پژوهش‌نامه حقوق اسلامی

مقاله پژوهشی | دریافت: ۱۴۰۴/۰۴/۱۷ | پذیرش: ۱۴۰۴/۰۸/۰۲
DOI: 10.30497/law.2026.248521.3808

تحلیل تقنینی جرم‌انگاری «هرزه‌نگاری مبتنی بر جعل عمیق» در ایران و اتحادیه اروپا

زهرا سادات میرهاشمی • دانشیار، گروه فقه و مبانی حقوق اسلامی، دانشکده الهیات، دانشگاه الزهرا (س)، تهران، ایران.
(نویسنده مسئول) z.mirhashemi@alzahra.ac.ir
فاطمه نجیب‌زاده • دانشجوی دکتری فقه و مبانی حقوق اسلامی، دانشکده الهیات و معارف اسلامی، دانشگاه فردوسی مشهد، مشهد، ایران. fatemeh.najibzadeh@mail.um.ac.ir

چکیده

پورنوگرافی دیپ‌فیک (Deepfake Pornography) یا بازنمایی جنسی فاقد رضایت با دست‌کاری واقع‌نمای تصویر و صوت، چالشی هم‌زمان برای کرامت انسانی و حکمرانی دیجیتال پدید می‌آورد؛ سرعت تکثیر، بازنشر نسخه‌های اساساً معادل، و زوال ادله از یک‌سو، و دشواری انتساب و جبران مؤثر از سوی دیگر. مسأله محوری پژوهش، سنجش توان چارچوب‌های موجود در ایران (مواد ۱۴ تا ۱۶ قانون جرائم رایانه‌ای) و اتحادیه اروپا (هم‌تنظیمی قانون هوش مصنوعی و قانون خدمات دیجیتال) برای تبدیل «هنجار مکتوب» به «حمایت مؤثر و بازدارنده» است. پژوهش با روش توصیفی تحلیلی و رویکرد تطبیقی، مبتنی بر منابع کتابخانه‌ای و تحلیل متون تقنینی و رویه‌های قضایی منتخب، و بر پایه سه شاخص مقایسه‌ای (چارچوب هنجاری، ضمانت اجرای عملی، و چالش‌های قضایی) سامان گرفته است. برآیند تحلیل نشان می‌دهد که منطق «طبقه‌بندی ریسک» و «شفافیت صرف» در اتحادیه اروپا، بدون «اثبات منشأ» و استاندارد گزارش کارشناسی چندنشانه‌ای، در مهار چرخه بازنشر و انتساب، کارایی محدودی دارد، و در ایران، اتکای تفسیری به مفاهیم عرفی، نبود تعریف مضیّق از جعل عمیق، فقدان خط‌کش‌های اجرایی برای سکوها و نبود پروتکل واحد ادله، اثر بازدارندگی را تضعیف می‌کند. مسیر بهینه، مستلزم تعریف قانونی آزمون‌پذیرِ دیپ‌فیک، استاندارد ملی گزارش کارشناسی (روش، آستانه تصمیم، حدود خطا)، هم‌ترازی کیفر با منطق اقتصادی شبکه (از جمله استرداد منافع نامشروع و جزای درصدی)، و الزام سکوها به حذف زمان‌مند، تحویل فراداده و جست‌وجوی «نسخه‌های اساساً معادل» است؛ ترتیبی که با اقتباس بومی‌شده از عناصر اروپایی و هماهنگ با ساختار حقوقی ایران، پاسخ حقوقی را از سطح اعلامی به سازوکار حمایتی و بازدارنده ارتقا می‌دهد.

واژگان کلیدی: هرزه‌نگاری، جعل عمیق، هوش مصنوعی، حریم خصوصی، کرامت انسانی.

مقدمه

«پورنوگرافی دیپ‌فیک» مرزهای عادت‌کرده به تمایز میان بازنمایی و واقعیت را درمی‌نوردد؛ محتوایی که با تولید یا دست‌کاری تصویر و صوت اشخاص، نسبت‌های خصوصی کاذب می‌سازد و به‌سرعت در سکوهای برخط تکثیر می‌شود. دامنهٔ زیان در این حوزه صرفاً حیثیتی نیست؛ پیامدهای روانی، اقتصادی و اجتماعی ماندگار برای بزه‌دیدگان و نیز هزینه‌های حکمرانی برای نظام دادرسی و تنظیم‌گر را در پی دارد. ازاین‌رو، «ارزیابی تقنینی» این پدیده، نه صرف مرور متون قانونی، بلکه سنجش پیوند هنجار با ادله و اجرا ضرورتی مضاعف یافته است. براین مبنا، پژوهش حاضر ابتدا مفاهیم کلیدی و مرزبندی‌های لازم را روشن می‌کند و سپس با تمرکز بر پیوند «هنجار، ادله و اجرا»، ظرفیت‌های حقوق ایران و امکان اقتباس گزینشی از چارچوب اتحادیه اروپا را در مواجهه با پورنوگرافی دیپ‌فیک بررسی می‌نماید.

–ادبیات پژوهش

–پورنوگرافی[1](هرزه‌نگاری)

اصطلاح «پورنوگرافی» ریشه‌ای تاریخی دارد که از زبان یونانی باستان به زبان‌های معاصر منتقل شده است. بخش نخست آن، «پورنو»، در زبان یونانی کلاسیک به معنای «فاحشه» یا «برده جنسی» بوده و در متون ادبی یونان باستان برای توصیف فعالیت‌های جنسی خارج از چهارچوب اخلاق غالب جامعه به کار می‌رفته است .(Löfgren-Mårtenson, 2008: 306) بخش دوم، «گرافی»، از فعل graphein به معنای «نوشتن» یا «ثبت‌کردن» مشتق شده و در زبان امروزی نشانگر «نمایش» یا «بازنمایی» است .(Hayes & Paul, 2007: 517) در گذار معنا، پورنوگرافی به نمایش صریح کنش‌های جنسی گفته می‌شود؛ صحنه‌هایی که هدف اصلی آن‌ها تحریک مخاطب است نه انتقال پیام هنری یا آموزشی. این هدف‌گذاری برانگیختگی جنسی، نقطهٔ تمایز این محتوا از آثار ادبی و علمی است و در بسیاری از نظام‌های حقوقی بر مبنای معیارهای «تمایل به شهوت» و «فاقد ارزش روشنگرانه» جرم‌انگاری یا محدود می‌شود (رضایی، ۱۳۹۶: ۱۰۵؛ موسوی، ۱۳۹۴:۵۰).

1. Pornography.

در ادبیات تقنینی ایران، واژه «پورنوگرافی» به طور صریح به کار نرفته است، اما قوانین مرتبط با آن در چارچوب‌هایی همچون «منافی عفت عمومی»، «مبتذل»، «مستهجن» و «جریحه‌دار کردن عفت و اخلاق عمومی» مطرح شده‌اند. به‌ویژه، در فصل هجدهم قانون مجازات اسلامی مصوب ۱۳۹۲ (ماده ۶۳۹)، نشر تصاویر یا مطالب مستهجن به‌عنوان «عمل مخالف عفت عمومی» شناخته شده و مرتکبان آن به حبس و شلاق محکوم می‌شوند. همچنین، ماده ۲۵ قانون جرائم رایانه‌ای مصوب ۱۳۸۸، هرگونه «محتوای مبتذل و مخل عفت عمومی» را جرم اعلام کرده و برای مرتکبان مجازات‌هایی از جمله حبس و جزای نقدی در نظر گرفته است (قانون مجازات اسلامی، ۱۳۹۲: ماده ۶۳۹؛ قانون جرائم رایانه‌ای، ۱۳۸۸: ماده ۲۵).

این قوانین در تلاش‌اند تا با برخورد قاطع با نشر محتوای مستهجن، حفاظت از اخلاق عمومی را تضمین کنند. در مقایسه، دیوان عدالت اتحادیه اروپا رویکردی متوازن در مواجهه با پدیده‌هایی چون پورنوگرافی دیپ‌فیک اتخاذ کرده است که به‌ویژه بر حقوق بشر و آزادی‌های فردی تأکید دارد. درحالی‌که قوانین ایران بر مقابله با محتوای مخل عفت عمومی تمرکز دارند، در اتحادیه اروپا، تلاش برای حفظ آزادی‌های فردی و حقوق بشر در کنار مقابله با چنین محتوای دیجیتال آسیب‌زا، موردتوجه قرار گرفته است. هر دو نظام حقوقی در صدد توازن میان حقوق فردی و منافع عمومی هستند.

-دیپ‌فیک (جعل عمیق)[1]

1.Deepfake.

دیپ‌فیک یک فناوری نوظهور است که با استفاده از روش‌های پیشرفته یادگیری و شبکه‌های عصبی مصنوعی، محتوای صوتی و تصویری غیرواقعی تولید می‌کند که بسیار مشابه به نمونه‌های واقعی است. این فناوری عمدتاً از الگوریتم‌های شبکه‌های مولد رقابتی و شبکه‌های عصبی پیچیده دیگر استفاده می‌کند که قادر به شبیه‌سازی دقیق ویژگی‌های انسانی، از جمله صورت، صدا، و حتی زبان بدن، به طور کاملاً واقعی هستند .(Korshunov & Marcel, 2023: 112) در فرایند تولید دیپ‌فیک، الگوریتم‌ها ابتدا از حجم وسیعی از داده‌های آموزشی، شامل تصاویر، ویدئوها و صداهای افراد مختلف، استفاده کرده و این داده‌ها را برای تولید محتوای جدید به کار می‌برند. این تکنولوژی به‌ویژه در تولید ویدئوهای جعلی از چهره‌های معروف، افراد عمومی، یا حتی چهره‌های ناشناس کاربرد دارد و می‌تواند تصاویر و ویدئوهایی بسازد که به‌راحتی با نمونه‌های واقعی اشتباه گرفته شوند (Garrido et al., 2024: 53).

استفاده از دیپ‌فیک در زمینه‌های مختلفی از جمله تولید اخبار جعلی، جعل هویت دیجیتال، و حتی پورنوگرافی انتقامی افزایش‌یافته است. این تهدیدات می‌توانند تأثیرات منفی گسترده‌ای بر حریم خصوصی افراد، امنیت سایبری و نظم اجتماعی داشته باشند .(Masood et al., 2021: 67) در نتیجه، دیپ‌فیک به یک چالش جدی برای سیستم‌های حقوقی و تقنینی در سراسر جهان تبدیل شده است. باتوجه‌به قدرت بالای این فناوری در جعل واقعیت، دولت‌ها و نهادهای بین‌المللی به دنبال راه‌حل‌هایی برای مقابله با تهدیدات آن از جمله توسعه الگوریتم‌های شناسایی، پیاده‌سازی قوانین سخت‌گیرانه‌تری در راستای حفاظت از حریم خصوصی و امنیت اطلاعات هستند (Franqueira et al., 2024: 92). قانون هوش مصنوعی اتحادیه اروپا ، «دیپ‌فیک» چنین تعریف شده در ماده ۳(۶۰) است: «محتوای تصویری، صوتی یا ویدئویی که با هوش مصنوعی تولید یا دست‌کاری شده و شباهتی گمراه‌کننده با افراد، اشیاء، مکان‌ها، نهادها یا رویدادهای واقعی پیدا می‌کند، به‌گونه‌ای که بیننده آن را اصیل یا موثق تلقی می‌کند» (قانون هوش مصنوعی اتحادیه اروپا، ماده ۳، بند ۶۰). از نظر ساختاری و محتوایی، این تعریف تقریباً با بند (۱) جزء (k) ماده ۳۵ «قانون خدمات دیجیتال[1]» همپوشانی دارد، هرچند در آن قانون به‌صراحت از اصطلاح «دیپ‌فیک» استفاده نشده است (قانون خدمات دیجیتال اتحادیه اروپا، ماده ۳۵، بند ۱، جزء(k)).

1.Digital Service Act.

بااین‌همه در تبصره ۱۳۴ همان قانون هوش مصنوعی اتحادیه اروپا، کلمه «شباهت محسوس»، به متن تعریف افزوده شده تا بر «میزان شباهت قابل‌رؤیت» تأکید کند: «سامانه هوش مصنوعی باید محتوایی را (به طور محسوس)شبیه واقعیت بسازد یا دست‌کاری کند تا بیننده آن را به‌اشتباه به‌عنوان محتوای اصیل بپذیرد» (قانون هوش مصنوعی اتحادیه اروپا، تبصره ۱۳۴).

ورود واژه «شباهت محسوس» به قانون هوش مصنوعی اتحادیه اروپا معانی متفاوت و گاه متضادی را به همراه دارد؛ از یک‌سو می‌تواند معیار سخت‌گیرانه‌تری برای شناسایی دیپ‌فیک تعیین کند و از سوی دیگر ممکن است تنها توصیفی توضیحی باشد که کمکی به تعریف اصلی دیپ‌فیک نمی‌کند. این تفاوت واژگانی میان ماده و تبصره، مرز روشن‌بین محتوای قابل‌قبول و ممنوعه را مبهم می‌سازد و نشان می‌دهد که قانون‌گذاران اروپایی درباره آستانه شباهت «دیپ‌فیک» نیازمند دقت و یکپارچگی بیشتری هستند(Labuz, 2024:787) .

علاوه بر این، فقدان معیارهای فنی مشخص برای اندازه‌گیری «شباهت محسوس» موجب شده تا رویه‌های قضایی و ابزارهای تشخیص در کشورهای عضو به‌شدت متنوع باشد؛ از تحلیل‌های بیومتریک تا الگوریتم‌های تشخیص چهره، هر یک با استانداردهای متفاوتی تعریف «شباهت محسوس» را دست‌کاری می‌کنند و این خود چالش جدیدی درباره شفافیت و اعتبارسنجش خلق می‌کند (Labuz, 2024:787).

باوجود هماهنگی نسبی میان قانون هوش مصنوعی اتحادیه اروپا و قانون خدمات دیجیتال، نبود یک تعریف مشترک و قابل‌اندازه‌گیری از «دیپ‌فیک» می‌تواند در عمل راه برای سوءاستفاده تولیدکنندگان محتوا باز کند و قضاوت نهایی درباره ماهیت ممنوعیت را به تشخیص موردی قاضیان و نهادهای فنی محول نماید؛ وضعیتی که نه‌تنها پیچیدگی‌های حقوقی را افزایش می‌دهد، بلکه اقتضای تصویب سریع‌تر و یکپارچه‌تر اصلاحات در چارچوب قوانین دیجیتال را بیش‌ازپیش آشکار می‌سازد.(Labuz, 2024:788)

-پورنوگرافی دیپ‌فیک[1]

پورنوگرافی دیپ‌فیک نوعی محتوای جنسی دیجیتال است که با استفاده از الگوریتم‌های پیچیده

یادگیری ماشین، به‌ویژه شبکه‌های مولد رقابتی و مدل‌های انتشار تولید می‌شود. در این نوع محتوا، ویژگی‌های چهره، بدن، زبان بدن و حتی صداهای افراد دست‌کاری یا شبیه‌سازی می‌شوند تا آنها را در صحنه‌های جنسی نشان دهند، بدون اینکه فرد اصلی در آن موقعیت‌ها حضور داشته باشد. هدف این فرایند تولید محتوای واقع‌گرایانه‌ای است که حتی توسط بینندگان یا ابزارهای شناسایی دیجیتال هم نتوان به‌راحتی جعلی بودن آن را تشخیص داد.(Franqueira et al., 2024: 22) در این پژوهش، «دیپ‌فیک» ناظر به محتوایی است که از حیث فنی و ادراکی، «قابلیتِ واقع‌نما شدن» و ایجاد تصور انتساب به فرد واقعی را برای ناظر متعارف دارد؛ بنابراین، «آگاهیِ قبلیِ بیننده از ساختگی بودن محتوا» (مثلاً به‌واسطه برچسب‌گذاری یا اطلاع‌رسانی) لزوماً وصف دیپ‌فیک بودن را منتفی نمی‌کند، هرچند می‌تواند در تحلیل شدت فریب و آثار مسئولیت مؤثر باشد. این تمایز با منطق الزام به افشا و برچسب‌گذاری نیز هم‌خوان است؛ به این معنا که برچسب‌گذاری ناظر به مدیریت ریسک و کاهش فریب است، نه نفیِ ماهیت واقع‌نمای محتوای دست‌کاری‌شده.

این محتوا از منظر حقوقی و تحلیلی، اغلب با چهار ویژگی اصلی شناخته می‌شود:

۱-عدم رضایت یا نبود اراده فردی :محتوای دیپ‌فیک به طور عمدی دست‌کاری یا ترکیب می‌شود تا فرد بدون رضایت وی در موقعیت‌های جنسی قرار گیرد. در این حالت، نه‌تنها تصویر یا ویدئو، بلکه صدا و بدن فرد هم ممکن است به طور مصنوعی تغییر کند تا ظاهراً به صحنه‌ای پورنوگرافی منتقل شود.(Garrido et al., 2025: 81)

۲-تنوع دست‌کاری‌ها و شدت آن‌ها :این دست‌کاری‌ها می‌توانند شامل تغییرات جزئی در ویژگی‌های صورت یا بدن، جایگزینی کامل چهره یا حتی ترکیب دقیق تصاویر واقعی و مصنوعی باشند. این تنوع در شدت دست‌کاری، قابلیت فریب دهی بیشتر و امکان تولید محتواهایی که واقعی به نظر می‌رسند را افزایش می‌دهد.(Masood et al., 2021: 110)

۳-قابلیت واقع‌نمایی و ایجاد تصور اصالت/انتساب برای ناظر متعارف: در پورنوگرافی دیپ‌فیک، هدف آن است که محتوا به نحوی تولید یا دست‌کاری شود که «می‌تواند» برای بیننده واقعی یا منتسب به فرد به نظر برسد؛ حتی اگر در برخی موارد، به‌واسطه برچسب‌گذاری یا اطلاع‌رسانی، مخاطب از ابتدا از ساختگی بودن آن آگاه باشد. این ویژگی به‌ویژه در فضای آنلاین، جایی که چنین محتوای دیجیتال به‌سرعت منتشر می‌شود، موجب پیچیدگی بیشتر شناسایی و مقابله با آن می‌شود .(McGlynn & Toparlak, 2025: 35)بدیهی است اگر محتوا اساساً واقع‌نما نباشد و برای ناظر

1.Deepfake pornography.

متعارف از ابتدا آشکارا مصنوعی و غیرقابل‌انتساب جلوه کند، بحث از «دیپ‌فیکِ واقع‌نما» در معنای مضیّق آن محل تردید خواهد بود.

۴-ایجاد فریب مؤثر :هدف اصلی در تولید پورنوگرافی دیپ‌فیک، ایجاد محتوای فریب‌دهنده‌ای است که تشخیص جعلی بودن آن برای بینندگان دشوار یا غیرممکن باشد. این ویژگی به‌ویژه در فضای آنلاین، جایی که چنین محتوای دیجیتال به‌سرعت منتشر می‌شود، موجب پیچیدگی بیشتر شناسایی و مقابله با آن می‌شود.(McGlynn & Toparlak, 2025: 37)

این نوع محتوا به‌ویژه در بازارهای غیررسمی آنلاین، نظیر «مستر دیپ‌فیکز»، به طور سیستماتیک تولید، توزیع و فروخته می‌شود. این بازارها دسترسی به محتوای پورنوگرافی دیپ‌فیک را آسان کرده و ممکن است در راستای سوءاستفاده از افراد خاص، به‌ویژه زنان و افراد مشهور، به کار روند (Han et al., 2024: 15).

