است (تصادفی یا عمدی) به آرشیوهای تاریخی همواره معتبر، مانند آرشیوهای خبری، راه یابند. اگر متولی این اسناد وجود آن‌ها را کشف نکند، این خطر وجود دارد که هنگام احراز اصالت ادله در دادرسی، ناخواسته صحت آن را گواهی کند. در رویه عملی هم معمولاً قضات، ادله‌ای مانند ویدئو را تا زمان اثبات خلاف، اصیل فرض می‌کنند و اثبات جعلی بودن بر عهده طرف مقابل است.

در حال حاضر، تلاش‌های شناسایی دیپفیک‌ها از توسعه و انتشار آن‌ها عقب‌تر است که تشخیص آن‌ها را روزبه‌روز دشوارتر می‌کند و دادگاه‌ها نیز فاقد استانداردها، رویه‌ها یا مقرراتی برای مواجهه با این نگرانی هستند. این امر به چالش‌هایی برای قضات، وکلا و دیگر ذی‌نفعان در تعیین اعتبار ادله و اختلال در فرایندهای تأیید مدرک منجر می‌شود و در نهایت به فرسایش شدید نظام‌های عدالت و حاکمیت قانون بینجامد. در واقع نظام عدالت با چالش اثبات اصالت دیجیتال در دادگاه‌ها مواجه است (کاسیا، گوارنا، کاسو، زنگارا و باتیاتو^۱، ۲۰۲۵).

مشکل واقعی آینده، نه فقط تهیه یک دیپفیک بلکه سهولت ایجاد یک اکوسیستم کامل اطلاعات نادرست است: یک ویدئوی جعلی، همراه با حساب‌های تویتر جعلی که به ویدئو لینک می‌دهند، حساب‌های جعلی در انجمن‌های

گفت‌وگو که محتوای ویدئو را مورد بحث قرار می‌دهند، وبسایت‌های جعلی که میزبان ویدئو هستند و گزارش‌های خبری جعلی در مورد آنچه در ویدئو نشان داده شده تولید می‌کنند، حساب‌های اینستاگرام جعلی که ویدئوی جعلی تولید می‌کنند و غیره. عبور از یک محیط فریب‌چندلایه بسیار دشوار خواهد بود (دلفینو، ۲۰۲۳). در چنین شرایطی فریب نظام عدالت کیفری چندان دشوار نخواهد بود. این مسئله به افزایش منازعات کارشناسی منجر می‌شود و روند اثبات جرم را کند و پرهزینه می‌کند.

دیپ‌فیک‌ها اکنون، یعنی در عصر پساحقیقت، پیچیده‌تر و شناسایی‌شان دشوارتر شده است. با دسترسی آسان به برنامه‌های ساخت دیپ‌فیک، استفاده از آن در پنج سال گذشته، به شدت افزایش یافته است. تشخیص محتوای دیپ‌فیک از محتوای واقعی نیازمند فناوری‌های پیشرفته‌ای است که در سیستم قضایی ایران کمتر در دسترس است. با فرض وجود چنین فناوری، دادرسی‌ها طولانی‌تر خواهند شد؛ زیرا طرفین همیشه می‌توانند ادعا کنند شواهد علیه آن‌ها جعلی و ساختگی است (دلفینو، ۲۰۲۳).

به هر ترتیب دیپ‌فیک‌ها این خطر را افزایش می‌دهند که دادگاه به اشتباه مدارک ارائه شده را واقعی فرض کند. فرد محکوم، همواره می‌تواند در ملأ عام ادعای بی‌گناهی کند و بگوید دادگاه شواهد جعلی را واقعی پنداشته است. در برخی دادگاه‌های آمریکا، مواردی گزارش شده که دیپ‌فیک به عنوان مدرک در پرونده‌های دیگر قضایی ارائه شده و قضات درباره اعتبار آن اظهار نگرانی کرده‌اند. در نمونه‌ای، یک فیلم دیپ‌فیک به عنوان مدرک در دادگاه مطرح شد، اما قاضی آن را رد کرد؛ زیرا تشخیص جعلی بودن آن دشوار بود و مدرک نمی‌توانست مبنای تصمیم‌گیری واقع‌گرایانه باشد. قضات گفته‌اند که سیستم قضایی برای پذیرش محتوای مصنوعی آماده نیست.^۱

در مورد برخی جرایم، صرف اتهام یا تعقیب قضایی، ممکن است تأثیر بزرگی بر زندگی و حرفه‌ای فرد بگذارد. حتی اگر بعدها کاملاً آشکار شود که فرد به دلیل جعلی بودن ویدئوها، به ناحق به دلیل اعمال جنسی با کودکان تحت تعقیب قرار گرفته، آسیب‌ها ممکن است پیش‌تر وارد شده باشد. در مورد اخبار جعلی که به عموم منتشر، ولی بعدها تکذیب می‌شود، پیام جعلی اولیه (که اغلب هیجان‌انگیز و جنجالی است) توجه بسیار بیشتری نسبت به تکذیب‌نامه بعدی جلب می‌کند. جعل یا دست‌کاری ادله الکترونیکی در تحقیقات قضایی، از یک سو و نیاز به ارائه قرائن معقول و اثبات قصد مجرمانه، از سوی دیگر برای تعقیب مرتکبان، چالشی اساسی ایجاد می‌کند؛ زیرا کسب شواهد معقول دشوار است. همچنین تمایز بین هنر و هجو، بیشتر ذهنی است و این ابزارها اغلب برای مقاصد ناخواسته استفاده می‌شوند.

