

딥페이크와 사생활 보호*

고 세 일**

【요 지】

한 사람이 쓰는 컴퓨터가 한 공간에 남아 있었다. 그 사람이 쓰는 컴퓨터는 다른 사람이나 다른 사람이 쓰는 컴퓨터와 연결되지 않았다. 시간이 흘러서 인터넷이 도입되었다. 전화선, 케이블선, 랜선을 통해서 인터넷으로 여러 컴퓨터와 연결되었다. 많은 사람 사이에 많은 정보가 오고 갔다. 인터넷 이전의 세상과 인터넷 이후의 세상이 많이 달라 보였다. 또한 시간이 흘러서 인공지능이 나타났다. 많은 사람은 인공지능이 새로운 산업혁명을 일으킬 것으로 여기고 열광하기도, 두려워하기도 했다. 또한 다른 시간이 흘러서 생성형 인공지능이 등장했다. 많은 사람은 생성형 인공지능의 똑똑함에 놀랐다. 많은 사람은 생성형 인공지능을 신기한 것으로 여겼다. 또한 생성형 인공지능이 만들어 내는 그림자에 대해서도 고민하기 시작했다.

생성형 인공지능이 만드는 어두운 그림자 가운데 한 영역이 딥페이크이다. 현실에서 딥페이크가 문제를 일으키는 영역은 크게 둘이다. 첫째는 ‘선거’와 관련된 영역이다. 둘째는 ‘음란물’이다. 딥페이크가 우리 사회에 부작용을 일으키는 두 영역의 공통점은 사람의 동일성(identity)을 침해한다는 점이다. 다른 사람이 나쁜 의도로 어떤 사람의 이미지에 대한 동일성을 깨뜨린다. 딥페이크 기술로 만든 나쁜 이미지가 그 사람의 진실한 이미지를 왜곡한다. 그러면서 그 사람에 대한 나쁜 이미지로 그 사람의 진실한 이미지를 대체한다. 따라서 대중은 딥페이크 기술로 만들어진 나쁜 이미지를 그 사람에 관한 참된 정보로 인식한다. 그런 면에서 딥페이크는 그 사람에 대한 동일성을 왜곡하고, 그 사람에 대한 자아상(self-image)을 허문다.

딥페이크가 일으키는 측면을 다루는 논문은 두 방향으로 접근한다. 첫째는 딥페이크에 대한 기술적 측면을 고찰하는 논문이다. 둘째는 딥페이크를 형사법적으로 규율하려고 탐구하는 논문이다. 그런 측면에서 아직 딥페이크에 대한 민사법적 접근의 논문은 시작 단계이다. 딥페이크를 법으로 규율하는 영역을 모두 망라하면, 이는 ① 플랫폼 책임, 잠재

* 이 논문은 2025년 4월 25일 고려대학교에서 열린 재산법학회 춘계학술대회의 발표문을 고친 글로, 2023년 충남대학교 교내연구비로 지원받은 연구임을 밝힙니다. 생각할 소중한 기회를 주신 세분의 심사위원 선생님에게 따뜻한 감사의 말씀을 전합니다.

** 충남대학교 법학전문대학원 교수, 법학박사

적 인공지능 개발자의 책임을 규율하는 행정법 영역, ② 형사법 영역, ③ 명예훼손과 사생활 침해를 규율하는 민사법 영역이다. 이 논문에서는 딥페이크에 대한 민사법적 접근으로, 딥페이크와 관련한 사생활 보호에 초점을 둔다.

글쓴이는 이 논문을 다음과 같이 구성한다. 먼저 II에서 딥페이크가 일으켰던 사례와 딥페이크 기술을 설명한다. 그 뒤에 III에서 딥페이크에 대한 법적 규율로 사생활 보호를 고찰한다. 마지막으로 IV에서 딥페이크가 일으키는 사생활 문제를 어떤 측면에서 규율하고 보호할 것인지에 대해서 제안한다.

[주제어]

딥페이크, 생성형 인공지능, 딥러닝, 사생활, 프라이버시, 동일성

I. 머리말

한 사람이 쓰는 컴퓨터가 한 공간에 남아 있었다. 그 사람이 쓰는 컴퓨터는 다른 사람이나 다른 사람이 쓰는 컴퓨터와 연결되지 않았다. 시간이 흘러서 인터넷이 도입되었다. 전화선, 케이블선, 랜선을 통해서 인터넷으로 여러 컴퓨터와 연결되었다. 많은 사람들 사이에 많은 정보가 오고 갔다. 인터넷 이전의 세상과 인터넷 이후의 세상이 많이 달라 보였다. 또한 시간이 흘러서 인공지능이 나타났다. 많은 사람은 인공지능이 새로운 산업혁명을 일으킬 것으로 여기고 열광하기도, 두려워하기도 했다. 또한 다른 시간이 흘러서 생성형 인공지능이 등장했다. 많은 사람은 생성형 인공지능의 똑똑함에 놀랐다. 많은 사람은 생성형 인공지능을 신기한 것으로 여겼다. 또한 생성형 인공지능이 만들어 내는 그림자에 대해서도 고민하기 시작했다.¹⁾

생성형 인공지능이 만드는 어두운 그림자 가운데 한 영역이 딥페이크이다.²⁾ 현실에서 딥페이크가 문제를 일으키는 영역은 크게 둘이다.³⁾ 첫째는 ‘선거’와 관련된 영역이다.⁴⁾ 둘째는 ‘음란물’이다. 딥페이크가 우리 사회에 부작용을 일으키는 두 영역의 공통

1) 예를 들어, 생성형 인공지능이 처음 등장했을 때, 생성형 인공지능을 써서 쉽게 만들어 낸 생성형 인공지능 책들이 그렇다. 또한 생성형 인공지능이 인간을 대체하면서, 실제로 인간이 일하는 곳이 줄어드는 현상이 여러 직업군에 나타나고 있다.

2) 딥페이크는 그림과 동영상을 만드는 생성형 인공지능의 한 종류이다. 김명주, “선거 속 딥페이크, 그 실태와 윤리,” 관훈저널 통권 제171호, 2024, 118쪽.

3) 인공지능과 불법행위책임 전반에 대해서는 고세일, “인공지능과 불법행위책임 법리,” 법학연구 제29권 제2호, 2018, 85-117쪽 참조.

4) 예를 들어, 심새롬, “후보가 같이 나라 말아먹자 . . . 초유의 딥페이크 대선 전쟁,” 중앙일보 2025년 4월 21일 기사, available at <https://www.joongang.co.kr/article/25329970> (마지막 확인:

점은 사람의 동일성(identity)을 침해한다는 점이다. 다른 사람이 나쁜 의도로 어떤 사람의 이미지에 대한 동일성을 깨뜨린다. 딥페이크 기술로 만든 나쁜 이미지가 그 사람의 진실한 이미지를 왜곡한다.⁵⁾ 그러면서 그 사람에 대한 나쁜 이미지로 그 사람의 진실한 이미지를 대체한다. 따라서 대중은 딥페이크 기술로 만들어진 나쁜 이미지를 그 사람에 관한 참된 정보로 인식한다. 그런 면에서 딥페이크는 그 사람에 대한 동일성을 왜곡하고, 그 사람에 대한 자아상(self-image)을 허문다.

딥페이크가 일으키는 측면을 다루는 논문은 두 방향으로 접근한다. 첫째는 딥페이크에 대한 기술적 측면을 고찰하는 논문이다.⁶⁾ 둘째는 딥페이크를 형사법적으로 규율하려고 탐구하는 논문⁷⁾이다. 그런 측면에서 아직 딥페이크에 대한 민사법적 접근의 논문은 시작 단계이다.⁸⁾ 딥페이크를 법으로 규율하는 영역을 모두 망라하면, 이는 ① 플랫폼 책임, 잠재적 인공지능 개발자의 책임⁹⁾을 규율하는 행정법 영역, ② 형사법 영역, ③ 명예훼손과 사생활 침해를 규율하는 민사법 영역이다. 이 논문에서는 딥페이크에 대한 민사법적 접근으로, 미국 불법행위법을 중심으로 딥페이크와 관련한 사생활 보호에 초점을 둔다.¹⁰⁾

글쓴이는 이 논문을 다음과 같이 구성한다. 먼저 II에서 딥페이크가 일으켰던 사례와 딥페이크 기술을 설명한다. 그 뒤에 III에서 딥페이크에 대한 법적 규율로 사생활 보호를 고찰한다. 마지막으로 IV에서 딥페이크가 일으키는 사생활 문제를 어떤 측면에서 규율하고 보호할 것인지에 대해서 제안한다.

2025년 4월 22일).

- 5) '왜곡하다'는 뜻의 영어 'distort'는 'dis+tort'로 이루어진다. 그런데 영어권에서 tort는 '불법행위'를 뜻한다. 따라서 '왜곡하다'는 말 자체가 '불법행위'를 내포하고 있다.
- 6) 예를 들어, 노하림, 유제혁, "Visual Saliency 기반의 딥페이크 이미지 탐지 기법," Journal of Platform Technology, Vol 12, No. 1, 2024, 128-140쪽.
- 7) 이는 주로 딥페이크가 만들어 내는 음란물에 대한 접근이다. 주현경, "딥페이크 기술의 사회적 위험과 규제 방안 - 딥페이크 이용 범죄를 중심으로 -, " 가천법학 제17권 제3호, 2024, 261-306쪽.
- 8) 이에 대해서는 석지운, "딥페이크 음란물로 인한 명예훼손과 민사책임," 연세법학 제46호, 2024, 303쪽과 노현숙, "딥페이크 기술에 의한 인격권 침해 검토 - 초상권을 중심으로 -, " 미디어와 인격권 제11권 제1호, 2025, 43-76쪽 참조.
- 9) 플랫폼 책임과 잠재적 인공지능 개발자 책임은 민사·형사책임으로 규율하기도 한다. 따라서 이 영역은 행정법 영역과 민사법 영역, 형사법 영역이 중첩하는 부분이다.
- 10) 미국 불법행위법 전반에 대해서는 고세일, "미국 불법행위법의 구조와 내용에 대한 연구 - 보통법전집의 과실을 중심으로," 민사법학 제59호, 2012, 87-138쪽 참조. 또한 미국 불법행위법의 인과관계 문제에 대해서는 "미국 과실불법행위와 인과관계 - 불법행위 보통법전집 제2판의 과실규정을 중심으로," 안암법학 제42호, 2013, 169-204쪽 참조.

II. 딥페이크 사례와 기술

1. 딥페이크의 빛과 그림자

딥페이크가 사회에 충격을 끼친 사례는 많다. 주된 사례를 꼽으면 미국의 경우는 2024년 1월 테일러 스위프트¹¹⁾의 성적 묘사가 담긴 딥페이크 사건¹²⁾이고, 한국의 경우는 2024년 텔레그램 딥페이크 음란물 유포 사건¹³⁾이다. 딥페이크 사건은 모두 사람을 대상(객체)로 여기는 것에서 출발한다.¹⁴⁾ 따라서 딥페이크의 사건에 비추어 볼 때, 사람은 인격적인 주체로서 존중하는 사회의 가치와 배려가 필요하다. 따라서 생성형 인공지능 기술을 써서 사람의 동일성을 침해하고, 인격적인 주체를 대상으로 삼는 것을 방지하도록 법 규율의 방향성을 삼아야 한다.

그런데 생성형 인공지능을 쓰는 딥페이크 기술은 가치중립적이다. 이는 ‘칼’의 경우와 같다. 사랑하는 가족이나 이웃을 위해서 과일을 깎거나 요리할 때 칼을 쓸 수 있다. 그러나 다른 사람을 해치는 목적으로 칼을 쓸 수도 있다.¹⁵⁾ 딥페이크의 경우, 일반적으로는 나쁜 사례를 쉽게 떠올린다. 그러나 좋은 목적으로 딥페이크 기술을 쓰는 사례도 많

11) 테일러 스위프트(Taylor Swift)는 미국의 인기 있는 가수이다. 테일러 스위프트가 공연할 때마다 지역경제가 살아난다고 해서, ‘테일러 스위프트 경제’(Taylor Swift Economy)라고 부른다. 테일러 스위프트의 최근 2년 반 동안 이러한 경제 효과는 22억 달러(2025년 4월 26일 환율 기준: 3조 1,653억 원)이다. Auzinea Bacon, “The end of an era: How Taylor Swift boosted the US economy,” CNN Business, available at <https://edition.cnn.com/2024/12/08/business/taylor-swift-eras-tour-economy/index.html> (마지막 확인: 2025년 4월 26일).