در این راستا، پورنوگرافی دیپ‌فیک نه‌تنها یک چالش فنی به شمار می‌آید، بلکه یک مسئله حقوقی و اخلاقی پیچیده نیز ایجاد می‌کند. از یک سو، لازم است که قوانین جدیدی در حوزه حریم خصوصی و حق مالکیت تصویر تدوین شوند و از سوی دیگر، باید راهکارهایی برای مقابله با آسیب‌های اجتماعی و روانی ناشی از چنین محتوای جعلی پیدا شود. این تهدیدات نیازمند تعادلی بین آزادی بیان و حفاظت از کرامت انسانی است، به‌ویژه هنگامی که محتوای دیپ‌فیک به ابزارهایی برای تهدید، اخاذی یا سوءاستفاده از شهرت تبدیل می‌شود.(Umbach et al., 2024: 95)

-پیشینه پژوهش

پیشینهٔ داخلی، عمدتاً خلأهای جرم‌انگاری و دشواری‌های دادرسی را توصیف کرده است. در یک مسیر، پژوهش اکبری و همکاران (۱۴۰۱) در مقاله‌ای با عنوان «تحلیل پدیده مجرمانه دیپ‌فیک‌ها (جعل‌های رایانه‌ای) پیچیده با نگاهی به سیاست کیفری ایران و چالش‌های حقوق بشری»، با تمرکز بر سیاست کیفری ایران و تبعات حقوق بشری، تصویری کلان از ظرفیت‌ها و کاستی‌های پاسخ کیفری به دیپ‌فیک ارائه می‌کند و بر نیاز به نوسازی قواعد و سازوکارهای حمایتی تأکید دارد. در مسیری نزدیک اما با زاویهٔ متفاوت، پژوهش شیری (۱۴۰۱) با عنوان «دیپ‌فیک یا همانندسازی صوتی یا تصویری غیرواقعی در حقوق کیفری»، جنبه‌های حقوق کیفری همانندسازی صوتی و تصویری غیرواقعی را واکاوی کرده و بر ضرورت بازنگری در عناصر قانونی و شیوه‌های مقابله تأکید می‌گذارد. باوجود ارزش این دو خط پژوهشی، دو حلقه همچنان کم‌رنگ مانده است: نخست، «تطبیق نظام‌مند» با منظومهٔ نوپدید اتحادیه اروپا که ترکیب تنظیم مبتنی بر ریسک، تعهدات شفافیت و مسئولیت سکوها را به‌صورت یک چارچوب عملیاتی پیاده می‌کند؛ دوم، «معماری ادله» که توضیح

دهد خروجی‌های فنی تشخیص چگونه به دلیل قابل‌استناد تبدیل می‌شود و نسبت خطای ابزار با بارِ اثبات چگونه برای دادرس تبیین می‌گردد. پژوهش حاضر باتکیه‌بر همین دو حلقه مغفول، چارچوبی سه سطحی هنجار، ادله و اجرا را به کار می‌گیرد و تفاوت خود را با پیشینه در همین «یکپارچه‌سازی تقنین، فناوری و دادرسی» و سنجش‌پذیر کردن پیشنهادها نشان می‌دهد.

–ضرورت پژوهش

منطق تطبیق نظام حقوقی ایران و اتحادیه اروپا از دو جهت توجیه‌پذیر است: از حیث هنجاری، هر دو نظام حقوقی در پی صیانت از کرامت و حریم خصوصی در بستر دیجیتال‌اند، اما ابزارها و زبان قانون متفاوت است؛ از حیث اجرایی، اتحادیه اروپا تلفیقی از تنظیم مبتنی بر ریسک، تعهدات شفافیت محتوا و سازوکارهای پلتفرم محور را به کار بسته، درحالی‌که ایران بیشتر بر قواعد کیفری عام (از جمله مواد ۱۴ تا ۱۶ قانون جرایم رایانه‌ای) و ظرفیت‌های قضایی–انتظامی تکیه دارد. سنجش امکان انتقال گزینشی برخی عناصر اروپایی به بستر ایران با درنظرگرفتن اقتضائات فرهنگی، فنی و ساختار دادرسی می‌تواند الگوی واقع‌بینانه‌ای برای بازطراحی پاسخ حقوقی فراهم آورد.

–روش پژوهش

روش پژوهش «توصیفی–تحلیلی با رویکرد تطبیقی» است؛ تحلیل اسنادی قوانین و رهنمودها، رویه‌های قضایی شاخص و گزارش‌های رسمی محور کار قرار می‌گیرد. شاخص‌های تطبیق‌پذیری و ابعاد مقایسه به‌صورت شفاف تعریف می‌شوند: در سطح هنجاری، تعریف و مرزبندی «محتوای مصنوعی» و جایگاه آن در جرم‌انگاری، و نیز تکالیف سکوها؛ در سطح ادله، معیارهای عینی انتساب (شباهت محسوس، قابلیت انتساب به شخص قابل‌شناسایی، فقدان رضایت معتبر) و استاندارد گزارش کارشناسی (روش، آستانه تصمیم، حدود خطا)؛ و در سطح اجرا، تناسب ضمانت اجرا با «شدت آسیب شبکه‌ای»، زمان‌بندیِ حفظ و حذف و همکاری برون‌مرزی. ملاک انتخاب اتحادیه اروپا، وجود بسته‌ای نسبتاً کامل از ابزارهای تقنینی نو و تراکم تجربه‌های رویه‌ای در مواجهه با محتوای دست‌کاری‌شده است؛ و ملاک انتخاب ایران، ضرورت سنجش ظرفیت‌های موجود و امکان هم‌ترازی با واقعیت فنی و رسانه‌ای. محدودیت‌ها نیز از ابتدا روشن می‌شود: پویایی مقررات اروپایی و زمان‌بندی اجرای کامل آن‌ها، و کمبود داده‌های تفکیک‌شده ملی دربارهٔ پرونده‌های دیپ‌فیک جنسی.

–سؤالات پژوهش

پرسش اصلی پژوهش آن است که چه الگوی پیوندی میان «هنجار، ادله و اجرا» می‌تواند پاسخ مؤثرتری به پدیده پورنوگرافی دیپ‌فیک در ایران بدهد و کدام مؤلفه‌های چارچوب اتحادیه اروپا قابلیت اقتباسِ گزینشی و کارآمد در بستر حقوقی ایران را دارند. بر همین مبنا، چند پرسشِ مکمل

مطرح می‌شود: نخست، مرزبندیِ قابل اتکا میان «محتوای جنسی دست‌کاری‌شده» و سایر اشکال هتک حیثیت بر چه معیارهای عینی باید استوار باشد. دوم، قواعد موجود ایران تا چه حد با شدت و گستره زیان‌های شبکه‌ای متناسب است و چه سازوکارهایی می‌تواند بازدارندگی را تقویت کند. سوم، استاندارد ادله و قالب گزارش کارشناسی چندنشانه‌ای چگونه باید تنظیم شود تا مرجع رسیدگی بتواند روش تشخیص، حدود خطا و نسبت آن با بار اثبات را روشن و قابل ارزیابی ببیند. چهارم، تکالیف سکوها در حذف به‌موقع، نگهداشت و تحویل فراداده و نیز شناسایی و حذف «نسخه‌های اساساً معادل» چگونه باید دقیق و سنجش‌پذیر تعیین شود. پنجم، همکاری برون‌مرزی و حفظ داده‌های زودزوال با چه ترتیباتی می‌تواند قابل اتکا و زمان‌مند شود. بر این اساس، ترتیب مباحث مقاله مسیر استدلال را از «تعریف و طبقه‌بندی» به «ضمانت اجرا و ادله» و سپس به «آسیب‌شناسی اجرا و پیشنهادهای عملی» هدایت می‌کند تا در پایان، امکان‌سنجیِ اقتباس گزینشی و بازطراحی حقوقی در بستر ایران به‌صورت یکپارچه روشن شود.

–ساختار پژوهش

مقاله ابتدا رویکرد تقنینی ایران در مواجهه با هرزه‌نگاری جعل عمیق را (در پیوند با ضمانت اجرا و ادله) بررسی می‌کند و سپس چارچوب اتحادیه اروپا را با تمرکز بر قانون هوش مصنوعی و قانون خدمات دیجیتال تحلیل می‌نماید. در گام بعد، چالش‌های اجرایی و فنیِ اجرای مقررات، پیامدهای اقتصادی–اجتماعی پدیده و نقش شرکت‌های فناوری در تولید یا مقابله با دیپ‌فیک‌ها واکاوی می‌شود. در پایان، جمع‌بندی و پیشنهادهای عملیِ قابل‌اجرا برای بهبود پاسخ حقوقی ارائه می‌گردد.

۱. بنیادهای نظری جرم‌انگاری محتوای دیجیتال جعلی: پیوند دیپ‌فیک با اصول حقوق بشر

در حوزه حقوق کیفری و حقوق بشر، جرم‌انگاری محتوای دیجیتال جعلی نظیر دیپ‌فیک به سؤالات مهمی پیرامون آزادی بیان، کرامت انسانی و حق حریم خصوصی برمی‌خورد. برای درک بهتر مبانی نظری جرم‌انگاری دیپ‌فیک، باید به دو جنبه اصلی توجه کرد: اول، پیامدهای حقوقی و اجتماعی دیپ‌فیک بر اصول حقوق بشر و دوم، منطق نظری پشت جرم‌انگاری این نوع محتوا.

۱.۱. پیوند دیپ‌فیک با اصول حقوق بشر

طبق اعلامیه جهانی حقوق بشر، حق بر حریم خصوصی در ماده ۱۲ به وضوح بیان شده است که «هیچ‌کس نباید مورد مداخله خودسرانه در امور خصوصی، خانوادگی، مسکن یا مکاتبات خود قرار گیرد» (اعلامیه جهانی حقوق بشر، ۱۹۴۸). با این حال، فناوری‌های جدید مانند دیپ‌فیک تهدیدی جدی برای این حق محسوب می‌شوند. با تولید محتوای جعلی از افراد در صحنه‌های جنسی بدون

رضایت آن‌ها، کرامت انسانی افراد به شدت آسیب می‌بیند. تولید و انتشار این محتوا به شکلی گسترده می‌تواند به تهدیدات روان‌شناختی و اجتماعی منجر شود که نه تنها به حقوق فردی آسیب می‌زند، بلکه تضعیف‌کننده اصول اجتماعی و امنیت افراد نیز است (رزمان، ۱۳۹۵: ۱۰۲).

در این راستا، از منظر حقوق بشر، می‌توان به حق آزادی از تعرضات غیرقانونی اشاره کرد که در کنوانسیون‌های بین‌المللی نظیر کنوانسیون اروپایی حقوق بشر (۱۹۵۰) به تصویب رسیده است. این اسناد، تهدیدات دیجیتال ناشی از دیپ‌فیک را به عنوان نقض اساسی حریم خصوصی و کرامت فردی به رسمیت می‌شناسند.(Garrido et al., 2025: 12) در دیپ‌فیکِ پورنوگرافیک، مسئله فقط «افشا» نیست؛ یک نسبتِ جنسیِ جعلی به شخص تحمیل می‌شود. به همین دلیل، تمرکزِ حقوقی صرف بر «مستهجن بودن» یا «انتشار» کافی نیست و دو مؤلفه نقش تعیین‌کننده دارند: اینکه محتوا به یک شخص قابل‌شناسایی قابل انتساب باشد و اینکه رضایت معتبر در تولید یا انتساب وجود نداشته باشد. هر پاسخ حقوقی که این دو مؤلفه را روشن نکند، یا در اجرا به ابهام می‌افتد یا در مرزبندی با موارد دامنه را بیش از حد گسترش می‌دهند.

۱.۲. تئوری‌های آزادی بیان و دیپ‌فیک

آزادی بیان یکی از اصول بنیادین در حقوق بشر است که در ماده ۱۹ اعلامیه جهانی حقوق بشر و بسیاری از اسناد ملی و بین‌المللی گنجانده شده است. با این حال، این آزادی مطلق نیست و در مواردی که به حریم خصوصی و کرامت انسان آسیب وارد می‌کند، می‌تواند محدود شود. در این زمینه، نظریات جان استوارت میل در کتاب «درباره آزادی» (۱۸۵۹) مهم است. میل بر این باور است که هر عمل یا بیان اگر منجر به آسیب رساندن به دیگر شود، باید محدود شود و دیپ‌فیک نمونه‌ای از چنین آسیب‌ها است که می‌تواند باعث خسارات جبران‌ناپذیر به فرد گردد. در نتیجه، محدودیت‌های قانونی برای جلوگیری از تولید و انتشار دیپ‌فیک‌ها ضروری است (Franqueira et al., 2024: 23). نقطه تعیین‌کننده در محدودسازی آزادی بیان، «قاعده‌گذاری دقیق» است: آنچه باید هدف قرار گیرد، بیانِ انتقادی یا طنز نیست، بلکه انتسابِ جنسیِ جعلیِ فاقد رضایت به شخص قابل‌شناسایی است. از همین رو، محدودیتِ موجه زمانی شکل می‌گیرد که (الف) زیان قابل پیش‌بینی و شدید باشد، (ب) نسبت دادن محتوا به فرد قابل اتکا باشد، و (پ) امکان توقف و حذف سریع وجود داشته باشد تا حمایت صرفاً در حد حکم روی کاغذ نماند.

۱.۳. فلسفه حقوق کیفری و جرم‌انگاری محتوای دیجیتال جعلی

در نظریه حقوق کیفری[۱]، جرم‌انگاری محتواهایی نظیر دیپ‌فیک به دلیل تهدید به نظم عمومی و آسیب به آسیب‌های روانی و اجتماعی افراد توجیه می‌شود. از این منظر، نقطه عزیمت این است که مداخله کیفری باید متناسب با «زیان قابل اعتنا» باشد؛ و در دیپ‌فیک‌های جنسیِ فاقد رضایت، زیان صرفاً مادی نیست، بلکه «زیان حیثیتی و کرامتی» و نیز نقض خودآیینی و حریم خصوصی جنسی قربانی را نیز در بر می‌گیرد. افزون بر این، گسترش‌پذیری و بازنشرپذیری دیجیتال می‌تواند زیان را از سطح فردی به سطح اجتماعی (فرسایش اعتماد عمومی به تصاویر و روایت‌های شخصی) تسری دهد و همین امر، منطق پیشگیرانه جرم‌انگاری را تقویت می‌کند (Feinberg, 1984: 12; Chesney & Citron, 2019: 1778).

طبق نظریه پیشگیری از جرم در فلسفه حقوق کیفری، مجازات و جرم‌انگاری محتوای دیجیتال جعلی می‌تواند باعث کاهش خطرات اجتماعی ناشی از انتشار این نوع محتوا گردد (عزیزی، ۱۴۰۲: ۸۵). در کنار پیشگیری، ادبیات جدید به «کارکرد بیانی» حقوق کیفری هم توجه دارد؛ یعنی جرم‌انگاریِ دقیق، این پیام هنجاری را تثبیت می‌کند که تعرض به حریم خصوصی جنسی و انتسابِ جعلیِ رفتار جنسی به اشخاص، صرفاً یک «سوءرفتار اخلاقی» یا «تخطی رسانه‌ای» نیست، بلکه نقض جدیِ حقِ فرد و کرامت انسانی است.(Citron & Franks, 2014, 104) در این راستا، قوانین باید به طور دقیق و شفاف در مورد استفاده غیرمجاز از تصاویر و محتوای شخصی در فضای دیجیتال تدابیر حمایتی اتخاذ کنند.(McGlynn & Toparlak, 2025: 40)

نظریه حمایتی دیگر در حقوق کیفری بر این تأکید دارد که افراد باید از آسیب‌های اجتماعی ناشی از سوءاستفاده از تصاویرشان محافظت شوند. به همین دلیل، ضروری است که در حقوق کیفری، تولید و توزیع دیپ‌فیک‌ها به‌ویژه آنهایی که حریم خصوصی افراد را نقض می‌کنند، جرم‌انگاری شوند تا عدالت اجتماعی و حقوق فردی افراد حفظ گردد.(Umbach et al., 2024: 96) البته همین مبانیِ فلسفی اقتضا می‌کند دامنه جرم‌انگاری «مضیّق و باضابطه» باشد؛ به‌گونه‌ای که محور جرم‌انگاری، دیپ‌فیک‌های زیان‌بارِ فاقد رضایت باشد و در عین حال، از تسری بی‌ضابطه به موارد رضایت‌مندانه

---

[۱].منظور از «نظریه حقوق کیفری» در اینجا، مجموعه مبانیِ فلسفه حقوق کیفری برای توجیه جرم‌انگاری است؛ یعنی این‌که قانون‌گذار بر چه اساسی می‌تواند رفتاری را جرم بداند. در این چارچوب، نخست «اصلِ زیان» مطرح است که مداخله کیفری را در جایی موجه می‌داند که رفتارِ مورد بحث به دیگران زیان قابل‌اعتنا وارد کند. در کنار آن، از کارکردهای حقوق کیفری مانند پیشگیری عمومی و بازدارندگی (کاهش وقوع جرم از طریق ایجاد هزینه و خطر برای مرتکبان) و نیز حمایت از خودآیینی، حریم خصوصی و کرامت انسانی (به‌ویژه در تعرضات جنسی و نقض تصویر/هویت دیجیتال) استفاده می‌شود تا ضرورت جرم‌انگاری دیپ‌فیک‌های زیان‌بار توضیح داده شود.

یا موارد کم‌خطر (مانند برخی اشکال طنز) پرهیز شود تا اصل تناسب و ملاحظه آزادی بیان رعایت گردد.(Chesney & Citron, 2019,1764)

در این حوزه، مداخله کیفری وقتی قانع‌کننده است که هم «دامنه» روشن باشد و هم «پشتوانه اجرایی» داشته باشد. دامنه روشن یعنی تمرکز بر دیپ‌فیک‌های جنسیِ فاقد رضایتِ قابل انتساب، نه تعمیم‌های مبهم. پشتوانه اجرایی یعنی کنارِ جرم‌انگاری، تکالیف حذف سریع، حفظ داده‌های ادله‌ای و استاندارد گزارش کارشناسی پیش‌بینی شود؛ وگرنه، حتی جرم‌انگاریِ شدید نیز در پرونده‌های واقعی به ضعف اثبات و فرسایش زمان می‌خورد.

۱.۴. پیوند دیپ‌فیک با تئوری‌های حقوق بشر و حفاظت از هویت

در نهایت، پیوند دیپ‌فیک با حقوق بشر نه‌تنها از منظر حریم خصوصی بلکه از حق بر هویت و حفظ کرامت انسانی نیز قابل‌تحلیل است. طبق کنوانسیون بین‌المللی حقوق مدنی و سیاسی (۱۹۶۶)، افراد حق دارند که شخصیت و هویت آنان مورد سوءاستفاده قرار نگیرد و از هرگونه تعرض به اعتبارشان محافظت شوند .دیپ‌فیک‌ها با دستکاری هویت افراد، این حق بنیادین را نقض می‌کنند. دستکاری‌هایی که در آن‌ها هویت فرد در صحنه‌های جنسی به‌طور جعلی جایگزین می‌شود، نه تنها اعتبار فرد را خدشه‌دار می‌کند بلکه از لحاظ روانی و اجتماعی نیز آسیب‌های جدی به فرد وارد می‌کند.

بنابراین، در راستای اصول حقوق بشر و حقوق فردی، قوانین باید به‌وضوح تولید و توزیع محتوای جعلی که باعث نقض هویت فرد می‌شود را محدود کنند. دیپ‌فیک‌ها به‌عنوان ابزاری برای تهدید، اخاذی و آسیب به شهرت و حقوق فردی شناخته می‌شوند و باید تحت قوانین ملی و بین‌المللی جرم‌انگاری شوند.(Masood et al., 2021: 114)

این موضوع به‌ویژه باتوجه‌به استفاده روزافزون از شبکه‌های اجتماعی و پلتفرم‌های آنلاین در سطح جهانی اهمیت می‌یابد، زیرا در این فضا، خطر سوءاستفاده از محتوای جعلی برای مقاصد مختلف به‌سرعت در حال افزایش است.

پیامد دستکاری هویت در دیپ‌فیکِ جنسی، صرفاً لطمه حیثیتیِ لحظه‌ای نیست؛ به‌دلیل ماندگاری دیجیتال، می‌تواند فرصت‌های شغلی و اجتماعی را محدود و فرد را به خودسانسوری و انزوا سوق دهد. بنابراین، حمایت از «هویت» در اینجا یعنی ایجاد امکان بازگردانیِ کنترل: توقف سریع انتشار، دشوارسازی بازنشر نسخه‌های هم‌ارز، و فراهم کردن ادله‌ای که در دادگاه قابل اتکا باشد.

۲. رویکرد تقنینی ایران در مواجهه با هرزه‌نگاری جعل عمیق

در پی گسترش بازنمایی‌های جنسی ساختگی با فناوری‌های نوین، سنجش کارآمدی پاسخ حقوقی ایران مستلزم فراتر رفتن از متن قانون و ارزیابی هم‌زمان بستر اجراست. این بخش با رویکرد توصیفی–تحلیلی–انتقادی، چارچوب هنجاری (مواد ۱۴، ۱۵ و ۱۶ قانون جرائم رایانه‌ای)، ضمانت‌های اجرای عملی، و ضعف‌های اجرا و چالش‌های قضایی را در پیوند با داده‌های میدانی و رویه‌های نزدیک بررسی می‌کند. تمرکز بر شاخص‌های عینی شباهت محسوس، انتساب قابل‌اعتماد، و فقدان رضایت معتبر در کنار الزامات حفظ داده، مسئولیت سکوها و استاندارد گزارش کارشناسی، امکان تبدیل ظرفیت قانونی به ادله قابل‌استناد و حمایت مؤثر از بزه‌دیدگان را فراهم می‌آورد. دستاورد نهایی، ترسیم معماری پاسخ ملی است که با اصلاحات تقنینی هدفمند و هم‌افزایی نهادی، شکاف میان هنجار و اجرا را کاهش می‌دهد.

