

# 생성형 AI 신기술 도입에 따른 선거 규제 연구

연구책임자 : 김주희 (국립부경대학교)

공동연구자 : 차재권 (국립부경대학교)

김현정 (동아대학교)

조성복 (국민대학교)

연구보조원 : 박서현 (국립부경대학교)

권수민 (국립부경대학교)

【2023년도 중앙선거관리위원회 정책연구용역 보고서】

# 『생성형 AI 신기술 도입에 따른 선거 규제 연구』

연구책임자 : 김주희 (국립부경대학교)

공동연구자 : 차재권 (국립부경대학교)

김현정 (동아대학교)

조성복 (국민대학교)

연구보조원 : 박서현 (국립부경대학교)

권수민 (국립부경대학교)

본 연구보고서는 2023년도 중앙선거관리위원회 정책연구용역 과제로서 연구내용은 중앙선거관리위원회의 공식 견해가 아님.

## 차 례

1. 서론 .....	1
1.1 연구의 필요성과 목적 .....	1
1.2 연구의 내용과 방법 .....	3
2. 생성형 AI 신기술 분류와 유형화 .....	7
2.1. 생성형 AI 규제 동향 .....	7
2.2. AI 규제 동향 유형화 .....	13
2.3. 선거 과정에서의 AI 활용 .....	18
3. 주요국의 생성형 AI의 선거 활용 문제 분석 .....	49
3.1 미국 .....	49
3.2 중국 .....	58
3.3 영국 .....	61
3.4 독일 .....	67
4. 주요국의 AI 및 생성형 AI 선거 규제 .....	73
4.1 미국 .....	73
4.2 영국 .....	89
4.3 유럽연합 .....	110
4.4 미국과 EU의 협력과 경쟁 .....	122
4.5 중국 .....	130
4.6 독일 .....	134
4.7 스위스 .....	159
4.8 핀란드 .....	169
5. 결론 .....	176
요약 .....	193
참고 문헌 .....	204

## 표 차례

〈표 2-1〉 생성형 AI가 최근 과학 논문, 뉴스 기사, 소셜미디어에서 나오는 허위 정보에 미치는 영향을 우려해야 하는 이유에 대한 네 가지 주장 .....	12
〈표 2-2〉 선거과정에서 AI 활용의 위험과 방향성 .....	19
〈표 2-3〉 핵심 평가에 대한 전반적인 요약 .....	48
〈표 4-1〉 미국의 ‘AI 로드맵’의 5가지 전략 .....	73
〈표 4-2〉 미 연방 AI 및 생성형 AI 규제 개요 .....	75
〈표 4-3〉 AI 행정명령 내 주요 사항 .....	78
〈표 4-4〉 알고리즘 책임법 2022 내 영향평가 주요 사항 .....	82
〈표 4-5〉 캘리포니아 주, 직장기술책임법의 주요 사항 .....	83
〈표 4-6〉 DC 알고리즘에 의한 차별금지법이 요구하는 새로운 사항 .....	85
〈표 4-7〉 미국 주 선거 및 선거운동에서의 AI 사용 적용 법령 .....	86
〈표 4-8〉 2023년 선거 및 선거운동에서 AI 관련 법령 .....	86
〈표 4-9〉 영국의 국가 AI 전략 개요 .....	103
〈표 4-10〉 영국의 국가 AI 전략의 세 가지 원칙과 시기별 계획 .....	105
〈표 4-11〉 유럽의회가 채택한 AI 관련 입법 이니셔티브 .....	111
〈표 4-12〉 EU 집행위원회에서 마련한 초안과 비교하여 2022년 5월 EU 의회 안에서 변경된 주요 사항 .....	112
〈표 4-13〉 EU의 AI 시스템 공급자 및 사업자의 법적 의무 .....	115
〈표 4-14〉 EU의 AI 시스템 수입업체 / 유통업체별 법적 의무 .....	116
〈표 4-15〉 DSA 내 온라인 중개 서비스 유형별 의무사항 .....	118
〈표 4-16〉 DMA 내 게이트키퍼의 경쟁제한 및 불공정행위 .....	119
〈표 4-17〉 유럽 각국의 가짜뉴스 관련 법적 규제 내용 .....	121
〈표 4-18〉 TTC 10개 작업그룹 .....	123
〈표 4-19〉 TTC ‘인공지능을 위한 공동로드맵’ 내 협력의 접근방식에 관한 제시 ..	125
〈표 4-20〉 EU AI법과 미국 AI 권리장전 내 위험관리 규제 격차 비교 .....	129
〈표 4-21〉 중국 생성형 AI 서비스 관리 잠정 방법 주요 내용 .....	132

<표 5-1> 생성형 AI 선거규제 국가별 비교분석 ..... 192

## 그림 차례

<그림 1-1> 연구 추진전략 방법 및 추진체계 ..... 5

<그림 1-2> 연구 추진체계 ..... 5

<그림 1-3> 연구진 구성 및 역할 분담 ..... 6

<그림 2-1> ChatGPT 출시 이후 주요 대형언어모델(LLM) 개발 일정 ..... 8

<그림 2-2> AI 환경에 적합한 LLM과 대화형 챗봇 ..... 10

<그림 2-3> 가트너, 2023년 이머징 테크놀로지 하이프 사이클 ..... 11

<그림 2-4> 생성형 AI 규제 유형화 ..... 15

<그림 2-5> 위스콘신 유권자 거주 확인용 발송 우편물 사례 ..... 21

<그림 2-6> 시설 위치에 대한 클러스터링 이미지 ..... 27

<그림 2-7> 인도 텔랑가나주 2020 지방선거에서 얼굴 인식 사용 ..... 36

<그림 2-8> 선거에 초점을 맞춘 CCTV 모니터링 서비스를 설명하는 VMukti의 페이지  
..... 42

<그림 2-9> 두 대의 카메라로 이루어지는 개인 재식별 :(싱가포르국립대, NTU)  
래피드 리치 객체 검색 연구실 ..... 44

<그림 3-1> 조 바이든 대통령이 트랜스젠더 혐오 발언하는 가짜 영상 ..... 49

<그림 3-2> 도널드 트럼프 전 대통령이 체포되는 가짜 영상 ..... 50

<그림 3-3> ‘Beat Biden’ 이라는 제목의 영상 내 중국이 대만을 공격하는 가짜 영상  
..... 52

<그림 3-4> 트럼프가 앤서니 파우치 전 소장과 포옹하는 허위 사진 ..... 54

<그림 3-5> 미 국방부 건물 폭발에 관한 가짜 이미지 트윗 ..... 55

<그림 3-6> 젤렌스키 우크라이나 대통령의 항복 선언 허위 영상 ..... 56

<그림 3-7> 푸틴 러시아 대통령의 항복 선언 허위 영상 ..... 57

<그림 3-8> 생성형 AI로 제작된 이스라엘-하마스 전쟁 이미지 ..... 58

<그림 3-9> 마이크로소프트 위기정보 보고서(Microsoft Threat Intelligence  
report)에서 게재된 중국 IO 자산이 게시한 AI 생성 이미지 ..... 59

<그림 3-10> 중국 내 지식재산권 분쟁 사례였던 AI 생성 이미지 .....	60
<그림 3-11> 미국 상원의원 테드 크루즈의 페이스북 정보 활용 스캔들 .....	61
<그림 3-12> 러시아 SNS 계정 15만여 개가 브렉시트 투표에 개입 .....	63
<그림 4-1> 페이스북이 영국 일간지에 낸 fake news 예방 캠페인(2017년 5월) ...	95
<그림 4-2> 영국 정부의 허위 정보 대응가이드 .....	98
<그림 4-3> EU AI 위험유형 분류 피라미드 .....	113
<그림 4-4> 2023년 미디어 리터러시 순위 .....	170
<그림 4-5> 핀란드 우수 사회 지표 .....	170
<그림 5-1> 생성형 AI 규제 유형화 .....	177
<그림 5-2> 2023 확장된 미디어 리터러시 순위 .....	190
<그림 5-3> Media Freedom 순위 .....	191

# 1. 서론

## 1.1 연구 필요성과 목적

### 가. 연구의 배경

- AI의 사용이 확산되며, 윤리·사회 문제에서부터 법적 문제에까지 다양한 이슈가 제기되고 있다. AI 사용에 대한 윤리·철학적 논란, 특정 콘텐츠에 AI 기술을 적용하였을 때의 인문사회학적 문제점을 넘어 인지조작, 생체 식별, 페이크 뉴스 생성 등 다양한 법적 문제점이 존재한다. 기술의 빠른 진보로 인해 문제점에 빠르게 대처하기 어렵다는 점에서 문제에 선행하여 대비하기 어려운 상황이다.
- 생성형 AI(Generative AI)의 사용이 확산되며, 이러한 기술이 선거에 오용 및 악용될 우려도 존재한다. 생성형 AI는 인터넷상의 광범위한 데이터를 통해 자체적으로 생성물을 창작함에 따라 진위 판별과 저작권 등의 법적 책임이 불분명한 상황이다. 다수의 국가가 AI 규제를 준비하는 가운데, 미국과 EU를 중심으로 생성형 AI에 대한 규제 입법에 나서고 있다.
- 본 연구는 생성형 AI 관련 제도를 주도하는 국가의 사례 연구를 통해 한국에서 발생할 수 있는 다양한 문제에 대비하는 정책적 방향성을 설정하는 것을 목적으로 한다.

### 나. 연구 수행 범위: 중범위 메타 분석

- 시간적 범위
  - 2000년대 이후 SNS(social media) 도입과 빅데이터의 선거 활용에 따른 선거 영향 및 문제점을 분석한다.

- 2014년 ~ 현재의 생성형 AI 기술 변화와 생성형 AI의 선거 영향 및 문제점 분석한다.
- 2020 ~ 최근까지 생성형 AI의 선거 개입 문제가 발생하는 국가의 생성형 AI 선거 규제를 분석한다.

### ○ 공간적 범위

- 미국과 유럽연합이 이 분야 규제를 가장 주도적으로 이끌고 있다는 점에서 미국과 유럽연합 그리고 유럽연합의 회원국 독일, 스위스, 핀란드 및 영국 사례를 분석하고자 한다. 또한 생성형 AI 선거 규제의 유형화를 위해 가장 먼저 이 분야의 규제를 만들어 가고 있는 중국 사례를 포함한다.
- 중국은 사회주의국가로 다른 민주주의 국가와는 다른 가치를 지향하고 있다는 점에서 미국과의 대척점에서 규제 유형의 차이를 구체화하기 위해 선택되었다.
- 유럽연합 및 독일, 그리고 핀란드를 포함하는 회원국과 유럽연합을 탈퇴한 영국과 회원국은 아니지만 유럽연합의 규제를 많은 부분 공유하는 스위스의 경우 유럽연합의 강력한 AI 규제에 어느 정도 부담을 느끼면서 대응하고 있다.
- 그러나 생성형 AI 규제가 주로 경제적 요인에 의해 결정된다면 생성형 AI 선거 규제의 경우 규제를 형성하는 동기와 요인이 국가별로 차이를 보일 것이라는 점에서 성숙한 민주주의 수준을 보여주는 국가를 선택하여, 이의 규제 유형을 파악한다. 이들이 어떠한 규제 방식을 추구하고 있는가에 대한 파악이 한국적 함의를 도출하는데 유의미할 것으로 본다.
- 한국의 생성형 AI 기술 규제 및 생성형 AI 관련 해외의 제도적 규제 동향 및 현황과 비교분석 그리고 한국적 함의 도출하고자 한다.



