세계는 지금 ②

생성형 AI의 발전과 선거에서 잠복된 위험

김현정 동아대학교 국제전문대학원 부교수

들어가며

생성형 인공지능(Generative AI : 이하 생성형 AI)이 등장하면서 우리의 일상생활에도 인공지능이 스며들기 시작했다. 2010년대 중반, 이른바 4차 산업혁명이 우리 일상의 삶을 크게 바꿔놓을 것이라는 예측이 많았다. 그러나 대중 입장에서 혁명에 비길 만한 변화는 그다지 체감하기 어려웠다. 물론 코로나-19의 장기화 상황에서는 대면접촉을 제한하고 비대면 중심의 온라인 문화가 확산되는 기술 진보를 실감하게 하였다. 그러나 팬데믹 시기가 종료되자 다시 관심이 줄어들었다.

생성형 AI가 가져온 변화는 산업 분야는 물론, 정치, 사회, 미디어 등 다양한 영역에 관심을 불러일으키며 일반 사람들의 생활 속으로 깊이 파고들고 있다. 생성형 AI는 산업 분야 전반에 혁신을 야기할 수 있는 강력한 잠재력을 내재하고 있다. 하지만 생성형 AI 활용이 급증하면서 이와 관련한 윤리·철학적 논란, 생성형 AI 기술 적용의 부작용 등 다양한 문제점이 제기되고 있다.

챗GPT 등에 의해 구현되고 있는 생성형 AI 기술은 텍스트 및 코드와 이미지 생성, 음성합성기술(speech synthesis), 영상 및 3D 모델, 오디오 및 음악 등 다양한

영역과 산업군에서 활용되고 있다.[01] 또한 생성형 AI를 활용해 빠른 속도로 다양한 형태의 디지털 재화를 제작하는 '수퍼 개인(Super Individuals)'이 등장했다. 자신이 상상하고 구현하고자 하는 디지털 재화를 생성형 AI에게 자연어 프롬프트로 요청하면 이를 제작할 수 있는 능력을 발휘할 수 있다.[02] 이제 매우 적은 비용으로 불특정 다수가 콘텐츠 제작 및 확산에 접근할 수 있게 된 것이다. 그만큼 부작용이 발생할 가능성 또한 높아졌다.

이러한 상황에서 미국, 유럽 등에서는 생성형 AI가 선거에 가져올 변화에 주목하고 있다. 특히 2024년은 미국의 대통령선거, 유럽연합(European Union)의 유럽의회선거, 우리나라의 국회의원선거 등 주요 선거가 밀집해 있다. 이제 생성형 AI 활용이 선거에 미치는 영향과 주요 국가들의 규제 동향을 검토할 때이다.

선거에 대한 생성형 AI의 오남용 위험

생성형 AI에 의해 선거를 방해하는 행위가 예측되는 상황에서 전 세계 주요 국가들은 허위정보, 허위영상(deepfake) 등에 대한 대비책에 고심하고 있다. 최근에 국제 정치에 영향을 줄 수 있는 허위영상이 속속 등장해 경각심을 주고 있다. 현재 전쟁을 치르고 있는 우크라이나의 젤렌스키 대통령의 항복 선언 영상(2022년), 미국 국방부 건물 펜타곤(Pentagon)의 일부가 폭발하는 영상(2023년), 트럼프 전 대통령이 폭력적으로 경찰에 체포되는 영상(2023년) 등은 전 세계에 큰 충격을 주었다. 나아가 미국의 공화당이 제작한 '비트 바이든(Beat Biden)' 영상은 선거 캠페인에서 생성형 AI의 활용을 어디까지 허용할 것인지에 대해 논란을 일으켰다.

이 영상 또는 사진이 소셜 미디어에 퍼져나갔을 때 반응은 즉각적이었고 거의 폭발적이었다. 이 같은 영상을 생성하는 데 간단한 명령어 한 문장이면 충분했다. 충격적인 이미지의 확산 초기에는 반응이 세계적이었다. 물론 ① 젤렌스키 대통령의 항복 선언 영상과 ③ 트럼프 전 대통령의 체포 영상의 경우, 내용상 진위 여부가 의심받기도 했지만 ② 펜타곤 폭발 사진은 일시적 주가 폭락을 이끄는 등 즉각적인 영향력을 보여주었다. 위 이미지들은 비교적 생성형 AI의 초기 생성물이어서 자세히 살펴보면 가짜임을 판별해낼 수 있었다. 그러나 이제는 전문가도 식별하

01 김현정·김주희(2024), 「생성형 인공지능에 관한 미국의 규제 및 선거에 미치는 영향 분석」, 《한국과 국제사회》 제8권 제1호, 한국정치사회연구소, p. 242.

