

June 9, 2025

인공지능 생성물 워터마크 기술 동향

I. 들어가며

최근 생성형 인공지능이 급속도로 발전하면서 인공지능 관련 안정성 및 보안 이슈에 대한 논의도 활발해지고 있습니다. 2026. 1. 22. 시행을 앞두고 있는 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법(이하 “**인공지능 기본법**”)에서도 인공지능 투명성 확보를 위한 표시의무가 규정되었고, 이러한 표시 방법으로 인공지능 식별무늬(“**워터마크**”)가 고려되고 있습니다.

과학기술정보통신부(이하 “**과기부**”)와 한국정보통신기술협회는 올해 초 ‘인공지능 생성물’의 안전한 생성과 활용을 지원하기 위해 ‘인공지능 식별무늬(AI 워터마크) 기술 동향 보고서’(이하 “**보고서**”)를 발간하였습니다([보고서 링크](#)). 본 뉴스레터에서는 인공지능 생성물의 오남용 사례와 이를 방지하기 위한 워터마크 기술의 활용 방안의 등 보고서의 주요내용을 간단히 소개하고자 합니다.

II. 인공지능 생성물 및 오남용 사례

1. 인공지능 생성물의 의미

“생성형 인공지능”은 기존 딥러닝¹ 방식의 인공지능 기술과 달리 사용자 요구를 기반으로 능동적으로 데이터나 콘텐츠 등의 결과물을 생성하는 한 단계 더 진화한 기술로, 사용자가 입력한 데이터의 구조와 특성을 모방하여 글, 소리, 그림, 영상, 그 밖의 다양한 결과물을 생성하는 인공지능 시스템입니다.

“인공지능 생성물”은 이러한 생성형 인공지능을 통해 생성된 정보를 의미하고, 인공지능 생성물은 인공지능을 포함한 알고리즘에 의해 크게 변경되거나 생성된 이미지, 동영상, 오디오 클립, 텍스트 등의 정보를 포함합니다.

¹ “딥러닝” 방식의 인공지능은 인간의 뇌 신경망을 따라하여 데이터를 계층적으로 학습한 패턴을 기반으로 추론할 수 있도록 하는 기술을 의미함

2. 인공지능 생성물 오남용 사례 증가

최근 생성형 인공지능이 고도화됨에 따라 헬스케어, 디지털 비서 등 인간 삶의 질 개선 및 제조 분야 등 다방면에서 활용되면서, 인공지능 생성물이 사람이 만든 것인지 아닌지를 구별할 수 없어 가짜뉴스와 딥페이크 등 사회 혼란을 야기하고 있습니다. 특히 인공지능 생성물의 오남용으로 인한 가짜뉴스 생성, 보이스 피싱, 디지털 성범죄, 저작권 분쟁, 인공지능 표절 등의 다양한 문제가 실제 법적 분쟁으로도 이어지는 사례가 국내 외에서 증가하고 있습니다.

일례로, 생성형 인공지능 서비스에 ‘나체’, ‘어린이’ 등의 텍스트를 입력하여 360개의 성착취물을 제작한 사건이 발생하였고, 법원은 해당 행위가 아동 청소년의 성보호에 관한 법률을 위반하였다고 판시하여 징역 2년 6개월이 선고된 바 있습니다 (서울고등법원 2024. 11. 1. 선고 2024노1823 판결).

III. 인공지능 생성물에 대한 워터마크 활용 사례

1. 워터마크 기술

워터마크란 이미지 등의 매체에 로고나 텍스트 형태의 식별자를 추가하거나 사람이 인지할 수 없는 패턴을 삽입하여 콘텐츠의 출처 및 소유권 확인에 사용하는 기술을 의미합니다. 워터마크 표시는 사람의 눈에 인식 가능한지 여부에 따라 **(1) 인지 가능 워터마크**와 **(2) 인지 불가능 워터마크**로 분류되며, 이미지, 영상, 음성 등 인공지능 생성 결과물의 유형에 따라 적용되는 기술에 차이가 있습니다.