## 1.2 연구의 내용과 방법

### 가. 연구 내용

- 첫 번째, 생성형 AI 신기술의 기술 분류에 따른 선거 활용 시 문제점을 유형화한다. 텍스트, 코드 생성, 이미지, 음성합성 기술(Speech synthesis), 영상 및 3D 모델, 오디오 등 각각 영역에서의 생성형 AI 기술이 선거에서 어떠한 유형의 문제를 일으킬 수 있을지 유형화하며 관련 사례를 제시하고, 2000년대 이후 빅데이터의 선거 활용에 따른 선거 영향과 문제점을 분석한다. 그리고 신기술의 선거 분야 내 활용 관련 국내외 선행연구를 검토하여 생성형 AI의 선거 과정에서 활용된 혹은 잠재적으로 활용될 가능성에 대한 검토와 방향성을 제시한다. 또한 이를 통해 생성형 AI의 규제 유형화와 생성형 AI의 선거 규제의 방향성에 대한 논의를 통해 국가별 특징을 구별하고자 한다.
- 두 번째, 미국, 영국, EU, 독일 등 주요 국가와 생성형 AI로 위협에 가장 잘 대응하고 있는 국가로 뽑히는 핀란드와 최근 이 영역에 관한 관심이 고조되고 있는 스위스의 AI 및 생성형 AI 기술 규제를 연구한다. 미국의 AI 연구개발·훈련에 관련한 법률과 행정명령을 분석한다. 또한, 생성형 AI 기술 규제 논의 동향을 연구하고, 최근 발의된 정치 광고에 AI 사용 출처를 표기하는 내용의 법안 내용을 분석한다. EU의 경우, 2021년 4월 제안된 EU 규제 프레임워크를 분석해 생성형 AI에 관한 제한 사항이 선거에 활용될 때 어떠한 제재 사안으로 작용할지에 관해 연구한다. 독일 정부는 독자적인 AI 규제안을 제시하기보다 EU의 규제 프레임워크 발효 시 이를 적용할 방침으로, 독일 내 AI 규제 제정에 관한 논의를 분석한다.
- 세 번째, 미국, 영국, EU, 독일, 핀란드 그리고 스위스 내의 생성형 AI 기술의 선거 영향 및 SNS, 빅데이터 기술의 선거 활용 당시 문제점을 분석한다. 이를 위해 2008년, 2012년 미국 대통령 선거에서 빅데이터 활용에 의한 선거 예측 사례를 분석하고, 유럽연합의 허위정보실천강령

(Code of Practice on Disinformation), 디지털 서비스법(Digital Services Act), 디지털 시장법(Digital Market Act)을 연구한다. 그리고 생성형 AI 선거 규제 유형화를 위해 가장 먼저 관련 규제를 마련한 중국의 사례와 비교한다.

- 네 번째, 사례 선정국의 생성형 AI 기술의 선거 영향 및 AI 빅데이터 기술의 선거 활용 당시 문제점을 비교 분석한다. 미국과 독일의 뉴스 수용 매체 등 가짜 뉴스에 영향을 받을 가능성이 높은 요인에 관해 비교 연구를 시도한다. 또한, 한국의 생성형 AI 관련 제도적 규제 동향과 현황을 분석하고 관련 법제도 및 선거 과정에 반영할 수 있는 대응책을 제안한다.

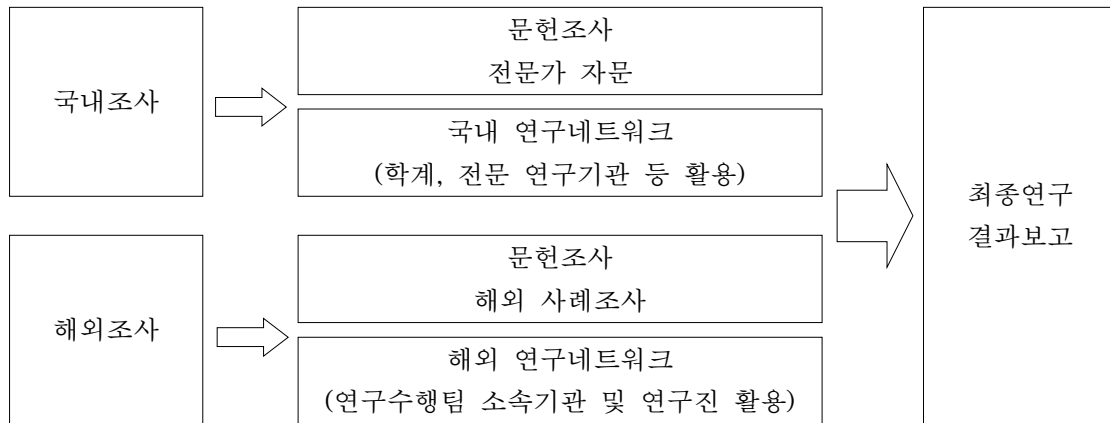
## 나. 연구 방법

- 자료 수집 및 문헌 분석
  - 주로 문헌 분석을 통한 연구로, 미국, EU, 독일의 국제 사례는 각 정부가 제공하는 자료, 전문 연구소의 보고서, 인터넷 검색 자료 등 원문 자료에 의존한다.
  - 이외에 필요한 통계자료는 정부와 국제기구에서 간행한 집계자료(aggregate data)를 분석하고 재구성한다.
- 자문 및 인터뷰
  - 생성형 AI 기술 관련 전문가에 대한 자문 면접을 시행한다.
  - 한국 내 ‘생성형 AI를 활용한 선거운동 등 관련 운용기준’ 담당자 및 관련 전문가와 자문 면접을 시행한다.
  - 사례 연구 선정 국가의 관련 전문가와 온라인 자문 면접을 시행한다.