حضور دیپ‌فیک‌ها الزام ایجاد بخش‌های حقوقی - فنی متخصص در دادگاه‌ها و نهادهای قضایی را ضروری می‌سازد. این بخش‌ها باید مسئول مستندسازی، کارشناسی فنی و احراز اصالت شواهد دیجیتال باشند. بدون چنین تخصصی، ادعاهای «جعلی بودن» شواهد به منازعات طولانی و پرهزینه تبدیل می‌شود. در آینده، دادگاه‌ها نیاز به کارشناسان خبره در پلیس علمی دیجیتال و ابزارهای تشخیص دیپ‌فیک خواهند داشت تا از پذیرش ناخواسته شواهد جعلی جلوگیری کنند (کاسیا و همکاران، ۲۰۲۳؛ دلفینو، ۲۰۲۳).

1. <https://www.forbes.com/sites/larsdaniel/2025/12/08/deepfakes-are-entering-us-courtrooms-judges-say-theyre-not-ready/?utm>

صلاحیت قضایی

یکی از موضوعات مهم حقوق بین‌الملل کیفری آن است که صلاحیت محاکم هر کشور برای اعمال قوانین خود محدود به قلمرو حاکمیت آن است که در ماده ۳ قانون مجازات اسلامی هم تبلور یافته است. جرایم سایبری و فراملی معمولاً این اصل سنتی را با چالش مواجه می‌کنند.

فقدان یک رژیم حقوقی بین‌المللی جامع برای مقابله با دیپفیک، موجب شده است تا همکاری‌های قضایی و پلیسی در این زمینه محدود باقی بماند. تفاوت‌های حقوقی در جرم‌انگاری دیپفیک‌ها در کشورها (مثلاً در یک کشور جرم است، در کشور دیگر آزادی بیان محسوب می‌شود)؛ مشکلات فنی و عملی در اجرای احکام قضایی فراسرزمینی؛ فقدان یک رژیم حقوقی بین‌المللی جامع و هماهنگ برای این حوزه از چالش‌های مهم رسیدگی قضایی است. دیپفیک‌ها معمولاً در یک کشور تولید می‌شوند، از طریق پلتفرم‌های بین‌المللی میزبانی می‌شوند (مثلاً یوتیوب، تیک‌تاک، اینستاگرام) و در کشورهای متعددی دیده و بازنشر می‌شوند. قربانی ممکن است در یک کشور، تولیدکننده در کشور دوم و پلتفرم در کشور سوم باشد. اگر دیپفیک در سروری در ایرلند میزبانی شده، اما قربانی در ایران است و تولیدکننده در روسیه قرار دارد، کدام دادگاه صلاحیت رسیدگی دارد؟ آیا صرف قابل مشاهده بودن دیپفیک در یک کشور برای ایجاد صلاحیت کافی است؟

حتی اگر دادگاهی صلاحیت خود را اعمال کند و حکمی صادر کند، اجرای آن علیه پلتفرم خارجی یا تولیدکننده در کشور دیگر دشوار است. همکاری‌های بین‌المللی (مثل معاهده بوداپست درباره جرایم سایبری) هنوز در همه کشورها به‌طور کامل اجرا نشده‌اند و پاسخ‌دهی پلتفرم‌ها به دستورهای قضایی ملی بسیار متفاوت است (مسکیز، کالپوکینه، داوری و لیاودانسکاس^۱، ۲۰۲۰). جرایم دیپفیک اغلب فراملی هستند و پلتفرم‌های جهانی مانند ایکس یا یوتیوب در ایران به‌طور کامل تحت نظارت نیستند.

تأثیر دیپفیک بر اثرگذاری رسانه‌ای و اعتماد عمومی

دیپفیک‌ها نه تنها چالش‌های فردی و قضایی ایجاد می‌کنند، بلکه اثرگذاری رسانه‌ای شبکه‌های اجتماعی و رسانه‌های سنتی را به‌شدت دگرگون کرده‌اند. این فناوری با محوریت واقعیت و جعل، اعتماد عمومی به محتوای رسانه‌ای را کاهش می‌دهد و متغیرهایی مانند اعتبار رسانه، تعامل مخاطبان و درآمدهای تبلیغاتی را تحت تأثیر قرار می‌دهد.

در شبکه‌های اجتماعی، انتشار سریع دیپفیک‌ها، به گسترش اطلاعات نادرست منجر می‌شود و کاربران را نسبت به تمام محتواها مشکوک می‌کند. مطالعات نشان می‌دهد که مواجهه با دیپفیک‌ها، اعتماد به اخبار آنلاین را ۳۰ تا ۴۰ درصد کاهش می‌دهد و این اثر در مصرف‌کنندگان اخبار شبکه‌های اجتماعی پایدارتر است. در رسانه‌های سنتی (تلویزیون، روزنامه‌ها)، دیپفیک‌ها اعتبار گزارش‌های خبری را زیر سؤال می‌برند و مخاطبان را به سمت منابع غیررسمی سوق می‌دهند که به کاهش مخاطب و درآمد تبلیغاتی منجر می‌شود.

