

딥페이크 기반 허위정보 대응 정책 비교 연구: 한국의 정책 보완 방향*

A Comparative Study on Deepfake-Based Disinformation Response Policies toward Recommendations for South Korea

이정은 (Jung-eun Lee) **
오효정 (Hyo-Jung Oh) ***

초 록

딥페이크 기술은 인공지능과 딥러닝을 활용하여 이미지, 음성, 영상 등을 조작하는 기술로, 정치적 선전, 금융 사기, 여론 조작 등 사회적 위험을 초래할 수 있다. 그러나 현재 한국의 대응은 디지털 성범죄 예방이나 선거 개입 방지 등의 협소한 범위에 집중되어 있어, 다양한 유형으로 확산되고 있는 딥페이크 허위정보의 폐해를 충분히 포괄하지 못하고 있다. 이에 본 연구는 딥페이크 기반의 허위정보, 특히 가짜뉴스에 대응하기 위한 주요 선진국의 정책을 법적·기술적·사회적 관점에서 비교·분석하고, 이를 바탕으로 한국의 실효성 있는 정책 보완 방향을 제시하는 것을 목표로 한다. 연구 결과, 주요 국가들은 플랫폼 운영자에게 딥페이크 콘텐츠 탐지 및 삭제 의무를 부과하고, 인공지능 기반 검증 기술을 활용해 허위정보 확산을 방지하는 것으로 나타났다. 반면, 한국은 개별적인 규제에 의존하고 있어 보다 포괄적인 법적 대응 체계 마련과 기술적 보완이 필요한 상황이다. 이에 본 연구는 딥페이크 허위정보의 범위 및 플랫폼 책임 명확화, 법적·제도적 지원을 통한 인공지능 탐지 기술 강화, 미디어 리터러시 교육 확대 및 시민 주도의 허위정보 모니터링 시스템 구축을 제안한다. 이를 통해 한국이 딥페이크 기반 허위정보 확산을 효과적으로 방지하고 신뢰할 수 있는 디지털 정보 환경을 조성하는 데 필요한 정책적 시사점을 제공하는 데 연구의 의의를 둔다.

ABSTRACT

Deepfake technology, which utilizes artificial intelligence and deep learning to manipulate images, audio, and video, poses significant societal risks, including political propaganda, financial fraud, and public opinion manipulation. However, South Korea's current response primarily focuses on preventing digital sex crimes and election interference, lacking comprehensive measures such as platform accountability, the application of artificial intelligence detection technology, and media literacy education. Therefore, this study examines and evaluates the legal, technical, and social responses of the United States, the European Union, and the United Kingdom to provide policy insights for South Korea. The findings indicate that major countries impose obligations on platform operators to detect and remove deepfake content while employing artificial intelligence-based verification technologies to prevent the spread of disinformation. In contrast, South Korea's approach relies on fragmented regulations, highlighting the need to establish a more comprehensive legal framework and implement artificial intelligence detection technologies. Based on these findings, this study suggests clarifying platform accountability, strengthening artificial intelligence detection technologies through legal and institutional support, expanding media literacy education, and developing a citizen-driven disinformation monitoring system. This study provides policy recommendations for South Korea to effectively combat the spread of deepfake-based disinformation and establish a trustworthy digital information environment.

키워드: 딥페이크, 허위정보, 가짜뉴스, 플랫폼 규제, AI 탐지기술, 디지털 리터러시

Deepfake, disinformation, fake news, platform regulation, AI detection technology, digital literacy

* 본 논문은 2023년도 전북대학교 연구교수 연구비 지원에 의하여 연구되었음.

본 논문은 2024년도 한국연구재단 연구비 지원에 의한 결과의 일부임

(과제번호: NRF-2021R1I1A3047435).

** 전북대학교 기록관리학과 연구교수(je.lee@jbnu.ac.kr) (제1저자)

*** 전북대학교 문현정보학과 교수, 문화융복합아카이빙연구소 공동연구원(ohj@jbnu.ac.kr) (교신저자)

■ 논문접수일자 : 2025년 2월 17일 ■ 최초심사일자 : 2025년 3월 7일 ■ 게재확정일자 : 2025년 3월 12일

■ 정보관리학회지, 42(1), 155-182, 2025. <http://dx.doi.org/10.3743/KOSIM.2025.42.1.155>

※ Copyright © 2025 Korean Society for Information Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited. the use is non-commercial and no modifications or adaptations are made.

1. 서 론

1.1 연구 배경 및 필요성

디지털 기술의 급속한 발전은 온라인에서 유통되는 정보의 양을 기하급수적으로 증가시켰다. 이와 함께 허위정보(disinformation)의 확산 문제가 심각한 사회적 이슈로 부각되고 있다. 허위정보는 단순한 오보(misinformation)와는 달리 특정한 정치적·경제적·사회적 목적을 달성하기 위해 의도적으로 조작된 정보로서 민주주의의 질서, 공공 안전, 그리고 경제적 신뢰를 심각하게 위협할 수 있다(이숙종, 2024). 특히, 온라인 플랫폼과 소셜 미디어를 통해 허위정보의 확산 속도가 빨라지면서 사실과 허위의 경계가 흐려지고, 이에 따른 여론 조작 및 사회적 갈등 유발 사례도 증가하고 있다. 한편, 최근 발표된 로이터 연구(Behre et al., 2024)에 따르면, 일반 국민들이 새로운 정보를 접하는 매체가 기존의 뉴스, 블로그, 카페 등 텍스트 기반의 미디어에서 유튜브(Youtube), 인스타그램(Instagram), 틱톡(TikTok)과 같은 동영상 중심의 시각 매체로 이동하고 있는 것으로 나타났다. 특히, 자극적이고 편향된 허위정보일수록 확산의 속도가 빠르고 파급력이 더욱 커지는 경향이 있다.

이러한 허위정보 확산을 더욱 심각하게 만드는 요인 중 하나가 딥페이크(Deepfake) 기술이다. 딥페이크는 인공지능(AI)과 딥러닝(Deep Learning) 기술을 활용하여 영상, 음성, 이미지 등을 정교하게 조작하는 기술로 원본과 거의 구별할 수 없는 콘텐츠를 생성할 수 있다. 기존 허위정보가 텍스트 중심으로 유포되었다

면, 딥페이크 기술은 시각적·청각적 요소를 결합하여 허위정보의 신뢰도를 더욱 높이는 방식으로 진화하고 있다. 이러한 기술적 특성으로 딥페이크가 악용될 경우, 정치적 선전, 명예훼손, 금융 사기 등 다양한 사회적 문제를 초래할 가능성이 크며(최현빈, 2024), 허위정보 확산 방식이 더욱 정교해질 수 있다.

최근 몇 년간 딥페이크 기반 허위정보의 사례가 급증함에 따라, 세계 각국은 이를 방지하기 위한 법적·기술적·사회적 대응책을 마련하고 있다. 미국, 유럽연합(EU), 영국 등 주요국은 딥페이크 기술의 확산을 방지하기 위해 다양한 법적 규제와 기술적 조치를 도입하고 미디어 리터러시 교육을 강화하는 등 사회적 대응에도 주력하고 있다. 그러나 한국의 대응체계는 주로 디지털 성범죄라는 협소한 범위의 대응에 초점이 맞추어져 있으며, 선거 개입, 경제적 사기, 언론 조작 등 보다 광범위한 허위정보 확산을 막기 위한 포괄적 대응책은 미흡한 실정이다. 2024년에 실시한 과학기술정보통신부의 국민 여론조사에 따르면, 응답자의 94.5%가 허위정보가 사회에 미치는 영향을 심각하게 인식하고 있으며, 딥페이크 기반 허위정보가 기존 허위정보보다 더 위험하다고 응답한 비율이 84.9%에 달했다(디지털공론장, 2025). 이는 딥페이크가 기존 허위정보보다 정교하게 조작될 수 있어 신뢰도를 높이기 때문이며, 기존의 단순한 성범죄 대응책을 넘어 보다 포괄적인 규제와 대책이 필요함을 시사한다. 특히 딥페이크 관련 기술과 이를 소비하는 디지털 미디어 환경이 매우 빠르게 발전하는 만큼, 현실에 부합할 수 있는 실효성 있는 정책 마련에 대한 고민이 필요하다.

1.2 연구 목적 및 연구 방법

본 연구는 딥페이크 기반의 허위정보, 특히 가짜뉴스에 대응하기 위한 주요 선진국의 정책을 비교·분석하고, 이를 바탕으로 한국의 실효성 있는 정책 보완 방향을 제시하는 것을 목표로 한다. 비교 대상국은 딥페이크 기반 허위정보에 대한 정부 차원의 정책뿐만 아니라 기술적 대응, 사회적 규범 등을 명문화한 나라로, 미국, 유럽연합(EU), 영국을 선정하였다. 분석 방법은 법적·기술적·사회적 관점의 분석 틀을 적용한다. 법적 관점은 플랫폼 사업자에게 콘텐츠 탐지 및 삭제 등 구체적 의무를 부여하는 법률의 유무와 법률의 적용 범위를 기준으로 분석한다. 기술적 관점은 국가 또는 플랫폼 차원에서 AI 탐지 기술을 활용하여 콘텐츠 진위를 판별하는 시스템의 실제 운영 여부 및 기술적 표준 채택 현황을 분석 기준으로 설정한다. 사회적 관점은 미디어 리터러시 교육의 의무화 여부와 시민 참여형 허위정보 감시 시스템의 운영 여부를 기준으로 분석한다. 이를 통해 각국의 대응 전략을 평가하고, 이를 기반으로 한국의 상황에 적합한 정책적 시사점을 도출할 것이다.

본 연구는 문헌 연구를 바탕으로 진행되며, 주요 법률, 정부 보고서, 학술 논문, 정책 문서, 언론 보도를 종합적으로 검토하여 국가별 대응 방안을 분석한다. 연구 방법으로는 질적 비교 연구를 채택하여 국가별 대응 전략을 분석하고, 한국 사회에 적용 가능한 정책적 시사점을 도출하여 효과적인 딥페이크 허위정보 대응 방안을 제시한다.

2. 이론적 배경

2.1 딥페이크 기반 허위정보의 개념 및 특성

딥페이크 기술의 발전과 정보 환경의 변화는 허위정보의 생성과 확산을 가속화하며 기존의 정보 검증 시스템으로 대응하기 어려운 새로운 형태의 정보 조작을 초래하고 있다(장미경, 2024). 과거에는 전문가만 가능했던 영상 조작이 이제는 일반 대중도 쉽게 활용할 수 있는 수준에 이르렀다. 특히, 오픈소스 딥페이크 소프트웨어가 등장하면서 복잡한 기술적 지식이 없는 비전문가도 몇 단계의 과정만 거치면 실제와 구분이 어려운 영상과 음성을 제작할 수 있는 환경이 마련되었다. 이로 인해 정치적·경제적·사회적 목적으로 딥페이크 기반 허위정보가 제작·유포되는 사례가 급증하고 있다.

딥페이크 기반 허위정보는 크게 세 가지 유형으로 구분할 수 있다. 첫째, 정치적 목적의 허위정보는 선거 개입, 여론 조작, 특정 인물이나 단체에 대한 허위정보 확산을 위해 사용된다. 대표적인 사례로는 정치인의 벌언이나 행동이 조작되어 유권자의 판단을 왜곡하거나 특정 정당 또는 후보에 대한 가짜뉴스가 유포되는 경우가 있다. 이는 선거 및 정책 결정 과정에서 사회적 신뢰를 저해하고 민주주의 시스템에 악영향을 미칠 가능성이 크다. 둘째, 경제적 목적의 허위정보는 금융 사기, 기업 명예 훼손, 주가 조작 등의 형태로 활용된다. 예를 들어, 기업 CEO의 음성을 딥페이크 기술로 변조해 투자자나 금융 담당자를 속이는 사례가 보고되고 있으며, 특정 기업이나 제품에 대한 허위정보를 유포하여 시장 혼

란을 조성하는 방식도 등장하고 있다(한국지능정보사회진흥원, 2025). 이처럼 딥페이크 기술이 악용될 경우, 경제적 피해뿐만 아니라 기업과 시장의 신뢰도에도 부정적인 영향을 미칠 수 있다.셋째, 사회적 혼란을 유발하는 허위정보는 가짜 범죄 영상, 허위 뉴스 리포트, 공포감을 조성하는 조작 콘텐츠 등의 형태로 제작된다. 최근에는 딥 페이크를 활용하여 유명 인물이 허위 발언을 한 것처럼 조작하는 사례나, 특정 사건을 왜곡하여 공포를 조성하는 영상이 확산되는 경향이 있다. 이러한 허위정보는 사회적 불안을 증폭시키고 대중이 실제 사건과 가짜뉴스를 구별하기 어렵게 만들어 혼란을 초래한다(한국경제, 2024).