한편, 사람의 인지 여부에 따라 워터마크의 장단점이 달라질 수 있습니다. (1) 인지 가능 워터마크(예: 원본 이미지에 로고나 텍스트 등의 표시를 추가하는 방법)의 경우, 도입 난이도 및 비용이 낮은 장점이 있으나 미적으로 사용자의 경험을 해칠 가능성이 있고 사용자의 의도적인 손상 등에 대해서는 신뢰성이 낮다는 단점이 있습니다. 반면, (2) 인지 불가능 워터마크(예: 공간 기반 LSB², 변환기반 DCT³)의 경우, 생성물의 품질 및 시청각적 경험에 영향을 주지 않으면서 보안이 강화된다는 장점이 있으나 복잡한 기술이 사용되어야 하고 추출과 검증을 위해서는 탐지 도구에 의존하여야 해서 일반 사용자의 접근성이 제한된다는 단점이 있습니다.

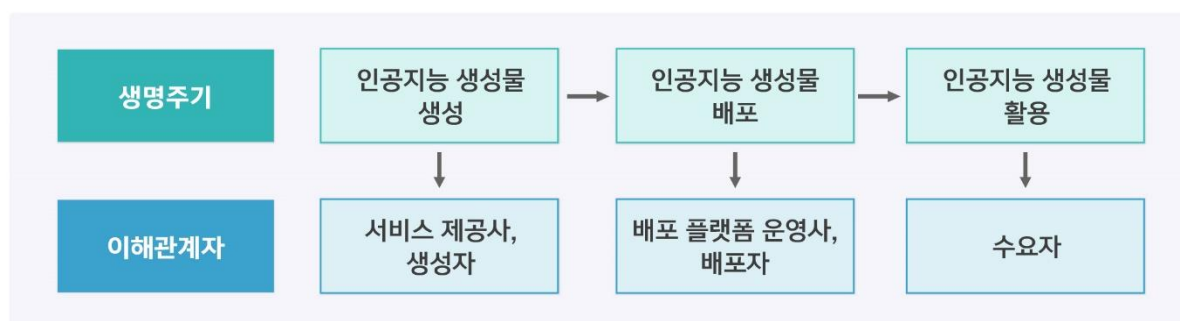
2 “LSB (Least Significant Bit)”란 이미지 픽셀에서 시각적으로 영향이 가장 적은 하위 비트(LSB)에 워터마크를 삽입하는 기술을 의미함

3 “DCT (Discrete Cosine Transform)”란 이미지를 블록 단위의 주파수 정보로 변환하여, 주파수 영역의 계수를 추출해 워터마크를 삽입하는 기술을 의미함

2. 인공지능 생성물 생명주기 및 이해관계자

인공지능 생성물의 생명주기는 ① 생성형 인공지능 서비스를 제공하거나 프롬프트를 통해 생성형 인공지능 서비스에 생성물 제작 요청을 하는 “**생성 단계**”, ② 생성된 인공지능 생성물을 플랫폼 또는 SNS 등을 통해 배포하는 “**배포 단계**” 및 ③ 플랫폼 등을 통해 인공지능 생성물을 활용하는 “**활용 단계**”로 구분할 수 있습니다. 이때, 생성 단계에서는 ‘서비스 제공사’ 및 ‘생성자’가, 배포 단계에서는 ‘배포 플랫폼 운영사’ 및 ‘배포자’가, 활용 단계에서는 ‘수요자’가 이해관계자로서 상호작용하게 되며, 이해관계자별 인공지능 생성물에 대한 워터마크 적용 방법을 구분하여 설정하는 방안이 논의되고 있습니다.

<인공지능 생성물 생명주기 및 이해관계자>



3. 워터마크 도입 현황

국내외 주요 기업(국내: 3개, 국외 12개)은 인공지능 생성물의 오남용 방지를 위해 자체적으로 워터마크 기술 도입하였습니다. 국내 기업의 경우 육안으로 인지 가능한 워터마크를 주로 활용하고 있고, 국외 기업들은 C2PA⁴ 및 자사 알고리즘 개발을 통하여 인지 불가능 워터마크의 도입을 추진하고 있습니다.