از منظر اقتصادی، کاهش اعتماد به‌طور مستقیم بر درآمد کسب‌وکارهای رسانه‌ای اثر می‌گذارد. برندها از تبلیغات در پلتفرم‌هایی که محتوای جعلی در آن‌ها شایع است اجتناب می‌کنند؛ زیرا ریسک آسیب به اعتبار خود را بالا می‌بینند. گزارش‌ها نشان می‌دهد که بحران اعتماد ناشی از دیپ‌فیک‌ها می‌تواند درآمد بخش رسانه دیجیتال را ۱۰ تا ۲۰ درصد کاهش دهد، به‌ویژه در حوزه تجارت الکترونیک و تبلیغات هدفمند. در ایران، با گسترش پلتفرم‌هایی مانند اینستاگرام و تلگرام، این پدیده می‌تواند چرخه معیوب کاهش اعتماد → کاهش تعامل → کاهش درآمد را تسریع کند.

از طرفی مقابله با دیپ‌فیک‌ها می‌تواند چالشی برای آزادی بیان تلقی شود؛ چنانکه در پرونده‌ای جدید^۱، چالش حقوقی برجسته در ایالات متحده مطرح شد که رسانه طنز سیاسی بابیلون بی^۲ و یک شهروند هاوایی را مقابل دادستان کل هاوایی قرار داد^۳. قضیه از اعتراض علیه قانون ایالتی هاوایی^۴ سرچشمه گرفت که «انتشار رسانه‌های همراه‌کننده انتخاباتی» را محدود می‌کرد و مجازات‌های کیفری و الزام به هشدار برای محتواهای تغییر یافته دیجیتال داشت. شاکیان استدلال کردند که این قانون آزادی بیان سیاسی محافظت‌شده توسط متمم اول قانون اساسی آمریکا را نقض می‌کند، به‌ویژه طنز و پارودی را که اساس کار رسانه است. قاضی فدرال در ژانویه ۲۰۲۶ قانون را «نامشخص» و دارای اثر خفه‌کننده بر آزادی بیان دانست و آن را غیرقانونی اعلام کرد و مانع اجرای آن شد. این حکم یک پیروزی برای دفاع از آزادی بیان سیاسی در فضای دیجیتال تلقی می‌شود و تأکید دارد دولت نمی‌تواند سخنرانی سیاسی، حتی طنز یا محتوای مصنوعی را صرفاً بر اساس احتمال همراه‌کننده بودن سرکوب کند.

چالش‌های جرم‌شناسی

دیپ‌فیک‌ها چالش‌های مهمی از جمله تغییر ماهیت بزه‌کاری، دشواری شناسایی مرتکبان به‌دلیل ناشناس‌ماندن در فضای مجازی، سرعت بالای انتشار محتوا در رسانه‌های دیجیتال به لحاظ جرم‌شناسی ایجاد کرده‌اند؛ همچنین، استفاده از هوش مصنوعی در تولید محتوای جعلی مرز میان واقعیت و جعل را تضعیف کرده و اعتماد عمومی به اطلاعات را کاهش می‌دهد. این وضعیت پیشگیری، کشف جرم و پاسخ کیفری مؤثر را با چالش‌های جدی مواجه ساخته است.

تغییر ماهیت بزه‌کاری

دیپ‌فیک موجب تحول در الگوهای سنتی بزه‌کاری شده است. در گذشته بسیاری از جرایم نیازمند حضور فیزیکی، تعامل مستقیم یا هزینه بالا بودند، اما امروز افراد می‌توانند تنها با یک رایانه و دسترسی اینترنتی، محتوای مجرمانه گسترده تولید کنند. دیجیتالی شدن ابزار ارتکاب جرم، محو مرزهای جرم سایبری و جرم سنتی، افزایش گمنامی بزه‌کار، کاهش هزینه و پیچیدگی ارتکاب جرم و ایجاد گونه‌های نوین بزه‌کاری، تحولی است که دیپ‌فیک‌ها در پدیده بزه‌کاری ایجاد کرده‌اند. ماهیت بزه‌کاری از «جرایم سنتی با عناصر انسانی» به سمت جرایم ماشینی و اتوماسیونی شده حرکت می‌کند.

1. The Babylon Bee, LLC v. Lopez (2026)

2. The Babylon Bee

۳. برای مطالعه در خصوص این پرونده، ر.ک:

https://dockets.justia.com/docket/hawaii/hidce/1%3A2025cv00234/174048?utm_source=chatgpt.com

4. (Act 191 / S2687)

چنین ماهیت بزه‌کاری ترکیبی سبب می‌شود تا الگوهای جرم‌شناسی کلاسیک برای تحلیل آن کافی نباشد (ساندوال، دآلمیدا وائو، سولاس و رودریگز^۱، ۲۰۲۴).

اگر در گذشته برای اخاذی جنسی^۲، بزه‌کار نیازمند تعامل واقعی با قربانی بود، امروزه می‌تواند فقط با تولید یک دیپ‌فیک مستهجن و ارسال آن به قربانی، تهدید و اخاذی کند. این امر به افزایش بزه‌کاری از راه دور^۳ منجر شده است (هریس^۴، ۲۰۱۹). گویی دیپ‌فیک‌ها، پلی میان جرایم سنتی (مانند افتراء، کلاهبرداری، تهدید) و جرایم سایبری ایجاد می‌کنند.