디지털 미디어의 확산, 특히 소셜 미디어의 대중화는 딥페이크 기반 허위정보의 유포를 더욱 가속화하는 핵심적인 요인 중 하나이다. 기존의 전통적 뉴스 매체는 정보 검증 과정을 거쳐 게시되지만, 소셜 미디어 플랫폼에서는 개인이 자체적으로 정보를 생성하고 공개하거나, 다른 출처를 통해 습득한 정보를 그대로 전달하는 등, 정보의 신빙성에 대한 판단 과정이 축소된다. 또한 자극적이거나 감성적인 콘텐츠가 더 빠르게 확산되는 알고리즘 구조를 가지고 있어 딥페이크 콘텐츠가 짧은 시간 내에 광범위하게 퍼질 가능성이 크다. 특히, 유튜브, 틱톡, 트위터(X) 등과 같은 이미지 및 동영상 위주의 시청각 미디어 플랫폼은 사용자의 관심사에 따라 콘텐츠를 자동 추천하는 방식을 사용하기 때문에, 흥미를 끄는 내용이라면 그 정보의 진위와 상관없이 더 많은 사용자에게 노출될 위험이 높다. 따라서 알고리즘 기반 추천 시스템은 허위정보의 확산을 더욱 촉진하며(변희원, 2024), 결과적으로 딥페이크 기반 허위정보가 바이럴

효과를 일으켜 대중의 인식에 영향을 미치는 주요 원인으로 작용하고 있다.

허위정보 확산에는 심리적 요인 또한 중요한 영향을 미친다. 사람들은 자신의 신념과 일치하는 정보를 더욱 쉽게 믿는 확증 편향(Confirmation Bias)을 가지며, 이는 딥페이크 기반 허위정보의 신뢰도를 높이는 결과를 초래한다. 또한, 현대 사회에서는 정보의 양이 폭발적으로 증가하면서 취득한 정보의 신빙성 여부를 깊이 검토하기보다는 표면적인 정보만으로 판단하는 경향이 강하다. 이러한 정보 과부하 환경에서는 딥 페이크 기반 허위정보가 빠르게 확산될 위험이 크며 정보 조작 기술이 더욱 정교해질수록 피해 규모도 커질 가능성이 높다(최성철, 2024).

결과적으로 딥페이크 기반 허위정보는 기술적 발전, 디지털 미디어의 확산, 그리고 인간의 인지적 특성이 복합적으로 작용하며 확산되는 구조를 지닌다. 따라서 이러한 요소를 종합적으로 고려한 대응 전략이 필요하며, 실제 사회 구성원들에게 적용 가능한 실효성 있는 정책 마련을 위한 법적·기술적·사회적 차원의 다변 접근이 필수적이다.

2.2 딥페이크 기반 허위정보의 주요 사례

딥페이크 기술이 빠르게 발전하면서 이를 악용한 허위정보 사례가 세계 각국에서 보고되고 있다. 정치적 목적의 선거 개입, 경제적 사기, 여론 조작, 명예훼손 등의 형태로 딥페이크가 활용되며, 이는 공공 신뢰를 저해하고 사회적 불안을 야기하고 있다. 따라서 최근 몇 년간 발생한 사례들을 분석하는 것은 딥페이크 기술의 악용 방식과 그 영향을 파악하고, 효과적인 정책 대응

방안을 마련하는 데 중요한 자료가 될 것이다.

먼저 정치적 목적으로 선거 개입과 여론 조작을 위해 딥페이크가 사용된 사례가 있다. 2022년 러시아-우크라이나 전쟁 중 볼로디미르 젤렌스키 우크라이나 대통령이 러시아에 항복을 선언하는 조작된 영상이 온라인에서 유포된 사건이 대표적이다. 해당 영상은 AI 기술을 이용해 젤렌스키 대통령의 얼굴과 목소리를 조작한 것으로 국민들에게 무기를 내려놓고 항복할 것을 촉구하는 내용을 담고 있었다. 이는 러시아의 선전 활동의 일환으로 추정되며 우크라이나 국민들의 사기를 저하시키고 혼란을 야기하려는 의도로 유포되었다(정성호, 2022). 또 다른 사건으로는 2023년 미국 공화당 대선 경선에서 론 디샌티스 캠프가 트럼프 전 대통령과 파우치 전 미국 국립 알레르기·전염병연구소(NIAID) 소장이 함께 있는 이미지를 조작하여 트럼프가 파우치를 끌어안거나 이마에 입 맞추는 듯한 모습을 연출한 딥페이크 이미지가 유포되었다(이민석, 2023). 이는 유권자들에게 특정한 인식을 심어주기 위한 정치적 딥페이크 활용 사례로 선거 캠페인에서 딥페이크 기술이 조작된 메시지를 전달하는 수단으로 활용될 수 있음을 보여준다.

경제적 이익을 목적으로 한 사기 범죄에도 딥페이크가 활용되고 있다. 2023년 홍콩에서는 한 다국적 기업의 최고경영자(CEO)의 음성을 조작한 전화 사기가 발생하였다. 범죄자들은 AI를 이용해 CEO의 음성을 변조한 후, 기업 회계 담당자에게 전화를 걸어 3,500만 달러(한화 약 460억 원)를 송금하도록 유도하였다. 해당 기업은 음성의 사실 여부를 즉각 검증하지 못한 채 송금을 진행하였고, 이후 자금은 해외

여러 계좌로 분산되어 회수하기 어려운 상황이 되었다(임지우, 2024). 딥페이크를 상업적 목적으로 활용한 또 다른 사례 중 하나로 2023년 할리우드 배우 톰 행크스(Tom Hanks)는 AI로 생성된 자신의 이미지가 본인 동의 없이 치과 보험 광고에 사용되었다는 사실을 공개적으로 밝혔다. 이는 딥페이크 기술이 초상권을 침해하는 방식으로 악용될 수 있음을 보여주는 사례로, 배우들의 디지털 권리를 보호해야 한다는 논의로 이어지고 있다. 이 사건은 단순한 허위 광고 문제가 아니라, '가상 배우(Virtual Actor)' 제작 논란과도 연결되었다. AI를 이용해 배우의 얼굴과 목소리를 복제할 경우, 배우들의 동의 없이 영상을 제작할 가능성이 커지면서 이에 대한 법적 보호 필요성이 제기되기도 하였다(최인준, 2023).

한편 딥페이크 기반 허위정보는 사회적 불안을 조성하고 가짜 범죄 영상 및 허위 뉴스 리포트 등의 형태로 제작되어 공포감을 조성하는 데도 사용되고 있다. 특정 종교집단이나 단체에서 딥페이크를 활용하여 가짜뉴스를 확산시킴으로써 대중을 혐혹하거나 불안감을 조성하여 포교 활동에 활용하는 사례가 있었다(김아영, 2024). 그러나 현재 보고된 사례들은 주로 정치적·경제적 목적이 강하며, 향후 사회적 불안을 조장하는 새로운 형태의 딥페이크 사례가 증가할 가능성이 있다.

3. 국내외 딥페이크 기반 허위정보 대응 정책 비교

본 장에서는 미국, 유럽연합, 영국 등 주요

국기들이 딥페이크 기반 허위정보 및 가짜뉴스 확산을 방지하기 위해 마련한 법적·기술적·사회적 대응 방안을 살펴본다. 구체적인 분석 대상 국가는 미국, 유럽연합(EU), 영국으로 선정하였는데, 그 이유는 이들 국가는 딥페이크 기반 허위정보 대응 정책에서 서로 차별성을 보이면서도 한국이 정책적으로 참고할 수 있는 사례를 제공하기 때문이다. 미국은 연방 및 주(州) 차원의 법적 대응을 통해 선거 개입 방지 및 플랫폼 규제를 시행하고 있으며, 유럽연합(EU)은 「디지털서비스법」과 「인공지능 법」을 기반으로 법적 책임 부과와 기술적 대응을 병행하고 있다. 영국은 기존 법체계를 활용한 대응책을 마련하는 한편, 공공 부문에서의 AI 탐지 기술 활용과 미디어 리터러시 교육 강화를 통해 사회적 대응을 강조하고 있다. 먼저 현재 국내에 수립된 대응 정책에 대해 기술해 보면 다음과 같다.

3.1 한국

한국은 딥페이크 기반 허위정보 확산을 방지하기 위해 법적 대응을 추진하고 있으나, 현재 법적인 규제는 성범죄, 선거 개입 등의 특정 영역에 집중되어 있다. 「성폭력처벌법」과 「청소년성보호법」 개정을 통해 딥페이크 성적 영상물의 제작, 유포뿐만 아니라 소지 및 시청까지 처벌 대상에 포함시켰지만(여성가족부, 2024), 가짜뉴스와 같은 정치적 허위정보 유포, 금융사기, 언론 조작 등 다양한 악용 사례를 포함하는 종합적인 법적 체계는 아직 미흡하다. 「공직선거법」은 허위정보 유포를 규제하고 있으나, AI 기술을 이용한 조작 콘텐츠에 대한 직접적

법적 규정이 미비하여 선거 개입을 효과적으로 방지하는 데 한계가 있다. 다행히 최근에는 딥페이크 허위정보와 관련한 규제를 강화하려는 움직임이 나타나고 있다. 2024년 개정된 「공직선거법」 제82조의8(딥페이크 영상 등을 이용한 선거운동)에서는 선거 전 90일부터 선거일 까지 생성형 AI로 제작된 조작 영상의 제작·편집·유포·상영·게시를 금지하는 조항이 추가되었다. 이러한 법적 조치는 AI 기반 허위정보가 선거에 미치는 영향을 고려한 예방적 대응으로 볼 수 있으며, 성범죄 중심의 기존 규제에서 벗어나 정치적 목적의 가짜뉴스 및 경제적 사기 등까지 포괄하는 방향으로 발전하는 과정으로 평가된다. 그러나 선거 외 영역에서의 AI 허위정보 대응은 여전히 미흡하며, 플랫폼 사업자의 책임을 명확히 하는 법적 조치 또한 필요하다. 이에 국회에서는 생성형 AI 기술의 부작용을 방지하기 위한 딥페이크 규제 관련 법안들이 논의되고 있다. 특히, 제22대 국회에서 발의된 인공지능 관련 법안들은 생성형 AI 기반 콘텐츠에 대한 출처 표기를 의무화하는 조항을 공통적으로 포함하고 있다(최진웅, 2024). 그러나 이러한 법안들은 아직 법제화되지 않았으며, 딥페이크 기반 허위정보의 확산을 체계적으로 규제할 수 있는 법적 근거는 여전히 미비한 실정이다. 〈표 1〉은 인공지능 법안 외에도 제22대 국회에서 발의된 정보통신망법 개정안 중 딥페이크 규제를 강화하기 위한 주요 법안들을 정리한 것이다.

기술적 대응 측면에서 한국 정부는 과학기술정보통신부를 중심으로 범정부 TF(Task Force)를 운영하며, 딥페이크 탐지 기술 개발을 지원하고 있다. 한국정보통신기술협회(TTA)는 AI

〈표 1〉 딥페이크 규제 관련 정보통신망법 일부 개정법률(안) 내용

| 발의 일자 | 딥페이크 규제 관련 내용 |
|------------|--|
| 2024.08.27 | 딥페이크 가해자에 대한 형사처벌 규정 (7년 이하 징역, 10년 이하 자격정지 또는 7천만 원 이하 벌금) 포함 |
| 2024.08.27 | 방송통신위원회·과학기술정보통신부가 합성 영상 피해·유통 실태 조사, 기술 개발 촉진, 교육·홍보 등 시책 추진을 통한 정보통신망 이용 환경 개선 규정 신설 |
| 2024.08.28 | 정보통신망에서 혐위정보 및 비동의 딥페이크 콘텐츠 유통 금지 및 위반 시 처벌 규정 신설 |
| 2024.08.29 | 정보통신서비스 제공자 및 이용자에 대한 AI 생성 정보 표시 의무 및 유통 방지 조치 강화 |
| 2024.08.29 | 정보통신서비스 제공자의 불법 촬영물 유통 방지 및 수사 협조 의무 강화 |

기반 영상 분석 및 워터마킹 기술을 포함한 콘텐츠 진위 판별 시스템을 연구하고 있으며, 이러한 연구를 바탕으로 AI 생성 콘텐츠의 신뢰도를 검증하는 기술을 개발 중이다(한국정보통신기술협회, 2025). 네이버와 카카오 같은 국내 IT 기업들도 자체적으로 AI 콘텐츠 식별 기술을 도입하여 플랫폼 내 딥페이크 감지 기술 개발에 속도를 내고 있다. 네이버는 딥페이크 콘텐츠나 이미지 등 콘텐츠를 신고할 수 있는 신고 채널을 개설하고 실시간 AI 이미지 필터링 시스템인 '클로바 그린아이(CLOVA Green-Eye)'를 네이버 카페, 블로그 등에 적용하여 24시간 모니터링하고 있다. 카카오는 오픈 채팅, 포털 다음 등에서 모니터링을 진행하고 있으며, 포털 다음에서 딥페이크 관련 검색어를 청소년 보호 검색어로 지정하고 있다. 카카오톡의 경우, 딥페이크 혐위정보를 배포·제공하는 행위의 적발 시 카카오톡 전체 서비스를 영구히 제한하고 있다(한국과학기술기획평가원, 2024). 그러나 현재 한국의 AI 탐지 기술은 딥페이크 기반으로 생성된 혐위정보 콘텐츠를 효과적으로 감지하는 역할에 더욱 집중할 필요가 있다.