02 이승환(2023), 「생성AI 확산과 저작권 이슈의 부상」, 《Futures Brief》 23-1호, 국회미래연구원, p. 1.

정치 및 선거에 영향을 미칠 수 있는 생성형 AI 악용 사례

① 젤렌스키 대통령의 항복 선언(허위영상)

③ 트럼프 전 대통령이 폭력적으로 체포되는 모습(허위영상)

④ 공화당이 제작한 '비트 바이든' 장면(허위영상)

② 펜타곤이 폭발로 화염에 휩싸인 모습(허위사진)

출처

① "Deepfake video of Volodymyr Zelensky surrendering surfaces on social media". The Telegraph. https://www.youtube.com/watch?v=X17yrEV5sl4 (검색일 : 2023. 11. 1.)

② "A tweet about a Pentagon explosion was fake. It still went viral". The Washington Post. https://www.washingtonpost.com/technology/2023/05/22/pentagon-explosion-ai-image-hoax/ (검색일 : 2023. 11. 1.)

③ "Fake Trump arrest photos: How to spot an AI-generated image". BBC. https://www.bbc.com/news/world-us-canada-65069316 (검색일 : 2023. 11. 1.)

④ "Republican Party Biden attack ad features attack on Taiwan". Taiwan English News. https://taiwanenglishnews.com/republican-party-biden-attack-ad-features-attack-on-taiwan/ (검색일 : 2023. 11. 1.)

기 어려운 영상이 생성되고 있다. 이 허위영상들은 각국 사회에 이에 대한 향후 대응책이 절실하다는 공감을 불러일으켰다.

최근에는 생성형 AI 생성물이 민주주의 가치를 훼손하는 사례가 빈발하고 있다. 2023년 9월 슬로바키아 국민의회선거를 이틀 남긴 시점에서 진보당 지도자와 기자가 우크라이나에 대한 슬로바키아의 군사 지원과 나토(NATO) 지원에 대한 암시 등을 논의하는 가짜 오디오 클립이 퍼져나갔다. 또한 2023년 아르헨티나 대통령선거 1차 투표 직전에 대통령 후보 불리치가 정치 파트너와 여성에 대해 거친 말을 뱉고 정부 직책을 제안하는 등의 내용을 담은 허위 오디오 파일이 급속하게 확산되었다. 이것은 선거 교란 행위로서 메타(Meta) 등 빅테크 기업이 생성형 AI 허위영상만을 명시적으로 감시하는 체제의 빈틈을 파고든 것이었다. 이에 세계의 주요 국가들은 선거의 특수성을 감안하여 생성형 AI 규제에 대해 세밀한 입법을 추진해가고 있다.

주요 국가들의 선거에 대한 생성형 AI 규제

그렇다면 생성형 AI 규제는 어떤 방향으로 나아가야 하는가? 앞의 이미지들 중, 특히 ④ '비트 바이든'은 공화당전국위원회(Republican National Committee)가 바이든 대통령이 재선 도전을 공표한 직후에 유포한 생성형 AI 생성물이다. 해당 영상은 바이든의 재선 이후, 중국이 대만을 공격하는 장면과 이후 대만이 전쟁의 참화에 사회경제적으로 황폐화된 이미지를 담고 있다. 다만, '비트 바이든' 영상은 "전부 AI 이미지로 제작되었다(built entirely with AI imagery.)"는 경고 문구와 워터마크를 포함하고 있어 법적 공방을 피하였다.

이에 대해 미국 내 반응은 신중하다. 만약 어떠한 창작물이든 이것이 허위임을 명시한 채 주장을 담아 제작한다면 과연 영화나 '페이크 다큐멘터리(fake documentary)'와 무슨 차이가 있는가? 또한 콘텐츠 생성도구로서 생성형 AI를 사용한 허위정보를 일반적인 허위정보와 구분하여 법적 제재를 가할 필요가 있는가? 그러나 또 다른 일부는, 생성형 AI의 무분별한 활용을 이대로 좌시해서는 안 된다는 입장이다. 생성형 AI는 온라인상의 광범위한 데이터를 스스로 학습하고 검색하여 자체적으로 창작물을 생성한다. 이에 따라 진위 판별과 창작자 판별, 저작권 등 법적 책임이 불분명한 상황이며, 이제는 국가 차원의 규제가 필요하다는 주장이다.