## 다. 연구 추진전략·방법 및 추진체계

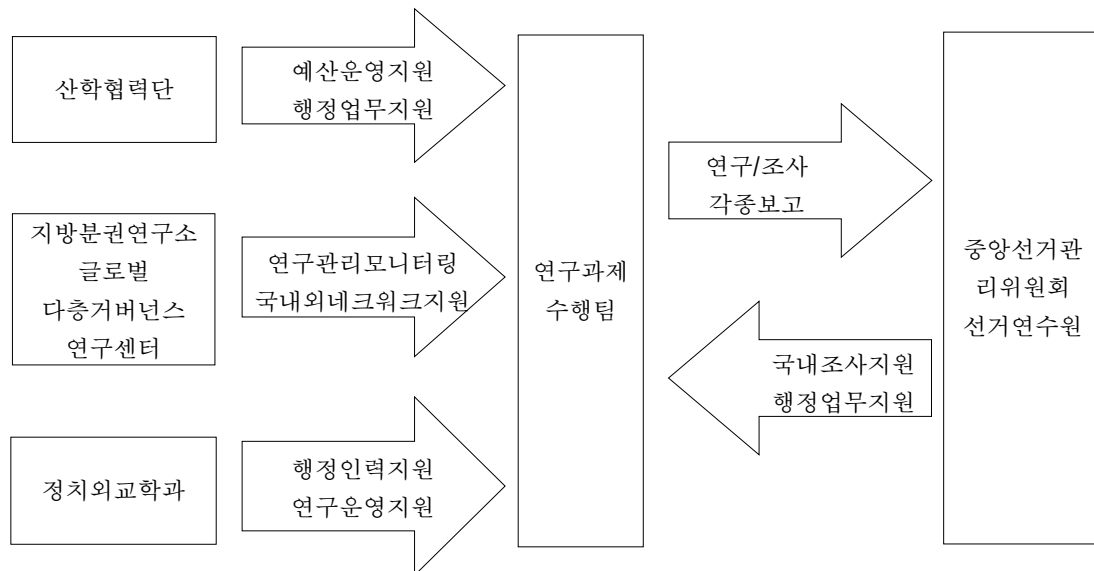
### ○ 연구 추진전략 방법 및 추진체계

<그림 1-1> 연구 추진전략 방법 및 추진체계



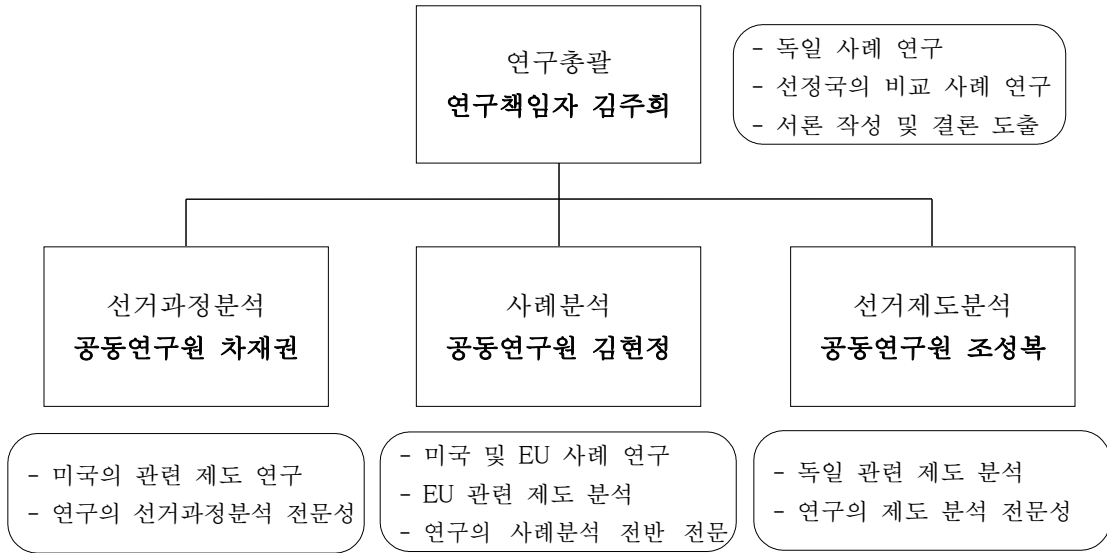
### ○ 연구 추진체계

<그림 1-2> 연구 추진체계



○ 연구진 구성 및 역할 분담

<그림 1-3> 연구진 구성 및 역할 분담



## 2. 생성형 AI 신기술 분류와 유형화

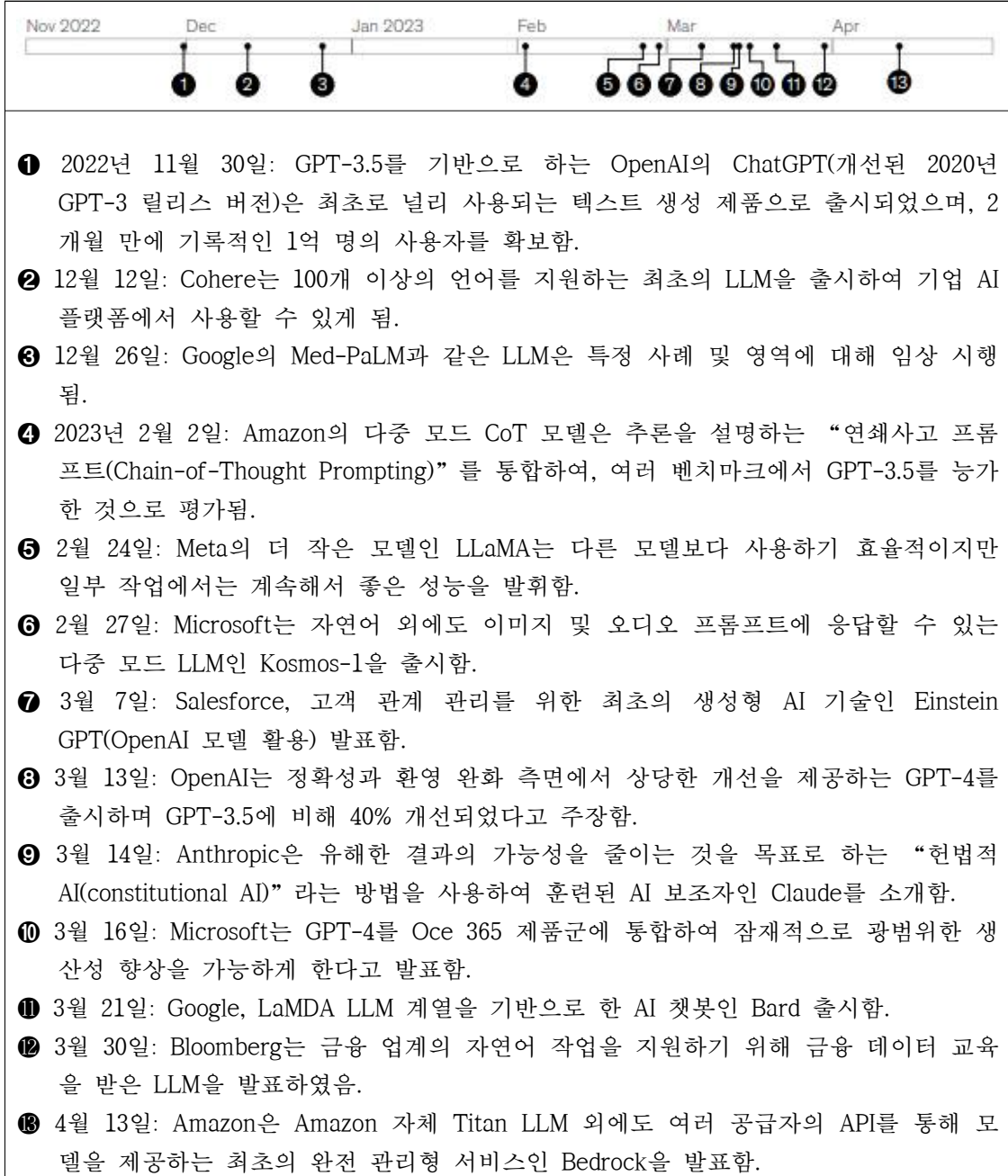
### 2.1. 생성형 AI 규제 동향

#### 가. 생성형 AI 기술 발전

- 인공지능(AI) 시스템의 활용은 사회적 혜택과 경제성장을 가져오고 있다. 또한 AI는 국가 혁신과 글로벌 경쟁력을 강화할 수 있는 강력한 잠재력을 내재하고 있다. 하지만 AI 활용이 급증하며 이와 관련한 윤리·철학적 사용 논란, 특정 콘텐츠에 대한 AI 기술 적용 시의 문제점이 제기되어 왔다.
- 특히 챗GPT 등으로 알려진 생성형 AI(Generative AI)의 사용이 확산되며 심각한 부작용이 예측되어 왔다. 현재 생성형 AI 기술은 텍스트, 코드 생성, 이미지, 음성합성 기술(Speech synthesis), 영상 및 3D 모델, 오디오 및 음악 등 다양한 영역과 산업군에서 활용되고 있다. 생성형 AI는 인터넷상의 광범위한 데이터를 통해 자체적으로 생성물을 창작함에 따라 진위 판별과 저작권 등의 법적 책임이 불분명한 상황이다.
- 2022년 11월 ChatGPT가 출시된 이후, 관련 업계의 경쟁이 치열해지고 있다. McKinsey 연구에 따르면, 해당 기술이 세계 경제에 연간 최대 4조 4천억 달러를 추가할 것이라 예측되었다.<sup>1)</sup> ChatGPT의 공개 버전은 단 두 달 만에 사용자 1억 명에 도달했다. ChatGPT 출시 이후, 주요 대형 언어모델(LLM: large language model) 2023년 4월까지의 개발 일정은 아래 <그림 2-1>과 같다.

1) McKinsey & Company. 2023a. "What's the future of generative AI? An early view in 15 charts." (August 25) <https://www.mckinsey.com/featured-insights/mckinsey-explainers/whats-the-future-of-generative-ai-an-early-view-in-15-charts#/> (검색일: 2023. 11. 01).

〈그림 2-1〉 ChatGPT 출시 이후 주요 대형언어모델(LLM) 개발 일정



출처: McKinsey & Company. 2023b. “What every CEO should know about generative AI.” (May 12) <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/what-every-ceo-should-know-about-generative-ai#/> (검색일: 2023. 11. 01).

- 생성형 AI를 지원하는 고급 기계적 학습은 수십 년에 걸쳐 제작되었다. 그러나 ChatGPT가 2022년 말에 출시된 이후 생성형 AI 기술의 새로운 버전이 한 달에 여러 번씩 출시될 만큼 시장은 경쟁체제에 돌입하였다.

2023년 3월에만 새로운 고객 관계 관리 솔루션 및 금융 서비스 산업 지원을 포함하여 6차례 주요 진전이 있었다(<그림 2-1> 참조).

- 생성형 AI는 이전에는 볼 수 없었던 방식으로 대중에 확산되었다. 즉시 사용 가능한 접근성으로 인해 생성형 AI는 이전의 모든 AI와 다르게 구현되고 있다. 사용자는 머신러닝과 상호작용하거나 머신러닝에서 가치를 도출하기 위해 관련한 기술적 지식을 습득할 필요가 없으며, 적절한 질문을 할 수 있는 거의 모든 사람이 생성형 AI를 사용할 수 있다.

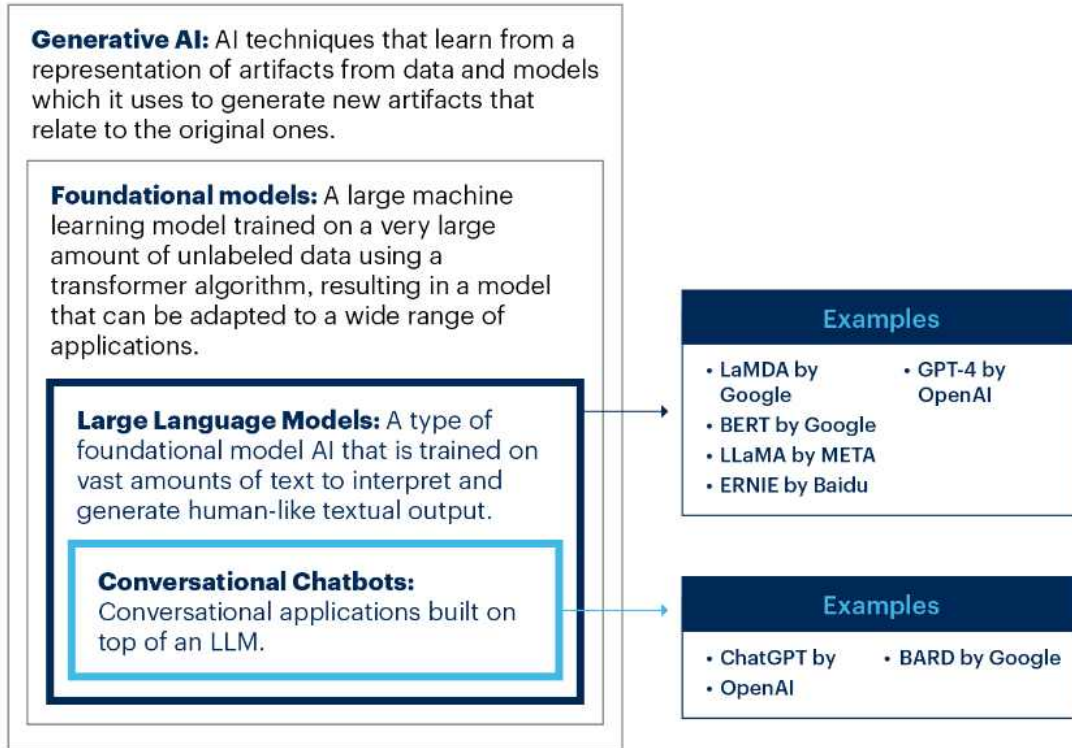
## 나. 생성형 AI와 기존의 AI 차이

- 구현할 수 있는 능력의 측면에서, 생성형 AI와 기존의 AI의 다른 점은 새로운 콘텐츠를 효율적으로 생성할 수 있다는 것이다. 자연적으로 표현되지 않는 ‘구조화되지 않은’ 형태(예: 서면 텍스트 또는 이미지)로 새로운 콘텐츠 및 행과 열이 있는 표를 생성할 수 있다.<sup>2)</sup>
- 학습 방식의 측면에서, 생성형 AI는 스스로 ‘생각하는’ AI이다. 주어진 데이터 기반 학습을 통해 발전하는 기존 AI와도 차별성이 있다. 글로벌 연구 조사 기업 ‘가트너(Gartner)’는 “콘텐츠나 사물에 대한 데이터를 통해 학습하고, 그것을 사용하여 창조적이며 현실적인, 완전히 새로운 아웃풋을 만들어내는 기계학습 방법”으로 정의한 바 있다.<sup>3)</sup>
- OpenAI의 ChatGPT는 LLM(대형 언어 모델)에 의한 텍스트 중심 생성형 AI이다. LLM은 공개 또는 비공개 데이터 컬렉션에서 시뮬레이션되거나, 가져온 대량의 텍스트(일반적으로 수십억 단어)에 대해 훈련되며, 이를 통해 텍스트 입력을 해석하고 인간과 유사한 텍스트 출력을 생성할 수 있다. 또한 LLM 분야의 획기적인 발전은 코드 생성에서 질문 답변에 이르기까지 이전에 인간이 수행했던 작업을 자동화하는 것을 포함하여 조직이 비즈니스를 수행하는 방식을 대폭 변화시킬 수 있는 잠재력을 가지고 있다(<그림 2-2> 참조).

2) McKinsey & Company. 2023b.

3) 전윤미. 2023. “생성형 AI가 기존 AI와 결정적으로 다른 점.” 『애플경제』 (3월 17일) <https://www.apple-economy.com/news/articleView.html?idxno=71123> (검색일: 2023. 11. 01).

<그림 2-2> AI 환경에 적합한 LLM과 대화형 챗봇



출처: Gartner. “What Is Artificial Intelligence?.” <https://www.gartner.com/en/topics/artificial-intelligence> (검색일: 2023. 11. 01).