در گذشته تولید محتوای جعلی به مهارت زیاد و هزینه سنگین نیاز داشت؛ امروز با ابزارهای رایگان یا ارزان دیپ‌فیک، می‌توان با چند کلیک محتوای مجرمانه تولید کرد. این موضوع باعث می‌شود که افراد بیشتری وارد فعالیت مجرمانه شوند؛ چون هزینه ارتکاب جرم کاهش یافته است.

دیپ‌فیک‌ها به دلیل قابلیت محو کردن مرز میان واقعیت و دروغ، نگرانی‌های جدی ایجاد می‌کنند. دیپ‌فیک‌ها ابزاری برای ارتکاب انواع بزه‌های نوظهور و سنتی شده‌اند؛ مانند دیپ‌فیک‌های مورد استفاده در پورن انتقام‌جویانه، کمپین‌های سیاسی. یک دیپ‌فیک می‌تواند در عرض چند دقیقه در میلیون‌ها گوشی و شبکه اجتماعی منتشر شود. اثرهای تخریبی و دامنه بزه‌دیدگی بسیار وسیع و فوری می‌شود، در حالی که در گذشته برای انتشار محتوای مجرمانه، زمان و زحمت زیادی لازم بود.

دشواری‌های پیشگیری از جرم

یکی از اهداف جرم‌شناسی، طراحی راه‌کارهای پیشگیری از وقوع جرم است. با این حال، پیشگیری از جرایم مبتنی بر دیپ‌فیک، به دلیل ماهیت فناورانه آن بسیار دشوار است و فناوری‌های شناسایی دیپ‌فیک هنوز از سرعت توسعه این پدیده عقب‌ترند. فناوری‌های تولید دیپ‌فیک روز به روز در دسترس‌تر و ارزان‌تر می‌شوند. مهارت فنی پیچیده‌ای برای تولید دیپ‌فیک نیاز نیست. شناسایی به موقع دیپ‌فیک‌های مضر، هنوز یک چالش فنی است؛ سامانه‌های خودکار اغلب نمی‌توانند محتوای جعلی را به سرعت و با دقت کافی شناسایی کنند (لانگا^۵، ۲۰۲۱). ابزارهای پیشگیری کلاسیک مثل فیلترینگ، ردیابی IP، نظارت پلیسی کافی نیستند و باید مدل‌های نوین پیشگیری وضعی طراحی شود. ماهیت فراسرزمینی و پویایی فناورانه امکان اجرای راه‌کارهای عملیاتی برای پیشگیری از جرم را نمی‌دهد. پیشگیری وضعی یک پیشگیری مکان محور است و شیوه‌های سنتی این نوع پیشگیری در دیپ‌فیک‌ها تقریباً بی‌فایده است. فناوری‌های تشخیص نیز، عموماً دسترسی همگانی ندارند و پیشگیری از این جرایم را دشوار می‌نماید.

سرگردانی سیاست کیفری

سیاست کیفری باید میان حمایت از امنیت عمومی و حفظ آزادی‌های اساسی تعادل برقرار کند. اتخاذ یک سیاست کیفری

1. Sandoval, De Almeida Vau, Solaas & Rodrigues
2. sextortion
3. Remote crime
4. Harris
5. Langa

هوشمند و اقتضایی برای مواجهه با دیپ‌فیک‌ها امری بس دشوار است. از یک سو، دیپ‌فیک می‌تواند موجب آسیب‌های گسترده اجتماعی، سیاسی و فردی شود و ضرورت مداخله کیفری را توجیه کند. از سوی دیگر، جرم‌انگاری گسترده ممکن است آزادی بیان، طنز سیاسی و خلاقیت هنری را محدود سازد.

توضیح اینکه دیپ‌فیک‌ها می‌توانند کاربردهای مثبت درخور توجهی برای اهداف تجاری و اجتماعی داشته باشند. مثل بازسازی صدای افراد مبتلا به بیماری‌های از دست‌دهنده صوت. همچنین باعث صرفه‌جویی چشمگیر در هزینه و زمان تولید ویدئوهای مصنوعی می‌شود. تولید محتوای ویدئویی امروز، فرایندی پُرهزینه و توسعه‌ناپذیر است: شامل تجهیزات فیزیکی، چندین دوربین، استودیو و بازیگر. با فناوری جدید، تولید ویدئوهای مصنوعی تنها با ۱۰ درصد هزینه معمول ممکن است.

دیپ‌فیک‌ها می‌توانند به‌عنوان هجو یا هزل (مثل میم‌ها) استفاده شوند. پیچیدگی هنگامی پدیدار می‌شود که مثلاً باید دیپ‌فیک‌های طنزآمیز مجاز باشند و تمایز روشنی بین استفاده هنری و مخرب ایجاد شود. آیا سازنده دیپ‌فیک طنز که بعدها برای اخاذی استفاده می‌شود، قابل مجازات است؟ چگونه می‌توان سازنده یک دیپ‌فیک سیاسی را که با انگیزه ایدئولوژیک عمل کرده، «اصلاح» کرد؟ هر قانونی که بخواهد استفاده از دیپ‌فیک مقامات سیاسی را ممنوع کند، نیازمند بازنگری کامل برای انطباق با حمایت از آزادی بیان است (اکبری و همکاران، ۱۴۰۱). سیاست کیفری سنتی در مواجهه با دیپ‌فیک‌ها، دچار سه بحران معنایی است:

- بحران کارآمدی، به این معنا که پاسخ‌های کیفری موجود نه بازدارنده‌اند، نه اصلاح‌کننده؛
- بحران مشروعیت به این صورت که جرم‌انگاری‌های گسترده ممکن است به امنیتی و پلیسی شدن سیاست کیفری منجر شود؛
- بحران عدالت که قربانیان را به‌دلیل دشواری اثبات، دچار بزه‌دیدگی می‌کند.