۲.۱. چارچوب هنجاری ایران در جرم‌انگاری تولید و اشاعه محتوای مستهجن دیپ‌فیک

ماده ۱۴ قانون جرائم رایانه‌ای ۱۳۸۸ (اصلاحی ۱۳۹۹) مقرر می‌دارد: «هر کس به‌وسیله سامانه‌های رایانه‌ای یا مخابراتی یا حامل‌های داده محتویات مستهجن را منتشر، توزیع یا معامله کند یا به‌قصد تجارت یا افساد تولید یا ذخیره یا نگهداری کند، به حبس از نود و یک روز تا دو سال یا جزای نقدی ... یا هر دو مجازات محکوم خواهد شد». تبصره ۴ همان ماده، «محتویات مستهجن» را شامل «هر تصویر، صوت یا متن (اعم از واقعی یا ساختگی)» می‌داند که «نمایانگر برهنگی کامل زن یا مرد، اندام تناسلی یا آمیزش جنسی باشد». به‌علاوه، مطابق تبصره ۱ ماده ۱۴ قانون مذکور، «محتویات مبتذل» نیز تحت شمول قانون قرار می‌گیرد. ماده ۱۵ همان قانون نیز «تحریک، ترغیب، تهدید، تطمیع یا فریب» افراد یا «آموزش شیوه دسترسی» به محتویات مستهجن را جدا از فعل تولید، جرم‌انگاری می‌کند.

از منظر تحلیلی، این چارچوب از نظر شمول ماهوی برای دیپ‌فیکِ پورنوگرافیک مناسب است؛ چون عبارت «اعم از واقعی یا ساختگی» در تبصره ۴، دفاعِ «واقعی نبودن محتوا» را عملاً بی‌اثر می‌کند و ملاک را «بازنمایی مستهجن» می‌گذارد. بااین‌حال، نقطه تعیین‌کننده در اجرا اثبات انتساب معتبر به شخص قابل‌شناسایی و احراز فقدان رضایت است؛ وگرنه تکیه بر مفاهیم کلیِ «مستهجن/مبتذل» می‌تواند به ناهمسانی رویه و تردید در ارزش ادله منجر شود. ماده ۱۵ هم مزیت مهمی دارد چون «تسهیل‌گری و توزیع» را مستقل از تولیدکننده هدف می‌گیرد، ولی کارآمدی‌اش به دسترسی سریع به داده‌های سکو و گزارش کارشناسی استاندارد وابسته است.

۲.۱.۱. انطباق ماهوی با دیپ‌فیک پورنوگرافیک

تصریح «اعم از واقعی یا ساختگی» در تبصره ۴ ماده ۱۴، محور ورود بازنمایی‌های «مصنوعی یا دستکاری‌شده» به قلمرو کیفری است. در نتیجه، ویدئوهای دیپ‌فیکی که در آن‌ها چهره فرد قابل‌شناسایی بر بدن برهنه یا صحنه آمیزشی الصاق می‌شود، حتی اگر شخص هرگز در آن موقعیت حاضر نبوده باشد، در تعریف «مستهجن» می‌گنجند؛ و سناریوهای نیمه‌عریان یا تحقیر جنسی فاقد برهنگی کامل نیز می‌توانند ذیل عنوان «مبتذل» (تبصره ۱) قرار گیرند، مشروط به اینکه عنصر جنسی غالب و فقدان رضایت معتبر در تولید، انتساب و یا انتشار محرز گردد (رزمان، ۱۳۹۵: ۵۳). از سوی دیگر، ماده ۱۵ «اکوسیستم توزیع» را هدف می‌گیرد: هدایت کاربران به کانال‌ها و سایت‌ها، آموزش ساخت یا دسترسی به دیپ‌فیک، و ترغیب به مشاهده و خرید محتوا، مستقلاً قابل تعقیب است؛ این انتقال تمرکز از «ماهیت محتوا» به «شیوه اشاعه»، برای مهار شبکه‌های تولید و توزیع دیپ‌فیک اهمیت راهبردی دارد.

از منظر تحلیلی، تصریح «اعم از واقعی یا ساختگی» در تبصره ۴ ماده ۱۴ یک «پُل مفهومی» برای ورود دیپ‌فیک‌های پورنوگرافیک به قلمرو کیفری ایجاد می‌کند؛ یعنی برای تحقق رکن مادی، «واقعی بودن» شرط نیست و کافی است محتوا از حیث عرفی واجد برهنگی و مضامین جنسی باشد و به شخص قابل‌شناسایی منتسب شود. بااین‌حال، در مقام اجرا، نقطه ثقلِ دعوا معمولاً روی دو مؤلفه می‌چرخد: (قابلیت انتساب قابل‌اعتماد) اینکه مخاطب متعارف، محتوا را به فرد مشخص نسبت می‌دهد و(فقدان رضایت معتبر)در تولید، انتساب یا انتشار.

هم‌زمان، ماده ۱۵ یک انتخاب سیاست کیفری مهم دارد: تمرکز را از «صرفِ ماهیت محتوا» به «زنجیره اشاعه» منتقل می‌کند. نتیجه این است که حتی اگر شناسایی سازنده دیپ‌فیک دشوار باشد، می‌توان حلقه‌های میانی (هدایت به کانال‌ها و سایت‌ها، آموزش دسترسی یا ساخت، ترغیب به مشاهده و خرید) را مستقل از فعل تولید تعقیب کرد. این سازوکار، از نظر کارآمدی، ظرفیت حقوقی مقابله با شبکه‌های توزیع و اقتصاد پنهانِ محتوا را بالا می‌برد؛ اما موفقیت آن در عمل به سرعت اقدام (حفظ داده، حذف سریع) و امکان جمع‌آوری ادله دیجیتال وابسته می‌ماند.

۲.۱.۲. رویه‌های نزدیک و قابلیت انطباق

اگرچه بخش عمده آراء منتشرشده ناظر به قالب‌های سنتی‌تر محتواست، قابلیت انطباق با بسترهای نوین قابل مشاهده است: دادنامه ۹۲۰۹۹۷۰۲۷۰۲۰۰۰۶۶ دادگاه تجدیدنظر تهران (۱۳۹۲) جرم‌انگاری انتشار محتوای مستهجن در شبکه‌های اجتماعی را ذیل ماده ۱۴ پذیرفته؛ و رأی

۹۴۰۹۹۸۲۹۲۵۴۰۱۴۰۷ دیوان عالی کشور (۱۳۹۵) نشر محتوای مستهجن در پیام‌رسان را قابل انطباق با همین ماده دانسته است. در نسبت با دیپ‌فیک، منطق این آراء که ماهیت مستهجن و شاکله توزیع را معیار قرار می‌دهند، در کنار تبصره ۴ (شمول ساختگی)، زمینه استنادپذیری را تقویت می‌کند. در سطح انتظامی نیز، گزارش‌های عملیاتی پلیس فتا درباره رصد گسترده حلقه‌های توزیع و برخورد با آموزش‌ها و پیوندهای دسترسی، نشان می‌دهد ماده ۱۵ می‌تواند حلقه واسط اشاعه را مهار کند، هرچند کارآمدی نهایی، منوط به سرعت حذف و همکاری سکوهاست (گزارش پلیس فتا، ۱۴۰۰).

از منظر تحلیلی، این دو رأی نشان می‌دهد «معیارِ انطباق» در عمل بیشتر بر ماهیت مستهجن محتوا و نحوه انتشار و دسترسی استوار است تا بر قالب فنی یا بستر انتشار؛ بنابراین، با ضمیمه‌شدن تبصره ۴ ماده ۱۴ (شمول محتوای ساختگی)، مسیر استناد به ماده ۱۴ برای دیپ‌فیک پورنوگرافیک از نظر منطقی هموارتر می‌شود. نکته مهم این است که این رویه‌ها به‌طور ضمنی یک پیام روش‌شناختی دارند: دادگاه‌ها وقتی با فناوری جدید روبه‌رو می‌شوند، معمولاً به جای ورود به جزئیات تکنیکی، از «قالب‌های حقوقی موجود» و شاخص‌های عرفی (مستهجن بودن و تحقق انتشار) استفاده می‌کنند؛ همین امر پیش‌بینی‌پذیری نسبی ایجاد می‌کند.

بااین‌حال، اتکای صرف به این «قیاس رویه‌ای» یک محدودیت جدی دارد: در دیپ‌فیک، محل اختلاف غالباً از «مستهجن بودن» عبور می‌کند و روی انتساب معتبر و کیفیت ادله دیجیتال متمرکز می‌شود؛ یعنی اگر دادگاه قانع نشود که محتوا واقعاً به شخص شاکی قابل‌انتساب است (یا زنجیره انتشار به متهم قابل‌نسبت)، همان قالب حقوقی هم به نتیجه نمی‌رسد. از این زاویه، اشاره به گزارش‌های عملیاتی پلیس فتا درباره مهار حلقه‌های توزیع ذیل ماده ۱۵ یک مزیت دارد: نشان می‌دهد نظام اجرا می‌تواند به‌جای تعقیب دشوار «سازنده ناشناس»، روی شبکه‌های اشاعه و آموزش/هدایت متمرکز شود. اما شرط کارآمدی این راهبرد، اقدام فوری برای حذف و حفظ داده و همکاری سکوها است؛ وگرنه حتی بهترین عنوان مجرمانه، در مرحله اثبات و اجرا فرسوده می‌شود.

۲.۱.۳. نکات انتقادی

نخست: اتکا به عرف و ناهمگونی پیش‌بینی‌پذیری

تکیه مواد ۱۴ و ۱۵ بر مفاهیم ارزشی «مستهجن یا مبتذل» مزیت انعطاف تفسیری دارد، اما در غیاب شاخص‌های راهنما، می‌تواند به ناهمگونی قضایی بینجامد و «امنیت قضایی» را تضعیف کند (نوری، ۱۳۸۶: ۴۲). گره در پرونده‌های دیپ‌فیک برجسته‌تر است: قاضی ناچار است میان «محتوای صرفاً

جنسی» و «بازنمایی تحقیر جنسی شخص مشخص» مرزگذاری کند. راهکار نظری قابل‌دفاع، افزودن شاخص‌های عینی آزمون‌پذیر در متن قانون یا تبصره تبیینی است: «شباهت محسوس برای ناظر متعارف»، «قابلیت انتساب به شخص قابل‌شناسایی» و «فقدان رضایت معتبر»؛ این سه شاخص، هم‌زمان مرز هنر، طنز و نقد را از مستهجن و مبتذل آسیب‌زا تفکیک و پیش‌بینی‌پذیری را افزایش می‌دهند.

دوم: تناسب ضمانت اجرا با منطق شبکه

مجازات‌های فعلی پیش‌بینی‌شده در مواد ۱۴ و ۱۵، حبس‌های کوتاه‌مدت و جزای نقدی سبک با «شدت آسیب شبکه‌ای» (سرعت انتشار، ماندگاری لکه حیثیتی، و انگیزه انتفاع) هم‌سنگ نیست و بنابراین بازدارندگی مؤثری ایجاد نمی‌کند (عزیزی، ۱۴۰۲: ۸۶). در پرونده‌هایی که بر «دانلودِ فروشی»، «عضویت ویژه» یا «تبلیغات» استوارند، سود اقتصادی معمولاً از ریسک کیفری بیشتر است. راه‌حل کارآمد آن است که نظام تشدید کیفر بر پایه سه شاخص قابل سنجش بازطراحی شود: «دامنه انتشار»، «میزان انتفاع اقتصادی»، و «تکرار سازمان‌یافته»؛ و در کنار آن، استرداد منافع نامشروع و جزای مالی درصدی بر مبنای گردش مالی شبکه توزیع اعمال گردد (تبریزی، ۱۴۰۳: ۹۷). این هم‌ترازی نو، نسبت «هزینه–فایده ارتکاب» را معکوس می‌کند و انگیزه اقتصادی جرم را به‌طور ملموس کاهش می‌دهد.

سوم: گذار از قاعده اعلامی به ابزار حمایتی

بدون تکالیف روشن برای سکوها (حذف در مهلت معین، ارائه داده ادله‌ای، پیشگیری از بازنشر نسخه‌های معادل) و بدون پروتکل ملی حفظ فوری داده و استاندارد گزارش کارشناسی دیپ‌فیک، ظرفیت هنجاری مواد ۱۴ و ۱۵ در محکمه به ادله ناپایدار فروکاسته می‌شود؛ بنابراین، پیشنهاد تکمیلی در چهار سطح باید در متن قانون و آیین‌نامه تصریح شود: (۱) الزام سکو به حذف در بازه ۲۴ ساعته برای محتوای جنسی ساختگی پرخطر؛ (۲) الزام به حفظ و ارائه فراداده بارگذاری و ویرایش؛ (۳) پایش فعال برای نسخه‌های اساساً معادل؛ (۴) مرکز تماس واحد ۲۴ساعته برای مکاتبات قضایی-انتظامی. این ملزومات، مواد ۱۴ و ۱۵ را از «اعلام جرم» به ابزار حمایت مؤثر نسبت به بزه‌دیدگان ارتقا می‌دهد.

۲.۱.۴. پیوستگی با ماده ۱۶ و پوشش سناریوهای فاقد برهنگی کامل

هر جا که بازنمایی دیپ‌فیک به‌لحاظ جنسی «غالب» نیست، اما لطمه اعتباری عرفی ایجاد می‌کند، مسیر تکمیلی ماده ۱۶ قابل استناد است: هم‌خوانی عنصر مادی «تغییر یا تحریف» با ساختار فنی جعل عمیق، و نتیجه‌محوری پیوستگی با ماده ۱۶ و پوشش سناریوهای فاقد برهنگی کامل

در مواردی که محتوای دیپ‌فیک هرچند بار جنسی دارد، اما «برهنگی کامل» یا نمایش صریح اعمال جنسی در آن وجود ندارد و به همین دلیل، انطباق آن با عنوان «مستهجن» محل تردید است، می‌توان به ماده ۱۶ قانون جرائم رایانه‌ای به‌عنوان مسیر تکمیلی استناد کرد .مبنای این استناد آن است که در دیپ‌فیک، عنصر مادیِ «تغییر یا تحریف» به‌طور مستقیم با سازوکار فنی جعل عمیق سازگار است و نتیجه‌ی آن نیز می‌تواند به «هتک حیثیت عرفی» بینجامد؛ بنابراین، سناریوهایی مانند نسبت‌دادن رفتار یا وضعیت جنسی تحقیرآمیز به شخص قابل‌شناسایی، حتی بدون نمایش برهنگی کامل، قابل پوشش خواهد بود. البته شرط کارآمدی این مسیر آن است که گزارش کارشناسی دو مؤلفه را صریح و مستند بیان کند: «شباهت محسوس» میان تصویرو ویدئوی جعلی و شخص واقعی، و «قابلیت انتساب قابل‌اعتماد» (یعنی توضیح روش و حدود خطا)، تا مرجع رسیدگی بتواند بار اثبات را منصفانه ارزیابی کند (عزیزی، ۱۴۰۲: ۱۰۱).«هتک حیثیت عرفی» امکان پوشش سناریوهای فاقد برهنگی کامل را فراهم می‌کند؛ به‌شرط آن‌که در گزارش کارشناسی، دو شاخص «شباهت محسوس» و «انتساب قابل اعتماد» روشن بیان شود تا بار اثبات متوازن گردد (عزیزی، ۱۴۰۲: ۱۰۱).

از منظر تحلیلی، نقطه‌ضعف مشترک مواد ۱۴ و ۱۵ در مواجهه با دیپ‌فیک آن است که بر مفاهیم ارزشیِ «مستهجن/مبتذل» تکیه دارند، بی‌آنکه شاخص‌های عینی برای تشخیص و انتساب ارائه کنند؛ پیامد طبیعیِ این خلأ، ناهمگونی رویه و کاهش پیش‌بینی‌پذیری در پرونده‌هایی است که مرز «محتوای جنسی» با «بازنمایی تحقیرآمیزِ منتسب به شخص معین» محل نزاع می‌شود. از سوی دیگر، حتی اگر انطباق ماهوی پذیرفته شود، ضمانت‌اجرای موجود با منطق شبکه هم‌تراز نیست: سرعت بازنشر، شکل‌گیری نسخه‌های «اساساً معادل» و انگیزه انتفاع باعث می‌شود حبس‌های کوتاه‌مدت و جزای نقدی ثابت اثر بازدارنده پایدار نداشته باشند. در این میان، ماده ۱۶ یک مسیر تکمیلی برای سناریوهای فاقد برهنگی کامل فراهم می‌کند، اما بدون استاندارد مشترکِ ادله، این مسیر می‌تواند اختلاف را از «عنوان مجرمانه» به «بار اثبات» منتقل کند. بر این مبنا، کارآمدسازی چارچوب موجود بیش از آنکه نیازمند جرم‌انگاری جدید باشد، محتاج سه اصلاح به‌هم‌پیوسته است: (۱) الحاق معیارهای قابل سنجشِ «شباهت محسوس»، «انتساب قابل‌اعتماد با ذکر حدود خطا» و «فقدان رضایت معتبر» برای کاهش ناهمگونی؛ (۲) بازطراحی بازدارندگی بر پایه دامنه انتشار، میزان انتفاع و تکرار سازمان‌یافته؛ و (۳) تبدیل قاعده اعلامی به ابزار اجرا از طریق تکالیف روشن سکوها (حذف زمان‌مند، حفظ و ارائه فراداده، و مدیریت نسخه‌های معادل) همراه با پروتکل ملیِ حفظ داده

و استاندارد گزارش کارشناسی.

۲.۲. ضمانت اجرای عملی: از «اعمال موردی» تا «معماری پاسخ مؤثر»

۲.۲.۱. نابسندگی الگوی فعلی پاسخ و لزوم «معماری چندلایه»

داده‌های عملیاتی پلیس فتا (رصد ۹٬۸۳۵ تارنما، شناسایی ۱٬۷۶۵ مورد و سهم بالای سکوهای خارجی) نشان می‌دهد که الگوی کنونی، عمدتاً موردی، پسینی و پلتفرم محور است؛ به بیان دیگر، از جنس «تعقیب پرونده به پرونده» و بدون معماری یکپارچه حفظ داده، تکلیف‌های روشن سکو و استاندارد ادله است (گزارش پلیس فتا، ۱۴۰۳). این الگو حتی با اجرای سخت‌گیرانه مواد ۱۴ و ۱۵، به دلیل بازانتشار سریع و «نسخه‌های اساساً معادل»، اثر بازدارنده‌ی پایداری ایجاد نمی‌کند. نیاز اصلی، انتقال از «اقدام پراکنده» به طرح ملی پاسخ است که چهار جزء هم‌بسته داشته باشد :(۱) حفظ فوری داده با دستورهای استاندارد و مهلت‌های معلوم، (۲) الزام سکو به حذف فوری و ارائه داده‌های ادله‌ای، (۳) گزارش کارشناسی استاندارد (روش، آستانه تصمیم، حدود خطا، چندنشانه‌ای)، (۴) پیوند پلیس، دادسرا و مرکز ملی برای اقدام هم‌زمان.

۲.۲.۲. تعارض «شدت آسیب شبکه‌ای» با «سبکیِ ضمانت اجرا»

منطق شبکه و سرعت انتشار، ماندگاری لکه حیثیتی و انگیزه انتفاع با حبس‌های کوتاه‌مدت و جزای نقدی سبک (مواد ۱۴–۱۶) هم‌تراز نیست؛ همین عدم تناسب، کارکرد بازدارندگی را تضعیف می‌کند (بابایی، ۱۳۹۷: ۹۴). برای هماهنگی با شدت آسیب شبکه‌ای، پیشنهاد می‌شود:

برای ارتقای بازدارندگی، سه محور به‌صورت روشن و قابل‌اجرا پیشنهاد می‌شود:

۱. طبقه‌بندی تشدیدی جرایم بر پایه سه شاخص آزمون‌پذیر («دامنه انتشار»، «انتفاع اقتصادی»، و «تکرار سازمان‌یافته»)، تا با افزایش هر یک از این شاخص‌ها، درجه مجازات به طور متناسب بالا رود؛

۲. بازگرداندن منافع نامشروع به همراه جریمه‌های درصدی متکی بر گردش مالی شبکه توزیع (نه جریمه‌های ثابت)، تا انگیزه اقتصادی ارتکاب عملاً از میان برود؛

۳. تکالیف تبعی برای سکو و گرداننده مانند اجرای اقدامات اصلاحی، گزارش‌دهی دوره‌ای به مرجع ناظر، و ممیزی منظم سازوکارهای تکثیر محتوا برای پیشگیری از بازنشر.

اجرای هم‌زمان این سه محور، مواد ۱۴ تا ۱۶ را از سطح یک «قاعده اعلامی» به ابزار مؤثر بازدارنده تبدیل می‌کند.