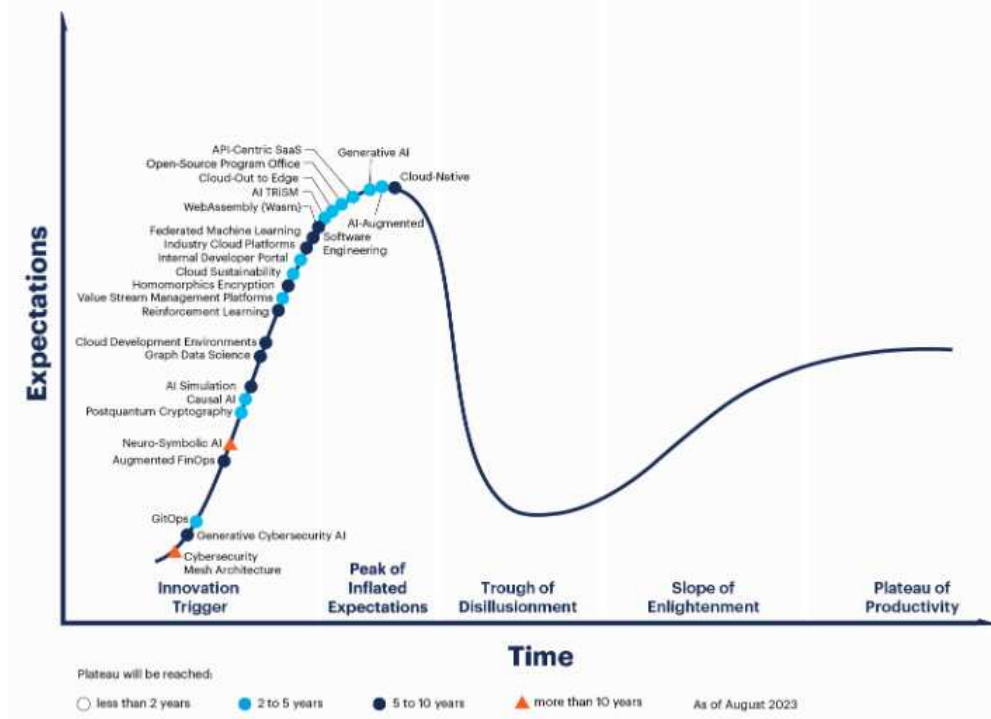
- 가트너에 따르면, <그림 2-3>과 같이, 생성형 AI는 하이프 사이클 5단계 중 현재가 기대감이 정점에 달한 상태(Peak of Inflated Expectations)로 분류되는 2단계에 포함돼, 앞으로 2년~5년 내에 혁신적인 성과를 달성할 것으로 예측됐다.<sup>4)</sup>
- 창발적 AI(emergent AI)의 한 예인 생성형 AI는 원본 소스 콘텐츠의 대규모 저장소에서 학습하여 콘텐츠, 전략, 디자인 및 방법의 새로운 파생 버전을 생성할 수 있으며, 이는 콘텐츠 및 제품 개발을 포함하여 계속해서 비즈니스에 중대한 영향을 미칠 것이다.<sup>5)</sup>

4) 김용우. 2023. “가트너 ‘생성형 AI 기대감 최고조’.” 『지디넷코리아』 (8월 17일) <https://zdnet.co.kr/view/?no=20230817104503> (검색일: 2023. 11. 01).

5) Perri, Lori. 2023. “What’s New in the 2023 Gartner Hype Cycle for Emerging Technologies.” Gartner (August 23) <https://www.gartner.com/en/articles/what-s-new-in-the-2023-gartner-hype-cycle-for-emerging-technologies> (검색일: 2023. 11. 01).



〈그림 2-3〉 가트너, 2023년 이머징 테크놀로지 하이프 사이클



출처: Perri, Lori. 2023.

#### 다. 생성형 AI의 잘못된 영향에 관한 원인 분류

- 생성형 AI가 정치, 경제, 전쟁 등 다양한 영역에서 사용되며, 이에 관한 문제점을 분석한 연구가 시작되고 있다.
- 사이먼은 관련 논문에서 생성형 AI가 허위 정보에 미치는 영향에 대한 두려움이 사회적으로 과장되고 있음을 지적하였다.<sup>6)</sup>

6) Simon, Felix, Sacha Altay, and Hugo Mercier. 2023. "Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown." Harvard Kennedy School Misinformation Review. 4(5): 2.

<표 2-1> 생성형 AI가 최근 과학 논문, 뉴스 기사, 소셜미디어에서 나오는 허위 정보에 미치는 영향을 우려해야 하는 이유에 대한 네 가지 주장

논쟁	주장	추정 효과	선행연구
허위 정보의 양 증가	접근 및 사용의 용이성으로 인해 생성형 AI는 개인 및 조직화된 행위자에게 거의 또는 무료로 대규모 잘못된 정보/허위 정보를 생성하는데 사용될 수 있음.	허위 정보의 양이 증가하면 악의적인 행위자가 부정확하거나 오해의 소지가 있는 정보로 확장시켜, 사실의 콘텐츠를 덮거나 혼란을 야기할 수 있음	Bell(2023), Fried(2023), Hsu & Thompson(2023), Marcus(2023), Pasternack(2023), Ordonez, et al. (2023), Tucker (2023), Zagni & Canetta (2023)
허위 정보의 질 향상	기술적 역량과 사용 용이성으로 인해 생성형 AI를 사용하여 고품질의 허위 정보를 생성할 수 있음.	잘못된 정보의 품질이 높아지면 설득력이 높아지며, 그럴듯하고 사실을 폭로하거나 검증하기 어려운 콘텐츠가 생성됨. 이는 허위 정보의 확산을 허용하거나 (잘못된 정보의 양 증가와 함께) 인식론적 위기, 즉 모든 유형의 뉴스에 대한 전반적인 신뢰 상실에 기여할 것임.	Epstein & Hertzmann(2023), Fried(2023), Goldstein, et al. (2023), Hsu & Thompson (2023), Pasternack (2023), Tiku (2022), Tucker (2023)
허위 정보의 개인화 증가	기술적 역량과 사용 편의성으로 인해 생성형 AI를 사용하면 사용자의 취향과 선호도에 맞춰 맞춤화된 고품질 허위 정보를 생성할 수 있음.	허위 정보에 대한 소비자의 설득력 증가, 위와 동일한 결과 발생	Benson(2023), Fried(2023), Hsu & Thompson(2023), Pasternack(2023)
타당하지만 허위 정보의 비자발적인 생성	생성형 AI는 유용한 콘텐츠(예: 코드를 생성하는 챗봇)를 생성할 수 있음. 그러나 완전히 부정확하면서도 그럴듯해 보이는 정보를 생성할 수도 있음. 의도하지 않은 경우 사용자는 잘못된 정보를 생성하여 잠재적으로 확산하는 주체가 될 수 있음.	생성적 AI 사용자와 잠재적으로 정보를 공유하는 사람들에게 잘못된 정보를 제공	Fried(2023), Gold & Fischer(2023), Ordonez, et al. (2023), Pasternack (2023), Shah & Bender (2023), Zagni & Canetta (2023)

출처: Simon, Felix, Sacha Altay, and Hugo Mercier. 2023.

- 해당 논문에서는 생성형 AI는 현실적이지만 허위이거나 오해의 소지가 있는 콘텐츠를 대규모로 쉽게 생성할 수 있게 해줄 것이며 사람들의 신념과 행동, 정보의 공공성 및 민주주의에 잠재적으로 치명적인 결과를 초래할 수 있을 것이라 지적하였다. 이에 사이먼은 관련 위험을 <표 2-1>과 같이 4개 범주로 구분하였다.
- 첫째, 생성형 AI 사용으로 인한 허위 정보의 양적 증가, 둘째, 기술 진보로 인한 생성형 AI 고품질 허위 정보의 증가, 셋째, 취향과 선호도에 맞춰 맞춤형 고품질 잘못된 정보 생성, 넷째, 의도하지 않은 경우의 사용자 허위 정보 생성 증가이다.
- 해당 연구에서, 필자는 미국이 풍부하고 경쟁적인 미디어 생태계를 갖춘 부유하고 민주적인 국가의 미디어 환경에 대한 증거에 주로 의존하고 있기 때문에, 위 4개 분류의 위험은 과대평가되는 사항이라 결론지었다. 다만, 다른 국가에서는 사용할 수 있는 데이터가 적고 생성형 AI가 그곳에서 더 큰 부정적인 영향을 미칠 수 있다는 점을 배제할 수 없음을 지적한다. 즉 미국 국내 환경에서 허위 정보 표시에 관한 투명성과 스크리닝을 통해 현재까지의 생성형 AI 콘텐츠에 관한 제어가 가능하나, 미국 외 지역에서의 콘텐츠 범람에 대한 우려를 지적하고 있다.

## 2.2 AI 규제 동향 유형화

### 가. 생성형 AI 규제 동향

- 주요국들은 생성형 AI가 선거에 미칠 방해 행위에 주목하고 있다. 최근 생성형 AI에 의한 거짓뉴스, 딥페이크(deepfake) 및 위조를 통한 허위 영상 생성 등의 선거를 방해하는 행위가 다수 발생해 왔기 때문이다. 젤렌스키 우크라이나 대통령의 항복선언 가짜영상, 미 국방부 폭발의 딥페이크 가짜영상, 트럼프 전 대통령 등 유명정치인에 대해 악의적으로 생성된 가짜영상 등은 사회에 충격을 주었다. 최근에는 전문가도 분별하기 힘든 가짜영상이 생성되며 이에 의한 파장이 우려되고 있는 것이다.

○ 이러한 가운데, 주요국들은 AI 관련 입법 추진과 관련해 두 가지 배경에 의해 추동된다.

- 첫째, 국가별로 AI 관련 입법 추진의 주요 관점이 경제 발전, 즉 AI 산업발전에 목표를 두고 있는 경우이다. 이는 AI 산업발전을 목표로 법률적 지원을 하기 위한 입법 추진 그리고 AI 산업이 줄 수 있는 위험을 방지하기 위한 규제 등으로 나뉜다. AI 산업 위험 방지를 위해 각 국가는 특정 영역에서의 AI 사용에 대한 고위험군 등으로 분류되어 관리된다. 또한 AI 산업 위험 방지를 위한 규제도 기술적 규제와 확산에 관한 규제로 분류된다.
- 둘째, AI가 정치 및 사회적 가치 위해를 방지하기 위한 규제로 분류된다. 관련 위험은 개인 혹은 사회에 대한 침해, 사회가 가지는 가치에 관한 침해 등에 해당한다.

○ AI에 대한 규제는 규제원칙 및 방식에서도 국가별 차이가 드러난다. 이는 규칙 기반(rule-based) 방식과 원칙 중심(principle-based) 방식으로 구분한다.

- 첫째, 국가별로 세부 규제를 구체적으로 적시한 규칙 기반 방식은 규제 방법을 자세히 규정하거나 일련의 규칙을 제공하는 것이다. 세부 규칙을 제정함에 있어서도 대상별 규제를 제정하기도 한다. 이는 AI를 통해 각종 콘텐츠 등을 활용하는 행위자에 대해 규제하는 방식이다. 이는 고위험 영역에서 사용할 AI에 대해 제어된 범위를 제공하거나, 생성된 콘텐츠에 AI 사용 식별을 강제하는 방식 그리고 AI의 확산의 통로가 되는 플랫폼에 대한 규제이다. 생성된 콘텐츠에 대하여 플랫폼 기업이 책임 있게 관리, 통제해야 할 의무를 부여하는 방식으로 나뉜다.
- 둘째, 원칙 기반 방식은 규제 결과와 원칙이 설정되고 해당 결과를 달성하는 방법에 대한 통제, 조치, 절차는 각 조직이 결정해야 하는 방식이다.<sup>7)</sup>

---

7) Conradie, Lizelle. 2019. "Rules-based approach vs Principle-based approach to regulation in the Financial Industry." Etude Risk Management <https://www.etude.co.za/article.php?article=32> (검색일: 2023. 11. 01).

## 나. 생성형 AI 규제 유형화

- 본 연구에서는 주요국의 생성형 AI 규제 동향을 ‘AI 관련 입법 목적’ 및 ‘AI에 대한 규제는 규제원칙 및 방식’에 따라 아래와 같이 분류한다.