تحول در بزه‌دیدگی

دیپ‌فیک‌ها شکل‌های جدیدی از بزه‌دیدگی ایجاد کرده‌اند. بسیاری از قربانیان بدون هیچ کنش یا ارتباطی با بزه‌کار، صرفاً به‌دلیل وجود تصاویر یا صدای خود در فضای مجازی هدف سوءاستفاده قرار می‌گیرند. زنان، کودکان، شخصیت‌های سیاسی و اقلیت‌ها بیش از سایر گروه‌ها در معرض آسیب‌های ناشی از دیپ‌فیک قرار دارند. پیامدهای این جرایم تنها به قربانی مستقیم محدود نیست، بلکه خانواده، محیط حرفه‌ای و اعتماد عمومی جامعه را نیز تحت تأثیر قرار می‌دهد.

این فناوری نه‌تنها دامنه بزه‌دیدگی را به گروه‌های بی‌سابقه گسترش داده، بلکه عمق آسیب‌ها را تا سطحی بی‌سابقه تشدید کرده است. قربانیان دیپ‌فیک را می‌توان در سه لایه متمایز تحلیل کرد:

- قربانیان غیرفعالی که صرفاً داده‌های بیومتریشان سرقت شده است؛
- گروه‌های آسیب‌پذیر تاریخی مانند زنان، کودکان و اقلیت‌ها که اکنون در خط مقدم حملات سایبری قرار گرفته‌اند (هریس، ۲۰۱۹)؛

• و بالاخره بزه‌دیدگان دیگری که در سایه‌ی این جرایم، متحمل آسیب‌های روانی، اجتماعی و اقتصادی می‌شوند.

هر یک از این سطوح، چالش‌های نظری و عمیقی را پیش روی جرم‌شناسی معاصر قرار می‌دهد.

سرقت هویت دیجیتال به مثابه نقض کرامت انسانی

دیپفیک‌ها مفهومی به نام «بزه‌دیدگی بدون کنش» را خلق کرده‌اند: افرادی که صرفاً به دلیل وجود تصاویر یا صداهایشان در فضای مجازی، به اشیائی برای جرایم سایبری تبدیل می‌شوند. یک شهروند معمولی ممکن است روزی کشف کند که چهره‌اش در ویدئویی مستهجن جعل شده، بدون آنکه کوچک‌ترین تعاملی با مجرم داشته باشد. اینجا، هویت بیولوژیک انسان به کالایی برای تخریب تبدیل می‌شود. پیامدهای این بزه‌دیدگی چندبعدی است:

- در سطح فردی، قربانی با بحران هویت مواجه می‌شود که انگ «قربانی جنسی» حتی پس از حذف محتوا، مانند داغ‌ننگی اجتماعی بر پیشانی فرد باقی می‌ماند.
- در سطح حقوقی، نظام‌های قضایی ناتوان از پیگیری سریع، بار اثبات بی‌گناهی را به دوش قربانی می‌گذارند.
- در اقتصاد دیجیتال، این افراد با محرومیت از فرصت‌های شغلی مواجه می‌شوند؛ چرا که جست‌وجوی نامشان در اینترنت با محتوای جعلی آلوده شده است. این شکل از بزه‌دیدگی، نظریه «حق بر هویت شخصی» فینبرگ را نقض می‌کند؛ حقی که اساس کرامت انسانی در عصر دیجیتال محسوب می‌شود.

گروه‌های آسیب‌پذیر: دیپفیک به مثابه ابزار بازتولید خشونت ساختاری

زنان، کودکان، اقلیت‌های نژادی و مذهبی و بازیگران سیاسی، هدف اصلی دیپفیک‌های مخرب قرار گرفته‌اند. برای زنان، این فناوری به ابزاری برای تشدید خشونت جنسیتی تبدیل شده است: بسیاری از این محتواها، بیشتر تحت عنوان «پورنوگرافی انتقام‌جویانه»^۱ منتشر می‌شوند؛ پدیده‌ای که در آن تصاویر خصوصی یا جعلی برای تحقیر و کنترل زنان استفاده می‌شود. آسیب‌های روانی این جرایم ویرانگر است: مطالعات بالینی نشان می‌دهند که ۴۵ درصد قربانیان افکار خودکشی را تجربه می‌کنند. برای اقلیت‌ها، دیپفیک‌ها به سلاحی برای تحریک خشونت جمعی تبدیل شده‌اند؛ مانند جعل سخنان توهین‌آمیز از رهبران مذهبی که می‌تواند به شورش‌های قومی بینجامد. سیاست‌مداران نیز با جعل سخنان فسادآمیز، اعتبارشان را پیش از انتخابات از دست می‌دهند؛ پدیده‌ای که طبق مطالعه دانشگاه آکسفورد (۲۰۲۴) تا ۳۰ درصد از اعتماد رأی‌دهندگان می‌کاهد. این الگوها تأیید می‌کنند که دیپفیک‌ها صرفاً یک فناوری نیستند، بلکه ابزارهای بازتولید نابرابری در جوامع هستند (هریس، ۲۰۱۹).