12) Kate Conger and John Yoon, “Explicit Deepfake Images of Taylor Swift Elude Safeguards and Swamp Social Media,” New York Times, Jan. 26, 2024, available at <https://www.nytimes.com/2024/01/26/arts/music/taylor-swift-ai-fake-images.html> (마지막 확인: 2025년 4월 21일).

13) 2024년 11월 4일 기준으로 피해자의 숫자는 920명(교육부 집계) 또는 1,200명(검찰청 집계)이다. 나무위키, “2024년 텔레그램 딥페이크 음란물 유포 사건,” (마지막 수정: 2025년 4월 15일), available at <https://namu.wiki/w/2024년%20텔레그램%20딥페이크%20음란물%20유포%20사건> (마지막 확인: 2025년 4월 22일).

14) 이러한 현상을 대상화(objectifying)이라고 한다. Chap Clark, Adoptive Youth Ministry - Integrating Emerging Generations into the Family of Faith 32-36쪽 (Baker Academic 2016) 참조.

15) 최근 교사가 초등학교 김하늘 학생을 칼로 살해한 사건은 우리 사회에 큰 충격을 주었다. 김명일, “내 돈으로 보상하나? 초등생 살해 교사, 범행 전 남편에 한 말,” 조선일보(2025년 4월 8일), available at https://www.chosun.com/national/national_general/2025/04/08/J4RDGZ2KSFELLJCGDIZITV4SEA/?utm_source=naver&utm_medium=referral&utm_campaign=naver-news (마지막 확인: 2025년 4월 25일).

다.¹⁶⁾ 독일 대학이 딥페이크 기술을 써서 질병 진단을 잘할 수 있는 인공지능을 개발했다.¹⁷⁾ 딥페이크 기술을 써서 영화배우가 다양한 나이의 연기¹⁸⁾를 할 수 있게 하는 것이다.¹⁹⁾

2. 딥페이크의 뜻과 기술

이곳에서는 딥페이크 기술을 설명하고자 한다. 딥페이크를 설명하는 방식은 두 가지이다. 첫째, 기술의 측면으로 설명하는 방식이다. 딥페이크에서 ‘딥’은 ‘딥러닝’(deep learning)을 가리키고, ‘페이크’(fake)는 ‘가짜’²⁰⁾를 가리킨다. 둘째, 역사의 측면으로 설명하는 방식이다. 딥페이크는 19세기에 사진 편집 기술로 사용했고, 그 뒤에 영화에 대한 편집 기술로 사용했다. 20세기 기술이 점진적으로 발전했고, 디지털 영상과 함께 영화 편집 기술이 급격하게 발전했다. 그런데 ‘딥페이크’라는 말이 주목을 받기 시작한 것은 2017년 소셜 미디어 플랫폼인 레딧(Reddit)²¹⁾에서 ‘deepfakes’라는 이용자 아이디를 갖은 사람이 유명인의 얼굴을 음란물 영화의 배우의 몸으로 합성하고 레딧에 유포하면

16) 이동욱, “AI가 만들어 낸 거울, 딥페이크의 양면성,” 카이스트신문(2024년 11월 5일), *available at* <https://times.kaist.ac.kr/news/articleView.html?idxno=21969> (마지막 확인: 2025년 4월 25일).

17) 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 277쪽. 의료정보가 민감하고 사생활 침해의 가능성으로 말미암아 인공지능을 훈련하는데 데이터가 충분하지 않다는 어려움이 있었다. 이러한 어려움을 딥페이크 기술을 써서 고해상도 이미지를 새롭게 만들어서 원본 영상과 구별할 수 없는 영상을 딥러닝 기술을 써서 해결한 경우이다.

18) 2019년의 ‘아이리시맨’에서 로버트 드니로, 조 페시, 알파치노가 20대에 시작해서 50년 정도의 세월을 연기하게 했다. 그리하여 젊은 시절의 대역을 쓰지 않고 딥페이크 기술로 모든 연령대의 연기를 가능하게 한 것이다. 우리나라의 경우에는 2022년 방영한 드라마 ‘카지노’에서 최민식씨가 30대의 얼굴과 목소리 연기를 하는데, 딥페이크 기술을 썼다고 한다. 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 275쪽. 딥페이크 기술을 쓴다면, 최근 화제가 된 “폭삭 속았수다”에서 오애순과 양관식을 연기한 아이유, 문소리, 박보검, 박해준의 경우, 딥페이크 기술을 쓰면 한 사람의 배우가 여러 세월의 연기를 모두 할 수 있게 하는 것을 가능하게 한다. 다만 이 경우에도 양면성이 있다. 어느 순간 식당과 커피숍에서 키오스크로 주문하게 되었다. 이로 말미암아 인건비 절감이라는 측면과 함께, 사람이 사라지고 있는 현상이 나타난다. 딥페이크 기술을 영화 산업에 쓸 때도 비슷한 결과가 나온다. 이런 문제 때문에 미국 할리우드에서 배우와 작가 동반 파업을 하기도 했다. Regan Morris, “AI was enemy No. 1 during Hollywood strikes. Now it’s in Oscar-winning films,” BBC (March 31, 2025), *available at* <https://www.bbc.com/news/articles/ce303x19dwgo> (마지막 확인: 2025년 4월 25일).

19) 이런 측면에서 딥페이크 기술을 써서 사망한 배우가 연기하도록 하는 것도 가능하다. 그런 측면에서 이에 대한 윤리적·법적 측면도 고찰할 필요가 있다.

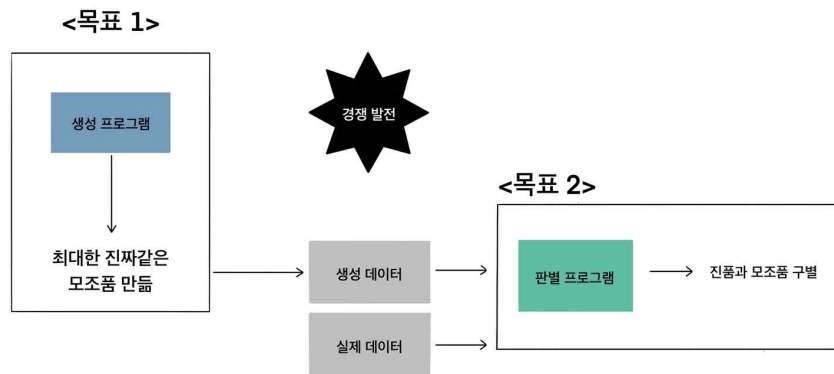
20) ‘가짜’라는 뜻으로 부정적인 속성을 말한다. 그러나 운동경기와 같이 ‘페이크’를 운동선수가 사용하는 기술의 뜻으로 쓰기도 한다.

21) <https://www.reddit.com/?rdt=51007> (마지막 확인: 2025년 4월 22일).

서 발생했다.²²⁾

유럽연합의 인공지능법(EU Artificial Intelligence Act) 제3조 60항은 ‘딥페이크’를 실존하는 사람, 사물, 장소, 법인, 사건을 닮아서, 거짓으로 진정하거나 진실된 것으로 보이게 하는 인공지능이 생성하거나 조작한 표상, 음성 또는 영상 콘텐츠로 정의한다.²³⁾

딥페이크 기술을 그림으로 표현하면 [그림1]과 같다.²⁴⁾



[그림 1]

딥페이크는 기계 학습 모델 가운데 하나인 생성적 적대 신경망(Generative Adversarial Network)²⁵⁾을 써서 기존에 있는 영상과 음성을 조작하여 새로운 가짜 데이터를 만드는 것이다. 생성적 적대 신경망은 훈련하기 어렵다. 또한 계산 자원이 많이 필요한 딥러닝 기술 가운데 하나이다. 생성적 적대 신경망은 두 개의 신경망을 갖고 있다. 첫 번째는 ‘생성 프로그램’(Generator)이다. 생성 프로그램은 원본 표상, 음성, 영상을 써서 원본과 차이가 없는 가짜 데이터를 만든다. 두 번째는 ‘판별 프로그램

22) Wikipedia, “Deepfake,” available at <https://en.wikipedia.org/wiki/Deepfake> (마지막 확인: 2025년 4월 22일).

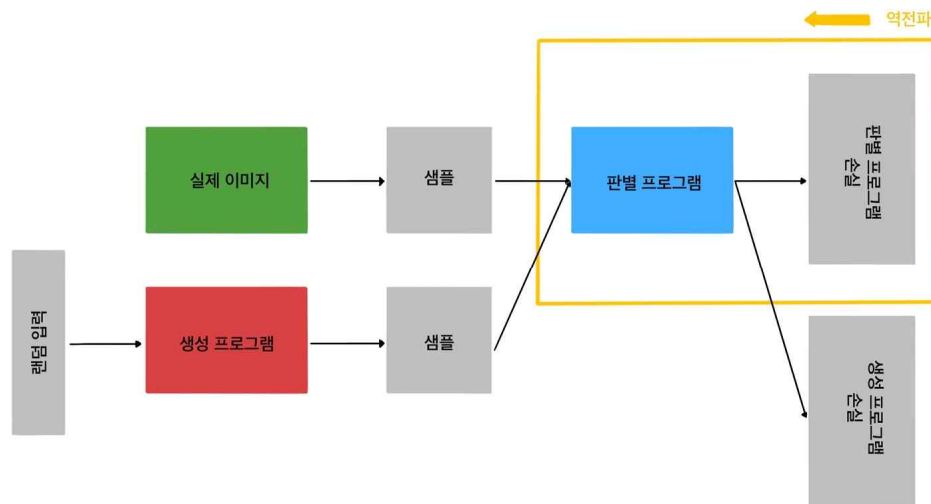
23) ‘deep fake’ means AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful. EU Artificial Intelligence Act, available at <https://artificialintelligenceact.eu/article/3/> (마지막 확인: 2025년 4월 22일).

24) 엄태웅, “[AI기획]경쟁 통해 배우는 인공지능 기술 GAN,” TechM 제46호(2017년 2월), available at <https://www.techm.kr/news/articleView.html?idxno=3617> (마지막 확인: 2025년 4월 22일).

25) 생성적 적대 신경망(GAN)은 기계 학습 프레임 워크의 한 종류인데, 생성적 인공지능에 접근하기 위한 중요한 프레임 워크이다. Wikipedia, “Generative adversarial network,” available at https://en.wikipedia.org/wiki/Generative_adversarial_network (마지막 확인: 2025년 4월 22일).

(Discriminator)이다. 판별 프로그램은 가짜 데이터를 써서 원본과 차이점을 찾는다.

[그림2]²⁶⁾



[그림 2 : 판별 프로그램 훈련에서 역전파]

생성적 적대 신경망은 생성 프로그램이 만든 데이터 가운데 일부 가짜 데이터를 판별 프로그램이 거부하는 과정을 되풀이하는 학습 과정으로 원본 데이터와 거의 똑같은 결과를 얻는다. 왜냐하면 생성적 적대 신경망을 사용한 딥페이크 생성 기술은 생성 프로그램이 판별 프로그램을 속여서 가짜 데이터를 실제 데이터와 좀 더 비슷하게 만들고, 사람이 눈으로는 구별할 수 없게 만들도록 하기 때문이다.²⁷⁾

3. 기술적 해결 방안

기술과 법이라는 주제에는 늘 고민거리가 있다. 문제는 기술이 대부분 앞서가고,²⁸⁾ 법이 빠른 기술에 늦게 대응한다는 점이다. 그런데 글쓴이는 기술의 문제를 기술로 막을 수 있다면, 가장 좋은 대처 방안으로 여긴다.²⁹⁾ 딥페이크의 핵심 문제는 ‘사실’과 ‘진

26) Google Machine Learning, “The Discriminator,” available at <https://developers.google.com/machine-learning/gan/discriminator> (마지막 확인: 2025년 4월 22일).

27) 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 274쪽.