미국과 EU는 현재, 생성형 AI 규제 법제화의 기준이 되고 있는 나라들이다. 미국의 경우, 연방정부는 AI 및 생성형 AI 규제에 관한 원칙적인 방향성을 제시하고 각 주정부는 차별 금지, 선거 규제, 알고리즘 규제 등에 관한 세부 법안을 제정하고 있다. 미국은 2019년 행정명령 제13859호를 통해 「미국 인공지능 구상」을 공표했다. 나아가 「미국인공지능진흥법」 등 지원책을 결의해 생성형 AI 산업 발전을 선두에서 이끌고 있다. 연방정부는 결국 생성형 AI 기술 규제 법안을 의결하지 못했고 다만, 정치 광고물에 AI를 사용할 때 출처 표기를 하는 것으로 규제 수준을 낮

미국 연방 AI 및 생성형 AI 규제 개요

법률 형태	분야	법안 제목	주요 내용
행정명령	권리보호	안전하고 신뢰할 수 있는 인공지능에 대한 행정명령 Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence	- 행정 효율화를 위한 연방 행정기관의 AI 이용 허용 - 해당 AI의 설계, 취득, 개발 과정에서의 미국 시민 보호 원칙 제시
지침	권리보호	AI 권리장전에 대한 청사진 Blueprint for an AI Bill of Rights	AI 개발 및 자동화 과정에서 개인, 기업, 정책 입안자가 AI 기술에 대한 책임을 인식하고, 시민권 보호, 평등 접근을 보장할 것을 촉구하는 핵심 원칙
법률	권리보호	AI 훈련법 AI Training for the Acquisition Workforce Act	미국 연방 공무원의 AI 이해도를 높이고 윤리적이고 안전한 사용을 목표로 하는 법률로, 일부 행정기관에 대해 행정기관 인력을 위한 AI 교육 프로그램을 수립하거나 제공하도록 요구함
법률	안보강화	2023 국방수권법 National Defense Authorization Act for Fiscal Year 2023	국방 및 정보기관이 AI 시스템과 기능을 기관 운영에 포함하도록 함
법안	위험관리	알고리즘 책임법 Algorithmic Accountability Act of 2022/2023	AI, 알고리즘 등을 사용하는 소프트웨어의 자동화된 의사결정 시스템에 대한 편향 평가 문제
법안	권리보호	미국 데이터 개인정보 및 보호법 American Data Privacy and Protection Act	개인 데이터와 관련하여 동의 없는 데이터 처리 및 전송에 대한 제한 및 보호를 제공하는 등 개인정보 처리 자체에 대한 보호
법안	권리보호	디지털 플랫폼 위원회법 Digital Platform Commission Act of 2023	법안이 통과되면 신설될 '디지털 플랫폼 위원회'는 온라인 플랫폼에 대하여 소비자 보호 등을 위해 청문회 개최, 조사, 연구 지원 등의 권한을 부여받으며, 광범위한 감독 권한을 가질 것으로 예상됨

출처 : 김현정·김주희(2024), 「생성형 인공지능에 관한 미국의 규제 및 선거에 미치는 영향 분석」, 《한국과 국제사회》 제8권 제1호, 한국사회정치연구소, pp. 255~256.

추어 발의했다. 미국의 경우, 워터마크 표기 등 투명성[03]을 보장하는 최소한의 제재를 통해 헌법에 명시된 표현의 자유를 보장하고 있기 때문이다.

미국 정부는 구글 등 플랫폼의 제어 영역에서 원칙만을 제시할 뿐, 허위영상 유포에 대한 기업의 책임은 면책 대상이 될 수 있다. 「통신품위법(CDA)」 제230조 조문 내, "서비스 제공자는 불쾌한 콘텐츠에 대한 접근을 제한하거나 제거할 수 있으며, 이러한 행동이 선량한 신념에 따른 행동일 경우 콘텐츠 소유자로부터 책임이 면제"[04]됨이 적시되어 있다. 다만 생성형 AI 생성물에 의한 선거 관련 오남용 사례의 발생을 방지하기 위해 플랫폼 기업들의 자발적인 노력이 필요하다는 사실을 정부가 독려하고 있다.