۲.۲.۳. خلأ پروتکل ادله و پیامد آن بر استنادپذیری

در عمل، فقدان «پروتکل واحد گزارش کارشناسی دیپ‌فیک» به تعارض نظرهای فنی، ناهمگونی در سنجش ادله و در نهایت کاهش وزن اثباتی گزارش نزد قاضی می‌انجامد؛ گزارشی که بخواهد استنادپذیر باشد باید به‌صراحت (الف) آستانه تصمیم‌گیری (نظیر سطح اطمینان و معیار پذیرش، رد)، (ب) حدود خطا و دامنه عدم قطعیت، و (پ) ترکیب نشانه‌ها در سه‌لایه تصویر، صوت و حرکت را توضیح دهد و نشان دهد نتیجه چگونه از همگرایی این نشانه‌ها به‌دست‌آمده است (شیری، ۱۴۰۱، ص ۱۵۱). افزون بر این، گزارش باید قابلیت بازتولید (ذکر نسخه ابزارها، مدل‌ها، پارامترهای به‌کاررفته، و داده‌های آزمون (زنجیره حفاظت از ادله)، ثبت زمان، محل، فراداده و هرگونه دست‌کاری مجاز)، و منطق انتساب به شخص قابل‌شناسایی را مرحله‌به‌مرحله مستند کند تا امکان ارزیابی قضایی مستقل فراهم شود (اکبری و دیگران، ۱۴۰۱: ۱۶۱). بر همین مبنا، تنظیم آیین‌نامه ملی گزارش کارشناسی محتوای مصنوعی زیر نظر قوه قضائیه با مشارکت پلیس فتا و دانشگاه‌ها، مشتمل بر پیوست‌های فنی، فرم‌های استاندارد (برای ثبت آستانه‌ها، حدود خطا و زنجیره حفاظت)، و دستورالعمل یکسان ارزیابی، شرط لازم برای تبدیل یافته‌های فنی به «ادله قابل استناد» است؛ در غیر این صورت، حتی عملیات موفق رصد و کشف در مرحله کارشناسی به ابهام ادله و کاهش قدرت اقناعیِ نتیجه منتهی خواهد شد (محمدی فردوئی، ۱۳۹۷: ۴۶۸).

۲.۲.۴. تکالیف سکو و خط‌کش‌های اجرایی قابل سنجش

اگرچه اسناد سیاستی مرکز ملی فضای مجازی بر «مسئولیت‌پذیری سکوها» و ضرورت اقدام هماهنگ تأکید دارند، چالش اصلی در سطح اجرا نه فقدان ایده، بلکه فقدانِ «معیار سنجش و امکان ممیزی» است؛ یعنی تکلیف‌ها تا زمانی که به شاخص‌های قابل‌اندازه‌گیری تبدیل نشوند، در عمل به استاندارد رفتاری الزام‌آور بدل نمی‌گردند و قابلیت پیگیری و ارزیابی هم پیدا نمی‌کنند (مرکز ملی فضای مجازی، ۱۴۰۳/۴/۴). بر همین مبنا، کارآمدسازی تکالیف سکوها نیازمند تعریف سه شاخص ممیزی‌پذیر است: نخست، «تعهد به سطح خدمتِ حذف»، یعنی تعیین بازه زمانی روشن برای اقدام، برای نمونه حذف ظرف ۲۴ ساعت در محتوای جنسی ساختگیِ پرخطر همراه با قرائن قوی؛ دوم، تعهد به حفظ و تحویل حداقلی و استانداردِ ادله دیجیتال، از جمله فراداده بارگذاری، نشانی مبدأ و سوابق ویرایش، به نحوی که زنجیره استنادپذیری مخدوش نشود؛ و سوم، تعهد به پیشگیری از بازنشر مؤثر، به این معنا که سکو به جای برخورد صرفاً واکنشی، ملزم به پایش و کنترل الگوهای بازتکثیر شود، به‌ویژه در قالب نسخه‌های اساساً معادل. در غیاب چنین شاخص‌هایی، اوامر قضایی عملاً به دستورهایی فاقد ابزار راستی‌آزمایی تبدیل می‌شوند و از سطح الزام به سطح توصیه تنزل می‌یابند (قناد، ۱۳۹۰: ۷۱). نتیجه آنکه، نقطه ثقل ضمانت اجرای مؤثر نه افزایش مداخله موردی، بلکه طراحی یک نظام پاسخِ قابل‌سنجش است؛ نظامی که در آن تکلیف سکو، استاندارد ادله، تناسب واکنش کیفری و مدنی، و هماهنگی نهادی به‌طور هم‌زمان و قابل ارزیابی پیش‌بینی شود.

از منظر تحلیلی، این مجموعه نشان می‌دهد چالش اصلی ایران در مواجهه با پورنوگرافی دیپ‌فیک، بیشتر از آن‌که «کمبود حکم قانونی» باشد، ضعف در سازوکار اجرا و اثبات است. آمار پلیس فتا هم مؤید همین نکته است که برخوردهای موردی و پرونده‌محور، حتی با اتکای کامل به مواد ۱۴ و ۱۵، در برابر بازانتشار سریع و تولید «نسخه‌های نزدیک به همان محتوا» اثر پایدار ندارد؛ چون چرخه انتشار در فضای مجازی از چرخه رسیدگی قضایی سریع‌تر است. بنابراین پاسخ مؤثر باید از حالت اقدام پراکنده خارج شود و به یک سازوکار هماهنگ و مرحله‌به‌مرحله تبدیل گردد؛ یعنی از لحظه کشف تا حفظ داده، از حفظ داده تا گزارش کارشناسی، و از گزارش کارشناسی تا دستور حذف و جلوگیری از بازنشر، همه در یک مسیر روشن و قابل اجرا قرار بگیرند.

هم‌زمان، ناهمخوانی ضمانت اجرا با واقعیت جرم در فضای مجازی باعث تضعیف بازدارندگی می‌شود. وقتی انگیزه مالی در توزیع محتوا بالاست، حبس کوتاه‌مدت یا جزای نقدی ثابت ممکن است در عمل به «هزینه قابل تحمل» تبدیل شود. به همین دلیل، منطقی‌تر است که شدت واکنش حقوقی بر اساس شاخص‌های روشن تنظیم شود؛ مثل گستره انتشار، میزان منفعت اقتصادی، و تکرار سازمان‌یافته، و در کنار آن بازگرداندن منافع نامشروع جدی گرفته شود تا تعقیب صرفاً جنبه ظاهری پیدا نکند.

نقطه حساس‌تر در عمل، مرحله دادگاه است: اگر شیوه‌نامه واحد برای کارشناسی دیپ‌فیک و «زنجیره نگهداری ادله» وجود نداشته باشد، نتیجه فنی به‌سختی به دلیل قابل اتکا تبدیل می‌شود. گزارش کارشناسی وقتی وزن اثباتی پیدا می‌کند که روش، میزان خطا، معیار تصمیم‌گیری و امکان بازبینی روشن باشد. از همین رو، دو اقدام مکمل اهمیت کلیدی دارد: نخست، تدوین دستورالعمل ملی برای گزارش کارشناسی محتوای مصنوعی؛ دوم، تبدیل تکالیف سکوها به معیارهای قابل سنجش و قابل پیگیری مثل مهلت حذف، تکلیف حفظ و ارائه داده‌های فنی، و اقدام برای جلوگیری از بازنشر نسخه‌های مشابه. جمع‌بندی اینکه اصلاح مسیر در ایران، کمتر به اضافه‌کردن عنوان مجرمانه جدید وابسته است و بیشتر به سامان‌دادن اجرای قانون و پیوند دادن «کشف فنی» به «اقناع قضایی» و سپس «حذف مؤثر» نیاز دارد.

۲.۳. آسیب‌شناسی ضعف‌های اجرا و چالش‌های قضایی: از «ابهام مفهومی» تا «ناهمسانی رویه‌ای»

۲.۳.۱. ابهام مفهومی و شکاف بین ماده ۱۴ (مستهجن ساختگی) و ماده ۱۶ (تحریف)

فقدان تعریف قانونی «دیپ‌فیک» سبب دوگانگی استناد شده است: برخی مراجع بر «تحریف» (ماده ۱۶) و برخی بر «مستهجن ساختگی» (ماده ۱۴) تکیه می‌کنند؛ نتیجه، ناهمگونی در احراز رکن مادی و اختلاف در بار اثبات است (شیری، ۱۴۰۱: ۱۴۹).

چاره تقنینی حداقلی آن است که تعریف مضیّق با سه سنجه آزمون‌پذیر «شباهت محسوس برای ناظر متعارف»، «انتساب قابل‌اعتماد به شخص قابل‌شناسایی»، «فقدان رضایت معتبر» به متن قانون الحاق گردد تا مرزگذاری روشنی میان دو ماده شکل گیرد.

۲.۳.۲. گسست ادله در بسترهای خارجی، زوال داده و نبود توافق زمان‌بندی همکاری

تمرکز توزیع بر سکوهای خارجی و زوال سریع داده (حذف، بارکدگذاری، نسخه‌های معادل) فرایند انتساب را فرساینده می‌کند؛ یافته‌های پلیس فتا درباره سهم بالای یک سکو، نمونه روشن این گره است (قناد، ۱۳۹۰: ۷۴). راه‌حل صرفاً حقوقی نیست؛ نیاز به تفاهم‌نامه‌های قضایی دوجانبه، مرکز تماس واحد ۲۴ساعته و توافق زمان‌بندی همکاری طبق زمان‌بندی مشترک دارد که در آن، مهلت حفظ داده، پاسخ به درخواست‌ها و حذف نسخه‌های معادل مشخص باشد.

۲.۳.۳. کمبود آمار تفکیک‌شده و پیامد آن برای سیاست‌گذاری و قضا

نبود آمار رسمی تفکیک‌شده درباره «دیپ‌فیک پورنوگرافیک» (در گزارش‌های عمومی پلیس فتا و اطلاعیه‌های مرکز ملی) وزن تجربی تصمیم‌سازی را پایین می‌آورد؛ نتیجه، سیاست‌گذاری واکنشی و فقدان ارزیابی اثربخشی است. باید کاربرگ استاندارد ثبت پرونده با کد اختصاصی این دسته شکل بگیرد و گزارش سالانه ملی منتشر شود تا پژوهش و قضا بتوانند بر داده‌های پایا تکیه کنند.

۲.۳.۴. خلأ حمایت ترکیبی از بزه‌دیدگان

بار اثبات و هزینه پیگیری، عملاً بر دوش بزه‌دیده می‌افتد. پیکربندی فعلی قانون، سازوکار جبران و حمایت فوری (حذف، مشاوره روانی و حقوقی، جبران هزینه حذف و بازیابی شهرت) را پیش‌بینی نکرده است.

برای حل این خلأ پیش‌بینی دستورهای حمایتی فوری (حفظ داده، حذف موقت، منع تماس)، صندوق حمایت از بزه‌دیدگان محتوای جنسی ساختگی (تأمین هزینه‌های فنی و حقوقی)، و الزام سکوها به همکاری حمایتی (واسطه‌گری در اطلاع‌رسانی به بازنشردهندگان، ابزارهای گزارش‌دهی سریع) پیشنهاد می‌گردد.

نکته نهایی آنکه تا زمانی که تعریف مضیّق، استاندارد ادله، تکلیف سکو، آمار تفکیک‌شده و ابزارهای حمایتی به‌صورت الزام‌آور در کنار هم قرار نگیرند، حتی اجرای دقیق مواد ۱۴–۱۶ نیز به «نتیجه دادگاهی پایدار و حمایت محور» تبدیل نخواهد شد.

از منظر تحلیلی، ضعف اصلی این حوزه بیش از آنکه «خلأ جرم‌انگاری» باشد، ناهماهنگی اجزای اجراست. نخست، چون تعریف قانونی روشنی از دیپ‌فیک وجود ندارد، مرجع رسیدگی گاهی به ماده ۱۴ (مستهجنِ ساختگی) و گاهی به ماده ۱۶ (تحریف) تکیه می‌کند؛ پیامدش ناهمسانی در احراز رکن مادی و جابه‌جایی بار اثبات است. دوم، در پرونده‌هایی که بستر انتشار خارجی است، زوال سریع داده و تأخیر در دسترسی به اطلاعات پلتفرمی روند انتساب را شکننده می‌کند؛ بدون «حفظ

فوری داده» و ترتیبات همکاری زمان‌مند، ادله زود از دست می‌رود و پرونده از نظر اثباتی تضعیف می‌شود. سوم، نبود آمار تفکیک‌شده درباره دیپ‌فیک پورنوگرافیک باعث می‌شود سیاست‌گذاری و حتی ارزیابی قضایی، مبتنی بر داده پایدار نباشد و امکان سنجش اثربخشی اقدامات کاهش یابد. چهارم، چارچوب فعلی حمایت فوری و جبرانِ عملی برای بزه‌دیدگان را کم‌رنگ گذاشته و در عمل هزینه و بار پیگیری را بر قربانی تحمیل می‌کند. نتیجه این است که تا زمانی که تعریف حداقلیِ قابل سنجش، استاندارد ادله، تکلیف روشن سکوها، داده‌های آماری قابل اتکا و سازوکارهای حمایتی فوری هم‌زمان دیده نشود، اجرای مواد ۱۴ تا ۱۶ به خروجی قضایی یکنواخت و حمایت‌محور منتهی نخواهد شد.

۳. رویکرد تقنینی اتحادیه اروپا در مواجهه با هرزه‌نگاری جعل عمیق

در این بخش، با محوریت «تناظر ریسک و الزام» در قانون هوش مصنوعی اتحادیه اروپا و هم‌افزایی آن با قانون خدمات دیجیتال، مسیر تقنین اتحادیه در قبال دیپ‌فیک به‌ویژه در قلمرو پورنوگرافی بدون رضایت صورت‌بندی می‌شود. ابتدا جایگاه دیپ‌فیک در چارچوب هنجاری و نقد طبقه‌بندی پیش‌فرض آن تبیین می‌گردد؛ سپس پیوندهای اجرایی با تکالیف سکوها و الگوی قضایی حذف «نسخه‌های اساساً معادل» روشن می‌شود. در ادامه، محدودیت اتکای صرف به شفافیت و ضرورت «اثبات منشأ و تمامیت محتوا» و نیز ارتقای گزینشی ریسک بر پایه معیارهای عینی بررسی می‌گردد. نهایتاً، ضمانت‌های اجرایی، الزامات شفافیت و آسیب‌شناسی ضعف‌های اجرا از ادله برون‌مرزی تا نابرابری ظرفیت نظارتی به همراه پیشنهادهای عملی قابل‌سنجش ارائه می‌شود.

۳.۱. چارچوب هنجاری و جایگاه «دیپ‌فیک» در قانون هوش مصنوعی اتحادیه اروپا

رویکرد اتحادیه اروپا در قانون هوش مصنوعی بر «تناظر ریسک–الزام» استوار است: تفکیک کاربری‌های هوش مصنوعی به سطوح «حداقلی/محدود»، «پرخطر» و «غیرقابل‌قبول» و الحاق الزامات متناسب. تولید و دست‌کاری محتوای تصویری، صوتی یا ویدئویی که «برای ناظر متعارف، صادق و واقعی می‌نماید»، ذیل الزامات شفافیت فصل مربوط قرار می‌گیرد (افشای مصنوعی بودن) و به طور پیش‌فرض در رده «ریسک محدود» دیده شده است (قانون هوش مصنوعی، تبصره ۱۳۴، ماده ۵۰) منتقدان نشان داده‌اند که این طراحی، عملاً تنظیم «دیپ‌فیک» را به آشکارسازی برچسب گونه فرومی‌کاهد و بر هم‌افزایی با اسناد افقی دیگر (مثل قانون خدمات دیجیتال) تکیه می‌کند. (BIock, 2024: 184)به بیان دیگر، هسته هنجاریِ فعلی، شفافیت را «شرط کف» می‌گیرد نه «قید

ماهوی» بر تولید و به‌کارگیری.

۳.۱.۱. نقد طبقه‌بندی پیش‌فرض و معیار ارتقای ریسک

قراردادن «دیپ‌فیک‌های پورنوگرافیک فاقد رضایت» ذیل «ریسک محدود» با شدتِ آسیب‌های حیثیتی و اقتصادی–اجتماعی آن هم‌تراز نیست. ادبیات تخصصی سه خلأ را برمی‌شمارد: (الف) اتکای بیش از حد به برچسب‌گذاری خود افشاگر؛ (ب) بی‌اعتنایی به «زیان‌های ساختاری» (ماندگاری لکه حیثیتی و بازنشر نسخه‌های اساساً معادل)؛ (ج) فقدان آستانه‌های روشن برای ارتقای ریسک (Block, 2024: 186–189)یک بازطراحی منسجم می‌تواند با اتکا به ابزارهای خود قانون هوش مصنوعی انجام شود: استفاده از اختیار ماده ۷ این قانون برای افزودن کاربری‌های خاص «دیپ‌فیک جنسی فاقد رضایتِ قابل‌انتساب به شخص مشخص» به پیوستِ مصادیق «پرخطر»؛ معیارهای ارتقا نیز بر سه سنجه آزمون‌پذیر بنا شود: «قابلیت انتساب به فرد قابل‌شناسایی»، «فقدان رضایت معتبر»، و «دامنه انتشار/قابلیت بازنشر معادل» (Becker, 2024: 360–364).

۳.۱.۲. هم‌تنظیمی با قانون خدمات دیجیتال و جهت رویه قضایی

در کنار قانون هوش مصنوعی، قانون خدمات دیجیتال، سازوکار «ارزیابی و مهار خطرات نظام‌مند» را برای سکوهای بسیار بزرگ تحمیل کرده و در رهنمودهای انتخابات ۲۰۲۴ نیز صراحتاً بر لزوم «برچسب‌گذاری روشن و پایدارِ محتوای تولیدشده با هوش مصنوعی ویا دست‌کاری‌شده (از جمله دیپ‌فیک)» تأکید کرده است؛ کمیسیون این الزامات را ذیل تکالیف کاهش ریسک پی می‌گیرد (European Commission, 2024: sec. Mitigation measures).در رویه قضایی نیز، رأی دیوان عدالت اتحادیه اروپا در پرونده شماره C-18/18، موسوم به «گلاویشنیک–پیشچک علیه فیس‌بوک ایرلند»، با تجویز حذف «نسخه‌های اساساً معادلِ» یک محتوای غیرقانونی، الگوی قضایی مهمی برای مهار بازتولید سریع دیپ‌فیک‌ها فراهم کرد (CJEU, 2019: paras. 45–49).همین الگو، امکان ترجمه خروجی‌های فنیِ تشخیص و یافتن «نمونه‌های معادل» به تکالیف اجرایی بر سکوها را تقویت می‌کند و پیوند چارچوب دوگانه هوش مصنوعی و خدمات دیجیتال را عینی می‌سازد (Block, 2024: 189–191).

۳.۱.۳. محدودیت «صرف شفافیت» و لزوم «اثبات منشأ»

پژوهش‌ها هشدار می‌دهند که اتکای صرف به خود برچسب‌زنی، نسبت به کنشگران بد نیت

بی‌اثر می‌ماند؛ اغلب تولیدکنندگان دیپ‌فیکِ زیان‌بار اصولاً برچسب نمی‌زنند و استثنائات نیز در عمل قابل سوءاستفاده است (Block, 2024: 190–192) ؛ بنابراین، کنار الزام افشا، به «معماری اثبات منشأ و تمامیت محتوا» (فراداده‌های مقاوم در برابر دست‌کاری و نشانگرهای ماشینی همراه محتوا) و «الگوی مثبت نشان» برای محتوای اصیل نیاز است؛ دومی دقیقاً پیشنهادی است که دکترین به‌عنوان مکمل منفی‌شان «دیپ‌فیک» طرح کرده تا بار اجرای عملی کاهش یابد و خطای نوع دوم کم شود (Becker, 2024: 353–361).