28) 독일의 경우, 기술에 앞서서 암호법을 제정하였다. 그러나 새로운 기술이 등장하면서 제정한 암호법이 기능을 잃어버린 사례도 있다.

29) 가장 중요한 출발점은 사회 공동체 구성원의 인식이고, 그러한 것을 하지 않는다는 부분에 대한

실'에 대한 도전이다. '가짜'를 '진짜'로 신뢰하게 하는 문제이다.³⁰⁾ 이런 경우에는 사회 전반에 불신을 일으켜서, '진짜'를 '가짜'로 여기는 부작용도 생긴다. 그런데 이 문제를 기술로 해결할 수 있으면, 법 정책이나 법 해석의 다른 과정으로 나아갈 필요성이 적다.

딥페이크로 만든 미디어의 양은 6개월마다 두 배씩 증가하고, 이 속도는 점점 빨라진다.³¹⁾ 딥페이크 문제를 해결하기 위한 기술 방안은 역설계(reverse-engineering)³²⁾처럼 딥페이크를 방지하는 방안이다. 이에 대한 출발점은 딥페이크 탐지 기술이다. AWS, 페이스북, 마이크로소프트와 학계는 힘을 합하여 딥페이크 탐지 프로젝트를 만들었다. 이 프로젝트에서 승자는 미화 1,000,000달러³³⁾를 상금으로 받는다.³⁴⁾ 이와 같은 딥페이크 탐지 기술에 대한 도전 프로젝트를 운영하고 있다는 것은, 아직 딥페이크 탐지 기술의 성과가 높지 않음을 말해준다.³⁵⁾

Ⅲ. 딥페이크에 대한 법적 규율과 사생활 보호

딥페이크는 '가짜'를 '진짜'처럼 보이도록 해서, 가짜를 진짜로 신뢰하게 하는 문제를 발생한다. 앞서 설명한 것처럼, 생성형 인공지능으로 만든 딥페이크 기술은 공정성을 훼손하고, 동일성에 혼란을 일으킨다. 공정성 훼손의 문제는 선거와 관련된 사항³⁶⁾이고,

‘공동체의 합의’가 중요하다.

30) 어떤 사람이 의사능력이 있는지를 확인하는 데 드는 시간과 비용이 커서, 이를 하나의 기준으로 만든 제도를 ‘행위능력’ 제도라고 설명한다. 이는 행위능력제도가 갖는 법경제학적 분석이다. 그런데 딥페이크의 경우에는 ‘가짜’를 ‘진짜’로 신뢰하게 하고, ‘진짜’를 ‘가짜’로 불신하게 하는 사회의 비용을 증가시키는 문제를 발생시킨다.

31) Golden, “Anti-deepfake technology,” available at https://golden.com/wiki/Anti-deepfake_technology-R9VDX65 (마지막 확인: 2025년 4월 22일).

32) ‘분해 공학’이라고도 한다. 예를 들어, 현대자동차 측에서 테슬라 자동차를 분해하여 전기차의 새로운 기술을 습득하는 방식이다.

33) 2025년 4월 22일 환율로 14억 2,660만 원이다.

34) MIT Media lab, “Detect DeepFakes: How to counteract misinformation created by AI,” available at <https://www.media.mit.edu/projects/detect-fakes/overview/> (마지막 확인: 2025년 4월 22일).

35) 이러한 딥페이크의 부작용을 기술로 해결할 수 있다면, IT 기업에 이러한 기술을 사용하는 것을 의무로 부과할 수 있을 것이다.

36) 공직선거법 제82조의8은 2023년 12월 28일에 신설했다. 현행(법률 제20902호, 2025년 4월 1일 시행) 공직선거법 제82조의8이 “딥페이크 영상 등을 이용한 선거운동”이라는 제목으로 다음을 규정한다. ① 누구든지 선거일 전 90일부터 선거일까지 선거운동을 위하여 인공지능 기술 등을 이용하여 만든 실체와 구분하기 어려운 가상의 음향, 이미지 또는 영상 등(이하 “딥페이크 영상 등”이라 한다)을 제작·편집·유포·상영 또는 게시하는 행위를 하여서는 아니 된다.

동일성 혼란에 대해서는 음란물, 명예훼손과 사생활 침해의 문제가 발생한다. 그런 측면에서 딥페이크에 대한 국내 연구는 딥페이크 기술과 형사법적 측면에 집중한다. 여기서는 앞에서 밝히는 것처럼 딥페이크 기술이 발생시키는 사생활 보호의 측면을 살핀다.

1. 사생활의 뜻과 유형

비교법 측면에서 사생활(privacy)은 두 가지 뜻을 갖는다. 첫 번째는 대륙법의 관점에서 ‘사생활’³⁷⁾은 여러 인격권 가운데 하나의 요소로 인식한다. 두 번째, 미국법의 관점에서 ‘사생활 권리’는 대륙법의 인격권 개념과 견줄 수 있는 넓은 개념이다. 그러므로 어떤 의미로 ‘사생활’을 쓰는지에 따라서, 그 인식의 범위가 달라진다. 이곳에서는 인공지능 기술이 대체로 미국의 거대 IT산업을 기반으로 발달했기 때문에, 딥페이크를 규율하는 측면에서 미국 ‘사생활 권리’에 바탕을 두어서 논의한다.

미국법에서 사생활권(Right of Privacy)은 쿨리(Cooley) 판사가 1888년에 정립한 ‘혼자 있을 권리’와 사무엘 워런(Samuel D. Warren II)과 루이스 브랜다이스(Louis Brandeis)가 1890년 논문에서 정립한 개념에 기초한다.³⁸⁾ 이에 따라서 미국 불법행위 보통법전집³⁹⁾ 제2판 제652A조는 “일반 원칙”이란 표제로 “사생활 침해의 네 가지 유형”을 규정한다.⁴⁰⁾

(1) 다른 사람의 사생활 권리를 침해한 사람은 그로 인해 다른 사람의 이익에 손해를 끼친 것에 대한 책임을 져야 한다.

(2) 다음과 같은 경우, 사생활 권리를 침해한 것이다.

(a) 제652B조⁴¹⁾가 규정한 바와같이 상당하지 않게 다른 사람의 은둔을 침해한 경

② 누구든지 제1항의 기간이 아닌 때에 선거운동을 위하여 딥페이크 영상 등을 제작·편집·유포·상영 또는 게시할 때는 해당 정보가 인공지능 기술 등을 이용하여 만든 가상의 정보라는 사실을 명확하게 인식할 수 있도록 중앙선거관리위원회규칙으로 정하는 바에 따라 해당 사항을 딥페이크 영상 등에 표시하여야 한다.

37) 법무부 민법개정안은 제3조의2(인격권) 제1항은 “사람은 생명, 신체, 건강, 자유, 명예, 사생활, 성명, 초상, 음성, 개인정보 그 밖의 인격적 이익에 대한 권리를 가진다”고 규정한다. 이에 대해서는 김훈주, “민법상 인격권 신설에 대한 검토”, 재산법연구 제40권 제4호, 2024, 198-212쪽 참조.

38) W. Page Keeton, Prosser and Keeton on Torts 849쪽 (West 5th ed 1988) 참조. 또한 고세일, “정신적 손해에 대한 배상청구권의 법적 기초와 인정영역,” 고려대학교 석사학위논문, 1995, 103-105쪽 참조.

39) 보통법(common law)를 정리했다는 측면에서 이 논문에서는 “보통법전집”으로 쓴다. 이에 대해서는 조규창, 비교법(상), 소화, 2005, 451쪽 참조.

40) Restatement (Second) of Torts § 652A (Am. Law. Inst. 1977). 자세한 내용에 대해서는 고세일, “미국 불법행위법상 사생활 보호에 대한 고찰,” 법학연구 제36권 제2호, 2025 참조.

41) 미국 불법행위 보통법전집 제652B조는 “은둔에 대한 침해”라는 표제로 다음을 규정한다. “다른

우, 또는

(b) 제652C조⁴²⁾가 규정한 바와같이 다른 사람의 이름·초상을 유용⁴³⁾한 경우, 또는

(c) 제652D조⁴⁴⁾가 규정한 바와같이 상당하지 않게 다른 사람의 사생활을 공개한 경우⁴⁵⁾, 또는

(d) 제652E조⁴⁶⁾가 규정한 바와같이 상당하지 않게 다른 사람을 대중에게 잘못 보이도록 공개한 경우.”

미국 불법행위 보통법전집 제652A조가 ‘사생활 침해’로 인정하는 유형은 네 가지이다. 제652A조 (a)항과 (c)항 일반적인 ‘사생활 침해’로 인식하는 영역이다. 이는 대륙법계 국가에서 인정하는 사생활 영역과도 같다.⁴⁷⁾ 반면에 제652A조 (b)항과 (d)항은 대륙

사람의 고독이나 은둔 또는 사적인 사항이나 관심사를 물리적 또는 그 밖의 방법으로 고의로 침해한 사람은 그 사생활 침해가 합리적인 인간에게 매우 불쾌감을 줄 수 있는 경우 사생활 침해에 대해서 그 다른 사람에게 책임을 져야 한다.” Restatement (Second) of Torts § 652B (Am. Law. Inst. 1977).

42) 미국 불법행위 보통법전집 제652C조는 “성명·초상의 유용”이라는 표제로 다음과 같이 규정한다. “다른 사람의 성명·초상을 자신의 사용하거나 이익을 얻으려고 유용한 사람은 다른 사람의 사생활 침해에 대해서 책임진다.” Restatement (Second) of Torts § 652C (Am. Law. Inst. 1977).

43) ‘유용’(appropriation)은 영업상 이익을 위하여 사용하는 영리적 이용을 뜻한다. 그런 측면에서 ‘퍼블리시권’과 연결될 가능성이 있다. 다만 미국 학계에서도 전통적으로 불법행위 학자들은 퍼블리시권을 인정하는데 소극적이다. 반면에 지적재산권 학계에서는 퍼블리시티권을 긍정한다. 퍼블리시티권에 대한 논의에 대해서는 권태상, 퍼블리시티권의 이론적 구성 - 인격권에 의한 보호를 중심으로, 경인문화사, 2013 참조.

44) 미국 불법행위 보통법전집 제652D조는 “사생활 공개”라는 표제로 다음을 규정한다. “다른 사람의 사생활에 관한 사항을 공개한 사람은 공개 사항이 다음과 같은 경우, 그 사람의 사생활 침해에 책임을 진다. (a) 합리적인 인간에게 매우 불쾌감을 주고, 또한 (b) 대중의 정당한 관심사가 아닌 경우.” Restatement (Second) of Torts § 652D (Am. Law. Inst. 1977).

45) 국가인권위원회는 의무 사관후보생 모집 신체검사항 때 병력(病歷)을 공개한 사건에 대해서 “개인의 질병은 지극히 사적인 영역으로 다른 사람이 알게 될 때 사생활의 비밀이 침해될 수 있음에도 불구하고, 정신질환 관련 질병 가능성을 다수의 동종업종 종사자에게 누설한 것은 사생활의 비밀 침해”라고 판단하였다(인권위 2009. 10. 26. 09진인590).

46) 미국 불법행위 보통법전집 제652E조는 “사람을 잘못 보이게 하는 공개”라는 표제로 다음과 같이 규정한다. “다른 사람에 관한 사항을 대중 앞에 잘못되게 보이도록 공개한 사람은, 다음의 경우에 사생활 침해에 대해서 책임진다. (a) 다른 사람을 잘못 보이게 한 것이 합리적인 인간에게 매우 불쾌감을 주고, 또한 (b) 피고가 공개된 사항이 잘못된 것임과 그 다른 사람을 잘못 보이게 한 것을 알았거나 현저하게 부주의하여 행동한 경우.” Restatement (Second) of Torts § 652E (Am. Law. Inst. 1977).