사회적 대응으로는 미디어 리터러시 교육과 디지털 성범죄 예방 교육으로 나누어 추진되고

있다. 한국언론진흥재단은 초·중·고등학교에서 AI 생성 콘텐츠 감별 프로그램을 운영하고 있으며, 방송통신위원회는 팩트체킹 기관과 협력하여 AI 기반 혐위정보 감시 시스템을 운영하고 있다(강준모, 2024). 또한, 공공기관과 언론사들이 협력하여 혐위정보 신고 시스템을 마련하고 있으나, 시민들이 이를 적극적으로 활용할 수 있는 체계적인 교육 및 홍보가 부족한 실정이다. 공교육 과정에서도 미디어 리터러시 교육이 필수적으로 포함되지 않아 시민들의 정보 검증 역량이 충분히 배양되지 못하고 있다.

한편 디지털 성범죄 예방에는 더 집중된 경향을 보인다. 여성가족부는 2018년부터 '디지털 성범죄 피해자 지원센터'를 운영하며 24시간 상담 서비스, 불법 촬영물 삭제 지원, 법률 지원, 심리치료 등 종합적인 피해 지원 체계를 구축하고 있으며, 경찰청은 '디지털 성범죄 특별수사단'을 운영하여 전문적인 수사를 지원하고 있다. 이렇듯 교육부와 여성가족부의 학교 및 공공기관 대상 교육 내용은 디지털 성범죄 예방에 집중되고 있으나(한국지능정보사회진흥원, 2025), 이러한 대응 방식은 딥페이크 기술 기반의 가짜뉴스 대응에 있어 특정 사례에만 집중되어 있어, 딥페이크를 활용한 다양한

형태의 허위정보에 포괄적으로 대응하지 못하는 한계가 있다. 따라서, AI 생성 콘텐츠 및 딥페이크 허위정보 대응 역량을 강화하는 방향으로 미디어 리터러시 교육을 확대하고, 기존의 디지털 성범죄 대응 체계를 허위정보 대응과 통합하는 노력이 필요하다.

3.2 미국

미국은 주(州) 차원의 법률 제정과 연방 차원의 기술적 대응 및 플랫폼 자율 규제를 통해 딥페이크 허위정보 확산에 대응하고 있다. 법적 측면에서 미국은 주(州) 별로 개별 법안을 마련하고 있으며, 특히 공직 선거와 성범죄 관련 규제를 강화하는 법안들이 도입되고 있다. 예를 들어, 텍사스주는 공직 선거와 관련하여 최초로 딥페이크 규제를 도입한 주로, 선거운동을 목적으로 딥페이크 정보를 제작·유포하는 행위를 금지하고 있다. 또한, 딥페이크 포르노그래피(Pornography)의 제작 및 유포 행위를 명확히 규정하고 처벌하는 법률을 시행하고 있다(최진웅, 2024). 캘리포니아주는 2019년부터 딥페이크 관련 규제를 도입하고 선거 기간 60일 이내에 후보자의 명예를 훼손하거나 유권자를 오도할 목적으로 조작된 음성·영상을 유포하는 행위를 엄격히 금지하고 있다. 다만, 해당 콘텐츠가 조작되었음을 명확히 알리는 경우 배포가 허용되며, 언론 보도, 풍자, 패러디 등은 규제 대상에서 제외된다(강준모, 2024). 2024년 9월에는 AI 관련 18개의 법률이 제정되면서 딥페이크 음란물의 제작 및 유포 금지, 선거 홍보물의 AI 활용 투명성 강화, 그리고 온라인 플랫폼의 딥페이크 및 가짜뉴스 콘텐츠 삭제의

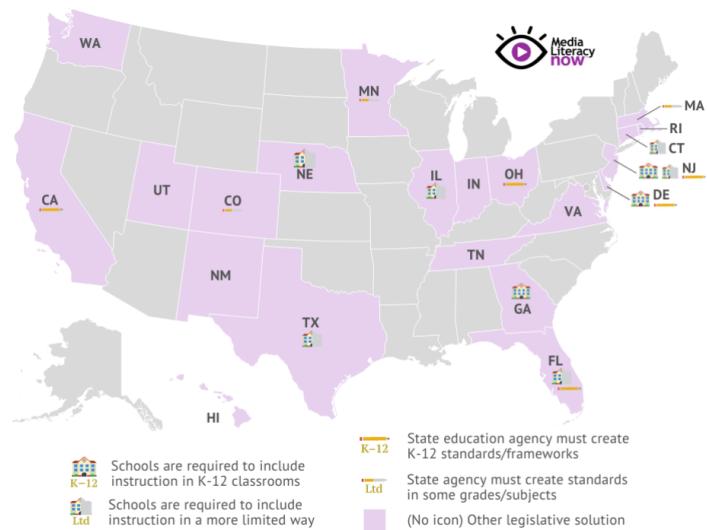
무화 등의 조치가 포함되었다. 특히, 선거 홍보물에 AI를 활용한 경우 반드시 이를 공개하도록 규정하였으며, 거대 온라인 플랫폼 사업자에게 허위정보를 포함한 AI 생성 콘텐츠의 차단 및 표시 조치를 의무화하였다(세계법제정보센터, 2024). 테네시주는 공직 선거와 성범죄뿐만 아니라 초상권 보호와 관련한 딥페이크 규제를 마련하고 있다. 일명 「엘비스법」(Elvis Act)으로 불리는 이 법은 특정인의 동의 없이 딥페이크 기술을 활용해 가상의 음성을 생성·조작하여 배포하는 행위를 초상권 침해로 규정하고 이에 대한 법적 처벌과 손해배상 청구를 가능하게 하고 있다(Elvis Act, §47-25-1105). 개별 주가 아닌 연방의 차원에서도 딥페이크 규제 법안이 논의되고 있는데, 그중 가장 대표적인 것이 「딥페이크 책임법안」(DEEPFAKES Accountability Act)이다. 이 법안은 딥페이크로 인해 피해를 입은 피해자에게 법적 지원을 제공하기 위한 목적으로 발의되었다. 더불어 딥페이크로 제작된 콘텐츠에 대하여 콘텐츠가 원본 영상과 변경되었음을 식별할 수 있도록 하는 공개 의무 부과, 형사처벌, 민사 금전적 제재, 피해자 지원 및 딥페이크 탐지 등에 관한 내용을 포함하고 있다(강준모, 2024). 그러나, 현재 이 법안은 입법 과정에서 논의 중이며, 아직 법제화되지 않은 상태이다. 또한, 연방 차원의 법적 규제와 개별 주의 입법 간에 차이가 존재하여, 미국 내 딥페이크 규제는 주마다 다르게 적용될 가능성이 있다. 이와 함께, 2021년 미 상원 의원 Rob Portman이 발의한 「딥페이크 특별전문위원회 법안」(Deepfake Task Force Act)도 주목할 만하다. 이 법안은 국토안보부와 백악관 기술정책실과 협력하여 '딥페이크 테

스크 특별전문위원회(Deepfake Task Force)'를 설립하는 것을 주요 내용으로 한다. 이 특별 전문위원회는 딥페이크 및 AI 기반 콘텐츠 조작이 초래하는 위험을 분석하고, 디지털 콘텐츠의 출처 및 이용 기록(history) 검증을 위한 기술과 표준을 개발하는 역할을 수행하도록 규정하고 있다(김희정, 2024).

기술적 대응 측면에서 미국은 AI 기반 탐지 기술을 활용하여 딥페이크 허위정보의 확산을 억제하고 있다. 국방고등연구계획국(Defense Advanced Research Projects Agency, 이하 DARPA)의 'MediFor(Media Forensics)' 프로젝트는 딥페이크 데이터 수집 및 탐지 기술 고도화를 위해 정부 차원에서 연구(DARPA, 2024)를 지원하는 대표적인 사례로 주목받고 있다. 이 프로젝트는 다양한 딥페이크 사례를 포함한 데이터 셋을 구축하여 탐지 모델을 학습시키는 데 활용되고 있으며, AI 탐지 기술의 정교성을 높이는 데 기여하고 있다. 또한, 미국 국립표준기술연구소(National Institute of Standards and Technology, 이하 NIST)는 'MediFor' 프로젝트의 일환으로 'Media Forensic Challenge (MFC)' 데이터 셋을 개발하여 공개했다(NIST, 2021). 이 데이터 셋은 연구자들이 딥페이크 탐지 모델을 학습시키는 데 활용할 수 있도록 설계되었으며, 안전하고 윤리적인 데이터 셋 구축의 한 예로 평가된다. 주요 IT 기업들도 AI 탐지 기술 개발에 적극적으로 나서고 있다. 유튜브는 2024년부터 AI 기반 탐지 알고리즘을 도입하여 허위정보가 포함된 AI 생성 콘텐츠에 'AI Generated' 태그를 자동 부착하는 기능을 추가하였다(YouTube Official Blog, 2024). 메타는 AI 기반 신고 시스템을 도

입하여 이용자가 딥페이크로 의심되는 콘텐츠를 신고하면 AI가 자동으로 분석하여 대응하는 알고리즘을 개발하였다(AI at Meta Blog, 2023). X(구 트위터)는 사용자 신고 시스템과 연계하여 허위정보가 포함된 딥페이크 콘텐츠를 감시하고, AI 기반 필터링 시스템을 도입하는 방식으로 대응하고 있다(X Help Center, 2025). 딥페이크 탐지 기술의 신뢰성을 높이기 위한 기술적 대응 방안으로 워터마킹(Watermarking)과 C2PA(Coalition for Content Provenance and Authenticity) 표준이 논의되고 있다. C2PA는 구글, 마이크로소프트, 어도비(Adobe) 등 글로벌 IT 기업들이 참여하는 프로젝트로, AI 생성 콘텐츠에 디지털 서명을 부착하여 출처를 검증할 수 있도록 하는 기술을 개발하고 있다(C2PA, 2025).

미국의 일부 주에서는 딥페이크와 가짜뉴스 대응을 위한 미디어 리터러시 교육을 적극적으로 도입하고 있으며, 이는 허위정보의 확산을 막기 위한 사회적 대응 전략으로 활용되고 있다. 최근 여러 주의 입법 기관에서는 미디어 리터러시 교육의 중요성을 인식하고 이를 법제화하는 움직임을 보이고 있는데, 현재까지 미국 19개 주 의회에서 미디어 리터러시 교육 관련 법안을 발의하거나 논의되고 있다. 이중 18개의 주는 유치원부터 12학년까지의 교육과정 내에 미디어 리터러시 및 디지털 시민 교육 관련 법안에 서명하였다(〈그림 1〉 참조). 특히, 미국에서 가장 많은 초중고생이 있는 캘리포니아주는 교사의 연수과정에 미디어 리터러시 교육을 포함하도록 하였다. 미디어 리터러시 교육을 공교육의 핵심 과제로 설정하고, 교사들을 위한 연수 프로그램을 마련하며, 교육 예산을 따로 배정



〈그림 1〉 미국 내 미디어 리터러시 교육 법제화 현황

출처: <https://chatgpt.com/c/67ad4c14-4578-800a-85ef-f9c477bca40c>

하고 있다(Media Literacy Now, 2024).

이러한 미디어 리터러시 교육 강화는 디지털 플랫폼을 통한 정보 확산 속도가 빨라지는 환경에서, 개인이 스스로 정보를 검증하고 허위 정보를 판별할 수 있는 역량을 갖추는 데 필수적이다. 같은 맥락에서 미국의 '미디어 와이즈 (MediaWise)' 프로그램은 대표적인 디지털 미디어 리터러시 교육 사례로 주목받고 있다. 이 프로그램은 비영리 언론 교육기관인 포이터 연구소(The Poynter Institute)가 운영하는 프로젝트로 시민들이 온라인 허위정보를 효과적으로 판별할 수 있도록 교육하는 것을 목표로 한다. 특히 청소년을 대상으로 한 '십 대 팩트체크 네트워크(Teen Fact-Checking Network)'는 중·고등학생들로 구성된 가상 뉴스룸으로, 소셜 미디어에서 확산되는 허위정보를 검증하는 역할을 수행하고 있다(MediaWise, 2025). 이 외에도 FactCheck.org, Snopes, PolitiFact

등 주요 팩트체킹 기관들은 AI 기반 분석 도구를 활용하여 뉴스 및 소셜미디어에서 유포되는 허위정보를 탐지하고 검증하는 시스템을 운영하고 있다. 또한 AI 기반 데이터베이스와 클라우드 소싱을 결합하여 시민들이 직접 허위정보를 신고하고 AI가 이를 분석해 신뢰도를 평가하는 방식으로 운영하고 있다(이정현, 박소영, 2024).