〈그림 2-4〉 생성형 AI 규제 유형화

		규제원칙		
		원칙 기반	규칙 기반	
입법 목적	경제 목적	미국	유럽연합	☞ AI 산업의 발전과 경쟁력 증진
	정치, 사회적 가치 목적	중국	유럽연합 회원국 스위스	☞ 민주주의, 사회주의 가치 수호

- 다수 국가들이 AI 규제를 준비하고 있는 가운데, 미국과 EU는 AI에 대한 규제 입법에 앞장서 왔다. 미국과 EU의 규제 내용 및 동향 대한 자세한 내용은 4장에서 다룰 것이다.

- 2019년 트럼프 전 대통령이 행정명령 제13859호에서 “미국 인공지능 구상”을 공표한 이후, 미국은 AI 관련 규제를 입법하였다. 본 연구에서는 AI 연구개발과 훈련과 관련한 법률인 「생성적 적대 신경망 출력물 확인법」, 「국가 인공지능 구상법 2020」을 분석한다. 또한 연방정부가 제정한 AI 사용에 관한 행정명령 제13960호, 「정부인공지능법 2020」, 「조달인력의 인공지능 역량강화법」, 「미국인공지능진흥법」 등과 더불어 최근 논의되고 있는 생성형 AI 기술 규제 논의 동향을 고찰한다. 또한 미국정부는 정치 광고에 AI를 사용했으면 출처를 표기하는 법안을 발의하였다. 최근 미국의 AI 관련 입법 동향은 생성형 AI 규제를 위한 행정명령(Execute Order), AI 권리장전(AI Bill of Rights), AI 훈련법(AI Training for the Acquisition Workforce Act)과 2023 국방수권법(National Defense Authorization

Act for Fiscal Year 2023) 등이 있다. 또한 위험관리 측면에서는 제시된 알고리즘 책임법(Algorithmic Accountability Act of 2022), 권리보호 측면에서의 미국 데이터 개인정보 및 보호법(American Data Privacy and Protection Act)과 디지털 플랫폼 위원회법(Digital Platform Commission Act of 2023)이 있다.

- 이와 같이 미국은 다양한 AI 관련 규제 입법 추진하고 있으며, 알고리즘 편향부터 개인정보 보호, 첨단 AI 모델의 안전 규정에 이르기까지 다양한 분야를 다룬다. 미국은 전체적인 AI 기술 및 산업발전, 개인권리 보호의 원칙 기반 아래, 관련 행정부 혹은 지방정부의 폭넓은 해석에 의한 규제 적용이 진행되고 있다.
- 미 정부의 AI 관련 입법 추진에도 불구하고, 생성형 AI 기술 규제 영역에 대한 준비는 미진한 상황이다. 미 상원은 관련 기업가(Elon Musk, Mark Zuckerberg, Bill Gates) 및 윤리학자와 학계를 포함한 17명의 비공개 AI 포럼을 구성하여, 규제가 미치지 못하는 영역을 파악한 후, 이를 FEC(연방선거관리위원회)가 이 문제를 검토할 수 있도록 할 방침이다. 즉 시장중심적 AI 규제의 원칙 기반 접근이 진행되고 있다.

○ EU는 2023년 6월 14일 유럽의회에서 인공지능 법안을 채택하였다. 이는 삼자입법기관의 정책결정과정을 거쳐 빠르면 2026년부터 시행될 예정이다.

- EU 집행위원회는 2021년 4월 AI에 대한 최초의 EU 규제 프레임워크를 제안하였으며, 최근 입법 과정을 완료하였다. 해당 규제안에는 AI의 위험을 허용할 수 없는 위험(Unacceptable risk), 고위험(High risk), 제한된 위험(Limited risk), 위험이 최소화되거나 없음(Minimal or no risk)으로 분류한다.
- 특히 해당 규제안에는 생성형 AI에 관한 투명성 요구 사항이 제시되었다. 이는 구체적으로 AI가 제작한 콘텐츠를 공개, 불법 콘텐츠 생성을 방지하도록 모델 설계, 훈련에 사용된 저작권 데이터의 요약게시에 해당한다. 연구팀은 이러한 생성형 AI에 관한 제한 사항이 선거에서의 활용에서 어떠한 제재 사안으로 작용할지에 관하여 분석한다.
- EU AI 규제가 3자 입법 과정에 진입했으나, ‘생체 인식 감시’ 등의

특정 문제에 대한 유럽의회와 각료이사회 간 상당한 입장 차이가 있어 난항이 예상된다.

- 생성형 AI 선거 규제 관련 선행연구는 대부분 가짜 뉴스 규제에 중점을 두고 있다. 본 연구는 생성형 AI 규제방식을 원칙-규칙 기반 규제로 유형화했다. 그러나 선거 분야에서 생성형 AI 규제 방식은 유사한 결과를 보이지 않는다. 또한 엄격한 구분이 어렵다.
- 생성형 AI 분야는 전 세계 주요 국가들의 주요 경제안보 영역으로 산업 경쟁력의 측면이 규제를 결정하는 주요한 요인이 될 것이다. 따라서 이 분야에 앞서 있는 국가의 경우 원칙 기반 규제를 선호할 것이며 이 분야에서 경쟁력이 떨어지는 국가의 경우 핀셋 규제가 가능한 규칙 기반 규제를 선호할 것이다. 이러한 경향은 이 분야 기술에 앞선 미국이 원칙기반 규제를, 이 분야 기술을 확보해야 하는 유럽연합의 경우 규칙 기반 규제를 활용한다는 점에서 주요한 사례가 될 수 있다.
- 그러나 두 행위자 모두 민주주의 가치를 수호한다는 점에서 언론의 자유를 포함한 주요한 민주주의 가치를 훼손할 가능성이 높은 강력한 법제화를 선호할 가능성이 낮을 것이다. 오히려 특정 이념에 근거한 국가의 경우 강력한 법제화를 선호할 것이다. 최근 신생민주주의 국가에서 생성형 AI 선거 규제의 법제화 경향이 두드러진다.
- 본 연구가 사례로 선택한 민주주의 국가들의 경우 사안의 심각성과는 별개로 엄격한 법제화 추구 움직임이 있으나 대부분 실패하고 있다. 앞에서 언급한 대로 대부분의 주에서 명확한 법제화는 실패하고 있다.
- 생성형 AI 선거규제 관련해서 규제 방식은 법제형과 자율규제형으로 나누어 볼 수 있다. 앞에서 생성형 AI에 대한 규제 접근과 구별될 수 있다.
  - 주요 국가의 사례를 통해 논란이 되고 있는 생성형 AI의 규제에 대한 강력한 법을 통해 문제를 해결할 것인지(규칙 기반) 아니면 신기술에 대한 국가 경쟁력과 시장 중심의 규제(원칙 기반)를 할 것인지에 대한 논란이 있다.

- 선거 캠페인에서 문제가 발생하며 국가의 정치 세력에 따라 강력한 법제화를 요구하는 목소리가 있으나 여전히 대부분의 사례 국가가 자율규제에 의존하고 있다.
- 선거운동과 선거 과정에서 문제 발생이 빈번해짐에 따라 법제화 규제 방식의 필요성이 높아지고 있다.
- 그러나 생성형 AI 규제에서 강력한 법제화 방식을 추구하고 있는 EU의 경우 산업경쟁력 확보를 동기로 한다는 점에서 선거 과정에 대한 생성형 AI 규제에서 회원국들의 규제 완화에 대한 요구가 빈번함에 따라 특정 기간, 즉 선거기간에만 법제화가 주요 규제방식으로 자리잡을 가능성이 있다. 그러나 대부분 국가에서 다양한 생성형 AI 선거법 제정이 실패하고 있다는 점에서 단기적으로는 대부분의 국가에서 자율규제를 유지할 것으로 보인다.

## 2.3 선거 과정에서의 AI 활용

### □ 선거 과정에서 핵심적인 다섯 가지 AI 활용 가능성<sup>8)</sup>

- 민주주의 국가 내에서 선거의 시간, 장소, 방식 등을 결정하는 선거 절차는 AI의 활용이 제한적인 몇 안 되는 분야 중 하나이다. 최근 많은 국가의 선거 관리 기관이 선거 과정에서 AI를 활용하는 방안을 모색하고 심의하기 시작했다.
- 선거 과정에서 AI 활용 가능성이 있는 다섯 가지 분야를 고려하고, 그 안에서 활용에 수반되는 과제를 발굴한다. 이 다섯 가지 분야는 유권자 명부 관리, 투표소 위치 결정, 투표소 보호 과정, 유권자 인증, 선거 영상 모니터링이다.
- 각 활용 분야마다 배경을 설명하고, 현재 또는 잠재적인 AI 사용 사례를 설명하며, AI 사용의 현재 또는 잠재적 파급 효과와 AI 사용을 고려할 때 위험을 완화할 수 있는 잠재적 방향에 대해 논의하고자 한다.

8) 선거과정에서 핵심적인 다섯가지 AI 활용 가능성에 관한 연구는 주로 다음 논문에 기반하여 작성했다. Simoes, Stanley and Muiris MacCarthaigh. 2023. "AI and Core Electoral Processes: Mapping the Horizons." AI Magazine.



- 그러나 현재 선거 과정에서의 AI 사용은 미미하므로 실제적이고 광범위한 AI 확산에 앞서 위험과 완화 가능성을 식별하는 정도로 제한된다는 점을 명확히 한다.
- 선거 과정에서 AI를 사용할 때 발생할 수 있는 위험과 기회의 지평을 파악하고 이 주제에 대한 논의를 형성하는 데 도움을 주기 위한 시도이다.
- 각 방안에 대한 분석에서는 현재 또는 잠재적인 AI 기술 사용 현황과 맥락을 비판적, 정성적으로 평가하는 데 초점을 맞춘다. 이후 잠재적 위험에 대한 분석과 위험 완화 또는 대체 경로에 대한 논의가 이루어진다. 분석의 개요는 다음 <표 2-2>와 같다.

<표 2-2> 선거과정에서 AI 활용의 위험과 방향성

수단	AI 사용	위험	방법
유권자 명부 관리	휴리스틱 중심의 근사적 접근	완결성-접근성 균형 이슈	접근성 중심 AI
	자료 연계	AI의 편향성	합리적 설명
	이상치 탐지	AI의 과도한 일반화	현지 조사
투표소 위치 결정	드롭 박스 위치 결정	비용 최적화	복수 결과
	시설 위치	변동성 및 검색 비용	검사형 AI
	클러스터링	편파적 조작	취약 유권자
투표 부스에서 발생할 수 있는 문제점 예측	예측가능한 치안	구조적인 인종차별	투명성
	시계열 모티프	폭력의 악화	엄격한 통계
		피드백 과정	공정한 AI
유권자 인증	얼굴 인식	인종/젠더(gender) 편견	대안
	생체 인식	알 수 없는 편견	편견 검사
		투표율 감독과 기타	극단적 사례 구상
영상 모니터링	비디오 기반의 득표 계산	선거의 완결성	가벼운 모니터링
	문제 파악	공동체의 주변화	개방형 데이터
	유권자 재식별	기존 모니터링 약화	

출처: Simoes, Stanley and Muiris MacCarthaigh. 2023: 220.