47) 현행(법률 제20908호, 시행 2025년 4월 8일) 형법 제316조는 “비밀 침해”라는 표제로, 제1항이 “불합 기타 비밀 장치한 사람의 편지, 문서 또는 도화를 개봉한 자는 3년 이하의 징역이나 금고 또는 500만원 이하의 벌금에 처한다”라고 규정한다. 또한 제2항이 “불합 기타 비밀 장치한 사람

법계 국가에서는 사생활 침해가 아니라 다른 인격권으로 보호하는 이익이다. 그런데 딥페이크 기술과 관련해서는 미국 불법행위 보통법전집 제652A조가 규정하는 네 유형과 모두 관련될 수 있다. 그러나 ‘가짜’를 ‘진짜’로 신뢰하게 한다는 측면에서 미국 불법행위 보통법전집 제652A조 (d)항의 사생활 보호와 관련성이 높다. 다시 말해서, 불법행위 보통법전집 제652E조가 규정하는 것처럼 상당히 많게 다른 사람을 대중에게 잘못 보이도록 공개한 경우가 ‘사생활 침해’가 된다.

2. 딥페이크와 관련한 사생활 보호

전통적인 관점에서는 미국 불법행위 보통법전집 제652A조 (d)항이 딥페이크와 관련한 사생활 보호 영역이다. 각국에서 딥페이크 기술의 위험성을 인식하고 이를 규율하려 한다. 다만 딥페이크 기술과 관련해서 이에 대한 법적 규율은 초기 상태이다. 미국도 크게 다르지 않다. 그러나 연방법과 각 주에서 딥페이크를 규율하려는 여러 법안을 제안한 상태이다.⁴⁸⁾

그런데 누구나 딥페이크 기술을 써서 콘텐츠를 조작할 수 있다.⁴⁹⁾ 또한 인터넷상에서 딥페이크 기술을 쓴 콘텐츠를 쉽게 배포할 수 있다. 또한 인터넷상 배포된 내용물에 대한 2차, 3차와 같은 끊임없이 가공할 수 있다.⁵⁰⁾ 그런 측면에서 딥페이크 기술을 써서 내용물을 만든 사람과 유포한 사람을 찾아서 책임을 귀속시키는 과정에 어려움이 있다.⁵¹⁾ 그리고 딥러닝과 같은 인공지능의 경우 결과물은 확인할 수 있지만, 데이터를 학습하고 결정하는 과정을 이해하거나 설명하기 어렵다.⁵²⁾ 그런 측면에서 설명 가능한 인공지능을 요구하고 있지만, 설명 가능한 인공지능의 기술도 성숙한 단계가 아니다. 그

의 편지, 문서, 도화 또는 전자기록 등 특수매체기록을 기술적 수단을 이용하여 그 내용을 알아낸 자도 제1항의 형과 같다”고 규정한다. 이는 개인의 사생활의 비밀(평온)을 침해하는 범죄로 처벌함으로써 형법상 사생활 보호를 추구하는 것이다. 배종대, 형법각론 제10전정판, 홍문사, 2018, 221쪽과 김일수·서보학, 새로 쓴 형법각론 제9판, 박영사, 2018, 190쪽 참조.

48) 딥페이크에 대한 미국 연방법과 주법의 제정 상황에 대해서는 Michelle M. Graham, “Deepfakes: Federal and state regulation aims to curb a growing threat,” (2024년 6월 26일) *available at* <https://www.thomsonreuters.com/en-us/posts/government/deepfakes-federal-state-regulation/> (마지막 확인: 2025년 4월 29일).

49) 그런 측면에서 딥페이크 기술로 쉽게 공정성과 동일성 침해를 하게 하는 개발자에게 법적 책임을 귀속하게 할 가능성이 있다.

50) 이런 측면에서 딥페이크 기술을 써서 내용물을 제작한 사람과 배포한 사람은 피해자에게 공동불법행위책임을 진다(민법 제760조).

51) 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 277쪽.

52) 인공지능의 블랙박스 측면때문에 이런 결과가 발생한다.

런 측면에서 이용자가 원하는 결과물을 낚는 인공지능과 설명 가능한 인공지능의 두 가지 모델을 쓰는 방식을 제안하기도 한다.⁵³⁾

3. 법적 규율

(1) 미국 연방법

1) 미국 연방 도전법

미국법은 사법(私法)인 경우, 연방 정부가 개입하지 않는 것이 연방 헌법의 원리이다. 따라서 딥페이크 관련 법도 주법(州法)을 기준으로 한다. 다만 사기, 기망 같은 경우에는 연방 정부가 개입한다. 딥페이크 관련해서는 ‘도전법’(The Defiance Act of 2024)⁵⁴⁾이 있다. 이는 ‘혼란스러운 명시적으로 위조된 표상과 비동의 편집법’(Disrupt Explicit Forged Images and Non-Consensual Edits Act)을 두문자(頭文字)로 줄인 것이다.⁵⁵⁾

이 법의 목적을 다음과 같이 규정한다. 2024년 양당이 합의한 동법은 동의 없이 성적으로 음란한 ‘딥페이크’ 표상과 영상을 배포한 것에 책임이 있는 사람에게 책임을 귀속시킨다. 또한 동법은 딥페이크를 다음과 같이 정의한다. “딥페이크는 인공지능을 통해 만들어진, 완전히 사실적으로 보이는 영상이나 표상이다. 다른 사람을 묘사한 실제 영상에 특정 인물의 모습을 합성하거나, 실제 행동이나 발언이 아닌 다른 행동을 하거나 말하는 것처럼 표현된 완전히 독창적인 콘텐츠로 구성될 수도 있습니다. 어떤 경우든 이러한 딥페이크는 거짓(hoaxes)⁵⁶⁾이다.”

도전법은 디지털 위조가 피해자를 나체 상태나 성적 행위에 참여하는 모습으로 묘사하면 민사소송을 제기할 수 있도록 규정한다. 민사상 구제책으로 딥페이크를 배포할 의도로 제작하거나 소유한 사람, 피해자의 동의 없이 딥페이크가 제작되었음을 알고 있거나 이를 무시한 채 배포하고 받은 사람에 대해서 책임을 부과한다.⁵⁷⁾ 이에 대한 소멸시

53) 이용자가 원하는 결과물을 제공하는 인공지능은 빠른 처리 속도를 갖고, 설명 가능한 인공지능의 경우에는 그 처리 속도가 매우 느리다. 김효은, “인공지능과 윤리,” in, 한국인공지능법학회, 인공지능과 법, 박영사, 2020, 64-65쪽 참조.

54) 영어 defiance는 ‘저항,’ ‘도전’이라는 뜻이 있다. 이는 위조된 표상과 동의를 받지 않는 상황에 대한 저항과 도전이라는 뜻을 담고 있다. 따라서 이 글에서는 ‘도전법’으로 옮긴다.

55) US Congress, S.3696 - Defiance Act of 2024, available at <https://www.congress.gov/bill/118th-congress/senate-bill/3696/text> (마지막 확인: 2025년 4월 22일). 이 법은 2024년 7월 23일에 상원에서 통과되었다.

56) 영어 hoax의 뜻은 “사실이 아닌 것을 사실처럼 거짓으로 꾸민다”라는 뜻의 우리 말 ‘날조’(捏造)에 가깝다.

효는 10년이다.⁵⁸⁾ 피해자가 비동의 성적 딥페이크 표상을 안 시점이나 18세가 된 시점부터 시작한다.⁵⁹⁾ 또한 소송 중 피해자의 사생활 보호를 위한 조치를 마련하고 있다.

2) 연방 도전법의 사생활 보호

동법은 사생활 보호와 관련해서는 다음 두 가지를 규정한다.

첫째, 연방의회는 (4) 피해자의 사생활이 침해되고 피해자는 자신의 모습과 신원에 대한 통제력을 잃기 때문에, 피해자가 자신의 의지와 상관없이 성적 디지털 위조물에 묘사될 때마다 피해자는 불안정해질 수 있음을 안다.⁶⁰⁾

둘째, (4) 개인정보 보호. - 본 조항에 따라 제기된 민사소송에서 법원은 다음을 포함하여 원고의 개인정보를 보호하기 위한 명령을 발부할 수 있다.

(A) 원고가 가명을 사용하도록 허용;

(B) 당사자들에게 공개 문서에서 원고의 개인 식별 정보를 삭제하거나 해당 문서를 봉인하여 제출하도록 요구; 또한

(C) 증거개시⁶¹⁾를 위한 보호명령 발부. 여기에는 모든 사적인 영상 묘사 또는 디지털 위조물이 법원의 관리, 보관과 통제 아래에 있어야 함을 명하는 명령을 포함할 수 있음.⁶²⁾

3) 연방 도전법의 민사 구제수단

미국 연방 도전법의 사생활 보호는 다음과 같은 민사 구제수단을 통해서 그 실효성을 확보하고자 한다. 동법은 구제수단으로 손해배상과 금지청구를 인정하면서, 징벌적 손해배상과 기타 구제수단으로 다음을 규정한다. “법원은 법률에 따라 가능한 기타 구

57) 여러 국가의 딥페이크 법제에 대해서는 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 279-282쪽.

58) 현행 민법 제766조는 가해자를 안 날로부터 3년(제1항), 불법행위를 한 날로부터 10년으로 규정한다. 그런데 미국 도전법은 불법행위를 안 날로부터 10년으로 규정한다. 또한 민법 제766조 제3항은 미성년자가 성폭력, 성추행, 성희롱, 그 밖의 성적(性的) 침해를 당한 경우에 이로 인한 손해배상청구권의 소멸시효는 그가 성년이 될 때까지는 진행되지 아니한다고 규정한다. 그 내용은 미국 도전법의 소멸시효와 취지가 같다.

59) (6) 소멸시효. - 본 조항에 따라 제기된 모든 소송은 다음 중 더 늦은 날짜로부터 10년 이내에 제기되지 않는 한 소멸한다. (A) 식별 가능한 개인이 청구의 근거가 되는 위반 사항을 합리적으로 발견한 날짜; 또는 (B) 식별 가능한 개인이 18세가 된 날짜.

60) S. 3696 제2조(Section 2) Findings (4).

61) Discovery를 뜻한다.

62) S. 3696 제3조(Section 3) 내밀한 표상 공개에 대한 민사소송(Civil Action Relating to Disclosure of Intimate Images).

제수단 외에도 징벌적 손해배상을 판정하거나 형평법상 구제수단을 명령할 수 있다. 여기에는 피고에게 내밀한 시각적 묘사나 디지털 위조물의 삭제, 파괴 또는 전시나 공개를 중단하도록 명령하는 임시 금지 명령, 가처분 명령 또는 영구 금지 명령이 포함한다.”⁶³⁾

또한 동법은 손해배상액으로 다음과 같이 규정한다.⁶⁴⁾

손해배상.-(A)(i)항의 목적상, 식별 가능한 개인은 다음을 배상받을 수 있다.

(i) 다음 금액의 예정 손해배상액⁶⁵⁾:

(I) 150,000달러⁶⁶⁾; 또는

(II) 청구에서 문제가 되는 행위가 다음과 같은 경우 250,000달러.⁶⁷⁾

(aa) 피고가 식별 가능한 개인을 상대로 한 실제 또는 시도된 성폭행, 스토킹 또는 괴롭힘과 관련하여 저지른 경우; 또는

(bb) 어떤 사람이 식별 가능한 개인을 상대로 한 실제 또는 시도된 성폭행, 스토킹 또는 괴롭힘의 직접적이고 직접적인 원인; 또는

(ii) 개인이 입은 실제 손해. 여기에는 청구에서 문제가 되는 행위로 발생한 피고의 이익 중 실제 손해액 계산에 고려되지 않는 이익을 포함한다.

(2) 미국 주법으로서 캘리포니아 법

1) 캘리포니아 형법

미국의 여러 주에게 딥페이크와 관련한 법을 제안하고 있다. 캘리포니아, 텍사스, 버지니아 등 소수의 주에서 딥페이크 관련 법률을 통과시켰다.⁶⁸⁾ 여기서는 딥페이크에 관

63) S. 3696 제3조(Section 3) 내밀한 표상 공개에 대한 민사소송 (b) 민사소송.

64) S. 3696 제3조(Section 3) 내밀한 표상 공개에 대한 민사소송 (b) 민사소송 (4)(C).

65) 원문은 “liquidated damages in the amount of”이다.

66) 2025년 4월 22일 환율로 2억 1,363만 원이다.