۳.۱.۴. مسیر اصلاح: ارتقای گزینشی ریسک و هم‌افزایی اجرایی

برای همسوسازی هنجار با واقعیتِ زیان، می‌توان سه‌گام حقوقی را هم‌زمان پیش برد: (۱) به‌روزرسانی پیوست‌های قانون هوش مصنوعی ذیل ماده ۷ و الحاق «دیپ‌فیک‌های جنسی فاقد رضایتِ قابل‌انتساب» به مصادیق «پرخطر»، با ارجاع به معیارهای عینی فوق؛ (۲) ترجمان رهنمودهای قانون خدمات دیجیتال به «استانداردهای اجرایی مشترک» نزد سکوها: برچسب‌گذاری پایدار که با بازنشر سفر کند، مخازن شواهد و دسترسی پژوهشی برای ارزیابی ریسک‌های نظام‌مند (European Commission, 2024: cf.AP News, 2024)؛ (۳) اتکای رویه به قاعده «نسخه‌های معادل» در «گلاویشنیک–پیشچک علیه فیس‌بوک ایرلند»، برای صدور اوامر فرامرزی حذف و هماهنگ‌سازی با تکالیف فصل شفافیت قانون هوش مصنوعی. منتقدانی که نگران «شکاف اجرا» هستند درست می‌گویند، اما همین قواعد اگر با ابزارهای اثبات منشأ و معیارهای ارتقای ریسک پیوند بخورند، می‌توانند از یک «قاعده اعلامی» به «ابزار بازدارنده مؤثر» دگردیسی یابند (AlgorithmWatch, 2025; Block, 2024, pp. 191–192)

از منظر تحلیلی، چارچوب اتحادیه اروپا در برخورد با دیپ‌فیک یک «منطق حداقلیِ اولیه» دارد: ابتدا با الزام به شفافیت، مرز میان محتوای اصیل و دست‌کاری‌شده را برای مخاطب روشن می‌کند و سپس بارِ مهارِ پیامدها را بر دوشِ سازوکارهای اجراییِ مکمل، به‌ویژه تکالیف سکوها و نظارت ریسک‌محور می‌گذارد. قوت این رویکرد در انعطاف‌پذیری آن است؛ یعنی در کاربری‌های مشروع، دست قانون را برای مداخله سنگین نمی‌بندد. اما ضعفِ کلیدی‌اش در پرونده‌های دیپ‌فیکِ جنسیِ فاقد رضایت آشکار می‌شود: چون کنشگرِ بدخواه اصولاً به برچسب‌گذاری تن نمی‌دهد و حتی اگر نشانه‌ای هم باشد، در چرخه‌های بازنشر و تغییر قالب می‌تواند بی‌اثر شود؛ بنابراین «اطلاع‌رسانی» به‌تنهایی بازدارنده نیست و برای حمایت واقعی از بزه‌دیده کافی نیست. بر همین مبنا، منطق تحلیل‌پذیرِ اصلاح این است که محور تنظیم‌گری از «اعلامِ مصنوعی‌بودن» به «مهارِ قابلیت

انتساب و انتشار» منتقل شود؛ یعنی در مواردی که محتوا قابل‌انتساب به شخص مشخص است و رضایت معتبر وجود ندارد، سطح الزام باید بالاتر برود و تکلیف سکو به حذفِ مؤثر، جلوگیری از بازنشر نسخه‌های مشابه، و فراهم‌کردن امکان ردیابی منشأ به‌صورت نتیجه‌محور طراحی شود. این جابه‌جایی، هم شکاف میان هنجار و واقعیتِ اجرا را کم می‌کند و هم بدون گسترش بی‌ضابطه جرم‌انگاری، تمرکز را روی پرخطرترین مصادیق نگه می‌دارد.

۳.۲. ضمانت‌های اجرایی و الزامات شفافیت

نقطه ثقل اجرای مقررات اتحادیه اروپا در حوزه دیپ‌فیک‌های پورنوگرافیک، سه لایه مکمل است: (۱) اختیارات تنبیهی، نظارتی و جریمه‌های سنگین، (۲) الزامات شفافیت و برچسب‌گذاری منشأ مصنوعی محتوا، و (۳) تبدیل خروجی‌های فنی تشخیص به «ادله قابل استناد» در فرایند دادرسی. مطابق قانون خدمات دیجیتال، مقام‌های ملی «هماهنگ‌کننده خدمات دیجیتال» و کمیسیون اروپا (برای سکوی بسیار بزرگ) می‌توانند از ابزارهای تحقیق، دستور توقف، و جریمه استفاده کنند؛ سقف جریمه‌های عدم‌تبعیت از تکالیف قانونی تا ۶ درصد گردش مالی جهانی سال پیش سکوی متخلف پیش‌بینی شده است (قانون خدمات دیجیتال، بند۳، ماده ۵۲) به موازات، قانون هوش مصنوعی، برای نقض‌های شفافیتی و تعهدات خاصاز جمله تعهدات مرتبط با محتوای دست‌کاری‌شده، جریمه‌های پلکانی مقرر کرده و برای نقض‌های جدی سقف ۳۵ میلیون یورو یا ۷ درصد گردش مالی جهانی را مجاز دانسته است(قانون هوش مصنوعی، بند۳، ماده۵۰). این طراحی بازدارنده، تنها هنگامی کارآمد است که همراه با سنجه‌های عملی نظارت و قابلیت اتکای ادله در محکمه پیش برود.

در وجه التزام‌های شفافیت، ماده ۵۰ «قانون هوش مصنوعی» افشای منشأ مصنوعی را برای محتوای تولید و دست‌کاری‌شده الزامی می‌داند، اعم از اعلان متنی، فراداده یا نشانه‌گذاری در خود فایل به‌نحوی‌که بیننده در نخستین مواجهه، مصنوعی بودن را به‌روشنی درک کند (قانون هوش مصنوعی، بند ۴، ماده۵۰). این الزام در مورد دیپ‌فیک‌های پورنوگرافیک اهمیت مضاعف دارد، زیرا کارکرد فریب‌کارانه (ایجاد نسبت جنسی کاذب به شخص قابل‌شناسایی) را مستقیماً هدف می‌گیرد. از حیث اجرا، دفتر اروپایی هوش مصنوعی به‌عنوان بازوی هماهنگی وراهنمایی، تدوین «آیین‌های عملیاتی» و ترویج روش‌های واحد برچسب‌گذاری را پیگیری می‌کند تا تشتت رویه در بین سکوها به حداقل برسد .(European AI Office, mandate) با این حال، تجربه میدانی نشان می‌دهد صرف پیش‌بینی تکلیف کافی نیست؛ باید کیفیت برچسب‌گذاری و «ماندگاری نشانه‌ها» در برابر تبدیل‌فرمت

و بازنشرهای زنجیره‌ای هم سنجش‌پذیر باشد. گزارش «آژانس امنیت سایبری اتحادیه اروپا» درباره تهدیدات هوش مصنوعی تصریح می‌کند که اتکاء تک‌بعدی به یک نشانه (مثلاً نشان دیجیتال نرم‌افزاری) در برابر تغییر رمزگذاری و بازانتشار، شکننده است و استفاده «چندنشانه‌ای» (ترکیب فراداده مقاوم، نشانگر درون‌تصویر، و ثبت زنجیره تولید) توصیه می‌شود.(ENISA, 2025: 15)

لایه دوم اجرا به «قابلیت اتکای ادله فنی» مربوط است: برچسب یا نشانه‌گذاری وقتی در دادگاه کار می‌کند که با گزارش کارشناسی استاندارد و آستانه تصمیم‌گیری روشن همراه شود؛ یعنی گزارش فنی به‌صراحت روش، حدود خطا و ترکیب نشانه‌ها را بیان کند تا دادرس بتواند نسبت میان خطای ابزار و بار اثبات را ارزیابی کند. در نبود استاندارد گزارش‌دهی، حتی محتوای دارای برچسب درست ممکن است به «دلیل کافی» بدل نشود. زیرا دادگاه باید بتواند مبنای فنیِ برچسب را ارزیابی کند: این‌که برچسب بر چه روش و داده‌ای تکیه دارد، میزان خطای محتمل آن چقدر است، و نتیجه با چه درجه‌ای از اطمینان بیان می‌شود. اگر این اطلاعات به‌صورت یکنواخت و قابل‌مقایسه ارائه نشود، برچسب عملاً به یک «ادعای فنیِ تأییدنشده» تنزل می‌یابد و در تعارض با اصل ارزیابی آزاد ادله و معیار اطمینان قضایی، کفایت لازم برای احراز انتساب و عنصر مادی را پیدا نمی‌کند؛ به‌ویژه وقتی طرف مقابل به خطاهای ابزار، دست‌کاری فراداده یا امکان جعلِ برچسب استناد کند. بنابراین، استانداردسازی گزارش کارشناسی، نقش «پل» میان تشخیص فنی و اقناع قضایی را دارد و مانع از آن می‌شود که نتیجه فنی، به‌دلیل ابهام در روش و حدود خطا، از درجه اعتبار ادله‌ای ساقط شود. اسناد راهنمای انتخاباتی کمیسیون اروپا نیز بر پیوند میان شفافیت محتوایی و قابلیت راستی‌آزمایی تأکید کرده و سکوی بسیار بزرگ را ملزم می‌کند که ریسک‌های نظام‌مند تحریف محتوایی را با شاخص‌های قابل ممیزی (ازجمله ردیابی منشأ محتوا و همکاری با «بازبینان مستقل») مدیریت کند (EC Guidelines on Electoral Processes, 2024: 8, 10). این منطق به‌روشنی در پرونده «گلاویشنیک–پیشچک علیه فیس‌بوک ایرلند»، نیز دیده می‌شود؛ دیوان عدالت اتحادیه اروپا سکو را ملزم دانست نه فقط مورد گزارش‌شده، بلکه «نسخه‌های اساساً معادل» را هم حذف کند، تا چرخه بازنشر هدفمند بی‌اثر شود؛ نکته‌ای که در مورد بازتولید پیاپی دیپ‌فیک‌های پورنوگرافیک تعیین‌کننده است.(CJEU, 2019: paras. 45–49)

در وجه «اعمال و همکاری»، داده‌های میدانی نشان می‌دهد که ظرفیت پیگیری برون‌مرزی و دسترسی به ادله دیجیتال، بر موفقیت اجرای مقررات اثر مستقیم دارد. گزارش مشترک یوروپل–یوروجاست از رشد ۲۲درصدی درخواست‌های ادله الکترونیک و نرخ پاسخ‌گویی حدود ۷۴ درصد

خبر می‌دهد؛ اما زمان‌بر بودن پاسخ‌ها و ناهمگونی قالب‌های داده، اثربخشی اقدامات فوری (مثل حذف محتوا) را محدود می‌کند .(Europol & Eurojust, 2024: 12–13) بر این مبنا، اگر دستور قضایی حذف سریع در یک کشور صادر شود، شکست همکاری به‌موقع ارائه‌دهنده خارجی می‌تواند کل سازوکار بازدارندگی را فرسوده کند؛ لذا هم‌ترازی سازوکار «دستورهای حذف» در قانون خدمات دیجیتال با تعهدات شفافیت در قانون هوش مصنوعی (برچسب‌گذاری، ثبت منشأ، نگهداری فراداده) برای قابلیت اجرای برون‌مرزی کلیدی است؛ قانون خدمات دیجیتال برای تحکیم این هم‌ترازی، علاوه بر اختیارات گسترده تحقیق و دستور توقف (قانون خدمات دیجیتال، بند ۲، ماده ۵۲)، جریمه‌های سنگین و جریمه‌های روزانه مقرر کرده است (قانون خدمات دیجیتال، بند ۳، ماده ۵۲).

نهایتاً، کارآمدی ضمانت‌های اجرایی در گرو «قابلیت تبدیل خروجی فنی به فرایندپذیرش قضایی» است؛ قاضی باید بداند یک محتوای برچسب خورده دقیقاً بر چه مبنای فنی و با چه حدود خطایی «جعلی» تشخیص‌داده‌شده است؛ سکوی میزبان باید فراداده بارگذاری، نشانی‌های اینترنتی، و لاگ‌های تغییر را در مهلت معین حفظ و ارائه کند؛ و همکاری برون‌مرزی باید شاخص‌های زمان‌مند داشته باشد (مهلت پاسخ، قالب مشترک، نقطه تماس دائمی). به همین دلیل، تکمیل ابزارهای تنبیهی قانون هوش مصنوعی اتحادیه اروپا (جریمه‌های ماده ۹۹) و اختیارات قانون خدمات دیجیتال (مواد ۵۱ و ۵۲) با سه اقدام تکمیلی توصیه می‌شود: (الف) تدوین استاندارد در سطح ملّی و اتحادیه‌ای برای «گزارش کارشناسی محتوای مصنوعی» شامل روش، آستانه تصمیم و حدود خطا؛ (ب) تعهد صریح سکوها به نگهداری و تحویل ادله فنی مربوط به منشأ و انتشار در مهلت‌های کوتاه؛ (ج) پیش‌بینی سازوکار «حذف نسخه‌های اساساً معادل» به‌صورت پیش‌فرض قابل‌اجرا در دستورها، با نظارت کمیسیون برای سکوهای بسیار بزرگ، همسو با منطق همسو با منطق رأی دیوان عدالت اتحادیه اروپا در پرونده شماره C-18/18. (دیوان عدالت اتحادیه اروپا، ۲۰۱۹، پرونده C-18/18، بندهای ۴۵–۴۹؛ قانون هوش مصنوعی اتحادیه اروپا، ماده ۵۰؛ قانون خدمات دیجیتال، مواد ۵۱–۵۲).

از منظر تحلیلی، ضمانت‌های اجرایی اتحادیه اروپا در این حوزه بر یک ایده ساده اما دقیق بنا شده است: «هزینه تخلف باید واقعی و قابل احساس باشد، و اثبات تخلف هم باید در دادگاه امکان‌پذیر و پایدار بماند». به همین دلیل، اتحادیه به یک ابزار بسنده نکرده و سه رکن را هم‌زمان پیش می‌برد؛ نخست، اهرم‌های نظارتی و جریمه‌های بازدارنده برای واداشتن سکو و ارائه دهنده به تبعیت، دوم، شفافیت و برچسب‌گذاری برای کاهش ظرفیت فریب، و سوم، تبدیل تشخیص فنی به

دلیل قابل اتکا، تا دعوا در مرحله دادرسی به بن‌بست «نمی‌شود ثابت کرد» نرسد. ارزش این طراحی در پیوند دادن «نظم تنظیم گری» با «قابلیت اثبات» است، یعنی همان نقطه‌ای که در پرونده‌های دیپ فیک جنسی تعیین کننده می‌شود.

با این حال، کارآمدی این بسته دقیقاً به ضعیف‌ترین حلقه آن گره می‌خورد. اگر برچسب‌گذاری پایدار نباشد، شفافیت به یک توصیه اخلاقی تقلیل پیدا می‌کند. اگر گزارش کارشناسی قالب و معیار مشترک نداشته باشد، خروجی فنی در دادگاه به ادعای غیر قابل ارزیابی تبدیل می‌شود. و اگر همکاری برون مرزی و تحویل داده زمان مند و استاندارد نباشد، پنجره کوتاه حفظ ادله از دست می‌رود و حتی شدیدترین جریمه‌ها هم اثر پیشگیرانه کافی نخواهند داشت. بنابراین، مزیت واقعی رویکرد اتحادیه زمانی بالفعل می‌شود که تکالیف سکو، نگهداری داده، زمان پاسخ، و معیارهای گزارش کارشناسی به شاخص‌های روشن و قابل سنجش تبدیل شوند، تا هم نظارت معنی پیدا کند و هم تصمیم قضایی بر پایه دلیل پایدار و قابل دفاع شکل بگیرد.

۳.۳. آسیب‌شناسی ضعف‌های اجرایی و چالش‌های قضائی

۳.۳.۱. شکاف ظرفیت فنی و استاندارد ادله دیجیتال

گزارش مشترک «سیریوس» نشان می‌دهد مقامات قضایی کشورهای عضو در جمع‌آوری و اعتبارسنجی ادله برخط به خدمات ارائه‌دهندگان متکی‌اند و فقدان رویه یکنواخت، زمان واکنش را طولانی و کیفیت ادله را ناهمگون می‌کند .(Europol & Eurojust, 2024: 53-61) در پرونده‌های پورنوگرافی دیپ‌فیک، همین ناهمگونی در گزارش کارشناسی (حدود خطا، آستانه تصمیم، ترکیب نشانه‌ها) مستقیماً بر قابلیت استناد و نسبت بار اثبات اثر می‌گذارد.

پیشنهاد عملی در این خصوص، استانداردسازیِ قالب و حداقل‌محتوای «گزارش کارشناسی تشخیص محتوای مصنوعی و دست‌کاری‌شده» در سطح اتحادیه اروپا است؛ یعنی تهیه یک الگوی مشترک برای گزارش‌هایی که در فرایند کیفری و مدنی به‌عنوان دلیل فنی ارائه می‌شوند و باید به‌طور شفاف روش به‌کاررفته، آستانه تصمیم، حدود خطا و عدم‌قطعیت، و نتایج چندرسانه‌ای (چهره، صوت و حرکت) را توضیح دهند. این الگوی مشترک می‌تواند در قالب دستورالعمل‌ها و رویه‌های اجراییِ مرتبط با قانون هوش مصنوعی و سازوکارهای هماهنگی قانون خدمات دیجیتال مستقر شود و هم‌زمان، دستورالعمل‌های واحد برای دادستانی‌ها و پلیس قضایی تصویب گردد تا پیوند میان کشف فنی و ارزیابی قضایی تقویت شود و از گسست میان «یافته فنی» و «اقناع دادگاه» جلوگیری شود.(Europol & Eurojust, 2024: 53–61)

۳.۳.۲. اجرای نابرابر الزامات شفافیت و ظرفیت نظارت بر سکوها

الزامات افشای منشأ مصنوعی محتوا در ماده ۵۰ قانون هوش مصنوعی، و تکالیف ارزیابی و کاهش ریسک نظام‌مند برای سکوهای بسیار بزرگ در قانون خدمات دیجیتال، در عمل به‌صورت نامتوازن اجرا می‌شوند؛ کمیسیون در «راهنمای کاهش ریسک‌های انتخاباتی» تصریح می‌کند که محتوای مولد هوش مصنوعی (از جمله دیپ‌فیک) باید در طرح‌های کاهش ریسک، آشکارا دیده و برچسب‌گذاری شود، اما کیفیت و سرعت اجرای این تمهیدات متفاوت است (European Commission, 2024: 1, 16–18). پیشنهاد عملیدر این خصوص، تبدیل الزامات افشا به «تعهد نتیجه‌محور» برای سکوها (نمایان‌سازی منشأ مصنوعی در لحظه نخست مواجهه کاربر، با نشان دیجیتال و فراداده غیرقابل زدایش) و پیش‌بینی بازه‌های زمانی کوتاه و قابل‌سنجش برای حذف محتوا و ارائه داده‌های ادله‌ای به مرجع رسیدگی است؛ همچنین هماهنگی این الزامات با اختیارات مواد ۵۱ و ۵۲ قانون خدمات دیجیتال در سطح اتحادیه به‌صورت دستورهای نمونه به سکوی خیلی بزرگ پیشنهاد می‌شود .(European Commission, 2024, pp. 1, 16–18) .

۳.۳.۳. ادله برون‌مرزی و «تاخیر بحرانی» در حفظ و تحویل داده

نقطه انسداد رایج، تأخیر در حفظ و تحویل داده‌های حیاتی است. «مقرره ادله الکترونیکی» (Reg. 2023/1543) امکان صدور «دستور تولید» و «دستور نگه‌داری» مستقیم به ارائه‌دهنده مستقر در عضو دیگر را پیش‌بینی کرده است، اما دوره‌ی گذار تا اجرایی‌شدن کامل و هماهنگی با قوانین اشخاص ثالث، خلأ زمانی معناداری ایجاد می‌کند(Reg. 2023/1543: arts. 10–12)؛ کاربرد این شکاف در جرایم با محتوای زودزوال مثل پورنوگرافی دیپ‌فیک، به ازدست‌رفتن ادله می‌انجامد. پیشنهاد عملی در این خصوص تا زمان اجرای کامل رژیم جدید، استفاده نظام‌مند از «دستورهای موقت حفظ داده» با ضرب‌الاجل‌های ساعتی است، همچنین تعریف نقطه‌ی تماس مشترک برای سکوهای فراملی و استفاده از «کانال تماس اضطراری» در چارچوب قانون خدمات دیجیتال به‌موازات درخواست قضایی پیشنهاد می‌گردد تا فاصله‌ی زمانی صدور تا اجرا به حداقل برسد(مقرره اتحادیه اروپا، ۱۵۴۳/۲۰۲۳، OJ L 191، ۲۸ ژوئیه ۲۰۲۳؛ خلاصه اورلکس، ۲۰۲۳).

۳.۳.۴. بازتولید «نسخه‌های اساساً معادل» و امتداد اوامر حذف

دیوان عدالت اتحادیه اروپا در پرونده «پرونده گلاویشنیک–پیشچک علیه فیس‌بوک ایرلند (C-18/18)» تصریح کرد که سکوی میزبان می‌تواند ملزم به حذف نه فقط محتوای دقیقاً یکسان، بلکه «نسخه‌های اساساً معادل» نیز بشود؛ این منطق در دیپ‌فیکِ پورنوگرافیک حیاتی است، زیرا بازنشرهای جزئی تغییریافته بلافاصله جایگزین نسخه حذف‌شده می‌شوند (CJEU, 2019: paras. 45–49). پیشنهاد عملی دراین‌خصوص «الزام حذف نسخه‌های معادل» در دستورهای قضایی و اداری، همراه با فهرست شاخص‌های معادل‌سنجی (هم‌نشانه‌ای چهره، صوت و زبان بدن به همراه زمینهٔ انتشار) و تکلیف به جست‌وجوی فعال؛ و اعمال سازوکار نظارت کمیسیون برای سکوهای بسیار بزرگ تا از اجرای یکنواخت اطمینان حاصل شود (دیوان عدالت اتحادیه اروپا، ۲۰۱۹: بندهای ۴۵–۴۹؛ قانون خدمات دیجیتال، ۲۰۲۲: مواد ۵۱–۵۲).