## 가. 유권자 명부 관리

- 오류가 없는 최신 유권자 명부는 선거의 완결성(integrity)을 보장하기 위해 중요하다.
  - 유권자 명부는 시, 구, 지방 또는 국가처럼 다양한 단위에서 관리한다.
  - 모든 유권자가 모든 선거에 투표할 수 있는 것은 아니기 때문에 유권

자 명부는 해당 선거의 특징을 반영할 수 있다.

▷ 예를 들어 아일랜드(Republic of Ireland)에서 아일랜드 시민이 아닌 자는 지역 선거에 참여할 수 있으나, 대선에는 참여할 수 없다. 또한, 유럽연합 시민은 유럽의회 선거에 참여할 수 있으나, 유럽연합의 회원국 선거에는 참여할 수 없다.

- 모든 유권자 명부는 최신화되어야 하며 이를 위해 유권자가 이주하거나 시민권 변경 혹은 사망했을 경우 명부에서 해당 유권자는 삭제 또는 추가되어야 한다.
- 명부에 추가신청 책임은 전입한 유권자에게 있으며, 조건에 부합하고 유권자가 등록되어 있지 않았을 때, 유권자는 명부 추가를 신청할 수 있다. 또한, 유권자가 같은 명부에 중복으로 등록되어서는 안 되며, 다른 주의 유권자 명부에 올라가 있어도 안 된다.

#### ○ 잠재적인 AI 활용

- 전 거주자가 유권자 명부에 남아있는 경우 선거의 완결성에 영향을 미칠 수 있으며(특히 허위 투표의 경우), 합법적인 유권자가 명단에서 누락되는 것은 선거 절차에 대한 신뢰를 떨어뜨릴 수 있다. 접근성과 완결성의 균형을 신중하게 유지하는 것이 유권자 명부 유지 관리 업무의 핵심이다.<sup>9)</sup>
- 학계에서는 아직 유권자 명부 유지와 관련하여 AI를 활용했다는 논의는 없지만, 미국에서 유권자 롤(rolls)을 없애기 위해 성명 확인(name matching)과 같은 자동화 시스템의 활용 사례는 존재한다.
- 유권자 명부 관리에 대한 비전통적인 접근 방식과 관련하여, 위스콘신주에서는 행정 데이터(운전면허 데이터, 우편 서비스 데이터)를 사용하여 유권자 명부 내 잠재적 부실 항목을 식별한 후 해당 주소에 여전히 거주하고 있는지 확인하기 위해 엽서를 보내는 사례가 있었다.<sup>10)</sup>

9) Merivaki, Thessalia. 2020. ““Our voter rolls are cleaner than yours” : Balancing access and integrity in voter list maintenance,” *American Politics Research* 48(5): 560-570.

10) Huber, Gregory, et al. 2021. “The racial burden of voter list maintenance errors: Evidence from Wisconsin’s supplemental movers poll books,” *Science Advances* 7(8).

- Simoes와 MacCarthaigh는 행정 자원(administrative sources)에서 근거 수집과 잠재적인 부실을 결정하는 패턴 활용이 AI에 의해 가속화될 패러다임이라고 주장한다.
- ▷ 데이터 패턴을 기반으로 휴리스틱을 개발하고 더 빠른 의사결정을 위한 패턴의 활용이 AI의 핵심이기 때문이다.
- ▷ 이러한 패러다임이 잠재적이고 실질적인 AI가 유권자 명부 관리에 접목될 가능성을 열어줄 것이다.
- ▷ <그림 2-5>는 위스콘신주에서 비활동 유권자에게 발송하는 카드의 형식을 보여주며, 유권자가 해당 주소에 여전히 거주하고 있는지 확인해야 하는 책임을 유권자에게 전가한다.

<그림 2-5> 위스콘신 유권자 거주 확인용 발송 우편물 사례



출처: Simoes, Stanley and Muiris MacCarthaigh. 2023: 221.

○ 유권자 명부 유지를 위해 활용될 수 있는 AI 기술

- 데이터 관리 및 분석 커뮤니티 내에서 자료 연계
  - ▷ 패턴 기반 휴리스틱을 사용하여 데이터 소스(예: 유권자 명부 데이터베이스) 내 또는 데이터 소스 전반(예: 유권자 명부 및 운전 면허 데이터베이스)에서 중복을 식별하거나 관련 기록을 비교할 수 있다.

- ▷ 유권자 명부에서 두 개의 유사한 항목을 식별하거나 다른 주의 운전 면허 기록과 일치하는 유권자 명부 항목을 식별한다면, 이는 잠재적으로 오래된 항목으로 이해할 수 있으며, 이는 추가 조사 또는 삭제가 필요한 사항이다.

- 이상치 탐지

- ▷ 이상치 탐지라는 AI 하위 분야의 기술<sup>11)</sup>을 적용하여 유권자 명단에서 “비정형” 항목의 유입을 식별할 수 있으며, 이러한 항목은 수동 조사를 위해 지정될 수 있다.
- ▷ 이러한 사용은 유권자의 다양성에 대한 인식 부족으로 인해 발생한다는 점에서 본질적으로 문제가 있지만, 또한 유사한 고려 사항이 배제된 데이터 기반 “스마트” 치안<sup>12)</sup>에서 사용되고 있는 이상치 탐지의 일반적인 사용은 필연적으로 유권자 명부 유지 관리로도 이어질 것으로 예상된다.

○ 유권자 명부 유지에 활용될 수 있는 AI 기술의 영향

- 일반적으로 유권자 명부 유지에 있어서 접근성(access)과 완결성(integrity) 간의 균형에 주목한다.<sup>13)</sup>
  - ▷ 접근성(access)에 주목하는 경우 누락된 유권자를 식별할 수 있으며, 완결성(integrity)에 집중하는 경우 유권자 항목에서 부정 혹은 만료된 유권자를 발견하고 이를 삭제할 수 있다.
  - ▷ 접근성(access)의 경우 유권자 등록은 개별 유권자가 특정 기간에 등록하는 데 의존하는 경우가 많으며 물론 대중의 인식을 높이기 위해 국가가 자금을 지원하는 캠페인이 있을 수도 있다. 그러나 유권자 명부 관리에 대한 기술적 개입은 무결성 측면에 집중될 가능성이 높다는 점에서 실질적으로 AI가 활용될 수 있는 영역은 완결성(integrity)이라고 할 수 있다. 그렇기에 AI 활용

11) Wang, Hongzhi, Mohamed Bah, and Mohamed Hammad. 2019. “Progress in outlier detection techniques: A survey.” *Ieee Access* 7: 107964-108000.

12) Yamin, Muhammad, Andrii Shalaginov, and Basel Katt. 2020. “Smart policing for a smart world opportunities, challenges and way forward.” *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC)* 1.

13) Merivaki, Thessalia. 2020.

이 증가할수록 접근성보다는 완결성(integrity)에 치중할 가능성이 높다.

▷ 데이터 기반 AI는 시나리오에서 유권자 명부와 같은 가용 데이터 소스를 활용하는 데 중점을 두기 때문에, 유권자 명부에 없는 부재자 유권자가 아닌 부정 유권자를 식별하기 위해 AI를 사용하는 개념을 더 쉽게 개념화할 수 있다. 다시 말해, 유권자 명부 유지 관리에 AI를 더 많이 사용하면 완결성-접근성 균형에서 완결성 쪽에 더 가까운 구성으로 이어질 가능성이 높다.

○ AI 활용에 따른 세 가지 문제가 존재한다.

- 인종 문제로 인한 부담(racial burden)

▷ Huber 외(2021)<sup>14)</sup>는 수작업을 통해서라도 부정하거나 오래된 유권자 명부 항목을 식별하기 위한 행정 데이터 사용 패러다임이 소수민족에게 오류가 더 많이 관찰되는 ‘인종적 부담’을 초래하는 오류를 발생시킨다고 주장한다.

▷ 이는 기술적 정당성이라는 분위기 속에서 동일한 종류의 인종적 편견을 낳을 수 있다는 점에서 AI가 단순히 인간의 의사결정을 재현하는 것 자체가 문제가 될 수 있음을 의미한다.

- AI의 편향성

▷ 패턴 기반 휴리스틱을 사용하여 오래된 유권자 명부 항목을 식별하는 것은 수동 프로세스에 비해 AI를 사용하면 필연적으로 더 빠르고 심층적으로 이루어진다. 이러한 접근 방식은 데이터 기반 AI 시스템에서 종종 관찰되는 것처럼 무수히 많은 종류의 편견을 낳을 수 있다.<sup>15)</sup>

- 과도한 일반화

▷ 유권자 명부 문제는 지역별로 상당한 차이가 있다는 연구가 존

14) Huber, Gregory, et al. 2021.

15) Ntoutsis, Eirini, et al. 2020. “Bias in data-driven artificial intelligence systems—An introductory survey.” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10(3).

재한다.<sup>16)</sup> 유권자 명부 관리에서 AI 활용은 효과적인 작업을 보장하기 위해 특정 지역에 대한 맞춤형 유권자 명부 유지 관리 기술을 암묵적으로 요구한다. 이러한 요구는 머신러닝 연구에서 두 번째로 중요한 가치로 여겨지는 일반화 가능성에 대한 AI의 초점과 상당한 긴장관계에 놓여 있다.<sup>17)</sup> 즉, 유권자 명부 관리를 위해 AI를 일반화하면 유권자 명부 유지 관리 문제가 국가 또는 글로벌 차원의 지배적인 추세와 다른 지역에는 상당한 불이익을 초래할 수 있다.

## ○ 유권자 명부 관리에 AI 활용에 대한 방향성 제시

### - 접근성 중심 AI

- ▷ 접근성-완결성 균형 중 완결성에 치중된 유권자 명부 관리에 AI 활용 관련 연구는 접근성을 위한 AI 활용 관련 연구와 격차를 보인다.
- ▷ 여기에는 행정 데이터와 같은 외부 소스를 사용하여 유권자 명단에서 누락을 식별하는 AI 기술을 개발하는 것이 포함한다.
- ▷ 그런 다음 이러한 잠재적 누락을 플래그 지정하여 유권자 명단 등록을 위한 팸플릿 기반 표적 요청을 보낼 수 있다. 이는 보다 포괄적인 유권자 명부를 통해 민주적 절차를 강화하는 데 도움이 될 수 있다.

### - 합리적인 설명

- ▷ 유권자 명부에서 부정 가능성이 높거나 오래된 항목을 식별하기 위한 AI 활용에서 합리적인 설명을 생성할 수 있는 AI를 사용하도록 요구하는 것은 편향된 작동에 대한 중요한 안전장치가 될 수 있다.
- ▷ 물론 유권자 명부 문제 관련 전문가로부터 합리적인 설명을 들어야 하며, 이러한 조건을 준수할 수 있는 AI를 개발하는 기술적

16) Merivaki, Thessalia (2020).

17) Birhane, Abeba, et al. 2022. "The values encoded in machine learning research." *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*.

과제는 이 분야의 AI 연구에 새로운 방향을 제시할 수 있을 것이다.

- 현지 조사

- ▷ 유권자 명부를 관리하는 각 관할 구역은 앞서 언급한 유권자 명부 문제의 지역적 차이를 고려하여 유권자 명부 관리 AI의 지역적 맥락에 대한 적합성 여부를 파악해야 한다.<sup>18)</sup>
- ▷ 이를 위해서는 유권자 명부 관리를 위한 AI 기술이 각 관할권에 적용 가능한지 면밀히 검토하고 특정 조건에 맞게 쉽게 조정할 수 있도록 운영의 투명성을 유지해야 한다.