인지 가능 워터마크를 적용하는 사례로는 삼성전자 갤럭시 스마트폰의 사진 앱에서 생성형 인공지능 기반의 ‘포토 어시스트’ 기능을 제공함에 있어 인공지능 생성물에 갤럭시 AI 로고 및 문구(‘AI로 생성한 생성물’) 삽입하도록 한 사례, SK텔레콤에서 제공하는 인공지능 에이전트 서비스인 ‘에이닷’에서 인공지능 생성물에 대해 좌측 하단에 로고 삽입하도록 한 사례가 있습니다.

4 “C2PA (Coalition for Content Provenance and Authenticity)”란 디지털 콘텐츠의 출처와 신뢰성을 확인하기 위한 메타데이터 표준을 제공하는 국제 연합을 의미하며, Adobe, Microsoft, BBC 등 여러 글로벌 기업이 참여하고 있음 (<https://c2pa.org/>)

인지 불가능 워터마크를 적용하는 사례로는 Microsoft의 CoPilot⁵ 서비스에서 생성된 인공지능 생성물에 생성물 생성 시기, 자격 증명, 이미지 생성 도구 등의 정보가 포함된 메타데이터를 활용하며 이는 AI가 생성한 이미지 픽셀 내에 내장하도록 하고, 'GUID' 고유 코드를 포함하여 이미지가 언제, 어디서, 어떻게 만들어졌는지 Content Credentials Verify를 통해 추적이 가능하도록 한 사례가 있습니다.

4. 워터마크 도입 시 고려사항

워터마크 기술 도입이 인공지능 산업의 스타트업, 중소기업 등에 새로운 진입장벽이 될 가능성도 있어 신중하여야 합니다.

비교적 도입 난이도가 낮은 인지 가능 워터마크를 자체 제작하는 경우(간단한 로고, 텍스트 삽입 등), 오픈소스를 활용하거나 다른 상용 도구를 활용하는 것보다 비용이 저렴하나, 이미지에 글씨 인지가 되어 사용성이 저하될 수 있고, 의도적으로 생성물을 손상시키는 문제가 발생할 가능성이 존재합니다.

반면, 자체 연구를 통해 워터마크 기술을 만들 경우, 이에 따른 R&D 비용 발생하게 되는 바 비용이 과도하게 많이 들거나 특정 기업이 표시 방법을 독점한다면, 자금이 충분하지 않은 스타트업과 중소기업의 입장에서는 새로운 진입장벽으로 작용할 가능성이 있습니다.

또한 기업의 입장에서 생성형 인공지능 기업의 상용 도구를 활용할 경우, 워터마크 관련 라이선스 비용과 계약 및 기술 지원이 종료되면 대체 기술을 찾아야 한다는 문제점도 고려하게 될 것입니다.

IV. 워터마크 도입 관련 해외 동향

미국, 유럽연합, 중국 등 주요 국가는 인공지능 생성물의 오남용을 방지하기 위하여 인공지능 기술을 사용해 제작되었음을 표시하도록 명시하는 내용의 권고안 및 법률안을 제정하고 있습니다.

(1) 미국: 인공지능 기술규제 담당 기관의 역할을 강조하며, 인공지능 생성물에 명시 의무 자격을 구현하도록 하는 조항을 포함한 법안을 주(州)별로 추진하고 있습니다. 인공지능 생성물을 이용한 정치적 이용이나 아동 성착취물 제작 등 범죄적인 이용을 금지한 조항은 다수 도입되었으나 구체적 기준은 아직 미비한 상황입니다

⁵ Microsoft CoPilot은 Microsoft에서 사용자가 입력한 텍스트를 기반으로 이미지를 생성하는 서비스를 가리킴

(2) 유럽연합: 인공지능 생성물 이용자 보호 및 안전보장을 위한 디지털서비스법(Digital Services Act)을 공표하였습니다. 해당 법령은 인공지능 생성물(이미지, 오디오, 영상)이 진짜처럼 보일 경우 그 인공지능 생성 또는 조작 여부에 대해 이용자들이 인지할 수 있도록 플랫폼 인터페이스 상에 명확하게 표시(마킹)할 것을 규정⁶하고 있으며, 이러한 의무는 2023. 8. 부터 대형 온라인 플랫폼 사업자들에 적용되고 있습니다. 위 표시 의무를 위반할 경우 전 세계 매출액의 6%에 해당하는 과징금⁷이 부과될 수 있습니다.