۳.۳.۵. شواهد وقوع، شدت زیان، و پیوند با توان نظارت

جمع‌بندی‌های تحلیلی اخیر نشان می‌دهد سهم غالب دیپ‌فیک‌های منتشرشده، محتوای جنسی بدون رضایت است و پیامدهای آن بر بزه‌دیدگان از آسیب‌های روانی تا زیان‌های اقتصادی و گسست‌های اجتماعی و نیز بر هزینه‌های حکمرانی محسوس است (خدمات پژوهشی پارلمان اروپا،

۲۰۲۵: ۱؛ آژانس امنیت سایبری اتحادیه اروپا، ۲۰۲۵: ۳–۴). پیشنهاد عملی دراین‌خصوص، پیونددادن سیاست افشای منشأ محتوا در قانون هوش مصنوعی اتحادیه اروپا با سازوکارهای گزارش‌دهی پسارویداد در قانون خدمات دیجیتال و بازبینی پس از انتخابات رویدادهای پرریسک است، همچنین همراه با الزام به ارائه سنجه‌های کارایی، از جمله نرخ کشف، نرخ حذف و نرخ بازنشر نسخه‌های اساساً معادل و فراهم‌کردن دسترسی پژوهشگران مستقل به داده‌های تجمیعی سکوها برای ارزیابی اثربخشی پیشنهاد می‌شود.

از منظر تحلیلی، دشواری اصلی اتحادیه اروپا در مهار پورنوگرافیِ دیپ‌فیک «کمبود قاعده» نیست، بلکه ناهمزمانیِ ظرفیت فنی، یکنواختیِ معیارهای ادله، و سرعتِ دسترسی به داده است. وقتی کیفیت گزارش کارشناسی از کشوری به کشور دیگر متفاوت باشد، یک پرونده واحد ممکن است در جایی «قابل اثبات» و در جایی دیگر «مبهم» تلقی شود و همین، بار اثبات و پیش‌بینی‌پذیری قضایی را مخدوش می‌کند. از سوی دیگر، شفافیت و برچسب‌گذاری تا زمانی که پایدار، ممیزی‌پذیر و نتیجه‌محور نشود، در برابر بازنشرهای زنجیره‌ای و دستکاری‌های جزئی کارایی کافی ندارد. گلوگاه سوم نیز «تاخیر در حفظ و تحویل داده» است؛ در جرایم زودزوال، چند ساعت تعلل می‌تواند کل زنجیره انتساب را از بین ببرد. نهایتاً، حتی اگر حذف انجام شود، بدون پوشش «نسخه‌های اساساً معادل» چرخه بازتولید سریع متوقف نمی‌شود. بنابراین، نقطه اتکای سیاستِ اجرایی باید بر یک بسته هم‌پیوند قرار گیرد :چارچوب واحد گزارش کارشناسی، تعهد زمانیِ حذف و تحویل داده، سازوکارهای فوری حفظ داده، و شاخص‌های سنجش کارایی)نرخ حذف، نرخ بازنشر، و کیفیت همکاری سکوها) تا نظارت و قضا بتوانند بر مبنای معیارهای قابل اندازه‌گیری تصمیم بگیرند.

۴. تحلیل تطبیقی رویکرد ایران و اتحادیه اروپا در مواجهه با پورنوگرافی دیپ‌فیک

در جمع‌بندی تطبیقیِ این پژوهش، رویکرد اتحادیه اروپا را باید نه یک «قانون واحد»، بلکه یک چیدمان چندابزاری دانست که هر کدام بخشی از چرخه‌ی تولید تا اشاعه‌ی پورنوگرافی دیپ‌فیک را پوشش می‌دهند: «قانون هوش مصنوعی» بیشتر ناظر به منطق تولید و دست‌کاری و الزامات شفاف‌سازی است، «قانون خدمات دیجیتال» بر رفتار سکوها در میزبانی، گزارش‌پذیری، حذف و مدیریت ریسک تمرکز دارد، و «مقررات عمومی حفاظت از داده‌ها» با صورت‌بندی حق‌محور (به‌ویژه حق پاک‌سازی/حق فراموش‌شدن) به قربانی امکان می‌دهد کنترل از دست‌رفته بر تصویر و داده‌ی شخصی را بازپس گیرد. مزیتِ این ترکیب آن است که پاسخ اتحادیه اروپا صرفاً کیفری یا صرفاً اخلاقی نیست؛ بلکه هم‌زمان دو هدف را دنبال می‌کند: از یک سو «توازن» میان آزادی بیان و حقوق

فردی، و از سوی دیگر حفاظت از منافع عمومی و کرامت انسانی. در مقام اجرا نیز همین معماری باعث می‌شود حمایت از قربانی صرفاً به نتیجه‌ی نامطمئنِ تعقیب مرتکب موکول نماند، بلکه از مسیرهای سریع‌تر و کارکردی‌تری مثل برچسب‌گذاری محتوای دست‌کاری‌شده، ابزارهای گزارش‌دهی درون‌سکویی، سازوکار «اعلام و اقدام»، و نقش «پرچم‌داران مورد اعتماد» تقویت شود؛ یعنی حق و ابزار، هم‌زمان کنار هم قرار می‌گیرند تا آسیب، قبل از تثبیت و بازتولید، متوقف شود.

بااین‌حال، همین الگوی اروپایی در عمل بی‌چالش نیست و اتفاقاً نقطه‌های اصطکاک آن برای تحلیل تطبیقی مهم است. پورنوگرافی دیپ‌فیک غالباً با سرعت بالا بازتولید می‌شود و عمدتاً در سکوهای آنلاین گردش می‌کند؛ بنابراین اتحادیه اروپا ناگزیر است «مسئولیت‌پذیری پلتفرم‌ها» را به‌صورت منظم و با ضمانت‌اجراهای سنگین پی بگیرد. این منطق، در طراحی سازوکارهای نظارتی و جریمه‌های بازدارنده دیده می‌شود، اما در مقام اجرا با سه مانع تکرارشونده روبه‌روست: محدودیت‌های صلاحیت سرزمینی (به‌خصوص در تعامل با بازیگران فرامرزی)، شکاف ظرفیت فنی در تشخیص و انتساب دیپ‌فیک، و ناهمگونی استانداردهای فنی و رویه‌ای میان کشورهای عضو. از سوی دیگر، اتکای زیاد به مدل «گزارش و اقدام» اگرچه مزیتِ واکنش‌پذیری سریع را دارد، اما می‌تواند به تأخیر در حذف محتوای غیرقانونی بینجامد؛ تأخیری که در پرونده‌های حیثیتی، عملاً معادل استمرار آسیب است. به همین دلیل، در دلِ تجربه اروپایی یک درس کلیدی وجود دارد: صرفِ پیش‌بینی تکلیف برای سکو کافی نیست؛ کیفیت رسیدگی باید ارتقا یابد، از تعیین بازه زمانی «اقدام به‌موقع» گرفته تا مستندسازی تصمیم پلتفرم و پیش‌بینی امکان اعتراض مؤثر برای قربانی، وگرنه همان سازوکار حمایتی به نقطه ضعف تبدیل می‌شود.

در سوی دیگر، مواجهه ایران با پورنوگرافی دیپ‌فیک، عمدتاً از مسیر قواعد کلاسیک‌ترِ مقابله با محتوای «مستهجن» و «منافی عفت عمومی» صورت می‌گیرد. مواد ۱۴ و ۱۵ قانون جرائم رایانه‌ای، با جرم‌انگاری تولید و انتشار محتوای مستهجن و نیز جرم‌انگاری تحریک یا ترغیب و تسهیل دسترسی، ظرفیت پوشش محتوای دست‌کاری‌شده دیجیتال، از جمله دیپ‌فیکرا در خود دارد و اتکای مفهومی بر «مستهجن/مبتذل» عملاً امکان واکنش کیفری را فراهم می‌کند. مزیت این رویکرد آن است که از منظر نظم عمومی و صیانت از ارزش‌های اخلاقی، «حساسیت هنجاری» بالایی دارد و نسبت به آثار اجتماعی این محتوا بی‌تفاوت نیست. اما همان‌جا که نظام اروپایی تلاش کرده شکاف

میان «متن» و «اجرا» را با ابزارهای دقیق پر کند، در ایران چالش‌ها پررنگ‌تر می‌شود: اتکای «مستهجن» به عرف، احتمال ناهمگونی برداشت‌های قضایی را بالا می‌برد؛ نبود استانداردهای روشن برای شناسایی و انتساب دیپ‌فیک و ضعف سازوکارهای اجرایی برای حذف سریع، باعث می‌شود حتی وقتی عنوان مجرمانه قابل اعمال است، مسیر توقف فوریِ آسیب و تثبیت ادله دشوار بماند؛ و از همه مهم‌تر، در سطح پلتفرم‌ها سازوکارهای الزام‌آورِ «اعلام و اقدام» و دسترسی پایدار به ادله دیجیتال هنوز به شکل نهادی و استاندارد جا نیفتاده است. نتیجه عملی این وضعیت آن است که بار اصلی پیگیری هزینه، زمان، و ریسک فرسایش ادله، عملاً روی دوش بزه‌دیده و فرایند کیفری می‌افتد؛ در حالی که ماهیت شبکه‌ایِ این جرم، «واکنش نقطه‌ای» را از ابتدا فرسایشی می‌کند.

در سطح سیاست‌گذاری، مقایسه دو نظام نشان می‌دهد اختلاف اصلی نه در «تعداد مواد قانونی» است و نه صرفاً در «شدتِ جرم‌انگاری»، بلکه در مسیرِ حل مسئله و نقطه تمرکزِ پاسخ حقوقی شکل می‌گیرد. در اتحادیه اروپا، کنار امکان تعقیب مرتکب، مرکز ثقلِ مداخله بر «مهار فوری آسیب» قرار دارد؛ یعنی حذف سریع محتوا، پاسخ‌گویی سکو به‌عنوان بازیگر کلیدیِ چرخه انتشار، و فراهم‌کردن ابزارهای حق‌محور برای بزه‌دیده تا بتواند توقف انتشار، محدودسازی دسترسی و پیگیری قصور پلتفرم را مطالبه کند. در مقابل، در ایران حرکت غالب بیشتر با محور «ممنوعیت محتوای مستهجن» و «تعقیب و مجازات مرتکب» پیش می‌رود؛ بنابراین پاسخ، ماهیتاً کیفری‌تر و پسینی‌تر می‌یابد و نقش سکو و سازوکارهای اقدام فوری، معمولاً در حاشیه و به‌صورت غیرمستقیم دیده می‌شود. همین تفاوتِ رویکرد، اولویت‌های اجرایی را تغییر می‌دهد: مجازات در برابر حذف فوری، واکنش پس از وقوع در برابر مهار چرخه بازنشر، و تمرکز بر مرتکب در برابر ورود فعال پلتفرم به‌مثابه یک مخاطبِ تکلیف.

نکته مهم آن است که این اختلاف مبنا، صرفاً یک بحث نظری میان «نظم‌محوری» و «حق‌محوری» نیست، بلکه ترجمه‌ای کاملاً عملی در پرونده‌های دیپ‌فیک دارد. نظامی که مسئله را نظم‌محور می‌بیند، غالباً مسیر را از کانال تعقیب کیفری و اثبات انتساب طی می‌کند؛ در نتیجه، حساس‌ترین نقاط یعنی سرعتِ توقف آسیب، مهلتِ حفظ داده‌ها، و جلوگیری از بازنشر نسخه‌های هم‌ارز، ممکن است در عمل به کُندی یا ناهمسانی دچار شود. در مقابل، نظامی که حق‌محور طراحی شده، از ابتدا می‌کوشد «کنترلِ قربانی بر تصویر و داده» را بازیابی کند؛ بنابراین حفظ سریع ادله دیجیتال، تکلیف روشن برای پلتفرم در حذف و پیشگیری از بازنشر، و امکان مطالبه‌گری مؤثر بزه‌دیده به اجزای

اصلی پاسخ تبدیل می‌شوند. بر همین اساس، اختلاف مبنایی تعیین می‌کند ادله چگونه و با چه استانداردی نگه‌داری می‌شود، مسئولیت سکو تا کجا امتداد می‌یابد، کدام نهاد نقش پیشران دارد (دادستانی و ضابطان در برابر نهادهای تخصصی نظارت و داده)، و نهایتاً ترمیم حیثیت با چه ابزارهایی و با چه سرعتی قابل تحقق است.

بر پایه همین منطق، دلالت‌های اجراییِ قابل استفاده برای سیاست‌گذاری داخلی بدون آن‌که به مبانی فرهنگی و اجتماعی ایران بی‌اعتنا باشد، بیش از آن‌که «تکثیر عنوان مجرمانه» باشد، به «قابل اجرا کردن قواعد موجود» مربوط می‌شود. تجربه اروپایی به‌صورت ضمنی نشان می‌دهد که سه حلقه اگر هم‌زمان تقویت شوند، مسیر حمایت از قربانی کوتاه و نتیجه‌مندتر می‌شود: نخست، جااندازی سازوکارهای الزام‌آورِ اعلام و اقدام در سطح سکوها با ضرب‌الاجل‌های روشن برای محتوای پرخطر، به‌همراه ایجاد یک نقطه تماس واحد و پاسخ‌گو برای مکاتبات قضایی و انتظامی؛ دوم، تضمین دسترسی به ادله دیجیتال از طریق پروتکل‌های نگهداشت، استانداردسازی زنجیره نگهداری و امکان صدور دستورهای فوری و الکترونیکیِ ساری‌الاثر برای جلوگیری از زوال داده؛ و سوم، سرمایه‌گذاری هدفمند بر ظرفیت فنی و تخصصی‌سازی رسیدگی (واحدهای تخصصی در پلیس فتا و دادسراها، آموزش قضات و کارشناسان، و ابزارهای تشخیص و انتساب). در کنار این سه محور، اندیشه‌ی «حق پاک‌سازی» نیز می‌تواند به‌عنوان یک ابزار مکمل دیده شود؛ نه به صورت نسخه‌برداری خام، بلکه در قالبی بومی‌شده که حدود آن را با موازنه میان حریم خصوصی، حیثیت، و ملاحظات اجتماعی و نیز با سازوکار دادرسی منصفانه روشن کند تا هم قابلیت اجرا داشته باشد و هم محل مناقشه نشود.

در نهایت پورنوگرافی دیپ‌فیک یک جرم «شبکه‌ای و زودزوال» است؛ بنابراین کارآمدیِ پاسخ حقوقی زمانی بالا می‌رود که از «واکنش صرفاً پرونده‌محور» عبور کنیم و هم‌زمان سه خروجی را تضمین کنیم: توقف سریع آسیب، حفظ پایدار ادله، و تعیین بازیگران پاسخ‌گو (از مرتکب تا سکو). اتحادیه اروپا با ترکیب ابزارهای پیشگیرانه و اصلاحی و با تقویت نقش پلتفرم‌ها در حذف و پیشگیری از بازنشر، یک تصویر نسبتاً عملی از این مسیر ارائه کرده است؛ ایران نیز اگر بخواهد با حفظ حساسیت‌های فرهنگی درباره حیثیت و حریم خصوصی، همان میزان کارآمدی را به دست آورد، بیش از هر چیز به «استانداردهای اجرایی»، «سازوکارهای فوری»، و «زیرساخت نهادی و فنی» نیاز دارد.

۵. چالش‌های عملی اجرای مقررات در زمینه پورنوگرافی دیپ‌فیک (هرزه‌نگاری جعل عمیق)

با آنکه پدیدهٔ پورنوگرافی دیپ‌فیک به‌تازگی در دستور کار حقوق فناوری قرار گرفته، هنوز واکاوی نظری و تجربی بسنده‌ای دربارهٔ بُعد اجرایی مقابله با آن صورت نگرفته است. شکاف میان «قاعدهٔ مکتوب» و «اثر عملی» عمدتاً از چهار چالش ناشی می‌شود: دشواری‌های شناسایی فنی، کاستی تجهیزات و مهارت‌های دادرسی، ناهماهنگی همکاری‌های فرامرزی، و موانع استرداد مرتکبان جرائم سایبری. این بخش با رویکردی عمل‌محور، هر یک از این گره‌ها را به‌صورت منسجم می‌کاود و راه‌حل‌های قابل پیاده‌سازی برای ارتقای کارآمدی مقررات پیشنهاد می‌کند.

۵.۱. شناسایی فنی دیپ‌فیک

در شرایط واقعی مانند فشرده‌سازی ویدئو، بازنمونه‌برداری، نور و سایهٔ نامعمول و انتشار نسخه‌های «اساساً معادل»، اتکای صرف به یک نشانهٔ محدود (برای نمونه تنها ردّ دیجیتال تصویر) دقت را کاهش می‌دهد و وزن اثباتی ادله را کم می‌کند. راه مطمئن، گذار از «تشخیص تک معیار» به «تشخیص چندبُعدی» است: سنجش هم‌زمان ریز حرکت‌های چهره، هماهنگی گفتار و حرکت لب، سازگاری نور و سایه و نیز به‌کارگیری الگوهای مقاوم در برابر دست‌کاری، در کنار «منشأسنجی» و «اصالت‌سنجی» محتوا که از لحظهٔ گردآوری تا ارائه در دادگاه همراه پرونده بماند (Europol, 2024: 14–16). در حوزهٔ صوت نیز ارزیابی‌های تازه نشان می‌دهد دقت تشخیص جعل در کانال‌های واقعی ارتباطی افت می‌کند؛ بنابراین به‌روزرسانی پیوستهٔ سامانه، تنوع‌بخشی داده‌های آموزشی و انجام «آزمون تعمیم بر نسل‌های جدید جعل» باید جزء تکالیف کارشناسی باشد تا گزارش فنی، ارزش اثباتی کافی پیدا کند (Shi et al., 2025: 1–3). افزون بر این، گزارش کارشناسی باید همراه با حدود عدم قطعیت، آستانهٔ تصمیم و نتیجهٔ آزمون تعمیم ارائه شود تا مرجع رسیدگی بتواند وزن ادله را به‌درستی بسنجد (Europol, 2024: 21)

۵.۲. کمبود تجهیزات و مهارت‌های دادرسی

اگر توان آزمایشگاهی و مهارت کارشناسان با شتاب تحول جعل‌های ترکیبی همگام نباشد، حتی یافتهٔ علمی نیز به «ادلهٔ قابل‌استناد» تبدیل نمی‌شود. دستور کار تحلیلی پلیس بین‌الملل تصریح می‌کند که ارزیابی اصالت رسانهٔ مصنوعی، به «ظرفیت‌سازی تخصصی»، «آزمایشگاه جرم‌یابی رسانه‌ای» و «شیوه‌نامهٔ روشن برای گزارش کارشناسی» نیاز دارد تا حدود خطا و عدم قطعیت برای قاضی شفاف شود (INTERPOL, 2024: 6-14). هم‌زمان، گزارش وضعیت ادلهٔ الکترونیک بر ضرورت «مرکز تماس داخلی واحد» برای دریافت و یکدست‌سازی گزارش‌های فنی، «چک فهرست

پذیرش ادلهٔ الکترونیک»، و «زنجیرهٔ نگهداشت یکپارچه»، شامل زمان مهر، شناسهٔ یکتا، مسیر انتقال و کنترل دسترسی تأکید می‌کند (Europol & Eurojust, 2024:10–12). چیدمان یادشده، تشتّت رویه را کم و امکان‌پذیرش گزارش‌های مبتنی بر یادگیری ماشین را در مرجع رسیدگی بیشتر می‌کند (INTERPOL, 2024: 14–15; Europol & Eurojust, 2024: 65–67).

۵.۳. همکاری بین‌المللی و نقش پلیس بین‌الملل

پراکنده بودن حلقه‌های تولید، میزبانی، پرداخت و توزیع در چند کشور، ناهمگونی پاسخ ارائه‌دهندگان خدمت و زوال سریع داده‌ها، پیگیری را کند و پرریسک می‌کند (Europol & Eurojust, 2024: 53–61). ساختار کارآمد از سه جزء هم‌زمان شکل می‌گیرد: نخست، فعال‌سازی «مرکز تماس واحد شبانه‌روزی» برای تعامل با سکوها، همراه با فرم‌های استاندارد درخواست و «حفظ فوری داده»؛ دوم، تعیین جدول‌های زمانی مشترک پاسخ و تشکیل «گروه‌های مشترک تحقیق» در پرونده‌های حساس؛ سوم، استانداردسازی فراداده و قالب‌های گزارش تا از دوباره‌کاری و ردّ درخواست‌ها جلوگیری شود (Europol & Eurojust, 2024: 58–67; Eurojust & Europol, 2025: 17–18). تجربه‌های عملی عملیات فراملی نیز نشان می‌دهد هماهنگی پلیسی و قضایی و یکپارچه‌سازی مسیرهای ارتباطی، «پنجرهٔ زود زوال شواهد» را می‌بندد و نرخ موفقیت را بالا می‌برد (Associated Press, 2024).