## 나. 투표소 위치 결정

### ○ 투표소 위치 결정에 AI 활용의 맥락

- 투표소는 유권자가 참정권을 행사하는 장소로서 선거에서 가장 중요한 장소로 간주할 수 있다.
  - ▷ 특히 게리맨더링, 즉 특정 정당에 유리하도록 선거구 경계를 재설정하는 관행에 대한 우려로 인해 지리적 위치의 중요성은 오래전부터 인식되어 왔다.
- 투표소 위치 결정은 선거구 내 개별 유권자가 투표권을 행사하기 위해 어디로 가야 하는지를 결정하는 보다 세분화된 작업이다.
  - ▷ 대부분의 투표소 위치는 어느 정도 고정적이지만 정치 심리학에서는 투표소의 특성(예: 학교 대 다른 장소,<sup>19)</sup> 교회 대 세속적 건물<sup>20)</sup>이 투표에 영향을 미칠 수 있다는 증거가 있으며, 이를 투표소 프레이밍 효과<sup>21)</sup>라고 한다.

18) Merivaki, Thessalia (2020).

19) Berger, Jonah, et al. 2008. "Contextual priming: Where people vote affects how they vote." *Proceedings of the National Academy of Sciences* 105(26): 8846-8849.

20) Rutchick, Abraham. 2010. "Deus ex machina: The influence of polling place on voting behavior." *Political Psychology* 31(2): 209-225.

21) Blumenthal, Jeremy and Terry Turnipseed. 2011. "The Polling Place Priming (PPP) Effect: Is Voting in Churches (Or Anywhere Else) Unconstitutional." *Boston University Law Review* 91(2): 561-599.

▷ 또한 투표소까지의 거리는 투표율과 상관관계가 있다는 주장이 있다.<sup>22)</sup>

- 투표소를 합리적으로 결정하는 것은 자유롭고 공정한 선거를 보장하기 위해 중요하다.

▷ 예를 들어, 기존 투표소는 학교(또는 서양에서 흔히 볼 수 있는 교회당)로, 여러 세대에 걸쳐 투표소로 계속 사용되어 온 곳일 수 있다. 정기적으로 사용하던 투표소가 운영이 중단되는 이례적인 경우, 가까운 학교(또는 교회)가 자연스럽게 대체 장소가 될 수 있다.

▷ 투표소 결정 프로세스를 자동화하는 주요 잠재적 상황은 내부 또는 외부 이동으로 인해 인구 이동이 많은 지역에서 발생할 수 있다.

▷ 대규모 내부 이동의 경우 수용 인원을 고려하여 새로운 투표소를 파악해야 할 수 있으며, 외부 이동의 경우 일부 기존 투표소의 운영을 중단해야 할 수 있다.

▷ 또한, 주기적으로 일부 관할권에서는 투표소를 일괄적으로 재결정/변경할 수 있는데, 인도에서는 투표소 합리화가, 미국에서는 텍사스와 조지아에서 편향성 논란을 일으킨 투표소 통폐합이 대표적이다.

#### ○ 투표소 위치 결정에서 잠재적인 AI 활용

- 기존 연구에서 투표소 위치 결정에 인공지능이 사용되었다는 증거는 거의 찾을 수 없었다.

- 이 분야에서 찾을 수 있는 유일한 컴퓨팅 방법은 투표용지 투입함 위치 문제에 대한 연구<sup>23)</sup>로, 특히 코로나 시대에 미국의 투표 인프라에 추가된 투표용지 투입함의 위치를 컴퓨터로 결정하는 것을 고려하고

22) Garnett, Holly and Sean Grogan. "I Came, I Saw, I Voted: Distance to Polling Locations and Voter Turnout in Ontario, Canada." *Canadian Journal of Political Science/Revue canadienne de science politique* 54(2): 316-334.

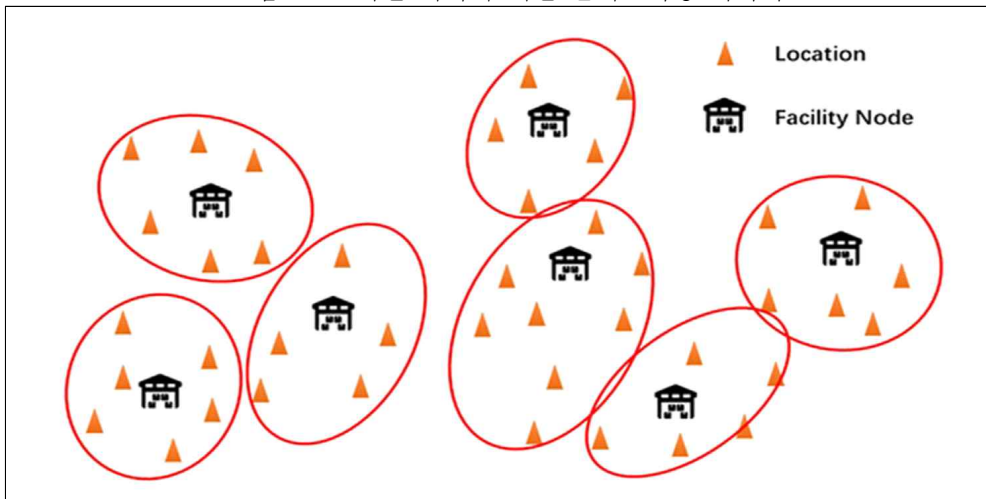
23) Schmidt, Adam and Laura Albert. 2022. "Locating ballot drop boxes." *arXiv preprint arXiv:2207.10707*.



있다.

- 반복적인 의사결정이 수반되는 경우, 합리적인 기간에 걸쳐 선행 투자를 정당화할 수 있도록 AI(또는 모든 종류의 자동화)가 우선시되어 선투자되는 경우가 많으나 투표소 위치 결정 문제처럼 반복적인 결정이 이루어지지 않는 경우 AI 연구가 부족한 것은 그리 놀라운 일은 아니다.
- 그러나 투표소 결정에 사용하기 위한 AI의 기술적 준비도는 관련 작업의 상당한 발전으로 인해 보통에서 높은 수준으로 간주될 수 있다.
  - ▷ 시설 위치에 대한 클러스터링 사용은 유사성이 공간적 거리와 반비례하는 것으로 판단될 수 있는 위치-근접성이 주된 배경 기준일 때 가장 적절하다.
  - ▷ 투표소의 경우, 유권자를 지리적 근거리 클러스터로 묶어 각 클러스터(즉, 유권자 그룹)에 투표소를 할당할 수 있다(그림 2-6) 참조).

〈그림 2-6〉 시설 위치에 대한 클러스터링 이미지



출처: Simoes, Stanley and Muiris MacCarthaigh. 2023: 224.

- ▷ 클러스터링 알고리즘은 1960년대까지 거슬러 올라가는 오랜 전통을 가지고 있다. 투표소 작업의 경우 이는 유권자가 투표소에 도착하기 위해 이동해야 하는 누적 거리를 최소화하는 것에 해당한다.
- ▷ 거리로 인해 일부 유권자는 여전히 투표소에 도착하기 위해 상

당한 거리를 이동해야 할 수 있다. 공정한 클러스터링에 대한 최근의 발전<sup>24)</sup>은 클러스터 대표(투표소)와의 거리 면에서 너무 불리한 대상(유권자)이 없도록 해야 함을 강조한다.

- ▶ 공정한 클러스터링에 대한 이러한 연구와 기타 발전<sup>25)</sup>은 투표소 위치 결정에 적용할 수 있는 고려사항에 더욱 부합하도록 클러스터링 문헌을 확장했다.

## ○ 투표소 위치 결정에서 잠재적인 AI 활용의 영향

- AI 활용에 따른 세 가지 위험이 존재한다.

- ▶ 비용 최적화: 비용을 투표소 위치 설정에 가장 중요한 기준으로 설정한다.
  - Celik Turkoglu와 Erol Genevois(2020)<sup>26)</sup>가 지적하는 투표소 위치결정은 수익성(profitability)과 지속가능성(sustainability) 사이의 균형에 있다고 주장한다.
  - 비용 최적화에 주요 기준으로 드롭 박스 위치 문제에 관한 기술이 고려된다.
  - 예를 들어, 경쟁에 비해 유동 인구가 많은 위치에 있는 시설을 배치함으로써 수익성을 향상할 수 있다.
  - 그러나 투표소는 각 유권자가 투표소에 고유하게 배정되고 선거 기관이 투표 과정을 독점하기 때문에 경쟁 시장에서 운영되지 않는다.
  - 따라서 비용 최적화가 투표소 결정에 영향을 미치지 않도록 상당한 주의를 기울여야 하며, 기성 기술이 해당 작업에 적합하도록 충분히 맞춤화되어야 한다.

---

24) P, Deepak and Savitha Abraham. 2020. "Representativity Fairness in Clustering." *In Proceedings of the 12th ACM Conference on Web Science 202-11*. Association for Computing Machinery; Jung, Christopher, Sampath Kannan and Neil Lutz. 2020. "Service in your neighborhood: Fairness in center location." *Foundations of Responsible Computing(FORC)*; Abbasi, Mohsen, Aditya Bhaskara, and Suresh Venkatasubramanian. 2021. "Fair clustering via equitable group representations." *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*; Stepanov, Nicola-Paul. 2022. "Fairness in Facility Location through Clustering." Master's thesis, University of Amsterdam.

25) Mahabadi, Sepideh, and Ali Vakilian. 2020. "Individual fairness for k-clustering." *International conference on machine learning*. PMLR.; Kleindessner, Matthäus, Pranjal Awasthi, and Jamie Morgenstern. 2020. "A notion of individual fairness for clustering." arXiv preprint arXiv:2006.04960.

26) Turkoglu, Celik and Erol Genevois. 2020. "A comparative survey of service facility location problems." *Annals of Operations Research* 292(1): 399-468.

▷ 투표소 변동성 및 투표소 검색 비용

- 투표소 위치 결정 기술의 가용성은 기술 개발 또는 조달에 드는 초기 비용으로부터 가치를 창출하기 위한 논리가 작동되는 현상을 발견할 수 있다.
- 이는 지정된 투표소가 정기적으로 바뀌는 현상을 초래하여 투표소 변동으로 투표소를 찾는데 발생하는 비용<sup>27)</sup>을 증가시켜, 투표율에 부정적인 영향을 미칠 수 있다.

▷ 편파적 조작: 게리맨더링(gerrymandering)

- 새로운 기술이 세밀한 게리맨더링을 가능하게 하여 전통적 게리맨더링처럼 선거구 경계가 아닌 투표소가 당파적 목적을 위해 조작될 수 있다는 것이다. 물론 아직 투표소 위치의 당파적 조작을 시사하는 증거는 없다.<sup>28)</sup>
- 그러나 당파적 목적을 위해 투표소의 지리적 조작 기술을 제안하는 최근 연구<sup>29)</sup>는 특정 강력한 가정하에서 특정 후보의 승리를 보장하기 위해 투표소 집합을 선택할 수 있는지이다.