67) 2025년 4월 22일 환율로 3억 5,605만 원이다. 우리 법원이 법익에 따라서 인정되는 위자료의 액수가 다르다. 판례 실무에서는 법익 사이에 순위의 차이가 발생한다. 이는 생명, 중한 신체·건강 침해, 중한 성적 자기결정권 침해, 중한 명예 침해, 경한 성적 자기결정권 침해, 사생활 침해, 생활이익과 개인정보 침해의 순서로 평가한다. 그런 측면에 사생활 침해에 대해서 인정되는 위자료 액수는 상대적으로 낮다. 이동진, 위자료 산정의 적정성에 관한 사법정책연구, 서울대학교 산학협력단, 2013, 58-59쪽. 이는 유럽연합 인공지능법이 인권과 민주주의에 잠재적인 위협이 될 수 있는 분야에 인공지능 기술을 쓰는 것을 금지하고, 이를 위반할 경우에 최대 3,500만 유로를 부과하는 현실과 대비된다. 3,500만 유로의 2025년 4월 22일 환율은 573억 7,655만원이다.

68) Lourdes Vazquez, *Recommendations For Regulation Of Deepfakes In The U.S.: Deepfake Laws Should Protect Everyone Not Only Public Figures*, Robert D. Reif Fellowship Article,

한 최초의 주법⁶⁹⁾인 캘리포니아 주법을 중심으로 설명한다. 캘리포니아 주는 선거 개입부터 성 착취에 이르기까지 다양한 문제에 대한 입장을 밝히며 미국 내 딥페이크 규제 의 새로운 선두 주자로 부상했다. 캘리포니아 주지사 개빈 뉴섬은 2024년에 여러 새로운 AI 법안에 서명했고, 그 가운데 다수는 인공지능을 이용한 딥페이크 제작을 억제하는 것을 목표로 한다.⁷⁰⁾

캘리포니아 상원 법안 제926호 제289장은 주로 형법 규정이다.⁷¹⁾ 그러나 동법은 사생활 보호에 관한 규정을 포함한다. 이 법안에 따라서 캘리포니아 형법 제647조가 개정되었다.⁷²⁾ 캘리포니아 형법 제647조가 규정하는 사생활 보호법익은 미국 불법행위법 제

14쪽 (2021) 참고.

69) K.C. Halm, Ambika Kumar, Jonathan Segal and Caesar Kalinowski IV, Two New California Laws Tackle Deepfake Videos in Politics and Porn (2. 28, 2020), *available at* <https://www.dwt.com/blogs/media-law-monitor/2020/02/two-new-california-laws-tackle-deep-fake-videos-in> (마지막 확인: 2025년 4월 22일).

70) 2024 Year in Review: Navigating California's Landmark Deepfake Legislation (12. 11, 2024), *available at* <https://www.mofo.com/resources/insights/241211-2024-year-in-review-navigating-california-s-landmark-deepfake-legislation> (마지막 확인: 2025년 4월 22일).

71) SB-926 범죄: 내밀한 표상의 배포(Crimes: distribution of intimate images), *available at* https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB926 (마지막 확인: 2025년 4월 22일). 이 법안은 만 18세 이상의 사람이 타인의 식별 가능한 신체 부위를 사실적으로 표현한 사진 이미지, 디지털 이미지, 전자 이미지, 컴퓨터 이미지, 컴퓨터 생성 이미지 또는 기타 그림으로 표현하거나, 표현된 사람이 성교, 남색, 구강성교, 성적 삽입 또는 자위행위를 하는 모습을 묘사한 이미지 또는 표현된 사람이나 표현된 사람이 참여하는 이미지를 의도적으로 제작 및 배포하거나 배포하도록 하는 행위를 범죄로 규정한다. 이러한 이미지는 합리적인 사람이 해당 이미지가 표현된 사람의 진짜 이미지라고 믿을 수 있는 방식으로 제작되었으며, 이미지를 배포하는 사람이 이미지 배포가 심각한 정서적 고통을 유발할 것임을 알고 있거나 알아야 하며, 표현된 사람이 그러한 고통을 겪고 있는 상황에서 제작되었다. 이 법은 범죄의 범위를 확대하여 주에서 의무화한 지역 프로그램을 부과한다.

72) 제647조 b항 5호 j호는 다음과 같이 규정한다. “(j) (1) 잠망경, 망원경, 쌍안경, 카메라, 영화용 카메라, 캠코더, 휴대전화, 전자 기기 또는 무인 항공기 시스템을 포함하되 이에 국한되지 않는 어떠한 기구를 사용하여 침실, 욕실, 탈의실, 탈의실 또는 테닝 부스의 내부 또는 점유자가 합리적으로 사생활을 보장받을 수 있는 기타 공간의 내부를 구멍이나 개구부를 통해 들여다보거나 다른 방식으로 관찰하는 사람. 이는 내부에 있는 사람의 사생활을 침해할 의도를 가지고 있다.

(2) 은닉형 캠코더, 영화 카메라 또는 모든 유형의 사진용 카메라를 사용하여 다른 식별 가능한 사람을 전자적 수단으로 비밀리에 비디오테이프, 필름, 사진 촬영 또는 녹화하여 그 사람이 입고 있는 옷 아래 또는 그 옷을 통해 그 사람의 신체 또는 속옷을 볼 목적으로 그 사람의 동의 또는 인지 없이 그 사람의 정욕, 열정 또는 성적 욕망을 자극, 호소 또는 충족시키고 그 사람의 사생활을 침해할 의도를 가진 사람. 이 단락의 목적상 “식별 가능한”은 피해자를 포함한 누군가가 피해자를 식별하거나 인식할 수 있음을 뜻한다. 피해자의 신원이 실제로 확인될 필요는 없다.

(3) (A) 은닉형 캠코더, 영화 카메라 또는 모든 유형의 사진 카메라를 사용하여, 다른 식별 가능한 사람이 옷을 전부 또는 부분적으로 벗은 상태일 수 있음을 비밀리에 비디오테이프, 필름, 사진 또는 전자적 수단으로 녹화하여, 다른 사람의 동의 또는 인지 없이, 침실, 욕실, 탈의실, 탈의

652A조 (2)(a)의 보호이익과 같다. 그런 측면에서 개정된 캘리포니아 형법 제647조 b항 5호 j호는 전통적인 의미의 사생활 보호 규정이다.

그런데 딥페이크 기술을 통한 사생활 침해에 대해서는 개정된 캘리포니아 형법 제 647조 (4)(B)~(D)가 규정한다.⁷³⁾ 그 가운데 특히 (4)(A)(ii)와 (B)가 중요하다.

“(4)(A)(ii) 타인의 식별 가능한 신체 부위를 사실적으로 표현한 사진 이미지, 디지털 이미지, 전자 이미지, 컴퓨터 이미지, 컴퓨터 생성 이미지 또는 기타 그림으로 표현된 이미지, 또는 묘사된 사람이 성교, 남색, 구강성교, 성적 삽입 또는 자위행위를 하는 모습을 묘사한 이미지 또는 묘사된 사람이 참여하거나 묘사된 사람이 참여하는 이미지를 고의로 제작·배포하거나 배포하도록 하는 사람.⁷⁴⁾ 이러한 이미지는 합리적인 사람이 해당 이미지가 묘사된 사람의 진짜 이미지라고 믿을 수 있는 방식으로 제작되었으며, 이미지를 배포하는 사람이 이미지 배포가 심각한 정서적 고통을 유발할 것임을 알고 있거나 알아야 하며, 묘사된 사람이 그러한 고통을 겪고 있는 상황에서 제작되었다. 이 조항은 해당 사람이 범죄를 저지를 당시 18세 미만이었던 사람에게는 적용되지 않는다.

(B) (i) (A)항에 명시된 이미지를 직접 배포하는 사람은 해당 이미지를 의도적으로 배포한다.

(ii) (A)항에 기술된 이미지가 배포되도록 의도적으로 유도하는 경우, 해당 이미지를 배포하도록 주선하거나, 구체적으로 요청하거나, 의도적으로 다른 사람에게 배포하도록 유도하는 경우.”⁷⁵⁾

실, 태닝 부스 내부 또는 다른 사람이 합리적으로 사생활을 기대할 수 있는 다른 공간의 내부에서 다른 사람의 사생활을 침해할 의도로 하는 사람. 이 단락의 목적상 “식별 가능한”이란 식별이 가능하거나 인식될 수 있음을 뜻하며, 피해자를 포함한 누군가가 피해자를 식별하거나 인식할 수 있음을 뜻한다. 피해자의 신원이 실제로 확인될 필요는 없다.”

73) 개정된 캘리포니아 형법 제647조 (4) (A) (i)는 다음과 같이 규정한다. “다른 식별 가능한 사람의 사적인 신체 부위 이미지 또는 성교, 남색, 구강성교, 성적 침투 또는 자위행위에 참여하는 모습을 묘사한 이미지를 의도적으로 배포하거나 배포하게 하는 사람. 이 경우 해당 인물들은 이미지가 비공개로 유지되어야 한다는 데 동의하거나 이해하며, 이미지를 배포하는 사람은 이미지 배포가 심각한 정서적 고통을 초래할 것이라는 사실을 알고 있거나 알아야 하며, 묘사된 사람이 그러한 고통을 겪는다.”

74) 중요한 법리는 고의로 딥페이크 기술을 써서 표상(이미지)를 제작하거나 배포하는 경우만을 처벌 대상으로 삼고, 과실의 경우는 처벌 대상이 아닌 것이다. 이는 미국 불법행위법상 사생활 침해의 경우에도 같다. 따라서 다른 사람의 사생활을 고의로 침해한 경우만을 사생활 침해에 대한 책임을 귀속시키고 있다. Restatement (Second) of Torts § 652B (Am. Law. Inst. 1977). 이러한 점에서 고의·과실을 불법행위의 성립 요건(민법 제750조)으로 하는 한국 민법과 다르다.

75) 개정된 캘리포니아 형법 제647조 (4) (C)은 다음과 같이 규정한다. “본 항에서 사용되는 다음 용어는 다음과 같은 의미를 가진다. (i) “배포”에는 공개적으로 전시하거나 소유하는 것을 포함한다. (ii) “식별 가능한”은 (2)항과 (3)항과 동일한 의미를 갖는다. (iii) “은밀한 신체 부위”란 생식기, 항문의 모든 부분을 의미하며, 여성의 경우 유륜 상단 아래의 유방의 모든 부분도 포함되며,

또한 캘리포니아주 개정 형법은 딥페이크 기술을 사용하여 표상(이미지)을 배포하는 경우, 처벌되지 않는 경우를 다음과 같이 규정한다.

“(D) 다음 중 어느 하나에 해당하는 경우 (A)항에 기술된 표상(이미지)을 배포하는 것은 본 항을 위반하지 않는다.

- (i) 불법행위를 신고하는 과정에서 배포하는 경우.
- (ii) 법적 절차에 사용하기 위한 소환장 또는 기타 법원 명령에 따라 배포하는 경우.
- (iii) 배포가 합법적인 공적 절차의 일환으로 이루어지는 경우.
- (iv) 배포가 공적 관심사 또는 공익과 관련된 경우. 묘사된 인물이 공인이라는 이유만으로 배포가 공적 관심사 또는 공익과 관련된 사안이 되는 것은 아니다.”

2) 노골적인 성적 디지털 표상법

캘리포니아 주는 소셜 미디어 플랫폼과 관련하여 2024년 9월 19일 주지사가 상원법안 제981호를 승인했다. 이는 노골적인 성적 디지털 표상(이미지)과 관련한 법이다. 동법의 목적은 소셜 미디어 플랫폼을 규율하며, 소셜 미디어 플랫폼은 사용자가 합리적으로 접근할 수 있는 메커니즘을 통해 캘리포니아 거주자가 아동 성 학대 자료라고 합리적으로 판단하는 자료를 소셜 미디어 플랫폼에 신고할 수 있는 수단을 제공하도록 요구하는 것이다. 또한 동법은 생성형 인공지능 콘텐츠를 통한 리벤지 포르노법을 확대하여 피해자와 법 집행기관에 이러한 콘텐츠를 단속할 수 있는 추가적인 수단을 제공한다. 그밖에 새로운 법적 구제수단은 명예훼손, 허위 정보에 따른 사생활 침해, 정신적 고통 유발, 퍼블리시티권 침해 등 피해자가 사용할 수 있는 기존 법적 수단을 보완하는 것으로 평가한다.⁷⁶⁾

동법은 제1조, 제22.7장(제22670조부터 시작)을 사업 및 직업법 제8부에 다음을 추가한다.