۵.۴. استرداد مرتکبان جرائم سایبری

در مرحلهٔ استرداد، اختلاف در عنوان مجرمانه، شرط دوگانه بودن جرم و ناهمسانی آستانهٔ پذیرش ادلهٔ الکترونیک، روند را کند و نامطمئن می‌کند (Europol & Eurojust, 2024: 65–67). مرور چالش‌های مشترک نیز نشان می‌دهد بدون «هم‌ترازسازی عناصر مادی و روانی جرم تصاویر صمیمی مصنوعی» و بدون «پیوست فنی استاندارد» برای درخواست استرداد، ارزیابی کشور مقصد دشوار است (Eurojust & Europol, 2025: 17–18). پیوست فنی باید شامل شناسه و شناسهٔ یکتای فایل، زمان مهر، زنجیرهٔ نگهداشت و شرح روش‌های تشخیص با حدود خطا باشد؛ و پیش از ورود به مسیر رسمی، با «حفظ فوری داده» و همکاری داوطلبانهٔ سکوها، ادلهٔ زود زوال تثبیت شود. این چینش، هم شرط دوگانه بودن جرم را آسان‌تر احراز می‌کند و هم اعتبار ادله را نزد مرجع مقصد افزایش می‌دهد (Europol & Eurojust, 2024: 65–67; Eurojust & Europol, 2025: 17–18).

۶. چالش‌های فنی و محدودیت‌های فناوری دیپ‌فیک

کارآمدی مقررات وقتی معنا دارد که خروجی‌های فنی تشخیص دیپ‌فیک در «صحنه واقعی» به ادله

قابل‌استناد تبدیل شود و روند رسیدگی را جلو ببرد. افت دقت ابزارها در فایل‌های واقعی شبکه‌های اجتماعی، ناپایداری برچسب‌های منشأ در رفت‌وآمد بین چند سکو، و امکان بازتولید نسخه‌های «اساساً معادل» نه فقط مسئله‌ای فنی که مستقیماً بر اعتبار ادله، بار اثبات، سرعت حفظ داده‌های زود زوال و امکان انتساب اثر می‌گذارد؛ بنابراین، صورت‌بندی حقوقی مؤثر باید بر درکی روشن از این محدودیت‌ها تکیه کند تا تصمیم قضایی بر پایه ادله روشن، مستند و تبیین‌پذیر گرفته شود.

در عمل، فایل‌های دیپ‌فیک معمولاً در همان شکلی که «اولین‌بار» تولید شده‌اند به دادگاه نمی‌رسند. یک ویدئو بعد از چند بار ارسال و بارگذاری مجدد در سکوها، غالباً فشرده می‌شود، کیفیتش افت می‌کند و بخشی از نشانه‌های ریزِ دست‌کاری از بین می‌رود؛ به همین علت، دقت سامانه‌های تشخیص هم پایین می‌آید. نتیجه‌ی حقوقیِ این وضعیت روشن است :اتکا به یک نشانهٔ واحد (مثلاً صرفاً ردّ دیجیتالِ تصویر) برای اثبات کافی نیست. گزارش کارشناسی وقتی برای مرجع رسیدگی «قابل اتکا» می‌شود که تشخیص بر پایه‌ی ترکیب چند نشانهٔ مکمل (چهره، صدا و حرکت) انجام شود و در متن گزارش، آستانه‌ی تصمیم و حدود خطا/عدم‌قطعیت به‌طور صریح ذکر گردد تا دادگاه بتواند ارزش اثباتی نتیجه را ارزیابی کند(Europol, 2024: 14-16; INTERPOL, 2024: 14-15)

همین مسئله در جعل صوتی هم دیده می‌شود. پیام‌های صوتی دست‌کاری‌شده وقتی از کانال‌های واقعی ارتباطی عبور می‌کنند، نشانه‌های جعل می‌تواند کم‌رنگ شود. اگر ابزارهای تشخیص به‌روز نشوند و روی نسل‌های جدید جعل «آزمون تعمیم» انجام نگیرد، احتمال خطا افزایش می‌یابد؛ و این یعنی در مقام دادرسی، احراز جعلی‌بودن گفتار دشوارتر می‌شود و بار اثبات عملاً سنگین‌تر خواهد شد.(Shi et al., 2025: 339–347)

از سوی دیگر، «گواهی‌های منشأ» شفافیت مسیر تولید و ویرایش را افزایش می‌دهند، اما بنا به طراحی درباره «قانونی/غیرقانونی» بودن داوری نمی‌کنند و در زنجیره‌های انتشار چندمرحله‌ای ممکن است برچسب‌ها ناقص یا حذف شوند؛ بنابراین، برای احراز زنجیره نگهداشت و امکان انتساب به‌تنهایی کافی نیستند و باید در کنار آن‌ها سنجش فنی چندنشانه‌ای و گزارش کارشناسی روشن قرار گیرد .(C2PA, 2025: 11–12) همچنین، «نشان‌گذاری درونی محتوا» حدومرز فنی دارد: برش ساده تصویر، تغییر اندازه یا بارکدگذاری می‌تواند نشان را تضعیف یا پاک کند بی‌آنکه افت کیفیت به چشم بیاید؛ پس اگر در گزارش به نشان‌گذاری استناد می‌شود، لازم است شرایط ازدست‌رفتن نشان و پایداری آن دقیقاً توضیح داده شود، وگرنه اعتبار ادله در محکمه با تردید روبه‌رو می‌گردد (Francati et al., 2025: 1–2).

افزون بر این، در بازارهای فعال پورنوگرافی دیپ‌فیک، سازندگان با انتخاب کُدِک‌های خاص،

زنجیره‌های چندمرحله‌ای پردازش و ساخت «نسخه‌های اساساً معادل»، الگوهای تک‌مسیره تشخیص را دور می‌زنند. پیامد حقوقی روشن است: دستورات حذف باید نسخه‌های «اساساً معادل» را هم در بر بگیرد و گزارش کارشناسی باید معیارهای هم‌ارزی اثباتی را روشن کند تا تکلیف حذف و مسئولیت میزبان قابل‌اجرا باشد؛ در غیر این صورت، باز انتشار سریع، حفظ داده زود زوال و اجرای مؤثر حکم را مختل می‌کند(Europol, 2024: 14–16, 21)؛ Han et al., 2025: 9–12.

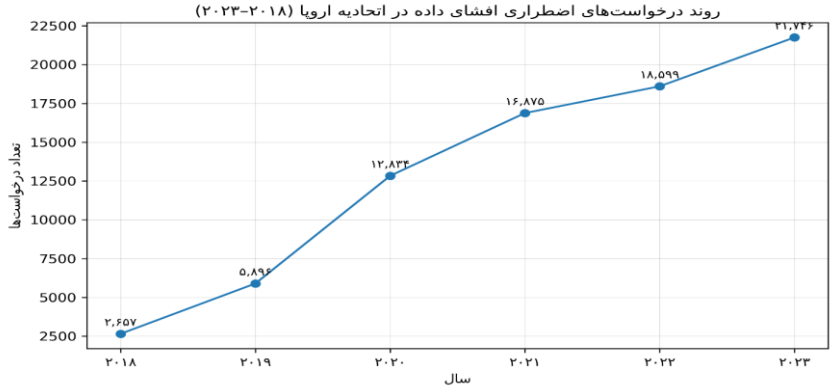
در سطح فرامرزی نیز، زوال سریع داده‌ها و ناهمگونی مسیرهای همکاری سبب می‌شود پنجره زمانی دسترسی به ادله از دست برود. راه عملی آن است که «مرکز تماس واحد شبانه‌روزی»، «حفظ فوری داده» و «قالب‌های استاندارد درخواست ادله الکترونیک» برقرار شود تا امکان انتساب و پاسخ به‌موقع فراهم گردد و ادله قابل‌استناد از بین نرود؛ این سازوکارها همسو باهدف کاهش اختلاف در معیارهای پذیرش ادله، بار اثبات را منطقی‌تر توزیع می‌کند.(Europol & Eurojust, 2024: 13)

۷تأثیرات اقتصادی و اجتماعی پورنوگرافی دیپ‌فیک بر قربانیان و جامعه

پورنوگرافی دیپ‌فیک برآیندی از زیان‌های هم‌زمان روانی، مالی و اجتماعی است که آثار آن از سطح فرد به محیط کار و شبکه‌های اجتماعی سرریز می‌کند: قربانی معمولاً با اضطراب، شرم اجتماعی و کناره‌گیری از فضاهای عمومی و برخط روبه‌رو می‌شود که به کاهش سرمایه اجتماعی، تضعیف شبکه‌های حمایتی و افت انگیزه و بهره‌وری می‌انجامد؛ در ادامه، هزینه‌های مستقیم حذف محتوا، دریافت مشاوره روان‌شناختی، پیگیری قضایی و ساعات کاری ازدست‌رفته، همراه با الگوهای اخاذی و تهدید به انتشار، زیان مالی را تشدید می‌کند و حتی می‌تواند به قطع همکاری کارفرما، توقف ارتقا یا خودسانسوری شغلی بینجامد .(Europol & Eurojust, 2024: 15) در سطح بازار کار، برچسب‌زنی و «بی‌اعتباری ادراکی» حتی وقتی جعلی بودن محتوا ثابت است، فرصت‌های استخدام و ترفیع را کاهش می‌دهد و شکاف دستمزد را برای گروه‌های هدف (به‌ویژه زنان و افراد جوان) تعمیق می‌کند؛ این چرخه با بازنشر نسخه‌های «اساساً معادل» و ماندگاری ردّ دیجیتال در سکوهای مختلف تقویت می‌شود و بازگشت قربانی به زندگی حرفه‌ای و اجتماعی را دشوارتر می‌سازد (EPRS, 2025: 2–3).

در سطح جامعه نیز، گردش مکرر این محتوا اعتماد عمومی به تصاویر و روایت‌های شخصی را فرسایش می‌دهد، به افزایش منازعات برخط و آفلاین دامن می‌زند و هزینه‌های تعدیل محتوا و رسیدگی انتظامی و قضایی را برای نهادهای عمومی و سکوهای خصوصی بالا می‌برد؛ در

نتیجه، منابع عمومی از سایر حوزه‌های رفاهی و امنیتی منحرف می‌شود و «هزینه اجتماعی» پدیده افزاینده می‌گردد(Europol & Eurojust, 2024: 15) . به‌عنوان شاخصی از افزایش بار رسیدگی انتظامی و قضایی در حوزه ادله الکترونیکی جرایم برخط (از جمله پرونده‌های مرتبط با انتشار محتوای جنسی بدون رضایت و ضرورت حفظ و دسترسی فوری داده)، گزارش سیریوس نشان می‌دهد «درخواست‌های اضطراری افشای داده» از ۲٬۶۵۷ مورد در سال ۲۰۱۸ به ۲۱٬۷۴۶ مورد در سال ۲۰۲۳ رسیده است(Europol; Eurojust; EJN, 2024: 57) ؛ نک. نمودار ۱.



نمودار ۱. روند درخواست‌های اضطراری افشای داده در اتحادیه اروپا۲۰۱۸–۲۰۲۳. منبع.(Europol; Eurojust; EJN, 2024: 57)

در حقوق ایران نیز تحلیل‌های بومی بر لطمه‌های حیثیتی، گسیختگی روابط خانوادگی، دشواری بازتوانی اجتماعی و اثرات شغلی ماندگار تأکید کرده و بر ضرورتِ حمایت ترکیبی حفظ فوری داده، مسیرهای روشن جبران خسارت، و خدمات روان اجتماعی برای جلوگیری از حاشیه‌رانی پایدار قربانی انگشت گذاشته‌اند .(Zhou et al., 2025: 61)جمع‌بندی آنکه، پیامدهای اقتصادی و اجتماعی پورنوگرافی دیپ‌فیک نه تک‌بُعدی و گذرا، بلکه پیوسته و تقویت‌کننده یکدیگرند؛ و بدون سازوکارهای حمایتی و جبران مؤثر، این آثار به کاهش مشارکت شغلی، فرسایش اعتماد و افزایش هزینه‌های حکمرانی در سطح کلان می‌انجامند (EPRS, 2025: 2–3; Europol & Eurojust, 2024:16).

۷. نقش شرکت‌های فناوری در تولید یا مقابله با دیپ‌فیک‌ها

شرکت‌های فناوری، به‌ویژه ارائه‌دهندگان پلتفرم‌های اجتماعی و توسعه‌دهندگان ابزارهای هوش مصنوعی، نقش اساسی در شکل‌گیری و پیشگیری از پدیده دیپ‌فیک‌ها دارند. این شرکت‌ها از یک‌سو، با طراحی و عرضهٔ فناوری‌های پیشرفته در زمینهٔ پردازش تصویر و ویدئو، ابزارهایی را در اختیار کاربران قرار می‌دهند که امکان تولید محتوای جعلی بسیار واقعی را فراهم می‌کند (Zhou et al., 2025: 55–57).به‌ویژه در پلتفرم‌های اشتراک‌گذاری ویدئو، بدون نظارت کافی، این ابزارها

می‌توانند به‌راحتی برای ایجاد و انتشار محتوای مضر استفاده شوند که در نتیجه به‌سرعت بین کاربران منتشر می‌شود .(Garrido et al., 2024: 76–78) این شرکت‌ها با فراهم‌کردن فناوری‌های مبتنی بر یادگیری عمیق و شبکه‌های عصبی، به تولید محتوای جعلی که قابلیت شبیه‌سازی دقیق ویژگی‌های انسانی را دارند، دامن می‌زنند و این مسئله چالش‌های جدیدی در زمینه حریم خصوصی، امنیت سایبری و اخلاق به وجود می‌آورد.(Franqueira et al., 2024: 32–35)

برای عینی‌سازی این نقش دوگانه (توانمندسازی تولید و تکلیف به مقابله)، می‌توان به نمونه‌های قضایی و اجرایی اشاره کرد که نشان می‌دهند چگونه مسئولیت پلتفرم‌ها و توسعه‌دهندگان ابزار، در عمل صورت‌بندی و اعمال می‌شود. از حیث «تکالیف سکوهای میزبان» در مهار انتشار و بازانتشار محتوای غیرقانونی، رأی دیوان عدالت اتحادیه اروپا در پرونده شماره C-18/18 موسوم به «گلاویشنیش–پیشچِک علیه فیس‌بوک ایرلند» یک مبنای رویه‌ای مهم فراهم می‌کند: دادگاه می‌تواند سکو را مکلف سازد نه فقط محتوای «عیناً یکسان»، بلکه در شرایطی «نسخه‌های دارای معنای اساساً معادل و هم‌معنای» همان محتوای غیرقانونی را نیز حذف یا دسترسی به آن را مسدود کند. ترجمان این منطق در حوزه دیپ‌فیک، به‌ویژه هرزه‌نگاری جعل عمیق فاقد رضایت، آن است که «تعهد سکو» صرفاً به حذف یک مصداق محدود نمی‌شود، بلکه می‌تواند به حذف نسخه‌های بازنشرشده و تغییرشکل‌یافته نیز تسری یابد تا چرخه بازتولید سریع، کارکرد بازدارنده‌ی دستور حذف را بی‌اثر نکند.(CJEU, 2019: paras. 45–50)

در سوی «تولید ریسک» نیز، نمونه‌ی نرم‌افزار «دیپ‌نود» (۲۰۱۹) نشان می‌دهد که نقش و مسئولیت شرکت‌های فناوری صرفاً به مرحله‌ی میزبانی و انتشار محدود نمی‌شود، بلکه در سطح «طراحی و عرضه‌ی نرم‌افزار» نیز معنا می‌یابد. «دیپ‌نود» با اتکا به یادگیری عمیق امکان تولید تصاویر برهنه‌ی جعلی از تصاویر معمولی را فراهم می‌کرد و پس از واکنش عمومی گسترده، ارائه‌ی آن متوقف شد؛ بااین‌حال، همین تجربه آشکار ساخت که در غیاب الزامات حداقلی کاهش خطر در مرحله‌ی طراحی، از جمله محدودسازی دسترسی، پیش‌بینی شروط سخت‌گیرانه‌ی استفاده و نظارت بر کانال‌های توزیع، ظرفیت تکثیر نرم‌افزار و انتقال آن به نسخه‌ها و بسترهای مشابه می‌تواند به تداوم «آسیب ساختاری» بینجامد. ازاین‌رو، تحلیل نقش شرکت‌های فناوری در دیپ‌فیک ناگزیر باید «مسئولیت در سطح محصول و خدمت» را نیز در کنار مسئولیت پلتفرمی لحاظ کند (Cole, 2019; Kastrenakes, 2019).

به‌عنوان نمونه‌ی ملیِ مشخص و متأخر، اقدام حقوقیِ مقام حقوقی شهر سان‌فرانسیسکو علیه وب‌سایت‌های موسوم به «برهنه‌ساز یا نودیفای» (۲۰۲۴) نشان می‌دهد که برخی نظام‌های حقوقی،

علاوه بر تکیه بر خودتنظیمی سکوها، از ظرفیت‌های حقوق عمومی و ضمانت‌اجراهای مدنی نیز برای مهار این خدمات استفاده می‌کنند. در این پرونده، تمرکز صرفاً بر «محتوا» نیست، بلکه «خدمتِ تسهیلِ تولیدِ تصاویر برهنه‌ی جعلیِ فاقد رضایت» در کانون دعوی قرار می‌گیرد و خواسته‌هایی مانند دستور منع و جریمه‌های مدنی، به‌منظور توقف ارائه‌ی خدمت و کاهش دسترسی و انتشارمطرح می‌شود. این الگو، به‌طور عملی ادعای «نقش دوگانه» شرکت‌های فناوری را عینی می‌کند: هم می‌توانند تولید را توانمند کنند و هم به‌موجب تکالیف حقوقی، مکلف به پیشگیری و کاهش مخاطره در زنجیره‌ی تولید تا توزیع خواهند بود (San Francisco City Attorney's Office, 2024; People of the State of California, 2024: 2–3).

این نمونه‌ها نشان می‌دهند نقش شرکت‌های فناوری در دیپ‌فیک‌ها صرفاً توصیفی نیست، بلکه در دو سطحِ «سکو» و «محصول و خدمات فناورانه»، به تکالیف و ضمانت‌اجراهای مشخصِ حذف، پیشگیری و کاهش مخاطره محقق می‌شود.

از سوی دیگر، این شرکت‌ها مسئولیت مقابله با این مشکل را نیز بر عهده دارند. بسیاری از پلتفرم‌ها در تلاش‌اند تا با پیاده‌سازی الگوریتم‌های شناسایی دیپ‌فیک، محتوای جعلی را تشخیص داده و از انتشار آن جلوگیری کنند. الگوریتم‌های شناسایی دیپ‌فیک به طور مداوم درحال‌توسعه هستند، اما همچنان در شناسایی جعل‌های پیچیده با مشکلاتی مواجه‌اند. این چالش به دلیل توانایی‌های روزافزون فناوری‌های دیپ‌فیک در ایجاد محتوای جعلی بسیار واقعی و به‌ویژه در برابر تغییرات جزئی در تصاویر و ویدئوها رخ می‌دهد. به طور مثال، برخی از سیستم‌های شناسایی قادر به شناسایی تغییرات پیچیده در چهره یا صدا نیستند و تنها در موارد ساده‌تر به‌درستی عمل می‌کنند (Wang et al., 2025: 49)؛ بنابراین، هرچند این شرکت‌ها درحال‌توسعه فناوری‌های شناسایی و مقابله هستند، اما هنوز نیاز به همکاری‌های بین‌المللی و ایجاد استانداردهای جهانی برای نظارت و جلوگیری از تولید دیپ‌فیک‌های پیچیده و مضر احساس می‌شود.(Kim et al., 2025: 3–6)

علاوه بر این، بسیاری از این شرکت‌ها در تلاش هستند تا با همکاری با نهادهای نظارتی، قوانین و مقررات سخت‌گیرانه‌تری را برای جلوگیری از سوءاستفاده از فناوری‌های دیپ‌فیک وضع کنند. در این زمینه، برخی از پلتفرم‌ها به طور داوطلبانه محتواهای دیپ‌فیک را حذف یا محدود می‌کنند و همچنین برخی از شرکت‌های فناوری مانند گوگل و فیس‌بوک الگوریتم‌های شناسایی و فیلتر محتوای جعلی را در پلتفرم‌های خود به کار می‌برند تا به‌صورت خودکار این محتواها را شناسایی و حذف کنند .(Masood et al., 2024: 13–15) بااین‌حال، چالش اصلی در این زمینه نبود هماهنگی کافی بین این شرکت‌ها و نهادهای قانونی است که در سطح جهانی می‌توانند مانع از تولید و انتشار دیپ‌فیک‌ها

شوند.

در نهایت، نقش شرکت‌های فناوری در تولید و مقابله با دیپ‌فیک‌ها باید به‌دقت مدیریت شود. اگر این شرکت‌ها تنها به تولید فناوری‌های پیشرفته بسنده کنند، می‌توانند در واقع به گسترش این پدیده کمک کنند. اما اگر در کنار توسعه فناوری‌های تشخیص و جلوگیری از دیپ‌فیک‌ها، از مسئولیت اجتماعی خود نیز بهره ببرند و با نهادهای نظارتی همکاری کنند، می‌توانند در کاهش آثار منفی این پدیده نقش مؤثری ایفا کنند. ایجاد قوانین و مقررات دقیق و استانداردهای بین‌المللی برای مقابله با دیپ‌فیک‌ها، امری ضروری است که باید در دستور کار این شرکت‌ها و نهادهای دولتی قرار گیرد (Masood et al., 2024: 13–15).