3.3 유럽연합

유럽연합은 디지털 플랫폼을 규제하기 위한 「디지털서비스법」(Digital Services Act, 이하 DSA)과 AI 기술을 규제하기 위한 「인공지능법」(Artificial Intelligence Act, 이하 AI 법)을 통해 딥페이크 기반 허위정보 및 가짜뉴스 확산에 대응하고 있다. 또한 미국과는 달리 플랫폼 사업자와 딥페이크 시스템의 공급자, 사용자 모두에게 일정한 의무를 부과하고 있다

(강준모, 2024). DSA 제35조에서는 대형 온라인 플랫폼(VLOPs)과 초대형 검색 엔진(Very Large Online Search Engine) 사업자에게 딥페이크 콘텐츠 관련 의무를 부과하고 있으며, 특히 허위정보(Misinformation)에 대한 위험 평가 및 감시 조치를 플랫폼 사업자가 수행하도록 하고 있다(최진웅, 2024). DSA 제34조에 따르면, 월평균 이용자 수가 4,500만 명 이상인 대형 온라인 플랫폼 사업자는 알고리즘, 추천 시스템, 광고 시스템이 불법 콘텐츠뿐만 아니라 허위 조작 정보와 같은 유해한 콘텐츠와 관련하여 위험을 초래하는지 평가해야 한다. 특히, DSA 제35조 제1항에 따르면, 실존 인물, 물체, 장소 등을 실제처럼 보이도록 생성·조작한 이미지, 오디오 또는 비디오가 플랫폼에 게시될 경우, 이용자가 쉽게 이를 식별할 수 있도록 ‘눈에 잘 띠는 표시’를 해야 하며, 서비스 이용자가 직접 표시 기능을 활용할 수 있도록 해야 한다. 이는 딥페이크 콘텐츠의 오인을 방지하고 정보 소비자가 AI 생성 콘텐츠를 명확히 인식할 수 있도록 하기 위한 조치라고 할 수 있겠다. AI 기반 기술을 규제하기 위한 법은 AI 법이 있다. 이 법은 2024년 유럽의회를 통과하여 2026년부터 시행될 예정이다. AI 법은 AI 시스템을 리스크 수준에 따라 4단계로 구분하고 각 단계별로 규제 강도를 달리 적용한다. 이 중 ‘고위험 AI 시스템(High-Risk AI Systems)’은 사회적 영향이 크고 규제 및 감독이 필요한 AI 기술을 의미하며, 딥페이크 콘텐츠는 AI 챗봇, AI 생성 텍스트·이미지·비디오 등과 함께 ‘제한된 위험(Limited Risk)’ 범주에 포함된다. 따라서 딥페이크 콘텐츠를 생성하는 AI 시스템 공급자는 AI 생성 또는 조작된 것임을 명확

히 표시하여야 하며, 콘텐츠에 기계가 판독할 수 있는 형태의 디지털 워터마킹 또는 메타데이터를 삽입하여야 한다. 또한 사용자가 AI 생성 콘텐츠임을 인지할 수 있도록 플랫폼에서 딥페이크 표시를 하여야 한다. 이 같은 법적 규제 체계를 통해 유럽연합은 AI 기반 콘텐츠 생성 및 탐지 기술의 표준을 정립하고, 플랫폼 운영자와 기술 개발자의 책임을 명확히 규정함으로써 딥페이크 허위정보 확산에 대응하고 있다.

유럽연합에서는 딥페이크 탐지 기술의 신뢰성을 강화하고, 미디어 환경에서 허위정보 확산을 최소화하기 위한 연구가 활발히 진행되고 있다. 유럽 내 주요 연구소에서는 딥페이크 콘텐츠를 효과적으로 탐지할 수 있는 AI 알고리즘을 개발하고 있으며, 이러한 기술적 대응은 유럽연합의 AI 관련 법제 및 정책과 연계되어 점진적으로 발전하고 있다. 대표적으로 독일의 응용과학 연구소인 프라운호퍼 연구소(Fraunhofer Institute)는 AI 기반 탐지 기술 연구를 선도하고 있으며, 특히 온라인 플랫폼과 뉴스 검증 시스템에 활용할 수 있는 탐지 기술 개발에 주력하고 있다. 이 연구소는 영상 내 인물의 얼굴 움직임을 분석하여 변조 여부를 판별하는 알고리즘을 개발하고 있으며, AI로 생성된 합성 음성을 탐지하기 위한 음성 패턴 분석 모델도 구축하고 있다. 이러한 연구는 유럽 내 뉴스 및 미디어 플랫폼이 AI 탐지 기술을 활용하여 허위정보를 실시간으로 감지할 수 있도록 지원하는 것을 목표로 한다(Fraunhofer Institute, 2025). 또 다른 연구기관인 막스 플랑크 연구소(Max Planck Institute)는 AI 및 머신러닝 알고리즘의 기초 이론을 연구하며, AI 기술의 신뢰성과 윤리적 사용을 보장하기 위한 연구를 수행하고

있다. 머신러닝을 이용한 데이터 패턴 분석, AI 알고리즘의 공정성 및 투명성 연구를 통해 딥페이크 탐지 기술의 이론적 토대를 마련하고 있으며, 이러한 연구 결과는 유럽연합의 AI 법 및 DSA의 기술적 규제와 연계되어 활용되고 있다. AI 법에서 규정하는 '고위험 AI 시스템(High-Risk AI Systems)'의 기준을 설정하는 과정에서 막스 플랑크 연구소의 연구 결과가 반영되고 있으며, 딥페이크 탐지 모델이 보다 신뢰성 높은 방식으로 운영될 수 있도록 기여하고 있다(European Commission, 2025). 유럽 내 주요 언론사들은 AI 기술을 활용하여 뉴스 검증 시스템을 강화하는 연구를 진행하고 있으며, 유럽 방송사 연합(EBU, European Broadcasting Union)은 BBC, AFP(Agence France-Presse), Deutsche Welle 등과 협력하여 AI 기반 뉴스 검증 시스템을 개발하고 있다. 이 시스템은 뉴스 영상 및 사진의 진위 여부를 AI가 자동 분석하는 기술을 적용하고 있으며, 예를 들어 뉴스 영상 속 인물의 입 모양과 음성이 일치하는지 확인하거나 이미지의 변조 흔적을 분석하여 허위정보를 조기에 감지하는 방식으로 운영된다. 이러한 기술을 도입함으로써 AI 기반 기술을 법적 규제와 연계하여 허위정보 확산을 방지하고, 미디어 콘텐츠의 신뢰성을 높이는 체계를 구축하고 있다(EBU, 2024). 이 같은 기술적 대응은 AI 법 및 DSA의 규제와 결합되어 미디어 신뢰성을 높이고, 딥페이크 기반 허위정보의 확산을 억제하는 역할을 하고 있다.

유럽연합은 딥페이크 기반 허위정보 확산을 효과적으로 방지하기 위해 법적·기술적 대응 뿐만 아니라, 시민들의 정보 판별 역량을 강화

하는 사회적 대응에도 주력하고 있다. 이를 위해 디지털 리터러시 교육을 확대하고, 시민 참여형 팩트체킹 시스템을 활성화하여 허위정보 및 가짜뉴스 대응 능력을 높이고 있다. 특히, 시청각미디어 서비스 지침(Audiovisual Media Services Directive, AVMSD)을 통해 회원국들에게 미디어 리터러시 교육을 의무화함으로써, 초·중등 교육과 일반 시민 대상 교육이 강화되도록 하고 있다. 지침에 따라 프랑스, 독일, 네덜란드 등 여러 국가에서는 초·중등학교 및 일반 시민을 대상으로 딥페이크 탐지 기술을 포함한 교육을 운영하고 있다(European Audiovisual Observatory, 2025). 교육 프로그램은 시민들이 AI 기술을 활용한 정보 조작을 인식하고 대응할 수 있도록 지원하며, 교육을 통해 딥페이크 허위정보를 판별하는 역량을 강화하는 것을 목표로 한다. 'EUvsDisinfo'는 유럽연합이 운영하는 공식 허위정보 대응 플랫폼으로, 러시아를 포함한 다양한 출처에서 유포되는 허위정보와 정보 조작에 대응하기 위해 2015년에 출범하였으며 'EU 외교·안보정책 서비스(European External Action Service, EEAS)' 산하에서 직접 운영하고 있다. 일반 시민이 허위정보를 신고할 수 있는 기능을 제공하고 있으며, 신고된 정보는 전문가 및 팩트체킹 기관의 검토를 거쳐 분석 후 공개하고 있다. EUvsDisinfo는 단순한 팩트체킹 프로젝트를 넘어, 허위정보 감시, 팩트체킹, 시민 참여형 신고 시스템, 미디어 리터러시 교육을 결합하여 허위정보 확산을 방지하고 신뢰할 수 있는 정보 환경을 조성한다. 이 플랫폼은 유럽연합이 허위정보 대응을 위해 조직적이고 체계적인 노력을 기울이고 있음을 보여주는 핵심적인 역할을 수행하고 있다

(EUvsDisinfo, 2025). 이와 함께, 유럽 내 공공 기관들은 디지털 플랫폼과 협력하여 혀위정보 대응 캠페인을 진행하고 있으며, 대표적으로 EU 집행위원회와 유네스코(UNESCO)는 'Think Before Sharing' 캠페인을 통해 혀위정보 판별 능력을 강화하는 공공 교육을 실시하고 있다 (Matthew Horwood, 2022). 이러한 캠페인은 시민들에게 혀위정보의 위험성을 인식시키고 신뢰할 수 있는 정보 판별법을 제공함으로써 온라인상에서의 정보 소비 역량을 높이는 데 기여하고 있다. 유럽연합의 사회적 대응은 법적·기술적 조치를 넘어 시민들의 참여와 교육을 기반으로 혀위정보 확산을 방지하는 다층적 전략을 구축하고 있다는 점에서 의의를 가진다. 특히, DSA를 통해 온라인 플랫폼의 책임을 명확히 하면서, AI 탐지 기술과 시민 참여형 대응 시스템을 결합하여 혀위정보를 차단하는 포괄적인 접근 방식을 적용하고 있다. 이러한 대응 방식은 AI 기술이 발전하면서 더욱 정교해지는 딥페이크 혀위정보에 대응하기 위한 효과적인 방안으로 평가되며, 다른 국가들에도 중요한 정책적 시사점을 제공할 수 있다.

3.4 영국

영국은 딥페이크 기반 혀위정보 확산을 방지하기 위해 기존 법률을 적용하는 동시에 새로운 법적 규제를 도입하여 대응하고 있다. 현재는 「명예훼손법」(Defamation Act, 2013), 「프라이버시 보호법」(Data Protection Act, 2018) 등을 활용하여 딥페이크 관련 범죄를 규제하고 있으나, 보다 직접적인 대응을 위한 입법 논의가 활발히 진행되고 있다. 2023년 10월 제정된

「온라인 안전법」(Online Safety Act)은 온라인 플랫폼의 책임을 강화하는 법적 틀을 제공하며, 소셜 미디어와 검색 서비스를 운영하는 기업이 딥페이크 관련 콘텐츠 신고 시스템을 구축하고, 신고된 불법 콘텐츠를 신속히 삭제할 것을 의무화하고 있다(최진웅, 2024). 해당 법을 준수하지 않을 경우, 영국 방송통신규제 기구(Ofcom)는 해당 기업에 최대 1,800만 파운드 또는 전 세계 수익의 최대 10%에 해당하는 벌금을 부과할 수 있도록 규정하고 있다. 한편, 영국 정부는 딥페이크 성적 이미지 제작 자체를 범죄로 규정하는 입법을 추진하고 있다. 2024년 발의된 「범죄 치안 법안」(Criminal Justice Bill)은 당사자의 동의 없이 딥페이크 기술을 이용하여 성적 이미지나 영상을 생성하는 행위를 불법화하고, 이를 위반할 경우 최대 2년의 징역형을 부과할 수 있도록 규정하고 있다(김지연, 2025). 이 법안은 기존의 「성범죄법」(Sexual Offenses Act, 2003) 및 「온라인 안전법」이 공유 또는 유포 행위를 중심으로 규제했던 것과 달리, 제작 단계에서부터 범죄로 규정한다는 점에서 보다 강력한 대응을 의미한다. 영국의 법적 대응은 온라인 플랫폼 사업자의 책임을 강화하고, 딥페이크 범죄에 대한 명확한 처벌 기준을 마련하는 방향으로 발전하고 있다. 그러나, 현재의 대응 방식이 기존 법률을 적용하는 방식에 의존하는 측면이 크며, 딥페이크 기반 혀위정보 확산을 보다 체계적으로 규제할 수 있는 독자적인 법률 제정이 필요하다는 지적이 제기되고 있다. 향후 영국은 법적 대응을 더욱 강화하고, AI 탐지 기술 및 공공 부문 협력을 확대하는 방향으로 개선해 나갈 필요가 있다.