بزه‌دیدگان غیرمستقیم: تأثیر موجی روانی - اجتماعی

پیامدهای دیپفیک‌ها به قربانیان مستقیم محدود نمی‌شود؛ آن‌ها شبکه‌ای از آسیب‌های ثانویه خلق می‌کنند که خانواده‌ها، محیط‌های کاری و کل جامعه را دربرمی‌گیرد. در خانواده‌ها، کودکان قربانیان با بحران‌های رشدی مواجه

می‌شوند: کشف اینکه تصویر والدین در محتوای جنسی جعلی استفاده شده، می‌تواند به اختلال‌های دل‌بستگی و بی‌اعتمادی مادام‌العمر بینجامد. همسران قربانیان نیز اغلب قربانی سرزنش‌های اجتماعی می‌شوند؛ یعنی جامعه آن‌ها را به دلیل عدم محافظت از همسر مورد سرزنش قرار می‌دهد.

در محیط‌های حرفه‌ای، انتشار یک دیپ‌فیک مخرب علیه یک پزشک یا قاضی، اعتماد عمومی به کل نهاد را تخریب می‌کند. هم‌صنفان این افراد با کاهش فرصت‌های شغلی، انزوای اجتماعی و افزایش فشار روانی مواجه می‌شوند و در سطح کلان، جامعه دچار فرسایش سرمایه اجتماعی می‌شود.

دیپ‌فیک‌ها نشان داده‌اند که در فضای سایبر، «بزه‌دیده» مفهومی سیال و چندلایه است. این پدیده نه تنها گروه‌های سنتی آسیب‌پذیر را هدف گرفته، بلکه شهروندان عادی را به قربانیانی تبدیل کرده است که کوچک‌ترین کنشی در فرایند بزه‌دیدگی نداشته‌اند. گستردگی دامنه قربانی‌شدن، نظام‌های حقوقی و حمایتی را با چالشی بنیادین مواجه ساخته است: چگونه می‌توان به قربانیانی که هویتشان ربوده شده، کرامت بازگرداند؟ چگونه می‌توان از جوامع در برابر فرسایش اعتماد محافظت کرد؟ پاسخ به این پرسش‌ها مستلزم تحولی پارادایمی در بزه‌دیده‌شناسی است؛ تحولی که در آن «حق بر هویت دیجیتال»، به عنوان یک حق بنیادین به رسمیت شناخته شود، سازوکارهای حمایت از بزه‌دیدگان ثانویه توسعه یابد، و مقابله با خشونت ساختاری در فضای سایبر در کانون توجه قرار گیرد. بدون این تحول، دیپ‌فیک‌ها به ابزارهای همیشگی برای بازتولید رنج انسانی در مقیاسی بی‌سابقه تبدیل خواهند شد.

ناکامی در انطباق نظریه‌های جرم‌شناختی با پدیده دیپ‌فیک

نظریه‌های کلاسیک عمدتاً بر کنش انسانی فیزیکی و روابط اجتماعی چهره‌به‌چهره متمرکز بوده‌اند. دیپ‌فیک‌ها جرم‌های فناورمحور، اتوماسیونی و فرامکانی هستند. در نتیجه بسیاری از نظریه‌های سنتی در توضیح جرم‌های مبتنی بر دیپ‌فیک دچار محدودیت می‌شوند. اگر در نظریه کنترل اجتماعی، ضعف در پیوندهای اجتماعی یا کنترل درونی زمینه‌ساز ارتکاب جرم شناخته می‌شود، اما تولید دیپ‌فیک به ارتباط ضعیف اجتماعی وابسته نیست؛ می‌تواند از سر تفریح، هیجان، اهداف مالی یا سیاسی توسط افراد به ظاهر «متصل به جامعه» نیز انجام شود.

اگر ساترلند^۱ (۱۹۴۷) بر یادگیری جرم از گذر تعامل اجتماعی تمرکز می‌کند، در دیپ‌فیک‌ها، بزه‌کاران اغلب از فرهنگ‌های آنلاین، انجمن‌های مجازی، یا منابع آموزش ماشینی (مثل فیلم‌های یوتیوب) یاد می‌گیرند. بنابراین یادگیری از گذر تعامل شفاهی و الزاماً در گروه‌های اجتماعی کلاسیک رخ نمی‌دهد. مرتون که بر فشارهای اقتصادی و اجتماعی در کشاکش اهداف و وسایل نیل آن‌ها تأکید می‌کرد، شاید امروز در تبیین پدیده دیپ‌فیک‌ها می‌بایست بیشتر بر فرصت فناورانه آن هم از سر هیجان تکنولوژیک یا اثبات مهارت، تمرکز می‌کرد.