반면에 유럽은 AI 사용에 대해 규제적 입장을 취하고 있다. EU 집행위원회는 2021년 4월 AI 규제 초안을 공표했다. 곧이어 2023년 6월, 「AI 법(The Artificial Intelligence Act)」을 위한 EU 3자(집행위원회, 유럽의회, 각료이사회) 입법 과정이 개시되었으며, 2024년 3월 14일 최종 승인되었다. 「AI 법」의 초안과 초안이 제시된 직후 논의에서는 생성형 AI에 관한 언급이 없었지만, 이후 상황이 완전히 바뀌어서 생성형 AI가 주요 쟁점이 되어 생성형 AI 영역이 「AI 법」의 최종안에 포함되었다.

EU의 「AI 법」 최종안은 다음 사항을 담고 있다. 첫째, 콘텐츠가 AI에 의한 제작물임을 공개할 의무(워터마크 표기), 둘째, 불법 콘텐츠 생성을 방지하는 모델 설계 의무, 셋째, 훈련에 사용된 저작권 데이터의 요약 게시 의무[05]가 부가되었다.

수정안에는 「생성형 AI에 대한 의무」 항목에 생성형 AI 시스템의 투명성 요구 사항과 학습에 사용된 저작권 데이터의 상세 요약까지 공개할 의무가 포함되었다. 하지만 저작권 전반에 대한 쟁점은 다루고 있지 않으며, 이후에도 해당 사항이 포함될 가능성은 낮다고 판단된다. 생성형 AI 금지 항목에 대해서는 '실시간 생체 인식' 사용 금지가 적시되었으나, 금지 사항 내 '상당한 피해'의 범위 또는 '해로운 대

---

03 「AI 규제」에서 '투명성(Transparency)' 원칙은 첫째, AI가 제작한 콘텐츠임을 공개할 의무(워터마크 제도), 둘째, 불법 콘텐츠 생성을 방지하는 모델 설계 의무, 셋째, 훈련에 사용된 저작권 데이터의 요약 게시 의무 등에 해당한다. EU의 경우 3개 원칙 모두 'AI 시스템의 투명성 요구 사항'에 포함시켰으며, 미국은 워터마크 표기만을 규제한다.

04 콘텐츠마케팅플랫폼, "[미국] 통신품위법", https://welcon.kocca.kr/ko/law/us/7 (검색일 : 2024. 2. 15.)

05 EU의 「AI 규제안」 내 저작권 관련 사항은 「유럽연합 디지털 단일 시장의 저작권 및 저작인접권 지침(CDSM)」의 법적 내용의 적용을 받는다. 「CDSM」 제4조 3항 '저작자 및 실연자의 이용 계약에서 공정한 보상'의 AI 규제 적용을 위해 AI 투명성 요구 사항의 2, 3절이 적시되었으나, 실제 '기계 판독이 가능한 형태'로의 생성물 내 요약 게시가 쉽지 않은 상황이다. 이에 EU 내 일부 디지털 전문가들은 규제 실행이 실천 가능하도록 해당 조항에 대한 「CDSM 4(3)」 선택 예외 조치(opt out)를 주장하고 있다(Paul Keller, "Generative AI and copyright: Convergence of opt-outs?" https://copyrightblog.kluweriplaw.com/2023/11/23/generative-ai-and-copyright-convergence-of-opt-outs/ (검색일 : 2024. 3. 18.)

우'의 개념에 대해 법률에서 정의하지 않은 상태로 제시되었다.[06] 생성형 AI 기업의 규모 및 위반 유형에 따라 750만~3,500만 유로의 과징금이 부과된다.

미국과 EU의 AI 관련 규제 비교

종류	미국	EU
AI 관련 법	2024. 5. 중 예상	2024. 3. 14. 최종 승인
워터마크	의무(기업별 표준화)	의무화, 표준화 추진
플랫폼 제재	면책 조항	제재 조항
생성형 AI 저작권	불인정	불인정
알고리즘 관련	제재	제재
허용되지 않는 AI 영역	무	유(8개 영역으로 제시)

출처 : 필자 작성

EU는 콘텐츠가 허위일 경우에 대해, 기존 미디어 혹은 디지털 관련 법안을 통해 제어하고 있다. EU는 허위정보의 소셜 미디어 확산으로 인한 선거 위해를 방지하기 위해 「허위정보 실천강령(Code of Practice on Disinformation)」을 강화하여, 책임자 처벌 및 수익에 대한 최고 6% 벌금 을 공표하였다. 나아가 2023년 8월 시행된 「디지털 서비스법(Digital Services Act)」을 통해, 디지털 플랫폼에 게시된 허위영상 콘텐츠 등 허위정보에 대한 1차 책임을 디지털 플랫폼 기업이 지도록 하였다.