نتیجه‌گیری

پیوند کارآمد میان قاعدهٔ ماهوی و سازوکار ادله، محور مهار «پورنوگرافی دیپ‌فیک» است؛ جایی که اعتبار حکم کیفری به قابلیت تبدیل خروجی‌های فنی به دلایلِ قابل‌استناد وابسته می‌شود. در چارچوب حقوقی ایران، مواد ۱۴ تا ۱۶ قانون جرائم رایانه‌ای با تصریح «واقعی یا ساختگی» و پوشش «تحریف»، قابلیت انطباق بر مصادیق فاقدرضایت جنسی را فراهم می‌کنند، اما کارآمدی به‌شرط تبیین سه شاخص عینی است: «شباهت محسوس برای ناظر متعارف»، «انتساب قابل‌اعتماد به شخص قابل‌شناسایی» و «فقدان رضایت معتبر». هنگامی که این سه شاخص در سطح تقنینی و آیین‌نامه‌ای تثبیت شوند و در گزارش کارشناسی همراه با آستانهٔ تصمیم و حدود خطا صورت‌بندی گردند، پیش‌بینی‌پذیری رسیدگی و حمایت از بزه‌دیده به طور معنادار ارتقا می‌یابد. در اتحادیهٔ اروپا نیز اتکای صرف به افشای منشأ، بدون استقرار معماری «اثبات منشأ و تمامیت محتوا» و بدون الزام سکوها به حذف «نسخه‌های اساساً معادل»، عملاً به یک الزام حداقلی و صوری تقلیل می‌یابد و چرخهٔ بازنشر را دست‌نخورده باقی می‌گذارد.

سنجش هزینه–فایده، کفهٔ اصلاحات را سنگین‌تر می‌کند. سرمایه‌گذاری اولیه در آزمایشگاه‌های جرم‌یابی رسانه‌ای، تربیت کارشناس و قاضی مسلط به سنجه‌های چندنشانه‌ای، و بار انطباق برای سکوها، در برابر کاهش زیان‌های حیثیتی و اقتصادی بزه‌دیدگان، افزایش بازدارندگی شبکه‌ای و کاستن از هزینه‌های حکمرانی ناشی از رسیدگی‌های فرساینده، توجیه‌پذیر است. بازتنظیم ضمانت اجرا بر مبنای «شدت آسیب شبکه‌ای» زمانی اثر واقعی دارد که منافع نامشروع مسترد و جزای مالی بر پایهٔ گردش شبکه‌ایِ توزیع محاسبه شود و مزیت اقتصادی ارتکاب عملاً از میان برود؛ به همراه این تغییر، ریسک کشف و حذف سریع، (با نگهداشت فراداده و پاسخ زمان‌مند) برای مرتکب محسوس می‌شود.

امکان اقتباس تجربهٔ اروپایی در بستر ایران، گزینشی و مرحله‌ای است. افشای منشأ هنگامی معنادار است که به زنجیرهٔ نگهداشتِ قابل ممیزی و قالب واحدِ گزارش کارشناسی متصل شود و به مرجع رسیدگی اجازه دهد نسبت میان خطای ابزار و بار اثبات را بسنجد. در سطح اجرا، زمان‌بندی کوتاه برای حفظ فوری داده، نقطهٔ تماس دائمی برای تعامل با سکوهای فراملی، و تعهد به جست‌وجوی فعال نسخه‌های اساساً معادل، فاصلهٔ میان کشف و اقدام را کاهش می‌دهد. در سطح هنجاری، تعریف مضیّق دیپ‌فیک بر اساس سه شاخص عینی و نظام تشدید پلکانی مجازات بر پایهٔ دامنهٔ انتشار، انتفاع اقتصادی و تکرار سازمان‌یافته، نسبت قاعده با واقعیت زیان را هم‌تراز می‌کند.

در سطح امکان‌سنجی اقتباس، بهره‌گیری از برخی عناصر اروپایی در ایران زمانی معقول است که «هزینه استقرار» با «فایده کاهش زیان و کاهش بار رسیدگی» هم‌خوان باشد. برای نمونه، تدوین یک «استاندارد ملی گزارش کارشناسی» و یکسان‌سازی زنجیره نگهداشت ادله، به هزینه‌هایی مانند آموزش و تجهیز حداقلی و تدوین قالب‌های واحد نیاز دارد؛ اما در مقابل، اختلافات کارشناسی را کم می‌کند، مسیر اثبات را کوتاه‌تر می‌سازد و از اطاله و تکرار رسیدگی می‌کاهد. همچنین، اگر سکوها به سازوکارهای مؤثرِ حذف و جلوگیری از بازنشر نسخه‌های «اساساً معادل» ملزم شوند، هرچند هزینه‌های اجراییِ تعدیل محتوا افزایش می‌یابد، ولی دامنه و مدت بزه‌دیدگی کاهش پیدا می‌کند و نیاز به پیگیری‌های مکرر (از جمله درخواست‌های حفظ داده و مکاتبات ادله‌ای) کمتر می‌شود؛ بنابراین، در مجموع می‌توان گفت حرکت مرحله‌ای به‌سوی این ابزارها به شرط بومی‌سازی و سنجش‌پذیر کردن تکالیف، از نظر هزینه–فایده قابل دفاع است.

چشم‌انداز کارآمد زمانی شکل می‌گیرد که سه سطح «هنجار، ادله و اجرا» به‌صورت یک چارچوب به‌هم‌پیوسته عمل کنند: تعریف روشن دیپ‌فیک و مرزبندی میان «مستهجن ساختگی» و «هتک حیثیت تحریفی»؛ استاندارد واحد گزارش کارشناسی چند نشانه‌ای با تصریح روش، آستانهٔ تصمیم و حدود خطا؛ و تکالیف سنجش‌پذیر برای سکوها در مهلت حذف، نگهداری و تحویل فراداده و حذف نسخه‌های معادل. با سنجه‌های ارزیابی مبتنی بر داده از جمله نرخ کشف، نرخ حذف و نرخ بازنشر معادل امکان می‌یابد که سیاست کیفری به طور مستمر بازبینی و اصلاح شود و پاسخ حقوقی، از واکنش‌های موردی به حکمرانی حمایت محور ارتقا پیدا کند.

در مقام جمع‌بندی اجرایی، مسیر اصلاحات را می‌توان به‌صورت مرحله‌ای چنین سامان داد: کوتاه‌مدت :تعریف مضیّق و روشن از «پورنوگرافی دیپ‌فیکِ فاقد رضایت» بر پایه سه شاخص عینی و پیش‌بینی و اجرای سریع «دستورهای حفظ داده» با مهلت‌های کوتاه و سازوکار ابلاغ روشن برای جلوگیری از زوال ادله.

میان‌مدت :تدوین و تصویب «استاندارد ملی گزارش کارشناسی محتوای مصنوعی» (روش، آستانه تصمیم، حدود خطا و زنجیره نگهداشت) و توانمندسازی کارشناسان و ضابطان برای یکنواختی و اتکاپذیری ادله.

بلندمدت :اصلاحات تقنینی و تنظیم‌گری از طریق بازنگری پلکانی ضمانت‌اجراها متناسب با دامنه انتشار و انتفاع و نیز تعیین تکالیف سنجش‌پذیر برای سکوها در حذف سریع، جلوگیری از بازنشر نسخه‌های اساساً معادل، و نگهداری و تحویل فراداده در مهلت معین.

منابع

۱) اکبری، عباسعلی، آقاپور، علی، و آقاپور، کمال. (۱۴۰۱). تحلیل پدیده مجرمانه دیپ‌فیک‌ها. نشریه آگاه، ۱۶(۵۹)، ۱۴۹–۱۶۹.

۲) بابایی، جواد. (۱۳۹۷). جرایم رایانه‌ای و آیین دادرسی حاکم بر آن. تهران: مرکز مطبوعات و انتشارات قوه قضائیه.

۳) پلیس فضای تولید و تبادل اطلاعات (فتا). (۱۴۰۰). راهنمای حمایتی و گزارش‌های اجرایی پیرامون انتشار تصاویر خصوصی.

۴) پلیس فضای تولید و تبادل اطلاعات (فتا). (۱۴۰۳). گزارش طرح مقابله با جرائم غیراخلاقی در سکوهای برخط (آمار رصد و سهم سکوها).

۵) رزمان، علی. (۱۳۹۵). بررسی کیفری جرائم جنسی و منافی عفت با نگاهی بر فضای سایبر. تهران: قانون‌یار.

۶) رضایی، مهدی. (۱۳۹۶). بررسی تطبیقی جرایم مستهجن در حقوق جزای ایران و اتحادیه اروپا. پژوهش‌های حقوقی، ۱۲(۳)، ۱۰۱–۱۲۰.

۷) شیری، عباس. (۱۴۰۱). دیپ‌فیک یا همانندسازی صوتی یا تصویری غیرواقعی در حقوق کیفری. فصلنامه تحقیقات حقوقی (ویژه‌نامه حقوق و فناوری)، ۲۵(۲)، ۱۴۳–۱۶۸.

۸) عزیزی، امیرمهدی. (۱۴۰۲). حقوق کیفری جرایم رایانه‌ای. تهران: مجد.

۹) قناد، فاطمه. (۱۳۹۰). جعل در بستر فناوری‌های اطلاعات و ارتباطات. آموزه‌های حقوق

کیفری، ۱(۲)، ۶۳–۸۳.

۱۰) محمدی فردوئی، عطاءالله. (۱۳۹۷). بررسی جزایی عناصر بزه جعل رایانه‌ای در حقوق ایران. نشریه قانون‌یار، ۲(۸)، ۴۵۹–۴۷۶.

۱۱) مرکز ملی فضای مجازی. (۱۴۰۳). اطلاع‌رسانی درباره موج‌های محتوای مخرب و ضرورت اقدام هماهنگ.

۱۲) موسوی، علی. (۱۳۹۴). ماهیت و مصادیق محتوای مستهجن در قوانین ایران. فصلنامه حقوق و سیاست، ۲۷(۱)، ۴۵–۶۸.

13) Aghaei, M., & Han, J. (2023). Combating deepfake technologies: Legal and ethical challenges in Europe and beyond. Springer.
14) AlgorithmWatch. (2025). A guide to the AI Act — Why gaps remain for fundamental rights.
15) AP News. (2024, March 14). EU asks major platforms how they are reducing generative-AI risks (including deepfakes) under the DSA.
16) Associated Press. (2024, May 30). Operation Endgame takes down ransomware networks across Europe. Associated Press News.
17) Becker, M. (2024). Generative KI und Deepfakes in der KI-VO — Für eine Positivkennzeichnung authentischer Inhalte. Computer und Recht, 40(6), 353–366.
18) Birrer, A., & Just, N. (2024). What we know and don't know about deepfakes: An investigation into the state of the research and regulatory landscape. New Media & Society, 00(0), 12.
19) Block, M. J. (2024). A critical evaluation of deepfake regulation through the AI Act in the European Union. Journal of European Consumer and Market Law, 13(4), 184–192.
20) C2PA (Coalition for Content Provenance and Authenticity). (2025). Content Credentials Explainer (Version 2.2), 4–5, 11–12.
21) Chesney, R., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1762–1768.
22) Citron, D. K., & Franks, M. A. (2014). Criminalizing Revenge Porn. Wake Forest Law Review, 49, 345–391. BU Law Scholarly Commons+2repository.law.miami.edu+2
23) Citron, D. K., & Franks, M. A. (2014). Criminalizing Revenge Porn. Wake Forest Law Review, 49, 345–391. BU Law Scholarly Commons+2repository.law.miami.edu+2
24) Cole, S. (2019, June 26). This Horrifying App Undresses a Photo of Any Woman With a Single Click. Motherboard (VICE).
25) Council of Europe. (2018, March 7). Recommendation CM/Rec(2018)2 of the Committee of Ministers to Member States on the roles and responsibilities of internet intermediaries.
26) Court of Justice of the European Union (CJEU). (2019). Eva Glawischnig-Piesczek v Facebook Ireland Limited, Case C-18/18, Judgment of 3 October 2019, ECLI:EU:C:2019:821 (paras. 45–50).
27) Court of Justice of the European Union (CJEU). (2019). Glawischnig-Piesczek v Facebook Ireland Limited, C-18/18 (paras. 45–49).
28) Court of Justice of the European Union. (2019). Judgment of 3 October

2019, Eva Glawischnig-Piesczek v Facebook Ireland Limited (C-18/18), ECLI:EU:C:2019:821 (paras. 45–49).
29) ENISA — European Union Agency for Cybersecurity. (2025). ENISA Threat Landscape 2025 (pp. 3–4).
30) Eurojust & Europol. (2025). Common challenges in cybercrime — 2024 review, 6–7, 17–18.
31) European Commission — European AI Office. (2024). The European AI Office: Mandate and tasks.
32) European Commission. (2020). EU Strategy on Victims' Rights (2020–2025).
33) European Commission. (2022). European Artificial Intelligence Strategy: Fostering research and investment through Horizon Europe and Digital Europe Programme.
34) European Commission. (2024). Guidelines for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes (OJ C 218/1, 14.6.2024), 8, 10.
35) European Commission. (2024, March 25). Guidelines under the DSA to safeguard election integrity (including labelling of AI-generated content such as deepfakes).
36) European Parliament Research Service. (2021). Tackling deepfakes in European policy.
37) European Union Agency for Cybersecurity (ENISA). (2025). AI threat landscape 2025: Standards and recommendations for trustworthy AI, 15, 43, 49.
38) European Union. (2022). Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act), OJ L 277. Arts. 51–52, 79–82.
39) European Union. (2023). Regulation (EU) 2023/1543 on European Production and Preservation Orders for electronic evidence.
40) European Union. (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act), OJ L 2024/1689. Art. 50, 82; Art. 99, 115.
41) Europol. (2024). Facing reality? Law enforcement and the challenge of deepfakes (Updated ed.). Publications Office of the European Union. Europol & Eurojust. (2024). SIRIUS EU electronic evidence situation report 2024, 8–12, 53–61, 65–67.
42) Europol; Eurojust; European Judicial Network (EJN). (2024). SIRIUS EU Electronic Evidence Situation Report 2024. p. 57.
43) Feinberg, J. (1984). Harm to Others: The Moral Limits of the Criminal Law (Vol. 1). Oxford University Press.
44) Francati, D., Goonatilake, Y. N., Pawar, S., Venturi, D., & Ateniese, G. (2025). The coding limits of robust watermarking for generative models, 1–2.
45) Franqueira, V., Altuncu, E., & Li, S. (2024). Deepfake: Definitions, performance metrics, and standards, datasets and benchmarks, and a meta-review. Journal of Artificial Intelligence Research, 70, 1–30.
46) Franqueira, V., Ribeiro, R., Silva, J. P., & Almeida, J. (2024). Preventing and detecting deepfake videos: Challenges and solutions. International Journal of Digital Law, 29(2), 32–35.
47) Garrido, A., Marzal, P., & Alvarado, E. (2024). The spread of fake news and deepfake content on social media platforms. Journal of Media Studies, 15(4), 76–78.

48) Garrido, M., Albuja, S., & Stojanovic, J. (2024). The rise of deepfake: Implications for digital content and the law. Cambridge University Press.
49) Ghial, R., Pundir, D., & Kaur, R. (2024). Right to be forgotten: A human rights-based approach for governance in generative AI. In T.
50) Gibson, Dunn & Crutcher LLP. (2023, August 1). EU strengthens cross-border access to e-evidence in criminal proceedings (application as of 18 Aug 2026).
51) Glawischnig-Piesczek v. Facebook Ireland Ltd (C-18/18) [2019] ECLI:EU:C:2019:821.
52) Google Spain SL v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González (C-131/12) [2014] ECLI:EU:C:2014:317.
53) Gosse, C., & Burkell, J. (2020). Politics and porn: How news media characterizes problems presented by deepfakes. Critical Studies in Media Communication, 37(5), 497–511.
54) Han, C., Li, A., Kumar, D., & Durumeric, Z. (2025). Characterizing the MrDeepFakes sexual deepfake marketplace. In Proceedings of the 34th USENIX Security Symposium, 9–12.
55) Han, Y., Kwon, J., & Lee, H. (2024). MrDeepFakes and the illegal distribution of deepfake pornography: A systematic review. International Journal of Digital Law, 29(2), 71–88.
56) Hayes, S. C., & Paul, P. (2007). The pornography "addiction" model: Evidence for false claims. Science and Engineering Ethics, 13(4), 515–533.
INTERPOL. (2024). Beyond illusions: Synthetic media and implications for law enforcement (INTERPOL Digital Forensics Series, 6), 14–15, 26.
57) Jacobi, S. (2023, September 6). AI and deepfakes: EU and Italian regulations. Jacobacci Law Firm.
58) Kastrenakes, J. (2019, June 27). Controversial deepfake app DeepNude shuts down hours after being exposed. The Verge.
59) Kim, D., Lee, T., & Park, Y. (2025). The role of AI platforms in tackling deepfakes: Challenges and innovations. Technology and Law Journal, 12(1), 3–6.
60) Korshunov, P., & Marcel, S. (2023). Deepfakes: Detection, challenges, and responses. Springer.
61) Krook, J., Winter, P., Downer, J., & Blockx, J. (2024). A systematic literature review of AI transparency laws in the EU and UK: A socio-legal approach to AI transparency governance. SSRN Preprint, 3.
62) Łabuz, M. (2024). Deep fakes and the Artificial Intelligence Act — An important signal or a missed opportunity? Policy & Internet, 783–800.
63) Leins, K., Williams, B., & Murray, G. (2023). AI and responsibility gaps in combatting deepfakes.
64) Löfgren-Mårtenson, L. (2008). Adolescents' perceptions of pornography. Journal of Sex Research, 45(3), 306–312.
65) Mania, K. (2022). Legal protection of revenge and deepfake porn victims in the European Union: Findings from a comparative legal study. Trauma, Violence & Abuse, 25(1), 117–129.
66) Masood, A., Hafeez, I., & Khan, S. (2024). Ethical implications of AI in media: Deepfake detection and prevention. Journal of AI Ethics, 5(2), 13–15.
67) Masood, M., Nawaz, M., Malik, K. M., Javed, A., & Irtaza, A. (2021). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. Journal of Visual Communication

and Image Representation, 77, 103–118.
68) McGlynn, C., & Toparlak, O. (2025). Digital voyeurism and the implications of deepfake pornography. ResearchGate.
69) Moreno, F. R. (2024). Generative AI and deepfakes: A human rights approach to tackling harmful content. International Review of Law, Computers & Technology, 38(3), 297–326.
70) People of the State of California, acting by and through San Francisco City Attorney David Chiu v. Sol Ecom, Inc., et al. (2024). Complaint for Injunctive Relief and Civil Penalties for Violations of Business and Professions Code § 17200 (Superior Court of the State of California, County of San Francisco).
71) Rigon, B., & Baratto, G. (2024). La deepfake pornography tra criminologia e diritto. Quaderni della facoltà di giurisprudenza, 86, 459–484.
72) San Francisco City Attorney's Office. (2024, August 15). City Attorney sues most-visited websites that create nonconsensual deepfake pornography (Press release).
73) Schuett, J. (2023). Risk management in the Artificial Intelligence Act. European Journal of Risk Regulation, 15(2), 367–385.
74) Senjyu, C. So- In, & A. Joshi (Eds.) , Smart Trends in Computing and Communications: Proceedings of SmartCom 2024, Vol. 5 (Lecture Notes in Networks and Systems, 949, 23–34). Springer Nature Singapore.
75) Shi, H., Shi, X., Doğan, S. M. Y., Alzubi, S. M., Huang, T., & Zhang, Y. (2025). Benchmarking audio deepfake detection robustness in real-world communication scenarios. Proceedings of EUSIPCO 2025, 1–3, 339–347.
76) Tortorelli, M. (2024). Image-based sexual abuse: Application limits and perspectives for reform of Article 612-ter of the Italian Criminal Code. Diritto Penale Contemporaneo. Rivista Trimestrale, 1(1), 207–240.
77) Truong, K. C. D. (2024). Reputation (Not Taylor's Version): Regulating AI-hallucinated deepfakes of public figures. University of Illinois Journal of Law, Technology & Policy, 449.
78) Umbach, L., Smith, K., & Hudson, D. (2024). Non-consensual synthetic intimate imagery: A growing threat to digital privacy. Journal of Digital Ethics, 12(1), 89–101.
79) Verdoliva, L. (2020). Media forensics and deepfakes: An overview. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910–932.
80) Wang, X., Zhou, S., & Zhang, Y. (2025). Detecting and mitigating deepfakes: A review of AI-based systems and challenges. Proceedings of the 2025 International Conference on AI and Media, 49.
81) Zhou, Z., Zhang, L., & Chen, J. (2025). Deepfake detection techniques and their limitations: A systematic review. Journal of Digital Forensics, 33(1), 55–57.