제22.7장. 디지털 신원 도용

22670. (a) (1) “적용 대상 자료”는 다음 모든 기준을 충족하는 자료를 뜻한다.

(A) 해당 자료는 디지털화를 통해 생성 또는 변경된 이미지 또는 비디오로서, 합리적인 사람에게는 다음 중 하나의 이미지 또는 비디오로 보일 수 있는 자료이다.

- (i) 식별 가능한 사람의 은밀한 신체 부위.

옷을 통해 가려져 있거나 명확하게 보이는 경우.”

76) California Legislative Information, SB-981 Sexually explicit digital images, available at https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB981 (마지막 확인: 2025년 4월 22일).

(ii) 성교, 남색, 구강성교 또는 성적 삽입 행위에 관여하는 식별 가능한 사람.

(iii) 자위행위에 관여하는 식별 가능한 사람.

(B) 신고자는 자료에 묘사된 인물이며, 신고자는 자료에 신고자의 초상이 사용되는데 동의하지 않았다.

(C) 해당 자료가 소셜 미디어 플랫폼에 표시, 저장 또는 호스팅된다.

(2) “해당 자료”에는 이미지의 밝기 또는 대비 변경, 이미지 또는 비디오의 내용에 영향을 미치지 않는 기타 사소한 변경을 포함하여 콘텐츠의 인식된 내용이나 의미에 중대한 변화를 초래하지 않는 사소한 변경만 포함된 이미지 또는 비디오는 포함하지 않는다.

(b) “신고 사용자”는 소셜 미디어 플랫폼이 제22671조에 따라 제공하는 메커니즘을 사용하여 소셜 미디어 플랫폼에 자료를 신고하는 자연인을 뜻한다.

(c) “노골적인 성적 디지털 신원 도용”은 소셜 미디어 플랫폼에 해당 자료를 게시하는 것을 뜻한다.

(d) (1) “소셜 미디어 플랫폼”은 (2)항에 규정된 경우를 제외하고 제22675조에 정의된 것과 동일한 의미를 갖는다.

또한 캘리포니아 주법은 소셜 미디어 플랫폼 사업자에게 딥페이크 기술을 기초로 한 사생활 침해 방지를 위한 의무 사항을 규정한다. 다음과 같이 규정한다.

“22671. 소셜 미디어 플랫폼은 다음의 모든 사항을 준수해야 한다.

(a) 소셜 미디어 플랫폼에 계정을 보유한 캘리포니아 거주자이자 신고 사용자가 성적으로 노골적인 디지털 신원 도용을 소셜 미디어 플랫폼에 신고할 수 있도록 합리적으로 접근 가능한 메커니즘을 제공한다.

(b) 소셜 미디어 플랫폼이 성적으로 노골적인 디지털 신원 도용 사례를 파악하고 신고 사용자에게 다음 두 가지 정보를 모두 제공하여 연락할 수 있도록 합리적으로 충분한 정보를 수집한다.

(1) 소셜 미디어 플랫폼이 신고 사용자의 신고를 접수한 후 48시간 이내에 신고를 접수했다는 확인.

(2) (1)항에 따라 요구되는 확인이 발급된 날로부터 7일 이내에, 신고된 성적으로 노골적인 디지털 신원 도용에 대한 소셜 미디어 플랫폼의 처리 현황을 신고 사용자에게 서면으로 업데이트한다.

(c) (1) (2)항에 규정된 경우를 제외하고, (b)항 (1)에 따라 요구되는 확인이 발급된

날로부터 30일 이내에, 신고된 성적으로 노골적인 디지털 신원 도용이 성적으로 노골적인 디지털 신원 도용이라고 믿을 만한 합리적인 근거가 있는지 여부를 판단해야 한다.

(2) (A) 소셜 미디어 플랫폼이 소셜 미디어 플랫폼의 합리적인 통제 범위를 벗어난 상황으로 인해 30일 이내에 (1)항을 준수할 수 없는 경우, 소셜 미디어 플랫폼은 해당 자료가 처음 신고된 날로부터 60일 이내에 (1)항을 준수해야 한다.

(B) (A)항이 적용되는 경우, 소셜 미디어 플랫폼은 (b)항에 따라 신고 사용자로부터 수집한 정보를 사용하여 지연 발생 가능성을 인지한 시점으로부터 48시간 이내에 지연 사실을 신고 사용자에게 서면으로 즉시 통지해야 한다.

(d) (1) (c)항에 따른 결정이 있을 때까지 신고된 성적으로 노골적인 디지털 신원 도용 사례가 소셜 미디어 플랫폼에서 공개적으로 표시되지 않도록 일시적으로 차단한다.

(2) 소셜 미디어 플랫폼에서 신고된 성적으로 노골적인 디지털 신원 도용 사례가 성적으로 노골적인 디지털 신원 도용이라고 믿을 만한 합리적인 근거가 있다고 판단하는 경우, 신고된 성적으로 노골적인 디지털 신원 도용 사례를 소셜 미디어 플랫폼에서 공개적으로 표시되지 않도록 즉시 삭제한다.

4. 우리 법에 대한 시사점

지금까지 딥페이크와 사생활 보호에 대한 미국법의 규율을 살펴보았다. 미국법에서도 딥페이크에 대한 법적 규율은 걸음마 단계로 평가할 수 있다. 미국법은 딥페이크와 관련하여 강력한 처벌을 수반하는 ‘형사책임’에 주안점을 둔다. 이는 딥페이크가 ‘고의로 행한 범죄’라는 사회 인식을 심어주고, 딥페이크로 다른 사람의 사생활을 침해하는 경우, 그 가해자의 범죄행위 처벌을 공동체 구성원들에게 강력하게 인식하게 한다.⁷⁷⁾

(1) 미국 연방 도전법의 시사점

미국 연방법인 ‘도전법’(The Defince Act of 2024)은 피해자에게 민사소송을 제기할 수 있도록 규정한다. 그리고 이에 대한 주요 법리는 ‘피해자의 동의’를 받지 않은 것이다. 이는 앞서 설명한 미국 불법행위 보통법전집 제652C조와 제652D조의 경우와도, 피해자의 동의를 핵심 법리로 설명한다.⁷⁸⁾ 또한 미국 불법행위 보통법전집 제652F조⁷⁹⁾가

77) 현행 우리의 양형기준은 허위영상물의 편집과 반포의 경우, 영리 목적이 아니면 양형기준을 6월-1년 6월로 설정한다. 따라서 아동·청소년성착물이 아닌 경우, 딥페이크에 대한 책임 인식을 낮게 설정한다. 양형위원회, 2024년 양형기준, 양형위원회, 2024, 200-203쪽.

78) Restatement (Second) of Torts § 652D cmt b illus. 6, c illus. 11 and f. (Am. Law. Inst.

규정하는 ‘절대적 면책특권’의 경우에도 피해자의 동의가 핵심 법리로 작동한다.⁸⁰⁾

그런데 미국 연방 도전법은 피해자의 보호를 최선책으로 규정하고 있다. ① 딥페이크의 피해자를 위해서 ‘가명’ 사용을 허용한다. ② 피해자의 개인 식별 정보를 삭제하거나 해당 문서를 봉인하여 제출하도록 요구한다. ③ 법원이 피해자의 증거개시를 위한 보호 명령을 발부하도록 한다. 이런 측면에서 피해자의 실질적 보호를 강화한다. 따라서 딥페이크 사진과 영상을 통해서 정신적 충격을 받는 피해자를 위한 실질적인 보호막을 마련하고 있는 셈이다. 우리 법에서도 피해자를 위한 실질적인 보호조치를 강화해야 한다.

또한 미국 연방 도전법은 딥페이크의 피해자를 실질적으로 보호하는 방안으로 가해자에 대한 징벌적 손해배상을 규정한다. 그 금액으로 150,000달러와 250,000달러를 법으로 정하고 있다. 딥페이크의 경우에는 징벌적 손해배상의 요건을 모두 충족하는 경우가 대부분이다.⁸¹⁾ 따라서 딥페이크의 경우에는 여러 특별법에서 인정하는 징벌적 손해배상액을 가해자에게 부과하여, 공동체 구성원들에게 딥페이크가 피해자에게 가슴 아픈 정신적 고통이 될 수 있다는 위험성을 알려야 한다.⁸²⁾

(2) 캘리포니아 형법의 시사점

미국은 여러 주법에서 딥페이크를 규율하는 법을 제정하는 과정에 있다. 그 가운데 최초의 주법인 캘리포니아 형법은 몇 가지 중요한 내용을 담고 있다. 미국법은 ‘주제별(topic) 접근 방식’을 채택하고 있다. 따라서 공법과 사법을 엄격하게 분리하지 않고 기능적으로 관련된 문제를 해결할 수 있는 모든 영역을 동원해서 해결 방안을 찾는다.⁸³⁾

1977).

79) 미국 불법행위 보통법전집 제652F조는 “제583조~제592조에서 규정한 명예훼손 사안을 공개할 수 있는 절대적 면책특권에 관한 규정은 사생활 침해 사안을 공개한 경우에도 적용한다”고 규정한다.

80) Restatement (Second) of Torts § 652F cmt b. illus. 2 (Am. Law. Inst. 1977). 또한 미국 캘리포니아 주 형법에서 규정하는 ‘노골적인 성적 디지털 표상법’에서도 ‘동의’를 핵심적인 법리로 규정한다.

81) 미국 불법행위법상 징벌적 손해배상 전반에 대해서는 고세일, “대륙법에서 징벌적 손해배상 논의 - 민법의 관점에서,” 법조 통권 제688호, 2014, 142-190쪽 참조.

82) 그런 측면에서 건전한 성윤리 교육과 함께 지속적으로 딥페이크의 위험성에 대한 공익광고와 처벌 내용에 대해서도 알려야 한다.

83) 예를 들어, 글쓴이가 공부한 미국 툴레인대학교(Tulane University)는 역사상 두 번 학교 문을 닫았다. 첫 번째는 미국 남북전쟁 때이고, 두 번째는 2005년 허리케인 카트리나로 뉴올리안즈 도시 전체가 침수된 때이다. 툴레인대학교 로스쿨은 2006년 1학기에 허리케인 카트리나가 일으킨 실제 문제를 해결하는 세미나를 개설했다. 허리케인 카트리나로 발생한 헌법, 행정법, 형법, 민법의 모든 법영역에서 구체적인 문제 해결을 위해서, 전미의 법률 전문가를 초빙하여 학생들과 토

그런 점에서 형법이지만 딥페이크 피해자 보호를 위한 실질적인 조치를 담고 있다.

미국은 다양한 역사 배경과 문화 배경을 갖는 다인종 국가이다. 따라서 미국법의 입법 특징 가운데 하나는 해석의 여지가 없도록 세밀하게 규정한다. 이는 대륙법 학자의 눈에는 법률이 아닌 시행령이나 시행규칙처럼 보인다. 그렇지만 이러한 입법 기술의 장점은 사회 구성원이 법을 추상적인 개념이 아니라, 구체적인 삶의 지침으로 받아들인다는 점이다.⁸⁴⁾

미국 캘리포니아 형법의 여러 규정은 “무엇이 딥페이크 기술을 통한 사생활 침해인지”를 명확히 밝힌다.⁸⁵⁾ 따라서 딥페이크로 처벌할 수 있거나, 피해자가 가해자에게 민법상 불법행위에 따른 손해배상을 할 때 명확한 기준을 제시한다. 이러한 구체적인 기준이 없는 경우에는 종래의 민법 제751조와 제750조의 요건을 갖추었을 때 딥페이크로 발생한 정신적 손해배상을 청구할 수 있다.⁸⁶⁾

미국 캘리포니아 형법이 딥페이크의 피해자를 위해서 소셜 미디어 플랫폼 사업자에 대한 구체적인 조치를 규정하고 있다.⁸⁷⁾ 이는 딥페이크 피해자를 위한 추상적인 보호장치라 아니라 실질적으로 피해자를 보호하고자 하는 강력한 입법 의사를 나타내는 것이다.⁸⁸⁾ 예를 들어, ① 신고 사용자의 신고를 접수한 후 48시간 이내에 접수를 확인하게 한다. ② 이에 따라서 7일 이내에 신고 사용자에게 처리 현황을 서면으로 업데이트하도

론하고, 실제 법해결 방안을 제시했다. 그런 면에서 특정 범영역에 한정하여 연구하는 전통을 갖고 있는 우리 학계와 견주어 좀 더 유연하게 접근할 수 있다. 글쓴이는 ‘딥페이크’가 ‘주제별 접근방식’으로 문제 해결을 할 수 있는 대표적인 영역으로 평가한다. 또한 법적 해결뿐만 아니라, 심리학, 교육학, 사회학, 경제학과 같은 다른 영역과 협업을 할 필요가 있다.