또한 유럽연합의 인공지능법(Artificial Intelligence Act)는 생성형 인공지능 시스템 공급자로 하여금 특정 시스템의 결과물이 인공지능 생성물임을 워터마크, 메타데이터 등 기계가 판독할 수 있는 형태로 표시해야 함을 규정⁸하고, 인공지능 생성물 표시의무 위반 시 전 세계 매출액의 최대 3% 또는 1,500만 유로의 과징금⁹을 부과할 수 있는 제재조항을 규정하고 있습니다.

(3) 중국: 생성형 인공지능 서비스 사업자에게 강력한 의무를 부과하고 있으며, 대중에게 생성형 인공지능 기술을 이용한 텍스트, 이미지, 오디오, 비디오 등의 콘텐츠를 제공한 경우 관련 내용을 표기하도록 규정¹⁰하고 있습니다.

(4) 기타 국가: 캐나다, 싱가포르, 프랑스, 독일 등은 인공지능 생성물을 불법적인 방법으로 사용하는 것을 처벌하거나 해당 인공지능 생성물의 삭제 조치를 강제할 수 있는 근거 규정을 마련하고 있습니다.

V. 시사점

국내에서는 2025. 1. 21. 제정된 인공지능 기본법이 2026. 1. 22. 시행을 앞두고 있으며, 해당 법에서 인공지능 생성물에 대한 투명성 확보 의무¹¹를 명문화하고 있습니다. 워터마크 도입을 포함한 표시의 방법에 관한 구체적인 사항은 하위법령인 인공지능 기본법 시행령에서 정해질 것으로 보이며, 2025. 6월 내지 7월경 고시 및 가이드라인과 함께 공개될 것으로 기대됩니다.

⁶ 유럽연합 디지털서비스법 제35조 (k)호 참조

⁷ 유럽연합 디지털서비스법 제52조 제3항 참조

⁸ 유럽연합 인공지능법 제50조 제2항 참조

⁹ 유럽연합 인공지능법 제99조 제4항 (g)호 참조

¹⁰ 중국 생성형 인공지능 서비스 잠정관리 방법(生成式人工智能服务管理暂行办法) 제12조 참조

¹¹ 인공지능 기본법 제31조에 따르면, 고영향 인공지능 또는 생성형 인공지능을 이용한 제품 또는 서비스를 제공하는 인공지능사업자는 해당 사실을 이용자에게 사전에 고지하여야 하며, 생성형 인공지능 또는 이를 이용한 제품 또는 서비스를 제공하는 경우 그 결과물이 생성형 인공지능에 의하여 생성되었다는 사실을 표시하여야 함

생성형 인공지능 시장이 확대되고 있는 만큼, 국내 인공지능 산업의 신뢰성을 높이고, 국제 경쟁력을 강화할 필요성이 있습니다. 워터마크 표시 활용 방안은 인공지능 생성물의 안전한 활용과 투명성 확보에 크게 기여할 것으로 보입니다.

과기부는 전담반 운영과 의견수렴을 통해 방침(가이드라인)을 마련하여 안전하고 신뢰할 수 있는 인공지능 산업 생태계를 조성하겠다는 입장이므로, 정부의 입장과 규제 동향을 계속해서 주시할 필요가 있습니다.

관련 구성원

강태욱

변호사

T 02.3404.0485

E taeuk.kang@bkl.co.kr

윤주호

변호사

T 02.3404.6542

E juho.yoon@bkl.co.kr

강정희

변호사

T 02.3404.6480

E jeonghee.kang@bkl.co.kr

법무법인(유한) 태평양의 뉴스레터에 게재된 내용 및 의견은 일반적인 정보제공만을 목적으로 발행된 것이며, 법무법인(유한) 태평양의 공식적인 견해나 어떤 구체적 사안에 대한 법률적 의견을 드리는 것이 아님을 알려드립니다. 뉴스레터와 관련된 문의사항이 있을 경우 위 연락처로 문의주시기 바랍니다.