همچنین دیپ‌فیک‌ها گاه توسط افراد ناشناس و در محیط‌های ناشناس تولید می‌شوند که اساساً در معرض فرایند برچسب‌گذاری اجتماعی قرار نمی‌گیرند. بنابر این نظریه، برچسب‌زنی در این خصوص چندان مفید واقع نخواهد شد. البته در بسیاری موارد، مرتکبان دیپ‌فیک معمولاً برای خود توجیهات خنثی‌کننده می‌سازند؛ اغلب شاید در برخورد

این گونه افراد این جمله‌ها را شنیده‌ایم که: «فقط یه شوخی بود»، «خودش مقصره که عکس‌هاش در دسترس بود»، «همه این کار رو می‌کنند». این توجیهات با نظریه فنون خنثی‌سازی متزا سازگاری بیشتری دارد.

همچنین نظریه فرصت درصدد است که ارتکاب جرایم را با استفاده از مؤلفه‌های انگیزه، آماج، موانع و کنترل‌ها تشریح کند؛ یعنی جرم زمانی رخ می‌دهد که یک بزه‌کار انگیزه‌مند، هدف مناسب و نبود کنترل مؤثر هم‌زمان شوند. در دیپفیک‌ها شاید این وضعیت حاکم باشد. حجم عظیم تصاویر/صداها در شبکه‌های اجتماعی (اینستاگرام، تلگرام) به عنوان اهداف مناسب و البته جذاب را در دسترس قرار می‌دهد. نبود محافظت کارآمد (فقدان فیلترهای مؤثر و ضعف مقررات‌گذاری پلتفرم‌ها به علاوه ناکارآمدی ابزارهای تشخیصی) و گسترش دسترسی به فناوری و کاهش هزینه و ناشناس ماندن، انگیزه مجرمان غیرحرفه‌ای را افزایش داده است.

نتیجه‌گیری

پژوهش حاضر نشان داد که فناوری دیپفیک به یکی از پیچیده‌ترین چالش‌های حقوقی و جرم‌شناختی عصر دیجیتال تبدیل شده است. این فناوری کاربردهای مشروع و مفید در حوزه‌های آموزشی، تجاری، رسانه‌ای و درمانی دارد؛ اما استفاده مجرمانه از آن، پیامدهای گسترده‌ای برای نظام عدالت کیفری، امنیت اجتماعی و اعتماد عمومی ایجاد کرده است.

چالش‌های پیچیده حقوقی مرتبط با فناوری دیپفیک‌ها، تنظیم مقررات آن‌ها را دشوار می‌کند. هرچند راه‌های دیگری در نظام حقوقی برای تعقیب استفاده‌ی مخرب از فناوری دیپفیک وجود دارد، نگرانی اصلی بر ضرورت تدوین قوانین یا مقررات خاصی متمرکز است که هم از حق هویت افراد محافظت کند و هم آزادی بیان را پاس بدارد. فناوری دیپفیک سلاحی قدرتمند باقی می‌ماند که اگر کنترل نشود، مرز بین واقعیت و خیال را محو می‌کند. این امر نیاز مبرم به ایجاد یک چارچوب قانونی یا نظارتی جامع‌الشمول را برجسته می‌سازد که باید تمام جنبه‌های فناوری دیپفیک و کاربردهای آن را پوشش دهد.

بررسی وضعیت حقوق ایران حکایت از این دارد که مقررات موجود، به ویژه قانون جرایم رایانه‌ای، تنها بخشی از رفتارهای مبتنی بر دیپفیک را پوشش می‌دهد و برای مقابله با همه اشکال این پدیده کافی نیست. همچنین فقدان تعریف قانونی مشخص، ابهام در مسئولیت پلتفرم‌ها، دشواری اثبات اصالت ادله دیجیتال و ماهیت فراملی جرایم، کارآمدی سیاست کیفری را با محدودیت مواجه ساخته است.

مطالعه تطبیقی نیز نشان داد که حتی نظام‌های حقوقی پیشرفته مانند ایالات متحده و اتحادیه اروپا هنوز در حال توسعه چارچوب‌های تنظیم‌گری مناسب برای مواجهه با دیپفیک هستند. تجربه این کشورها بیانگر آن است که صرف جرم‌انگاری، نمی‌تواند پاسخ کافی به این پدیده باشد.

در حوزه جرم‌شناسی، دیپفیک موجب تحول در الگوهای بزه‌کاری، بزه‌دیدگی و سیاست جنایی شده است. کاهش هزینه ارتکاب جرم، افزایش ناشناس ماندن مرتکبان، امکان انتشار گسترده محتوا و دشواری شناسایی ادله جعلی، موجب شده بسیاری از نظریه‌های کلاسیک جرم‌شناسی کارایی محدودتری پیدا کنند. برنامه‌های تشخیص قادر به شناسایی همه

دیپ‌فیک‌ها نیستند و تضمین اینکه تمام محتوای آنلاین از طریق یک سیستم ایمن و قابل اعتماد احراز هویت شود، دشوار یا غیرممکن خواهد بود. در واقع سرنوشت دیپ‌فیک‌ها منوط به توسعه فناوری‌های تشخیص دیپ‌فیک و پابندی اخلاقی پلتفرم‌های توزیع محتوا برای شناسایی محتویات مخرب خواهد بود.