84) 미국의 입법과정 전반에 대해서는 김영희, 미국의 입법과정, 한국학술정보, 2007 참조.

85) 예를 들어, 미국 캘리포니아 형법 제647조가 그렇다.

86) 요건 분석을 하면 다음과 같다. ① 정신적 손해를 발생시킨 인격권(사생활) 침해가 있어야 한다. ② 인격권(사생활) 침해로 피해자에게 정신적 손해가 발생해야 한다. 이는 가해자의 피해자에 대한 사생활 침해로, 피해자의 사생활 향유를 잃은 것이다. ③ 가해자의 피해자에 대한 인격권(사생활) 침해와 피해자의 정신적 손해 사이에 인과관계가 있어야 한다. ④ 가해자의 인격권(사생활) 침해가 위법해야 한다. 이는 결과불법론, 행위불법론, 결과·행위불법론에 따라서 미세하게 달라질 수 있다. 그러나 딥페이크의 경우에는 가해자가 의식적으로 피해자의 이미지를 침해하는 것이다. 따라서 위법성 판단에 결과불법론, 행위불법론, 결과·행위불법론의 평가 차이가 거의 발생하지 않는다. ⑤ 가해자의 유책성 판단은 객관적 과실설과 주관적 과실설에 따라 차이가 있다. 객관적 과실설은 평균적인 보통 사람을 기준으로 유책성을 판단하는 것이다. 반면에 주관적 과실설은 구체적인 행위자(가해자)를 기준으로 한다. 현행 민법인 책임능력을 규정하는 측면에서는 ‘주관적 과실설’이 좀 더 타당한 것으로 평가한다. 캘리포니아 형법도 딥페이크 범죄를 저지를 당시에 18세 미만의 사람은 처벌 대상이 아님을 분명히 한다. 이는 ‘주관적 과실설’을 전제한 것이다.

87) 캘리포니아 형법 22671. “소셜 미디어 플랫폼 사업자의 준수 의무”의 내용이다.

88) 또한 캘리포니아 형법은 딥페이크로 만들어진 이미지를 배포하는 경우에도 신고, 법적 절차, 공익 차원의 경우에는 처벌 대상이 아님을 명확히 규정한다. 캘리포니아 형법 제647조(4)(D)항 참조.

록 한다. ③ 신고 접수된 후 30일 이내에 디지털 신원 도용에 합리적인 근거가 있는지를 평가하도록 한다. ④ 플랫폼 사업자가 수집한 정보의 지연 가능성을 인지한 시점에서는 48시간 이내에 지연된 사실을 신고자에게 바로 통지하도록 한다. ⑤ 딥페이크에 따라서 피해자의 디지털 신원 도용 사례가 플랫폼에서 공개적으로 표시되지 않도록 일시 차단하도록 한다. ⑥ 최종적으로 피해자의 디지털 신원 도용 사례가 발생했다고 상당한 근거가 있는 경우는 그 내용을 플랫폼에서 즉시 삭제하도록 한다.⁸⁹⁾

이와 같은 규정 내용은 대륙법 전통을 갖는 우리 법에서는 생경하다. 플랫폼 사업자의 내부 업무지침처럼 보인다. 그런데 법의 실질적 정의를 추구한다는 측면에서 오히려 캘리포니아 형법은 “명확성의 원칙”을 철저하게 반영하고 있다. 필자는 이러한 내용은 우리 법의 경우 딥페이크와 관련된 시행령과 시행규칙에 피해자 보호를 위해서 구체적으로 고려할 수 있는 내용이라고 판단한다.⁹⁰⁾

2026년 1월 22일 시행 예정인 “인공지능 발전과 신뢰 기반 조성 등에 관한 기본법”(인공지능기본법) 제31조는 “인공지능 투명성 확보의무”를 규정하고, 제32조는 “인공지능 안전성 확보의무”를 규정한다. 이러한 인공지능의 투명성과 안전성을 규정하는 측면에서 진일보한 내용으로 평가한다. 다만 인공지능의 투명성과 안전성 확보가 구체적으로 무엇을 뜻하는지에 대한 여러 해석이 가능하다. 인공지능기본법은 이와 관련된 내용을 대통령령으로 정하도록 하고 있다. 또한 과학기술정보통신부장관은 구체적인 이행방식에 대해서 고시하도록 규정한다.⁹¹⁾

IV. 맺음말

딥페이크에 대한 민사법적 접근으로, 딥페이크와 관련한 사생활 보호의 여러 측면을

89) 이러한 캘리포니아 형법의 피해자를 위한 실질적인 조치는 현행 “정보통신망 이용촉진 및 정보 보호 등에 관한 법률”(정보통신망법, 2025년 7월 22일 시행 예정인 법률 제20678호) 제44조의 2(정보의 삭제요청 등), 제44조의 3(임의의 임시조치), 제44조의 5(게시판 이용자의 본인 확인), 제44조의 6(이용자 정보의 제공청구), 제44조의 7(불법정보의 유통금지) 규정과 상당히 비슷한 측면이 있다. 다만 2024년 텔레그램 딥페이크 음란물 유포 사건 이전에서 이러한 법규정은 제정된 상태였다. 따라서 법규정과 법현실의 차이를 어떻게 극복하는 것이 중요한 쟁점이 된다.

90) 이러한 미국법의 규율 태도는 더 이상 미국법에 한정되지 않는다. 예를 들어, 90년대에 독일은 자국의 민법전에 등기를 시행규칙과 같은 아주 상세한 내용을 규정하고 있다. 이러한 입법방식은 현실에서 구체적으로 기능하는 측면이 있기 때문에, 이런 입법방식을 채택하고 있다.

91) 그러나 인공지능기본법 시행령은 아직 제정되지 않은 상태이다. 따라서 미국 법에서 규정하는 피해자 보호를 위한 실질적인 내용을 구체적으로 규정했으면 한다.

살폈다. 딥페이크가 문제된 사례와 딥페이크 기술을 설명했고, 딥페이크에 대한 법적 규율로 사생활 보호 측면을 살폈다. 딥페이크 기술은 인공지능의 기술과 더불어 발전하고 있다. 기술 발전에 따른 빛과 그림자는 늘 있다. 인공지능 기반의 딥페이크 기술에도 빛과 그림자가 있다. 의료, 영화 산업과 같은 영역에서 좋은 목적으로 쓸 수 있지만, 나쁜 목적으로 쓸 경우 그 부작용이 크다.

딥페이크 기술에 대한 글쓴이의 생각을 정리하면 다음과 같다.

첫째, 기술 이전에 딥페이크 기술을 쓰는데 대한 사회 구성원의 인식이 중요하다고 평가한다. 밀레니엄 세대와 그 이후 세대는 디지털 환경에서 자랐다. 따라서 이전 세대와 견주어서 디지털 환경이 삶의 중심이 되고, 그에 따른 부작용도 크다.⁹²⁾ 통계에서도 딥페이크 음란물 범죄와 관련하여 10대가 차지하는 비중이 2021년 65.4%에서 2023년 75.8%로 증가하고 있다.⁹³⁾ 그런 점에서 딥페이크 기술을 나쁜 목적으로 쓰지 않는 방향으로 사회 공동체 구성원의 인식을 이끄는 것이 중요하다. 딥페이크와 관련하여 공익광고 또는 사회 공동체 구성원의 인식을 변화시킬 공익광고와 가이드라인을 만들고 홍보·교육을 할 필요가 있다.⁹⁴⁾ 이를 통해서 실질적인 교육을 하고 그 효과를 점검하는 것이 필요하다. 이런 측면에서 인공지능의 윤리가 중요하다.

둘째, 딥페이크도 인공지능 기술의 한 종류이다. 따라서 기술의 문제를 기술로 막을 수 있거나 보완할 수 있다면 가장 좋은 방안이라고 평가한다. AWS, 페이스북, 마이크로소프트와 학계는 힘을 합하여 딥페이크 탐지 프로젝트를 제안하여 고액의 상금을 제안하는 이유로 그런 측면으로 이해할 수 있다. 다만 기술 발전에 따른 창과 방패의 역할을 계속될 것이다.

셋째, 그 다음으로 딥페이크에 따른 사생활 침해를 방지하기 위해서 법적 규율이 필

92) Tim Clydesdale and Kathleen Garces-Foley, Meet the Twentysomethings - The Twentysomething Soul: Understanding the Religious and Secular Lives of American Young Adults 22쪽 (Oxford Academic 2019) 참조.

93) 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 294쪽.

94) 이는 요식적인 전시 행정이 아니라, 실제 도움을 줄 수 있는 교육이 필요하다. 학생에게 딥페이크로 음란물을 만들고 배포하고 시청하는 행위는 명백한 범죄라는 사실을 명확하게 인식시킬 필요가 있다. 석지윤 박사는 현재 학습교재의 상황을 다음과 같이 설명한다. “정보통신교육원의 초·중·고 AI 윤리교재를 보면 초등 교재에는 딥페이크로 되살아난 ‘독립운동가들’과 선생님도 딥페이크 피해라는 내용에 대한 영상을 시청한 후 딥페이크 기술을 법으로 규제해야 하는가에 대한 생각을 쓰도록 하고 있다. 중등 교재에는 딥페이크 기술로 만든 버추얼 휴먼이나 영화 산업에서의 활용 등 인공지능의 긍정적인 내용과 범죄에 악용되기도 한다는 서술이 있으며, 고등 교재에는 딥페이크는 가짜 뉴스를 만들어 여론을 조작하기도 한다는 내용과 전쟁에 딥페이크 영상이 활용되는 것에 대한 비판적 수용 필요성에 대한 언급이 있을 뿐이다.” 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 295쪽.

요하다. 법적 규율의 측면에서는 딥페이크 기술을 운용하는 제조사 또는 소셜 미디어 플랫폼 사업자에 대한 행정법상의 의무 부과가 필요하다. 또한 형법상 처벌 대상임을 법으로 규정하고, 이로써 형법상 제재를 가할 필요가 있다. 그러나 형법은 죄형법정주의 사상에 따라서 제한적으로 운영될 측면이 크다. 또한 현실적으로 딥페이크에 따른 사생활 침해가 미성년자로 발생한 경우가 많다. 그러나 현행 형법 제9조는 형사미성년자인 14세 미만인 자에게는 형사처벌 하지 않는 것으로 규정한다. 그러므로 현실에서 형사미성년자가 일으키는 딥페이크 관련 문제를 형법으로 규율하기 어려운 문제가 발생한다.

따라서 현실에서는 딥페이크를 민사법적으로 규율할 가능성이 크다. 따라서 딥페이크에 따른 사생활 침해의 경우에 민사책임으로는 재산상 손해배상, 위자료, 게시물 삭제, 게시물 금지를 고려할 수 있을 것이다.⁹⁵⁾ 그러나 사생활 침해의 경우에는 재산상 손해배상 보다는 정신적 손해에 따른 위자료 청구를 할 경우가 많다.⁹⁶⁾ 그렇지만 현재 우리 법원에서 운영하는 위자료 액수의 기준이 낮다. 생명침해의 경우, 한국 법원은 위자료의 기준 금액을 1억 원으로 설정한다. 그러므로 식물인간이나 전신마비의 경우, 1억 원 이하로 설정한다. 저산소성 뇌손상에 따른 식물인간으로 7천만 원 정도를 인정한다.⁹⁷⁾ 그런데 사생활 침해의 경우에는 그 금액이 훨씬 낮은 수준이다. 다만 딥페이크 기술을 써서 악의적으로 다른 사람의 동일성(identity)을 침해한 경우, 위자료액을 높일 수 있다.⁹⁸⁾ 그러나 현재의 법원 위자료 실무에 있어서는 피해자를 실제로 보호하기에는 부족한 금액으로 평가한다.