영국은 정부 및 연구기관이 주도하는 딥페이크 탐지 기술 개발과 온라인 플랫폼의 책임 강화 정책을 추진하고 있다. 영국 국립 사이버보안센터(National Cyber Security Center, 이하 NCSC)는 AI 기반 딥페이크 탐지 기술 연구를 수행하고 있으며, 주요 기술 기업과 협력하여 딥페이크 콘텐츠 감지 시스템을 개발하고 있다 (NCSC, 2024). NCSC는 딥페이크 탐지 기술을 국가 안보, 선거 보호, 금융 사기 방지 등 다양한 분야에 적용하고 있으며, 특히 공공 안전과 관련된 영역에서의 활용에 중점을 두고 있다. 한편, BBC, ITN, 스카이 뉴스 등 영국 내 주요 언론사들은 AI 기반의 딥페이크 검증 기술을 도입하여 뉴스 영상의 진위를 검토하는 시스템을 운영하고 있다(김미경, 2023). 이러한 시스템은 뉴스 영상 속 인물의 음성 패턴과 입 모양이 일치하는지 분석하거나, 이미지의 픽셀 변화를 감지하여 변조 여부를 판별하는 방식으로 운영된다. 특히, AI가 영상의 원본 출처를 추적하여 허위정보 여부를 판별할 수 있도록 설계되었으며, 허위정보가 확산되기 전에 신속하게 차단하는 것이 목표이다(EBU, 2024). 영국 금융권에서도 딥페이크 기술을 악용한 금융 사기에 대응하기 위한 AI 탐지 시스템 도입이 이루어지고 있다. 영국 결제시스템 규제기관(PSR)은 푸시 결제 사기(Authorized Push Payment Fraud) 증가에 대응하기 위해 금융 기관이 AI 기반 탐지 시스템을 도입하도록 권고하고 있으며, 이는 딥페이크 기술이 신원 도용 및 금융 사기에 악용되는 것을 방지하는 역할을 수행할 수 있다(박지홍, 2024). 그러나, 이러한 금융권의 대응은 정부 주도의 기술적 대응과는 구별되며, 개별 금융기관 및 규제 기관

의 자율적 조치라는 점에서 한계를 가진다.

영국의 사회적 대응은 미디어 리터러시 교육 강화, 시민 참여형 팩트체킹 시스템 도입 등을 포함한 전략을 기반으로 이루어지고 있다. 정부와 비영리 단체들은 협력하여 초·중·고등 학생 및 일반 시민을 대상으로 딥페이크 및 가짜뉴스 콘텐츠를 식별하고 대응하는 능력을 배양하는 디지털 리터러시 교육을 운영하고 있다. 특히, 공립학교에서는 미디어 리터러시 교육을 필수 교과로 지정하여 학생들이 AI 기반 조작 정보를 판별하는 능력을 배양할 수 있도록 지원하고 있으며, 시민들이 온라인에서 접하는 정보의 신뢰성을 스스로 평가할 수 있도록 교육을 제공하고 있다(이정현, 박소영, 2024). 영국의 대표적인 팩트체킹 기관인 'Full Fact'는 딥페이크 허위정보 대응을 위한 시민 참여형 팩트체킹 활동을 확대하고 있다. Full Fact은 2010년 설립된 영국의 비영리 팩트체킹 기관으로 특히, 공공 정책, 언론 보도, 선거 관련 허위 정보 및 가짜뉴스를 검증하며, AI 기술을 활용하여 정치인의 발언과 미디어 콘텐츠를 자동 분석하는 시스템을 운영하고 있다. Full Fact은 페이스북 및 구글과 협력하여 소셜미디어에서 유포되는 딥페이크 영상과 변조된 이미지를 탐지하고 교정하는 작업을 진행하고 있으며, 시민들이 딥페이크 콘텐츠를 판별할 수 있도록 가이드라인을 제공하는 'Full Fact Toolkit'을 운영하고 있다. 이를 통해 허위정보 및 가짜뉴스가 빠르게 확산되는 것을 방지하고, 시민들이 딥페이크 기술을 악용한 정보 조작을 인식하고 대응할 수 있도록 지원하고 있다(시청자 미디어재단, 2022).

4. 한국의 딥페이크 혀위정보 대응 정책과 개선 방향

4.1 딥페이크 혀위정보 대응 현황 및 정책적 시사점

앞장에서 분석한 주요국의 대응 현황을 정리해 보면, 각국은 자국의 사회적·정치적 환경에 맞추어 다양한 방식으로 대응책을 마련하고 있으며, 일부 국가는 법적 규제를 강화하고 플랫폼의 책임을 명확히 하는 등 적극적인 조치를 취하고 있음을 알 수 있었다. 본 장에서는 상기 분석을 바탕으로 한국의 대응 현황을 평가하고, 주요국들의 정책을 참고하여 보완해야

할 점을 논의하고자 한다.

〈표 2〉는 한국을 비롯한 주요국들의 딥페이크 혀위정보 대응 정책을 정리한 표로, 각 나라마다 각국의 상황에 따라 법적, 기술적, 사회적 대응을 유기적으로 결합하여 딥페이크 혀위정보 확산을 방지하고 있다. 특히, 미국과 유럽연합은 플랫폼 사업자의 책임을 명확히 하고, AI 탐지 기술 적용을 법제화하는 등 강력한 규제 체계를 구축하고 있다. 반면, 한국의 대응은 특정 법률을 통한 개별적 규제와 기술 연구에 집중되어 있으며, 전반적인 대응 체계의 통합적 정비가 미흡한 실정이다. 따라서 한국이 주요국과 비교하여 어떠한 차이를 보이며, 대응 체계에서 보완해야 할 점이 무엇인지 면밀히 살

〈표 2〉 국가별 딥페이크 혀위정보 대응 정책 비교 및 시사점

| 비교 항목 | 법적 대응 | 기술적 대응 | 사회적 대응 |
|-------|---|---|--|
| 한국 | <ul style="list-style-type: none"> - 성범죄 중심, 선거 개입 규제 미비 - AI 콘텐츠 출처 표시 규제 부재 | <ul style="list-style-type: none"> - AI 탐지 기술 개발 중, 플랫폼 적용 제한적 - IT기업 자체 탐지기술 도입 - 법적 의무화 부재로 기술 적용 제한 | <ul style="list-style-type: none"> - 미디어 리터러시 교육 확대 중 - 혀위정보 신고 및 감시시스템 미흡 - 시민 참여형 팩트체킹 시스템 부족 |
| 미국 | <ul style="list-style-type: none"> - 주(州) 단위 규제만 마련 - 선거 개입·성범죄 중심 법안 강화 - 연방 차원 총괄 규제는 부재 | <ul style="list-style-type: none"> - 빅테크 기업 자율 규제 - 정부 주도의 AI 탐지 기술 개발 - C2PA 표준 도입 논의 | <ul style="list-style-type: none"> - 미디어리터리시 교육 일부 주(州) 의무화 - Think Before Sharing 등 대국민 인식 개선 중심 - 시민신고 시스템 및 혀위정보 감시 강화 |
| 유럽연합 | <ul style="list-style-type: none"> - DSA를 통한 플랫폼 규제 - AI으로 기술 대응 - 플랫폼 사업자의 책임 강화 | <ul style="list-style-type: none"> - AI 탐지 및 워터마킹 적용 - AI 기반 뉴스 검증 시스템 도입 - 공공 연구기관 주도 탐지 기술 개발 | <ul style="list-style-type: none"> - 미디어 리터러시 교육 필수화 - 시민 참여형 프로젝트 - 디지털 플랫폼을 통한 대중 캠페인 진행 |
| 영국 | <ul style="list-style-type: none"> - 기존 법 적용, 온라인 안전법 논의 중 - 온라인 플랫폼의 딥페이크 혀위정보 콘텐츠 삭제 의무화 | <ul style="list-style-type: none"> - 정부 주도 딥페이크 탐지 연구 (NCSC) - AI 기반 뉴스 검증 기술 적용 - 금융권 AI 탐지 기술 도입 | <ul style="list-style-type: none"> - 공교육 과정 내 미디어 리터러시 포함 - 팩트체킹 기관과 연계하여 검증 과정 강화 - 공교육 및 미디어 정책과 연결된 시민 참여형 캠페인 |
| 시사점 | <ul style="list-style-type: none"> - 딥페이크 혀위정보에 대한 법제적 범주 정립 - 플랫폼 사업자에 대한 책임 강화 | <ul style="list-style-type: none"> - 플랫폼 사업자의 AI 탐지 기술 적용 의무화 및 국제 표준 도입 | <ul style="list-style-type: none"> - 공교육으로의 확대 - 디지털 리터러시 강화 - 시민 참여형 감시시스템 구축 |

펴볼 필요가 있다.

먼저, 법적 대응 측면에서 한국은 개별 법률을 기반으로 딥페이크 허위정보를 규제하고 있으나, 전반적인 규제 체계는 아직 체계적으로 정비되지 않은 상태이다. 현재 「공직선거법」, 「성 폭력처벌법」, 「정보통신망법」 등 개별 법률을 통해 일부 대응이 이루어지고 있지만, 딥페이크 기술이 야기할 수 있는 사회적·경제적·정치적 위협을 종합적으로 관리할 수 있는 법적 틀은 마련되지 않았다. 예를 들어, 「공직선거법」은 AI 기반 조작 콘텐츠의 유포를 선거운동 기간 중에만 제한하고 있어 선거 외 기간에는 규제가 적용되지 않는다. 반면, 미국 일부 주에서는 선거 개입 목적으로 제작된 딥페이크 콘텐츠를 불법화하는 법안을 통과시켰으며, 유럽연합은 「디지털서비스법」을 통해 플랫폼 사업자에게 딥페이크 탐지 및 삭제 의무를 부과하는 등 강력한 규제를 시행하고 있다. 이처럼 미국과 유럽연합은 선거 개입을 포함하여 플랫폼의 책임을 강화하는 법적 대응을 적극적으로 추진하고 있지만, 한국의 경우 플랫폼 사업자의 자율 규제에 의존하고 있어 법적 강제력이 부족한 상태이다. 따라서 보다 포괄적인 법적 규제 체계를 마련할 필요성이 있다. 특히, 2024년 개정된 「공직선거법」 제82조의8을 통해 선거 전 90일부터 선거일까지 생성형 AI로 제작된 조작 영상의 제작·편집·유포·상영·게시를 금지하는 조항이 추가되었으나, 이는 특정 기간과 선거 관련 콘텐츠에 한정된 규제에 불과하다(최진웅, 2024). 즉, 비선거 기간에 발생할 수 있는 딥페이크 허위정보 문제에 대한 대응책이 부족하며, 이는 경제적 사기, 기업 명예 훼손, 허위 뉴스 확산 등의 문제로 이어질

수 있다. 따라서 딥페이크 불법 콘텐츠에 대한 명확한 정의와 이를 생산, 배포하는 플랫폼까지의 규제로 확대하는 방안을 모색해야 한다.

한국의 기술적 대응은 딥페이크 탐지 기술이 연구 단계에 머무르고 있어, 이를 플랫폼 사업자에게 적용하도록 의무화하는 법적 제도적 기반 마련이 필수적이다. 또한, 미국과 유럽연합이 추진 중인 C2PA 표준 및 워터마킹 기술 도입을 적극적으로 참고할 필요가 있다. 현재 한국은 과학기술정보통신부와 한국정보통신기술진흥센터(KICT)를 중심으로 AI 기반 딥페이크 탐지 기술을 연구하는 방식으로 이루어지고 있다. 한국정보통신기술협회(TTA)는 AI 기반 영상 분석 및 워터마킹 기술을 포함한 콘텐츠 진위 판별 시스템을 개발하고 있으며, 국내 주요 IT 기업인 네이버와 카카오도 자체적인 AI 콘텐츠 식별 기술을 도입하여 플랫폼 내 딥페이크 감지를 위한 연구를 진행 중이다. 예를 들어, 네이버는 '클로바 그린아이' AI 시스템을 적용하여 플랫폼 내 딥페이크 콘텐츠를 실시간으로 모니터링하고 있으며, 카카오는 딥페이크 허위정보가 포함된 콘텐츠를 신고할 수 있는 시스템을 운영하고 있다. 그러나 이러한 기술적 연구에도 불구하고, 현재 AI 탐지 기술이 실제 플랫폼에 의무적으로 적용되지 않는 한계를 지닌다. 한국의 경우 AI 기반 딥페이크 탐지 기술이 법적 강제력 없이 일부 기업 및 연구기관 차원에서만 이루어지고 있어, 실질적인 효과를 거두기 어려운 실정이다. 반면, 미국은 국방고등연구계획국(DARPA)의 'MediFor' 프로젝트를 통해 정부 주도로 AI 탐지 기술을 개발하고 있으며, 유럽연합은 'C2PA' 표준을 도입하여 AI 생성 콘텐츠의 출처를 추적하는 기술을

플랫폼 사업자에게 법적으로 의무화하고 있다. 이와 달리 한국은 AI 탐지 기술의 연구 개발이 진행되고 있음에도 불구하고, 이를 실제 플랫폼에 적용할 수 있는 법적·제도적 기반이 부족한 상황이다. 따라서 AI 탐지 기술이 연구 단계에 머무르지 않고 실질적으로 활용될 수 있도록 법·제도를 정비하고, 플랫폼 사업자의 책임을 강화하는 조치가 필요하다.