بنابراین مقابله مؤثر با دیپ‌فیک مستلزم اتخاذ رویکردی چندبعدی است که علاوه بر اصلاح قوانین کیفری، شامل توسعه فناوری‌های تشخیص، ارتقای سواد رسانه‌ای، ایجاد سازوکارهای تخصصی احراز اصالت ادله دیجیتال، تقویت همکاری‌های بین‌المللی و تنظیم‌گری مسئولانه پلتفرم‌های دیجیتال باشد.

محدودیت‌های پژوهش

مهم‌ترین محدودیت این پژوهش، نوظهور بودن فناوری دیپ‌فیک و فقدان رویه قضایی گسترده در ایران است. همچنین بسیاری از مقررات و پرونده‌های مرتبط در سایر کشورها هنوز در حال تحول بوده و دسترسی به برخی داده‌های قضایی و آماری محدود است.

پیشنهادهای پژوهشی آتی

۱. مطالعه تطبیقی مسئولیت پلتفرم‌های دیجیتال در قبال دیپ‌فیک؛
۲. بررسی فقهی و حقوق بشری دیپ‌فیک؛
۳. تحلیل ادله الکترونیکی و استانداردهای اثبات در جرایم مبتنی بر هوش مصنوعی؛
۴. مطالعه جرم‌شناختی بزه‌دیدگان دیپ‌فیک در ایران؛
۵. بررسی نسبت میان آزادی بیان و جرم‌انگاری دیپ‌فیک سیاسی؛
۶. تحلیل مسئولیت مدنی ناشی از خسارات دیپ‌فیک؛
۷. بررسی امکان توسعه سازوکارهای خودتنظیم‌گری پلتفرم‌های دیجیتال.

منابع

- اکبری، عباسعلی، آقاپور، علی و آقاپور، کمال (۱۴۰۱)، تحلیل پدیده مجرمانه دیپ‌فیک‌ها (جعل‌های رایانه‌ای پیچیده) با نگاهی به سیاست کیفری ایران و چالش‌های حقوق بشری، فصلنامه علمی کارآگاه، ۱۶(۵۹)، ۱۴۹-۱۶۹.
- شیری، عباس (۱۴۰۱)، دیپ‌فیک یا همانندسازی صوتی یا تصویری غیرواقعی در حقوق کیفری، فصلنامه تحقیقات حقوقی، ۲۵، ۱۴۳-۱۶۸.

References

- Akbari, A.A., Aqapoor, A. & Aqapoor, K. (2022). Analysis of the criminal phenomenon of deep fakes (complex computer forgeries) Looking at Iran's criminal policy and human rights challenges. KARAGAH (THE DETECTIVE), 16(59), 149-169. SID. <https://sid.ir/paper/963293/en> (in Persian)

- Cassia, M., Guarnera, L., Casu, M., Zangara, I. & Battiato, S. (2025, October). Deepfake forensic analysis: Source dataset attribution and legal implications of synthetic media manipulation. In *2025 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRINE)* (pp. 687-692). IEEE. DOI: 10.48550/arXiv.2505.11110
- Citron, D. K. (2019). Sexual Privacy. *Yale Law Journal*, 128(7), 1870–1960.
- De Seta, G. (2021). Huanlian or changing faces: Deepfakes on Chinese digital media platforms. *Convergence: The International Journal of Research into New Media Technologies*, 27(4), 935-953.
- Dehghani, A. & Saberi, H. (2025). Generating and Detecting Fake Image and Audio: Tools & Technologies. *IEEE Communications Surveys*, 27(1), 110–128.
- Delfino, R. A. (2023). Deepfakes on trial: A call to expand the trial judge's gatekeeping role to protect legal proceedings from technological fakery. *Hastings Law Journal*, 74(2), 291–344.
- Harris, D. (2018–2019). Deepfakes: False pornography is here and the law cannot protect you. *Duke Law & Technology Review*, 17, 99–127.
- Łabuz, M. (2023). Regulating deepfakes in the artificial intelligence act. *ACIG*, 2(1), 1–42.
- Langa, J. (2021). Deepfakes, real consequences: crafting legislation to combat threats posed by deepfakes. *Boston University Law Review*, 101, 761–810.
- Meskys, E., Kalpokienė, J., Jurcys, P. & Liaudanskas, A. (2020). Regulating deep fakes: Legal and ethical considerations. *Journal of Intellectual Property Law & Practice*, 15(1), 24–31. Available at SSRN: <https://ssrn.com/abstract=3497144>
- Ramluckan, T. (March 2024). Deepfakes: The legal implications. *Presented at the International Conference on Cyber Warfare and Security, Reading*, pp. 282–288.
- Sandoval, M. P., De Almeida Vau, M., Solaas, J., & Rodrigues, L. (2024). Threat of deepfakes to the criminal justice system: A systematic review. *Crime Science*, 13 (1), 41.
- Shiri, A. (2023). Deepfake or unreal audio or video simulation in the criminal law. *Legal Research Quarterly*, 25 (Special Issue of Law & Technology), 143–168. <https://doi.org/10.52547/jlr.2023.230618.2451> (in Persian)
- Sutherland, E. H. (1947). Differential association theory. *Criminology theory: Selected classic readings*, 225-227.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology innovation management review*, 9(11).
- Yamaoka-Enkerlin, A. (2024). Disrupting disinformation: Deepfakes and the Law. *Legislation and Public Policy*, 22, 725–749.