그런 측면에서 캘리포니아 주법에서 규정하고 있는 소셜 플랫폼 사업자에 대한 의무 부과가 우리 법에 대한 실제적인 시사점을 주는 것으로 평가한다. 판례 법리를 통해서 불법행위에 따른 금지청구를 인정한다.⁹⁹⁾ 그러나 이는 전통적인 사법(私法)의 사후적

95) 석지윤 교수는 딥페이크 음란물에 따른 명예훼손에 대한 민사책임으로 손해배상, 게시물 삭제, 게시물금지를 든다. 이는 사생활 침해의 경우에도 같다고 평가한다. 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024, 291쪽.

96) 미국 불법행위법상 정신적 손해에 대한 새로운 경향에 대해서는 고세일, “인신손해에 대한 불법행위의 유형과 책임 내용에 대한 연구 - 미국 불법행위 보통법전집 제3판 재산 손해와 정신적 손해 규정을 중심으로 -,” 법학연구 제30권 제1호, 2019, 397-436쪽 참조.

97) 인천지방법원 2023. 12. 8. 선고 2020가합56871 판결 (사고 당시 39세). 위자료 70,000,000원. 일일수입 394,184,939원, 적극적 손해(기왕치료비 108,345,936원, 향후치료비 86,241,120원, 기왕보조구입비 20,208,386원, 향후보조구입비 15,867,726원, 기왕개호비 36,069,740원, 향후개호비 178,278,063원) × 피고의 책임제한(60%)=573,517,546원.

98) “악의적으로 의도했다는 측면에서” 징벌적 손해배상이 인정될 가능성도 있다. 우리 법에서 징벌적 손해배상에 대해서는 고세일, “대륙법에서 징벌적 손해배상 논의 - 민법의 관점에서 -,” 법조 제688호, 2014, 164-165쪽 참조.

구제수단이다. 따라서 캘리포니아 주법에서 규율하는 것처럼, 딥페이크 기술에 따르는 사생활 침해가 발생했을 때, 이를 신고할 수 있는 수단을 마련할 필요가 있다. 또한 사생활 침해 신고가 상당한 경우, 소셜 플랫폼 사업자는 게시물 게시에 대한 임시조치를 하고, 최종적으로 사생활 침해로 판단한 경우, 그 게시물을 삭제할 의무를 부과할 필요가 있다.¹⁰⁰⁾

미국법은 다문화, 다인종 국가여서, 제정법의 특징이 가능한 해석의 여지가 없도록 법을 상세하게 규정하는 것이다. 그런 측면에서 대륙법의 시각으로 보면, 미국 제정법은 시행규칙처럼 보인다. 따라서 미국법의 내용을 온전히 우리 법에 담아내는 것은 쉽지 않다. 그러나 왜 그런 내용을 미국법에 담고 있는지 그에 대한 함의와 시사점은 우리에게도 여전히 중요하다.¹⁰¹⁾

■ 투고일: 2025. 04. 30 심사일: 2025. 05. 11 게재확정일: 2025. 05. 19

99) 이에 대해서는 김태선, “위법한 차별의 구제와 금지청구 - 인격권 관련 실무에 대한 몇 가지 측면의 평가를 겸하여 -,” 민사법학 제58호, 2012, 71-108쪽 참조.

100) 측면은 조금 다르지만 우리나라에서 운영하는 ‘지우개(잊힐 권리) 서비스’도 이와 비슷한 측면을 갖고 있다. 개인정보보호위원회, 지우개(잊힐 권리) 서비스, *available at* <https://www.privacy.go.kr/front/contents/cntntsView.do?cntsNo=260> (마지막 확인: 2025년 4월 22일).

101) 이에 대해서는 권영준, “불법행위법의 사상적 기초와 그 시사점 - 예방과 회복의 패러다임을 중심으로,” 저스티스 제109호, 2009, 102-104쪽 참조.

[참고문헌]

- 고세일, “정신적 손해에 대한 배상청구권의 법적 기초와 인정영역,” 고려대학교 석사학위논문, 1995.
- _____, “미국 불법행위법의 구조와 내용에 대한 연구 - 보통법전집의 과실을 중심으로,” 민사법학 제59호, 2012.
- _____, “미국 과실불법행위와 인과관계 - 불법행위 보통법전집 제2판의 과실규정을 중심으로,” 안암법학 제42호, 2013.
- _____, “대륙법에서 징벌적 손해배상 논의 - 민법의 관점에서 -,” 법조 제688호, 2014.
- _____, “인공지능과 불법행위책임 법리,” 법학연구 제29권 제2호, 2018.
- _____, “인신손해에 대한 불법행위의 유형과 책임 내용에 대한 연구 - 미국 불법행위 보통법전집 제3판 재산 손해와 정신적 손해 규정을 중심으로 -,” 법학연구 제30권 제1호, 2019.
- _____, “미국 불법행위법상 사생활 보호에 대한 고찰,” 법학연구 제36권 제2호, 2025.
- 권영준, “불법행위법의 사상적 기초와 그 시사점 - 예방과 회복의 패러다임을 중심으로,” 저스티스 제109호, 2009.
- 김명주, “선거 속 딥페이크, 그 실태와 윤리,” 관훈저널 통권 제171호, 2024.
- 권태상, 퍼블리시티권의 이론적 구성 - 인격권에 의한 보호를 중심으로, 경인문화사, 2013.
- 김영희, 미국의 입법과정, 한국학술정보, 2007.
- 김일수·서보학, 새로쓴 형법각론 제9판, 박영사, 2018.
- 김태선, “위법한 차별의 구제와 금지청구 - 인격권 관련 실무에 대한 몇 가지 측면의 평가를 겸하여 -,” 민사법학 제58호, 2012.
- 김효은, “인공지능과 윤리,” in, 한국인공지능법학회, 인공지능과 법, 박영사, 2020.
- 노하림, 유제혁, “Visual Saliency 기반의 딥페이크 이미지 탐지 기법,” Journal of Platform Technology, Vol 12, No. 1, 2024.
- 노현숙, “딥페이크 기술에 의한 인격권 침해 검토 - 초상권을 중심으로 -,” 미디어와 인격권 제11권 제1호, 2025.
- 배종대, 형법각론 제10전정판, 홍문사, 2018.
- 석지윤, “딥페이크 음란물로 인한 명예훼손과 민사책임,” 연세법학 제46호, 2024.
- 양형위원회, 2024년 양형기준, 양형위원회, 2024.
- 조규창, 비교법(상), 소화, 2005.
- 주현경, “딥페이크 기술의 사회적 위험과 규제 방안 - 딥페이크 이용 범죄를 중심으로 -,” 가천법학 제17권 제3호, 2024.
- American Law Institute, Restatement of the Law Second Torts 2d, §§ 504-707A, 1977.
- Chap Clark, Adoptive Youth Ministry - Integrating Emerging Generations into the Family of Faith, Baker Academic, 2016..
- Kenneth S. Abraham, The Forms and Functions of Tort Law, 5th ed., Foundation Press, 2017.
- Lourdes Vazquez, Recommendations For Regulation Of Deepfakes In The U.S.: Deepfake Laws Should Protect Everyone Not Only Public Figures, Robert D. Reif Fellowship

- Article, 2021.
- Sang Joong Kim and Seil Ko, Tort Law in South Korea, Wolters Kluwer, 2024.
- Tim Clydesdale and Kathleen Garces-Foley, Meet the Twentysomethings - The Twentysomething Soul: Understanding the Religious and Secular Lives of American Young Adults, Oxford Academic, 2019.
- W. Page Keeton, Prosser and Keeton on Torts, West 5th ed., 1988.
- 2024 Year in Review: Navigating California's Landmark Deepfake Legislation (12. 11, 2024), *available at* <https://www.mofo.com/resources/insights/241211-2024-year-in-review-navigating-california-s-landmark-deepfake-legislation>.
- California Legislative Information, SB-981 Sexually explicit digital images, *available at* https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB981.
- Golden, "Anti-deepfake technology," *available at* https://golden.com/wiki/Anti-deepfake_technology-R9VDX65.
- Google Machine Learning, "The Discriminator," *available at* <https://developers.google.com/machine-learning/gan/discriminator>.
- Kate Conger and John Yoon, "Explicit Deepfake Images of Taylor Swift Elude Safeguards and Swamp Social Media," New York Times, Jan. 26, 2024, *available at*, <https://www.nytimes.com/2024/01/26/arts/music/taylor-swift-ai-fake-images.html>.
- K.C. Halm, Ambika Kumar, Jonathan Segal and Caesar Kalinowski IV, Two New California Laws Tackle Deepfake Videos in Politics and Porn (2. 28, 2020), *available at* <https://www.dwt.com/blogs/media-law-monitor/2020/02/two-new-california-laws-tackle-deepfake-videos-in>.
- Michelle M. Graham, "Deepfakes: Federal and state regulation aims to curb a growing threat," (2024년 6월 26일) *available at* <https://www.thomsonreuters.com/en-us/posts/government/deepfakes-federal-state-regulation/>.
- MIT Media lab, "Detect DeepFakes: How to counteract misinformation created by AI," *available at* <https://www.media.mit.edu/projects/detect-fakes/overview>.
- SB-926 Crimes: distribution of intimate images, *available at* https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB926.
- US Congress, S.3696 - Defiance Act of 2024, *available at* <https://www.congress.gov/bill/118th-congress/senate-bill/3696/text>.
- Regan Morris, "AI was enemy No. 1 during Hollywood strikes. Now it's in Oscar-winning films," BBC (March 31, 2025), *available at* <https://www.bbc.com/news/articles/ce303x19dwgo>.
- Wikipedia, "Deepfake," *available at* <https://en.wikipedia.org/wiki/Deepfake>.
- Wikipedia, "Generative adversarial network," *available at* https://en.wikipedia.org/wiki/Generative_adversarial_network.

<Abstract>

Deepfakes and Privacy Protection

Ko, Se-il*

A person uses their computer in their space. Their computer is not connected to other computers. As time passed, the Internet was introduced. Many computers became connected to the Internet through telephone, cable, and LAN lines. A wealth of information was exchanged among many people. The world before and after the Internet was different. Additionally, as time progressed, artificial intelligence emerged. Many were excited or afraid that artificial intelligence would bring about a new industrial revolution. Moreover, as time continued, generative artificial intelligence arrived. Many people were astonished by the capabilities of generative artificial intelligence. Many thought generative artificial intelligence was incredible. However, they also began to worry about the shadows that generative artificial intelligence makes. One aspect of the dark shadow that generative artificial intelligence casts is “deepfakes.”

In reality, there are two significant areas where deepfakes cause problems. The first is related to “elections.” The second is “pornography.” The commonality between the two regions is that deepfakes adversely affect our society because the new technology violates human identity. Tortfeasors intend to undermine a person’s image and identity. Deepfakes create harmful images and distort a person’s actual representation. Subsequently, these distorted images replace the authentic representation. As a result, the public perceives the harmful images created using deepfake technology as accurate information about the person. In this sense, deepfakes distort a person’s identity and damage their self-image.

Thus, various academic journals have addressed deepfakes using two main approaches. One focuses on the technical aspects of deepfakes, while the other examines their regulation under criminal law. In this context, academic articles discussing civil responsibility related to deepfakes are still in the early stages. When considering all areas that legislate deepfakes, there are (1) administrative law regulating the responsibilities of

* Professor, Ph.D., Chungnam National University School of Law.

platform providers and potential AI developers, (2) criminal law, and (3) civil law governing defamation and invasion of privacy. In this article, the author concentrates on privacy protection concerning deepfakes, based on civil responsibility. This article is organized as follows: First, in section II, the author explains the cases arising from deepfakes and deepfake technology. Then, in section III, the author examines privacy protection as a legal regulation for deepfakes. Finally, in section IV, the author proposes measures to protect privacy from deepfakes.

[Key-Words]

Deepfakes, Generative Artificial Intelligence, Deep Learning, Privacy, Identity