중국은 최초로 생성형 AI에 대한 규제 사항을 입법화한 국가이다. 2023년 8월 중국은 「생성형 AI 법규제(生成式人工智能服务管理暂行办法)」를 도입, 시행하였다. 중국 정부는 「생성형 AI 법규제」에 대해 콘텐츠 관리를 중점적으로 다루며, "사회주의적 가치를 구현해야 한다"는 중심 원칙 아래에 당국의 허가를 받아야 생성형 AI 서비스를 제공할 수 있는 '라이선스 제도'를 도입하였다.[07]

올해 4월에 우리나라도 제22대 국회의원선거를 치를 예정이다. 생성형 AI의 오용으로 선거가 위기에 빠지는 일이 없도록 우리 사회도 기민한 대응이 필요하다. 중앙선거관리위원회는 올해 2월 국회의원 총선거를 앞두고 '인공지능(AI) 기반 딥페이크 영상'을 이용하는 선거운동을 금지시켰다.

---

06 Center for Democracy & Technology, "EU AI Act Brief — Pt. 1, Overview of the EU AI Act", https://cdt.org/insights/eu-ai-act-brief-pt-1-overview-of-the-eu-ai-act/ (검색일 : 2024. 3. 15.)

07 김다운 기자, "각국의 생성 AI 규제 및 정책 동향", 《투이컨설팅》 https://www.2e.co.kr/news/articleView.html?idxno=302827 (검색일 : 2023. 12. 1.)

나가며

어느 날 우리 주위로 성큼 다가온 생성형 AI는 빠르게 일상을 파고들고 있다. 경제, 사회뿐만 아니라 정치와 선거 영역에도 생성형 AI는 심원한 영향을 미칠 수 있다. 덴마크의 '신세틱(Synthetic)', 일본과 핀란드의 AI 정당 등 이들은 AI와 공존하는 세상을 표방하고 정치 활동을 펼치고 있다. 뉴질랜드의 '샘(SAM)'과 러시아의 '알리사(Alisa)'는 AI 정치인으로 활약 중이다.

국내에서도 생성형 AI 규제안의 도입 필요성이 논의되고 있다. 중앙선거관리위원회의 '인공지능(AI) 기반 딥페이크 영상'을 이용하는 선거운동 금지 방안 보다 근본적 대책이 필요하다. 이와 관련해 우리에게 시급한 규제 사항을 아래와 같이 제시한다.

첫째, 생성형 AI에 의한 창작물임을 표식하는 표준 워터마크 규제가 필요하다. 과학기술정보통신부는 2023년 8월 민간 기업과 전문가들과 함께 '제4차 인공지능 최고위 전략대회'를 열고 이런 내용의 「인공지능 윤리·신뢰성 확보 추진계획」을 발표했다.[08] 그런데 지금까지도 워터마크 표준안이 도입되지 않고 있다. 표준안이 없을 경우, 콘텐츠 내 기만적인 표식으로 규제를 우회할 가능성이 남게 된다.

둘째, 허위 사실 유포에 대한 기존의 규제가 존재하기 때문에 정치 활동 및 선거와 관련한 생성형 AI에 대한 규제는 강화될 필요가 있다. 물론 표현의 자유는 보장되어야 한다. 하지만 선거를 앞둔 시점에서는(흔히 3개월 이내) 생성형 AI의 무분별한 사용에 의한 사회 혼란을 방지하기 위해 미디어 관련 법이 좀 더 촘촘하게 재정비될 필요가 있다.

셋째, 국내 플랫폼 기업의 자발적인 공동대응 노력이 필요하다. 이미 메타 등 해외 빅테크 기업들은 스스로 규제안을 내놓고 있다. 국내 플랫폼 기업도 선거에 대한 공동의 노력을 기울여야 할 때가 되었다.

넷째, 생성형 AI와 같은 혁신기술은 규제를 통해 제어하는 데 한계가 있다. 규제는 항상 기술 발전을 뒤쫓아갈 수밖에 없기 때문이다. 생성형 AI를 직접 사용하거나, AI 생성물을 접하게 될 경우를 대비해 시민 교육이 필요하다. 성숙한 시민의 자세로 생성형 AI 활용을 제어하고, 민주시민으로서 생성형 AI 창작물을 분별하여 접근할 수 있도록 민관 협력의 시민 교육이 제시되어야 할 것이다.

---

08 김은성 기자, "정부, AI 생성물에 '워터마크' 도입 추진", 《경향신문》
https://m.khan.co.kr/economy/economy-general/article/202310251245001#c2b (검색일 : 2023. 12. 1.)