법적·기술적 대응의 미흡할수록 시민들이 허위정보를 인식하고 대응하는 능력을 키우는 사회적 대응이 더욱 중요해진다. 이를 위해 각국은 미디어 리터러시 교육과 시민 참여 시스템을 강화하는 방향으로 정책을 발전시키고 있다. 유럽연합은 공교육 과정에서 미디어 리터러시 교육을 필수화하고 있으며, 시민들이 딥페이크 기반 허위정보를 직접 신고할 수 있는 체계를 구축하고 있다. 미국 또한 팩트체킹 기관과 협력하여 AI 기반 허위정보 감시 시스템을 운영하고 있으며, 일부 주에서는 미디어 리터러시 교육을 필수 교과로 지정하고 있다. 반면, 한국은 미디어 리터러시 교육과 시민 참여 시스템이 미흡한 상태이다. 현재 한국언론진흥재단과 시청자미디어재단이 일부 학교에서 AI 기반 미디어 리터러시 교육을 제공하고 있으나, 이는 특정 대상에 한정되어 있어 전 국민을 대상으로 한 체계적인 교육 프로그램이 부족하다. 또한, 시민들이 직접 허위정보를 신고할 수 있는 체계적인 시스템이 마련되지 않았으며, 공교육 과정에서도 미디어 리터러시 교육이 필수적으로 포함되지 않아 대응력이 미흡한 상황이다. 이러한 한계를 해결하기 위해 미디어 리터러시 교육을 공교육 과정에 필수적으로 포함하고, 시민 참여형 허위정보 감시 시스템을 구축하는 등의 조치

가 요구된다. 또한, 허위정보 감지 및 대응을 위한 팩트체킹 기관과의 협력을 강화하여, 공공과 민간이 공동으로 허위정보를 탐지하고 이를 신속히 차단할 수 있도록 정책적 지원을 확대할 필요가 있다. 따라서 시민들이 딥페이크 허위정보를 비판적으로 분석하고 적극적으로 대응할 수 있도록 미디어 리터러시 교육을 강화하고, 시민 참여형 감시 시스템을 구축하는 것이 필수적이다. 다음 장에서는 이러한 시사점을 바탕으로 법적 대응을 강화하고, AI 탐지 기술의 실효성을 높이며, 시민 참여를 확대하는 정책적 개선 방안을 구체적으로 논의하고자 한다.

4.2 딥페이크 기반 허위정보 대응 정책 개선 방향

딥페이크 기술의 발전으로 인해 허위정보 조작 방식이 더욱 정교해지고 있으며, 이에 대한 대응 역시 개별적인 접근을 넘어 종합적인 전략이 요구된다. 한국의 대응은 법적·기술적·사회적 측면에서 각각 진행되고 있으나, 이들 조치가 유기적으로 결합되지 못해 실효성이 충분히 확보되지 않은 상태이다. 이에 법적 대응의 강화, 기술적 대응의 보완, 그리고 사회적 대응의 확대를 중심으로 보다 효과적인 정책 개선 방안을 제시하고자 한다.

4.2.1 법적 대응의 강화

한국은 AI 기반 허위정보를 포괄적으로 규율할 수 있도록 법적 체계를 보완해야 한다. 현재 딥페이크를 활용한 허위정보가 사회적으로 큰 문제를 야기하고 있음에도 불법 콘텐츠로 명확히 규정되지 않을 경우는 제작 및 유포 행위를

금지·처벌하는 것은 어렵다. 이는 헌법상에 보장된 표현의 자유와의 충돌 가능성 때문이며, 미국과 유럽연합 등의 주요국에서도 음란물 및 선거 관련 내용을 제외하면 딥페이크 기반 허위 정보의 유포 자체를 처벌하고 있지 않다. 따라서 한국도 딥페이크 콘텐츠 제작 자체보다는 온라인 플랫폼을 통한 유포 및 확산 규제에 초점을 맞춘 정책을 마련하는 것이 더욱 효과적일 수 있다. 미국과 유럽연합의 사례를 참고하여 플랫폼 사업자에게 딥페이크 탐지 및 삭제 의무를 부과하는 법안을 검토할 필요가 있다. 고려할 내용으로는 일정 규모 이상의 온라인 플랫폼 사업자에게 허위 조작 정보 대응 의무를 부과하여 콘텐츠 접근 차단 및 삭제 등의 조치를 강화하는 방안 등이 있다. 다만, 포털의 뉴스 기사 편집 논란과 같이 언론의 자유를 침해하거나 플랫폼이 자의적으로 콘텐츠를 제한한다는 논란이 발생하지 않도록 독립적인 외부 위원회를 두어 심의하는 방안이 논의되어야 한다. 「공직선거법」의 적용 범위를 보완하는 조치도 필요하다. 현재 이 법은 선거 기간에만 한정해 딥페이크 허위정보 유포를 규제하고 있으나, 이를 확대해 평상시에도 허위정보를 지속적으로 차단할 수 있는 법적 대응이 마련되어야 한다.

특히 온라인 플랫폼에서 허위정보가 빠르게 확산될 가능성을 고려하여 플랫폼 사업자의 책임을 강화하는 법적 조치가 요구된다. 유럽연합은 「디지털서비스법(DSA)」을 통해 온라인 플랫폼 사업자에게 AI 탐지 기술 적용을 의무화하고 있다. 한국도 이를 참고하여 딥페이크 콘텐츠를 신속히 감지하고 차단할 수 있도록 법적 규제를 정비할 필요가 있다. 또한, 미국과 유럽연합은 AI 생성 콘텐츠의 신뢰성을 확보

하기 위해 워터마킹과 디지털 서명 기술을 도입하고 있으며, 한국 역시 이러한 조치를 마련하여 정보의 신뢰성을 강화해야 한다. 아울러, 글로벌 플랫폼과 국내 플랫폼 간 규제 형평성을 확보하는 것도 중요한 과제이다. 유럽연합은 DSA를 통해 해외 플랫폼에도 동일한 법적 의무를 부과하고 있으며, 한국도 국내외 플랫폼 사업자 모두가 법적 규제를 준수하도록 정책적 정비가 필요하다.

4.2.2 기술적 대응의 보완

딥페이크 기반 허위정보를 효과적으로 차단하기 위해서는 탐지 기술의 고도화와 실시간 대응 시스템 구축이 핵심 과제이다. 이를 위해 탐지 기술의 정밀성을 높이고, 실시간 대응 체계를 강화하며, 콘텐츠 인증 기술을 도입하는 한편, 국제 협력을 확대하는 것이 중요하다. 딥페이크 제작 기술이 정교해질수록 탐지 기술도 발전하고 있지만, 새로운 조작 기법이 끊임없이 등장하면서 탐지 기술이 이를 따라잡기 어려운 상황이 지속되고 있다. 이를 해결하기 위해서는 대규모 학습 데이터 구축과 탐지 모델의 지속적인 업데이트가 필수적이다.

그러나 이러한 데이터를 수집하는 과정에서 몇 가지 중요한 한계점이 존재한다. 첫째, 고품질 딥페이크 데이터 셋 구축에는 막대한 비용이 소요된다. 원본 영상과 음성을 확보한 후 AI 모델을 활용해 변조하는 과정에서 대량의 연산 자원과 전문적 기술이 요구된다. 둘째, 데이터 수집 과정에서 윤리적·법적 문제가 발생할 가능성이 크다. 딥페이크 탐지를 위해서는 실제 인물의 얼굴, 음성, 영상 데이터를 포함해야 하지만, 이는 개인정보 보호 및 초상권 침해의 논

란을 야기할 수 있다. 특히, 유명인의 얼굴을 동의 없이 합성하거나 허위정보를 포함한 조작 콘텐츠를 학습 데이터로 활용할 경우, 법적 문제뿐만 아니라 사회적 신뢰도 저하로 이어질 수 있다. 이러한 문제를 해결하기 위해 일부 연구기관에서는 윤리적 문제를 최소화하면서도 탐지 모델의 학습에 필요한 데이터를 확보할 수 있도록 ‘합성 데이터(Synthetic Dataset)’를 활용하는 방안을 고려하고 있다(김민진, 2022). 즉, 실제 인물이 아닌 가상 인물(AI가 생성한 얼굴) 데이터를 이용하여 딥페이크 탐지 모델을 훈련하는 방식이다. 이 접근 방식은 실제 인물의 초상권 및 개인정보 보호 문제를 해결하면서도 탐지 모델이 정교한 딥페이크 사례를 학습할 수 있도록 지원하는 효과적인 방법으로 평가된다. 일례로 미국 DARPA의 ‘MediFor’ 프로젝트는 정부 차원에서 딥페이크 데이터 수집 및 탐지 기술을 고도화하는 대표적인 사례로 다양한 딥페이크 사례를 포함한 데이터 셋을 구축하고 이를 탐지 모델 학습에 활용하고 있다. 따라서 한국도 이에 준하는 공공 AI 데이터 셋 구축 프로젝트를 추진할 필요가 있다. AI 탐지 기술이 보다 효과적으로 활용되려면 법적 강제력을 강화하고 플랫폼의 자율 규제에 의존하는 현행 방식에서 벗어나 공공기관과 협력하여 실질적인 운영 방안을 마련하는 것이 중요하다.

또한 AI 탐지 기술을 실시간으로 적용할 수 있는 인프라를 구축하고, 온라인 플랫폼에서 허위정보가 확산되기 전에 차단할 수 있도록 실시간 탐지 시스템을 강화하는 것이 필수적이다. 딥페이크 탐지 기술이 발전하더라도 조작된 콘텐츠를 원천적으로 방지하는 인증 시스템이 함께 마련되지 않으면 실효성이 떨어질 수 있

다. 이를 해결하기 위해 콘텐츠 제작 단계에서부터 디지털 워터마킹 기술을 적용하여 진위를 확인하는 방안이 필요하다. 유럽연합은 C2PA 표준을 통해 AI 생성 콘텐츠에 디지털 서명을 삽입하도록 규정하고 있으며, 미국에서는 주요 IT 기업들이 블록체인 기술을 활용한 인증 시스템을 개발하고 있다. 한국도 이를 참고하여 미디어 콘텐츠의 원본성을 보증하는 기술적 인프라를 구축하고, 공공기관과 민간 IT 기업 간 협력을 통해 워터마킹 기술을 표준화하는 작업을 추진해야 한다.

딥페이크 기술은 국경을 초월하여 확산될 가능성이 크므로, 개별 국가의 대응만으로는 한계가 있다. 따라서 국제적인 딥페이크 탐지 기술 표준을 마련하고, 각국 정부 및 연구기관이 협력하는 글로벌 대응 체계를 구축해야 한다. 유럽연합은 C2PA 표준과 디지털서비스법(DSA)을 통해 AI 탐지 및 검증 기술 적용을 의무화하고 있으며, 미국도 주요 IT 기업들과 협력하여 AI 기반 탐지 기술의 표준을 정립하는 작업을 추진하고 있다. 이에 맞춰 한국도 국제기구(UNESCO, OECD 등) 및 글로벌 IT 기업들과 협력하여 글로벌 탐지 기술 표준 수립에 적극 참여해야 하며, 국내에서 개발된 탐지 모델이 해외에서도 활용될 수 있도록 기술 인증 체계를 마련할 필요가 있다.

4.2.3 사회적 대응의 확대

딥페이크 기반 허위정보의 확산을 효과적으로 방지하기 위해서는 시민들의 대응 역량을 높이는 것이 필수적이며, 이를 위해 디지털 리터러시 교육 강화 및 시민 참여형 대응 체제 구축이 필요하다. 특히 기존의 단순한 미디어 리

터러시 교육을 제공하는 것을 넘어, AI 기반 콘텐츠의 신뢰성을 검증할 수 있는 비판적 분석 능력을 배양하는 디지털 리터러시 교육으로 확대할 필요가 있다. 디지털 리터러시란 단순히 디지털 기기나 인터넷을 활용하는 능력을 넘어, 디지털 환경에서 제공되는 정보를 비판적으로 이해하고 분석하며, 이를 책임감 있게 활용하는 능력을 의미한다. 특히, 딥페이크로 조작된 정보를 사실로 받아들이지 않고 이를 판별할 수 있는 역량을 갖추는 것은 허위정보 확산을 방지하는 핵심 요소가 된다. 디지털 리터러시가 부족할 경우, 사용자는 딥페이크 기반 허위 정보를 사실로 오인할 가능성이 크고, 나아가 이를 무분별하게 확산시켜 사회적 혼란과 경제적 피해를 초래할 수도 있다. 이에 따라, 영상이나 사진을 단순히 소비하는 것이 아니라 출처를 확인하고, 정보를 교차 검증하며, 신뢰할 수 있는 매체와 전문가의 분석을 참고하는 능력을 함양하는 것이 중요하다.

현재 한국에서도 4차 산업혁명, 디지털 전환, AI 및 데이터 관련 정책에서 디지털 리터러시의 중요성이 지속적으로 강조되고 있지만, 실제 교육 현장에서 체계적으로 실행되지 못하고 있는 실정이다. 특히, 온라인을 통해 유통되는 정보의 출처 및 신뢰도를 분석하고 정보 조작 여부를 판단하는 비판적 사고 역량을 기르는 교육이 부족한 상황이다. 이에 반해, 유럽연합은 공교육 과정에서 디지털 리터러시 교육을 필수화하고 있으며, 미국 일부 주에서도 AI 기반 허위정보 대응을 위해 디지털 리터러시 교육을 강화하는 정책을 추진하고 있다. 한국도 이를 참고하여 디지털 리터러시 교육을 정규 교과 과정에 포함하고, 정보취약계층을 대상으

로도 교육을 확대할 필요가 있다.

단순히 미디어를 소비하는 입장에서 넘어 미디어를 제작하는 교육 과정에서도 딥페이크에 대한 윤리 교육을 포함해야 한다. AI 기반 딥페이크 영상과 딥보이스 기술이 일반화됨에 따라, 단순한 영상 편집 기술뿐만 아니라, AI 및 소프트웨어를 활용한 콘텐츠 제작 과정에서 고려해야 할 법적·윤리적 기준을 다루는 교육이 필수적이다. 특히, AI 생성 콘텐츠가 보편화되면서, 딥페이크 기술의 오남용을 방지하고 제작자의 책임과 윤리를 강조하는 가이드라인 마련이 요구된다. 이를 위해, 미국과 유럽연합에서 시행하는 공교육 과정 내 미디어 리터러시 교육의 무화 및 시민 참여형 팩트체크 시스템 사례를 참고하여 한국도 공교육 내에서 딥페이크 콘텐츠 판별 교육을 의무화하고, 시민 참여형 허위 정보 감시 시스템 도입을 추진할 필요가 있다.

딥페이크 및 AI 생성 허위정보를 비판적으로 분석할 수 있도록 인식 개선 캠페인 및 세미나를 개최하는 것도 효과적인 대응 방안이 될 수 있다. 공공기관, 미디어 기관, AI 전문가가 협력하여, 딥페이크 기술과 허위정보의 위험성, 판별 방법 등을 시민들에게 교육하고 홍보하는 체계를 구축하는 것이 중요하다. 이를 통해, 시민들의 허위정보에 대한 감수성을 높이고, 정보 검증 능력을 강화할 수 있도록 지원해야 한다.

또한 딥페이크 기반 허위정보를 판별하는 것에 넘어, 이러한 허위정보가 발견될 경우 신속히 차단하고 확산을 방지할 수 있는 적극적인 대응 체계 마련이 필요하다. 예를 들어, 시민 참여형 신고 시스템을 구축하여 허위정보 대응력을 높이는 방안도 효과적인 방안이 될 수 있다. 이를 통해 개인이 허위정보에 노출되는 것을

막을 뿐만 아니라 부정확한 정보의 확산을 방지할 수 있도록 신고 체계를 강화해야 한다. 또한, 관련 당국이 신속하고 적극적으로 대응할 수 있도록 하고, 이러한 사례를 교육 및 홍보에 활용하여 지속 가능한 선순환 체제를 마련하는 것이 바람직하다.

결론적으로, 한국이 딥페이크 기반 허위정보 확산을 방지하기 위해서는 법적·기술적·사회적 대응이 유기적으로 결합되어야 한다. 딥페이크 허위정보에 법적 적용 범위와 이를 생성, 유통하는 플랫폼 책임을 명확히 하고, AI 탐지 기술을 의무 도입하며, 디지털 리터러시 교육을 강화하는 등의 다각적인 접근을 제언하는 바이다.

5. 결 론

본 연구는 주요국의 딥페이크 기반 허위정보 대응 정책을 비교·분석하고, 이를 바탕으로 한국의 정책적 보완 방향을 제언하였다. 연구 결과, 미국, 유럽연합, 영국 등 주요국들은 법적·기술적·사회적 대응을 종합적으로 추진하며, 특히 플랫폼 책임 강화와 AI 탐지 기술 적용을 의무화하는 정책을 발전시키고 있음을 확인하였다. 반면, 한국은 디지털 성범죄와 선거 개입 방지에 초점을 맞춘 개별적 대응에 머물러 있어, 가짜뉴스에 의한 경제적 사기나 여론 조작 등 다양한 악용 가능성에 대한 포괄적 대응이 미흡한 상황이다.

이에 본 연구는 한국의 대응 방안을 법적 대응 강화, 기술적 대응 보완, 사회적 대응 확대라는 세 가지 방향으로 제언하였다. 첫째, 선거 개

입과 금융 사기 등 딥페이크 기반 허위정보 확산을 규율할 법적 근거를 마련하고, 플랫폼 사업자의 책임을 명확히 해야 한다. 둘째, AI 탐지 기술의 실효성을 높이기 위해 플랫폼에 적용하도록 법적 의무를 부과하고, 워터마킹 및 메타데이터 표기 의무화를 검토해야 한다. 셋째, 디지털 리터러시 교육을 정규 교과 과정에 포함하고, 시민 참여형 허위정보 감시 시스템을 구축하여 사회적 대응을 강화해야 한다.

본 연구는 주요국의 사례를 통해 국내 상황에 실질적으로 적용할 수 있는 정책 방향을 도출하였다는 점에서 의의를 갖는다. 특히, 법적·기술적·사회적 측면을 종합적으로 분석함으로써 한국 실정에 맞는 딥페이크 대응 정책에 대한 구체적인 방향을 제시하였다. 또한 정책적 분석을 넘어 문헌정보학적 관점에서 디지털 정보의 신뢰성 및 정보 유통구조 개선 방안을 제시한다는 학술적 의의를 가진다. 특히 정보의 신뢰성과 투명성을 확보하기 위한 플랫폼의 책임성과 콘텐츠 검증 시스템의 중요성을 분석함으로써 디지털 정보관리 및 이용자 역량 강화를 위한 정책적 방향을 제공한다는 점에서 문헌정보학 분야에 기여할 수 있다. 그러나 본 연구는 정책적 분석에 초점을 맞추었기 때문에 실증적 연구가 부족하다는 한계를 가진다. 향후 연구에서는 AI 탐지 기술의 효과성 평가, 플랫폼 규제의 실효성, 딥페이크 대응 정책이 사회적 신뢰에 미치는 영향 등을 실증적으로 검토하는 작업이 필요할 것이다. 본 연구의 논의가 향후 한국의 딥페이크 대응 정책 수립과 법·제도 정비에 실질적으로 기여하고, 나아가 신뢰할 수 있는 디지털 정보 환경 조성의 밑거름이 되길 기대한다.

참 고 문 헌

- 강준모 (2024). 딥페이크 관련 국내외 규제 동향 분석. 정보통신정책연구원, KISDI AI Outlook, 2024(18), 1-17. 출처: <https://eiec.kdi.re.kr/policy/domesticView.do?ac=0000190899>
- 김미경 (2023. 11. 23.). 정치권 '딥페이크' 비상령...영국 사이버보안센터 AI기술, 영국 총선 위협할 수 있다. AI라이프경제. 출처: <https://www.aifnlife.co.kr/news/articleView.html?idxno=22298>
- 김민진 (2022). 합성데이터의 부상. 정보통신정책연구원, AI Trend Watch, 1-10. 출처: <https://www.kisdi.re.kr/report/view.do?key=m2101113025339&masterId=4311435&arrMasterId=4311435&artId=706876>
- 김아영 (2024. 9. 9.). 로잔대회 서울선언문, '동성애 · 딥페이크 시대' 속 도전한 선교 과제는? 국민일보. 출처: <https://www.kmib.co.kr/article/view.asp?arcid=0020507444>
- 김지연 (2025. 1. 8.). 영국, 딥페이크 성적 이미지 만들기만 해도 처벌. 연합뉴스. 출처: <https://www.yna.co.kr/view/AKR20250108001800085>
- 김희정 (2024). 딥페이크 기술을 이용한 신종범죄에 대한 법정책적 시사점: 외국의 법정책 대응을 중심으로. 서강법률논총, 13(3), 97-128. <http://doi.org/10.35505/slj.2024.10.13.3.97>
- 디지털공론장 (2025). 딥페이크를 활용한 가짜뉴스. 출처: https://survey.beingdigital.kr/survey/deepfake_results.php
- 박지홍 (2024). 영(英)은행권, 딥페이크 사기 확산 추세에 대응. 하나금융경영연구소 Bi-Weekly Hana Financial Focus, 14(4), 8-9. 출처: <https://eiec.kdi.re.kr/policy/domesticView.do?ac=0000182043>
- 변희원 (2024. 8. 23.). 가짜뉴스 밟혀져도 알고리즘 추천…“소셜미디어, 英폭동 로켓 부스터”. 조선일보. 출처: https://www.chosun.com/economy/tech_it/2024/08/07/UUIDAFMQDZBA3KZP4SADWM2YPY/?utm_source=chatgpt.com
- 세계법제정보센터 (2024. 10. 16.). 미국 캘리코니아주, 인공지능 관련 18개 법률 제정. 출처: https://world.moleg.go.kr/web/dta/lgs/TrendReadPage.do?CTS_SEQ=53860&AST_SEQ=315&ETC=2
- 시청자미디어재단 (2022). [팩트체크 플랫폼 소개] 영국 팩트체크 기관 Full Fact. 팩트체크 동향리포트 'FACT(팩트)', 2022(5), 17-22. 출처: https://www.dbpia.co.kr/pdf/pdfView.do?nodeId=NODE11112210&googleIPSandB0x=false&mark=0&minRead=10&ipRange=false&b2cLoginYN=false&icstClss=010000&isPDFSizeAllowed=true&accesssgl=Y&language=ko_KR&hasTopBanner=true
- 여성가족부 (2024. 9. 27.). 딥페이크 성착취물로 아동 · 청소년 협박 시 '징역 3년'…강요 '5년'. 대한민국 정책브리핑.

- 출처: https://www.korea.kr/news/policyNewsView.do?newsId=148934495&pWise=sub&pWiseSub=C7&utm_source=chatgpt.com
- 이민석 (2023. 8. 15.). 대선 뒤덮는 딥페이크… 美정부, 'AI 가짜뉴스'에 칼 뽑았다. 조선일보.
출처: https://www.chosun.com/international/us/2023/08/15/B4WFPERKAJD63HAMB5WLWY2SP4U/?utm_source=chatgpt.com
- 이숙종 (2024. 3. 6.). [가짜뉴스와 민주주의 시리즈] 허위조작정보에 대한 대응: 국제적 규제 추세와 한국의 대응 방안. 동아시아연구원 워킹페이퍼.
출처: https://eai.or.kr/new/ko/pub/view.asp?board=kor_workingpaper&intSeq=22419&keyWord=&keyword_option=&more=&utm_source=chatgpt.com
- 이정현, 박소영 (2024). 인공지능 팩트체크와 '사실성'의 기술사회적 정의. *한국언론학보*, 68(4), 119-157.
<https://doi.org/10.20879/kjjcs.2024.68.4.004>
- 임지우 (2024. 2. 5.). [영화 속 얘기 아냐... 금융사 직원, 딥페이크에 속아 340억원 송금]. 연합뉴스.
출처: <https://www.yna.co.kr/view/AKR20240205050200009>
- 장미경 (2024). 인공지능 시대의 허위정보와 가짜뉴스. *한국과학창의재단 KOSAC 이슈페이퍼*, 2024-7호.
출처: https://www.kosac.re.kr/menus/249/boards/476/posts/40443?contentYn=Y&page=1&utm_source=chatgpt.com
- 정성호 (2022. 3. 17.). [우크라 침공] 우크라 대통령 항복선언?…페북·유튜브, 딥페이크 동영상 삭제.
연합뉴스. 출처: <https://www.yna.co.kr/view/AKR20220317057500091>
- 최성철 (2024. 9. 9.). 딥페이크란 무엇이고, 기업에 어떤 영향을 미치는가?. 삼성SDS 인사이트 리포트.
출처: https://www.samsungsds.com/kr/insights/what-is-a-deepfake.html?utm_source=chatgpt.com
- 최인준 (2023. 10. 4.). 톰 햄크스도 당했다… 美 할리우드 습격한 'AI 딥페이크'. 조선일보.
출처: https://www.chosun.com/economy/tech_it/2023/10/03/2HZPQJKDWVEFTMKQUB5MYHRHPU/
- 최진웅 (2024). 인공지능 기반 딥페이크에 대한 해외 법제 및 시사점. 국회입법조사처, NARS현안분석 제338호.
- 최현빈 (2024. 9. 13.). 성범죄, 명예훼손, 사기, 선거법...딥페이크는 모든 범죄로 확산중. *한국일보*.
출처: https://www.hankookilbo.com/News/Read/A2024091114330005356?utm_source=chatgpt.com
- 한국경제 (2024. 2. 10.). [이광빈의 플랫폼S] 스위프트 음란 딥페이크…'N번방 쓰나미' 불길한 전조?
출처: https://www.hankyung.com/article/202402106212Y?utm_source=chatgpt.com
- 한국과학기술기획평가원 (2024. 10. 11.). 글로벌 주요국, 커지는 '딥페이크' 우려 속 규제 마련 속도.
출처: https://www.kistep.re.kr/gpsTrendView.es?mid=a30200000000&list_no=3374&act

=view

한국정보통신기술협회 (2025). 인공지능(AI) 워터마크 기술 동향 보고서.

출처: <https://www2.korea.kr/briefing/pressReleaseView.do?newsId=156671542&pWise=sub&pWiseSub=C3>

한국지능정보사회진흥원 (2025). [AI REPORT 2024-13] AI 일상화 시대의 위협, 딥페이크 대응 방안.

출처: https://www.nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=82618&bcIdx=27675&parentSeq=27675

AI at Meta Blog (2023, October 6). Stable Signature: A New Method for Watermarking Images

Created by Open Source Generative AI. Available:

<https://ai.meta.com/blog/stable-signature-watermarking-generative-ai/>

Behre, J., Hölig, S., & Möller, J. (2024). Reuters institute digital news report 2024: Ergebnisse für deutschland (Working paper). Verlag Hans-Bredow-Institut.

<http://doi.org/10.21241/ssoar.94461>

Coalition for Content Provenance and Authenticity (2025). About. Available:

<https://c2pa.org/about/about/>

Data Protection Act (2018). Available: <https://www.legislation.gov.uk/ukpga/2018/12/contents>

Defense Advanced Research Projects Agency (2024). MediFor: AI-based Media Forensics Project.

Defense Advanced Research Projects Agency. Available:

<https://www.darpa.mil/search/results?q=Media+Forensics>

European Broadcasting Union (2024, June 11). Generative AI and Public Service Media. Available:

https://www.ebu.ch/files/live/sites/ebu/files/Publications/Reports/open/2024.06.11_Gen-AI-and-PSM_EN.pdf

Elvis Act. TN Code § 47-25-1105. Available:

<https://law.justia.com/codes/tennessee/title-47/chapter-25/part-11/section-47-25-1105>

European Audiovisual Observatory (2025). A Unique Information Source on the Audiovisual

Sector in Europe. Available: <https://www.obs.coe.int/en/web/observatoire/about>

European Commission (2025). Artificial Intelligence (AI) in Science. Available:

https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/artificial-intelligence-ai-science_en

EUvsDisinfo (2025). About. Available: <https://euvsdisinfo.eu/about/>

Fraunhofer Institute (2025). About Fraunhofer. Available:

<https://www.fraunhofer.de/en/about-fraunhofer.html>

Matthew Horwood (2022, August 05). UNESCO Launches 'Think before Sharing' Campaign to

- Stop Conspiracy Theories. Trending News. Available:
https://www.westernstandard.news/news/unesco-launches-think-before-sharing-campaign-to-stop-conspiracy-theories/article_5db389ee-141e-11ed-8510-332e8d8f6c1a.html
- Media Literacy Now (2024). U.S. Media Literacy Policy Report. Available:
<https://medialiteracynow.org/impact/current-policy/>
- MediaWise (2025). How to Spot Misinformation Online. Available:
<https://www.poynter.org/shop/fact-checking/how-to-spot-misinformation-online-july-2021/>
- National Cyber Security Center (2024). The Near-Term Impact of AI on the Cyber Threat. Available: https://www.ncsc.gov.uk/report/impact-of-ai-on-cyber-threat?utm_source=chatgpt.com
- National Institute of Standards and Technology (2021, June 06). User Guide for NIST Media Forensic Challenge (MFC) Datasets. Available:
<https://www.nist.gov/publications/user-guide-nist-media-forensic-challenge-mfc-datasets>
- Online Safety Act (2023). Available: <https://www.legislation.gov.uk/ukpga/2023/50/contents>
- Sexual Offences Act (2003). Available:
https://www.legislation.gov.uk/ukpga/2003/42/pdfs/ukpga_20030042_en.pdf
- X Help Center (2025, February). Violent Content. Available:
<https://help.x.com/en/rules-and-policies/violent-content>
- YouTube Official Blog (2024, March 18). How We're Helping Creators Disclose Altered or Synthetic Content. Available: <https://blog.youtube/news-and-events/disclosing-ai-generated-content/>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Byun, Hee-won (2024, August 23). Algorithm Recommendation even if Fake News is Revealed... “Social Media, 英 Riot Rocket Booster”. The Chosun Daily. Available:
https://www.chosun.com/economy/tech_it/2024/08/07/UUIXAFMQDZBA3KZP4SADWM2YPY/?utm_source=chatgpt.com
- Choi, Hyun Bin (2024, September 13). Sex Crimes, Defamation, Fraud, Election Law... Deepfake Is Spreading to All Crimes. Available:
https://www.hankookilbo.com/News/Read/A2024091114330005356?utm_source=chatgpt.com
- Choi, In-jun (2023, October 4). Even Tom Hanks Fell Victim... ‘AI Deepfake’ Takes Hollywood

- by Storm. The Chosun Daily. Available:
https://www.chosun.com/economy/tech_it/2023/10/03/2HZPQJKDWVEFTMKQUB5MYHRHPU/
- Choi, Jin-eung (2024). Overseas Legislation on AI-Based Deepfake and Implications. National Assembly Research Service, NARS Current Issues and Analysis No. 338.
- Choi, Seong-cheol (2024, September 9). What is Deepfake and How Does It Impact Companies?. Samsung SDS Insight Report. Available:
https://www.samsungsds.com/kr/insights/what-is-a-deepfake.html?utm_source=chatgpt.com
- Digital Public Forum (2025). Fake News Using Deepfake. Available: Available:
https://survey.beingdigital.kr/survey/deepfake_results.php
- Jang, Mikyung (2024). Disinformation and Fake News in the Age of Artificial Intelligence. Korea Foundation for the Advancement of Science and Creativity, KOSAC Issue Paper, 2024(7). Available:
https://www.kosac.re.kr/menus/249/boards/476/posts/40443?contentYn=Y&page=1&utm_source=chatgpt.com
- Jung, Sung-ho (2022, March 17). [Invasion of Ukraine] Ukrainian President Declares Surrender?... Facebook and YouTube Delete Deep Fake Videos. Yonhap News Agency. Available:
<https://www.yna.co.kr/view/AKR20220317057500091>
- Kang, Joon-mo (2024). Analysis of domestic and foreign regulatory trends related to Deepfake. Korea Information Society Development Institute, KISDI AI Outlook, 2024(18), 1-17. Available: <https://eiec.kdi.re.kr/policy/domesticView.do?ac=0000190899>
- Kim, Ah Young (2024, September 9). The Lausanne Conference Seoul Declaration: Mission Challenges in the Era of Homosexuality and Deepfake?. Available:
<https://www.kmib.co.kr/article/view.asp?arcid=0020507444>
- Kim, Hee-jung (2024). Legal and policy implications of new types of crime using deepfake technology: Focusing on legal and policy responses in foreign countries. Sogang Law Journal, 13(3), 97-128. <http://doi.org/10.35505/slj.2024.10.13.3.97>
- Kim, Ji-yeon (2025, January 8). UK: “Creating Deepfake Sexual Images Alone Is Punishable”. Yonhap News Agency. <https://www.yna.co.kr/view/AKR20250108001800085>
- Kim, Mi-kyung (2023, November 23). Political ‘Deepfake’ Emergency Decree... UK Cyber Security Center “AI Technology Could Threaten the UK General Election.” Life in Artificial Intelligence. Available: <https://www.aifnlife.co.kr/news/articleView.html?idxno=22298>

- Kim, Min-jin (2022). The Rise of Synthetic Data, Korea Information Society Development Institute, AI Trend Warch, 1-10. Available:
<https://www.kisdi.re.kr/report/view.do?key=m2101113025339&masterId=4311435&arrMasterId=4311435&artId=706876>
- Korea Foundation for Broadcast and Media Audience Rights (2022). [Introduction to fact-checking platforms] UK fact-checking organization Full Fact. Fact-Checking Trends Report 'FACT', 2022(5), 17-22. Available:
https://www.dbpia.co.kr/pdf/pdfView.do?nodeId=NODE1112210&googleIPSandBox=false&mark=0&minRead=10&ipRange=false&b2cLoginYN=false&icstClss=010000&isPDFSizeAllowed=true&accessgl=Y&language=ko_KR&hasTopBanner=true
- Korea Institute of S&T Evaluation and Planning (2024, October 11). Global Major Countries Speed up Regulation Amid Growing Concerns over 'Deepfake'. Available:
https://www.kistep.re.kr/gpsTrendView.es?mid=a30200000000&list_no=3374&act=view
- Lee, Jeonghyun & Park, Soyoung (2024). Artificial intelligence fact-checking technology and the sociotechnical definition of 'factuality'. Korean Journal of Journalism & Communication, 68(4), 119-157. <https://doi.org/10.20879/kjcs.2024.68.4.004>
- Lee, Min-seok (2023, August 15). Deepfakes Flooding the Presidential Election...The U.S. Government Drew a Sword for 'AI Fake News'. The Chosun Daily. Available:
https://www.chosun.com/international/us/2023/08/15/B4WFPERKAJD63HAMB5WLY2SP4U/?utm_source=chatgpt.com
- Lee, Sook-jong (2024, March 6). [Fake News and Democracy Series] Responding to Disinformation: International Regulatory Trends and Korea's Countermeasures. East Asia Institute Working Paper. Available:
https://eai.or.kr/new/ko/pub/view.asp?board=kor_workingpaper&intSeq=22419&keyw=ord=&keyword_option=&more=&utm_source=chatgpt.com
- Lim, Ji-woo (2024, February 5). [Not Just a Movie... Bank Employee Fooled by a Deepfake and Sent 34 Billion Won]. Yonhap News Agency. Available:
<https://www.yna.co.kr/view/AKR20240205050200009>
- Ministry of Gender Equality and Family (2024, September 27). If Children and Adolescents are Threatened with Deepfake Sexual Exploitation, '3 Years in Prison'... Forced "5 Years". Republic of Korea Policy Briefing. Available:
https://www.korea.kr/news/policyNewsView.do?newsId=148934495&pWise=sub&pWiseSub=C7&utm_source=chatgpt.com

- National Information Society Agency (2025). Threats in the Era of AI Normalization, Deepfake Countermeasures. The AI Report 2024. Available:
https://www.nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=82618&bcIdx=27675&parentSeq=27675
- Park, Ji-hong (2024). UK (英) banking sector responds to trend of spreading Deepfake fraud. Hana Financial Management Research Institute Bi-Weekly Hana Financial Focus, 14(4), 8-9. Available: <https://eiec.kdi.re.kr/policy/domesticView.do?ac=0000182043>
- Telecommunications Technology Association (2025). Report on Trends in Artificial Intelligence (AI) Watermarking Technology. Available:
<https://www2.korea.kr/briefing/pressReleaseView.do?newsId=156671542&pWise=sub&pWiseSub=C3>
- The Korea Economic Daily (2024, February 10). [Lee Kwang Bin's Platform S] Swift Deepfake Porn... "Tsunami in Room N." Ominous signs?. Available:
https://www.hankyung.com/article/202402106212Y?utm_source=chatgpt.com
- World Laws Information Center (2024, October 16). U.S. State of California Enacts 18 Laws on Artificial Intelligence. Available:
https://world.moleg.go.kr/web/dta/lgslTrendReadPage.do?CTS_SEQ=53860&AST_SEQ=315&ETC=2