○ 투표소 위치 결정의 AI 활용에 대한 방향성 제시

- 자유롭고 공정한 선거를 위한 최선의 이익에 부합하도록 투표소 결정에 기술을 사용하는 위험을 완화하고 진보적인 방법을 고려해야 한다.
  - 투표소 위치 집합을 식별하는 것은 여러 기준을 동시에 최적화하는 집단 최적화 작업이다. 따라서 최적화 결과, 즉 제안된 투표소 위치 집합이 주어졌을 때 전문가(예: 선거 관리인)가 그 위치 중 하나가 부적절하다고 판단하여 그 위치를 약간 변경하는 것은 정당화될 수 없다.
- ▷ 유권자가 투표소의 위치가 부적절함을 느낀다고 해서 이를 수정할 수 없기에, 이와 같은 문제를 해결하기 위해 AI 기술은 다수

27) Brady, Henry and John McNulty. 2011. "Turning Out to Vote: The Costs of Finding and Getting to the Polling Place." *American Political Science Review* 105(1): 115-34.

28) Shepherd, Michael, Adriane Fresh, Nick Eubank, and Joshua Clinton. 2021. "The politics of locating polling places: race and partisanship in North Carolina Election Administration, 2008-2016." *Election Law Journal: Rules, Politics, and Policy* 20(2): 155-177.

29) Fitzsimmons, Zack, Omer Lev. 2020. "Selecting Voting Locations for Fun and Profit." In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, 224(30).

의 선택지를 생성할 수 있어야 한다.

▷ 특정 투표소 위치의 부적절성에 관한 판단은 수학적으로 추상화하여 최적화에 반영할 수 없는 다양한 외적 이유(예: 쓰레기 처리장과의 근접성, 적절한 화장실 시설의 부재 등)에서 비롯될 수 있어 선거관리기관의 판단과 더불어 다양한 외부적인 요인이 함께 결합되어 결정된다.

· 투표소의 위치를 약간만 변경해도 다른 투표소의 위치를 변경해야 전체 최적화에 부합하는 결과를 얻을 수 있으므로 특정 투표소의 위치를 변경하는 프로세스는 부적절하다.

· 하나의 클러스터 중심을 변경하면 다른 클러스터 중심이 변경되고 멀리 떨어진 물체도 클러스터에 할당될 수 있는 경우와 유사하게 보일 수 있다.

· 인공지능의 결과를 조금이라도 합법적으로 수정할 수 없다는 것은 인간 전문가에게 곤욕스러운 일이다.

- 투표소에 대한 기술적 조작을 가능하게 하는 AI 도구의 출현<sup>30)</sup>으로 인해 이러한 조작을 방지할 수 있는 기술이 요구되고 있다.

▷ 이러한 안전장치는 그 자체로 감사 지향 AI로 작동하여 당파적 조작을 발견하는데 유의미하다.

▷ 이러한 감사 AI가 적대적 학습 프레임워크 내에서 사용되어 보다 정교한 AI 기반 악성 조작을 개발할 수 있는 위험이 있다.<sup>31)</sup>

- 투표소의 위치는 사회적 취약 계층에게 불리하게 작용할 수 있기에, 포괄적인 투표 촉진하기 위한 노력이 필요하다.

▷ 어떤 형태의 투표소 위치 결정이든, AI 기반이든 사람이 직접 결정하든, 불리한 유권자가 발생할 수 있다,

▷ 장애가 있는 유권자를 식별하는 새로운 AI 흐름은 보다 포용적인 투표를 촉진하고 투표소 구성의 결함을 상쇄하기 위한 노력(예: 교통편 지원 우선순위 지정 또는 원격 투표 옵션 제공)을 지원할 수 있다.

· 혜택을 받지 못하는 유권자를 식별하는 새로운 AI 스트림(novel

30) Fitzsimmons, Zack, Omer Lev. 2020.

31) Gui, Jie, Zhenan Sun, Yonggang Wen, Dacheng Tao, Jieping Ye. 2021. "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications." IEEE Transactions on Knowledge and Data Engineering 35: 3313-3332.

stream of AI)은 포괄적인 투표와 선택한 투표소 구성의 적자를 상쇄하는 것을 촉진한다.

#### 다. 투표 부스에서 발생할 수 있는 문제점 예측

##### ○ 투표 부스에서 발생할 수 있는 문제점과 AI 활용의 맥락

- 유권자의 참정권 행사를 위협하는 현상 중에는 유권자 협박이 있으며, 이는 점점 더 전 세계적인 현상으로 인식되고 있다,
  - ▷ 러시아 일부 지역에서는 강력한 법적 보호 장치가 마련되어 있음에도 불구하고 유권자에 대한 협박은 여전히 존재한다.<sup>32)</sup>
  
- 다른 형태의 투표권 침해에는 폭력적인 수단을 통한 부스 점거<sup>33)</sup>가 포함된다.
  - ▷ 이러한 우려를 불식하기 위해 잠재적 범죄자에 대한 예방적 구금<sup>34)</sup>과 특정 투표소를 문제 투표소로 지정하여 경찰의 추가 보호를 제공하는 것을 정당화하는 데 사용할 수 있다.
    - 투표소 경찰 배치 및 순찰의 이러한 분류와 밀도 변화는 선거 계산에 영향을 미칠 수 있다.
    - 또한, 투표소에 필요한 보호 수준은 위치에 따라 영향을 받기 때문에 두 가지 요인이 밀접하게 연결되어 있다.
  
- 투표소 보호의 우선순위 결정을 잠재적인 AI 사용의 한 방법으로 고려할 수 있다.

##### ○ 투표 부스 보호의 우선순위 결정을 위한 잠재적 AI 활용

- 기존연구에 따르면 투표소에서의 협박과 폭력에 대한 우려는 전통적으로 공공 부문의 AI 보급률이 낮은 저개발국<sup>35)</sup>에서 주로 강조되고 있다.

32) Woodruff II, James. 2011. "Where the Wild Things Are: The Polling Place, Voter Intimidation, and the First Amendment." *University of Louisville Law Review* 50: 253-285.

33) Sharma, S. 1999. "Booth Capturing: Judicial Response." *Journal of the Indian Law Institute* 41(1): 44-55.

34) Verma, Arvind. 2005. "Policing Elections in India." *India Review* 4(3-4): 354-376.

35) Besaw, Clayon. 2021. "Election Violence Spiked Worldwide in 2020-Will this Year be Better?"

▷ 이는 투표소 보호 우선순위 지정을 위한 맞춤형 AI 기술이 존재하지 않는 배경이 될 수 있다.

- 선거에서 결과의 안정성을 보장하기 위해 ‘유권자 그룹’의 보호 우선순위를 정하는 이론적 AI 연구가 진행되고 있지만,<sup>36)</sup> 투표소 보호 우선순위를 정하는 작업에 적용하려면 많은 조정이 필요할 수 있다.
- 문제가 있는 투표소의 식별 기술은 범위가 발생하기 쉬운 장소를 식별하는 핫스팟 치안과 매우 유사한 작업에서 AI 개발이 급성장했다는 점에서 이 분야의 AI 활용 방안으로 선택할 수 있다.
- 이러한 전제들이 바로 문제 투표소를 예측하는 아이디어의 기반이 된다. 예측적 치안 기술은 기업 소프트웨어 공급업체가 선거 AI 시장으로 확장 가능성을 열어준다.
- 투표소 문제 예측은 고유한 뉘앙스가 있는 작업이므로 이를 위한 맞춤형 방법은 데이터 기반 AI의 관련 하위 분야에서 발전된 기술을 활용할 수 있다. 예를 들어, 임박한 투표소 폭력을 나타내는 미리 컴파일된 패턴 세트를 시계열 스트림에서 자동으로 발견하는 데 활용할 수 있으며, 이 작업은 시계열 모티브 발견의 프레임워크 내에서 수행될 수 있다.<sup>37)</sup>

#### ○ 투표 부스 보호를 위한 잠재적 AI 활용의 파급효과

- 문제가 있는 투표소 예측에 AI를 사용한 사례가 없다는 점을 감안할 때, 향후 이 작업에 AI를 활용할 시 발생할 수 있는 잠재적 파급 효과에 초점을 맞춘다.
- 투표소의 밀집도는 직간접적인 영향을 미칠 수 있다.

▷ 직접적인 위협의 예로, 경찰이 투표소를 엄격하게 단속하면 유

---

The Conversation (February 18) <https://theconversation.com/election-violence-spikedworldwide-in-2020-will-this-year-be-better-153975>. (검색일: 2023. 08. 25).

36) Dey, Palash, Neeldhara Misra, Swaprava Nath, Garima Shakya. 2019. “A Parameterized Perspective on Protecting Elections.” In Proceedings of the Twenty Eighth International Joint Conference on Artificial Intelligence, IJCAI-19: 238-244.; Dey, Palash, Neeldhara Misra, Swaprava Nath, Garima Shakya. 2021. “A Parameterized Perspective on Protecting Elections.” Theoretical Computer Science 874: 15-31.

37) Mueen, Abdullah, et al. 2009. “Exact Discovery of Time Series Motifs.” In Proceedings of the 2009 SIAM International Conference on Data Mining (SDM): 473-484.

권자들이 투표소를 기피하여 투표율이 감소할 수 있다.

· 2017년 스페인에서 경찰의 급습으로 투표율이 감소한 것으로 관찰되었다.<sup>38)</sup>

- 선거 정책이 선거 결과에 영향을 미칠 수 있다는 광범위한 관찰이 있다.<sup>39)</sup>

▷ 따라서 지지율이 낮을 것으로 예상되는 지역을 문제가 있는 지역으로 지정하는 인공지능을 장려하여 선거 결과를 간접적이고 조용하게 자신에게 유리하도록 유도하는 것이 현 정당의 이익에 부합할 수 있다.

- 예측 및 핫스팟 폴리싱의 사용 패턴을 반영하면, 현재 해당 영역의 위험에 대한 이해는 문제가 있는 투표소를 예측하는 간접적인 위험에 대해 생각해 볼 수 있는 유리한 지점을 제공한다.

▷ 핫스팟 치안, 그리고 일반적으로 새로운 치안은 소수인종 지역에 더 높은 수준의 치안 서비스를 제공함으로써 미국 내 체계적인 인종차별을 악화시킨다는 주장이 제기되고 있다.<sup>40)</sup>

▷ 예를 들어, 치안 유지는 경찰의 폭력성 악화와 연결되어. 이에 따라 이 기술의 선구자들조차도 회의적인 반응을 보였다.<sup>41)</sup>

- 응용 AI는 종종 모델링하고자 하는 바로 그 현상을 재구성하는 수행적 측면을 가지고 있다는 주장이 있다.<sup>42)</sup>

- 예측 치안에서 폭주하는 피드백 루프를 관찰한 결과<sup>43)</sup>는 치안 AI의 이러한 수행적 측면이 문제가 있는 투표소를 예측하는 데 유사한 기술을 사용하고 무수히 많은 방식으로 확산할 수 있음을 시사한다.

---

38) Rodon, Toni, and Marc Guinjoan. 2022. "Beaten Ballots: Political Participation Dynamics Amidst Police Interventions." *Political Science Research and Methods* 10(1): 153-170.

39) Verma, Arvind. 2005.

40) Braga, Anthony, Rod Brunson, and Kevin Drakulich. 2019. "Race, Place, and Effective Policing." *Annual Review of Sociology* 45(1): 535-555.

41) Weisburd, David. 2016. "Does Hot Spots Policing Inevitably Lead to Unfair and Abusive Police Practices, or Can We Maximize Both Fairness and Effectiveness in the New Proactive Policing?." *University of Chicago Legal Forum* 2016(16): 661-689.

42) McQuillan, Dan. 2022. *Resisting AI: An Anti-fascist Approach to Artificial Intelligence*. Bristol: Bristol University Press.

43) Ensign, Danielle, et al. 2018. "Runaway Feedback Loops in Predictive Policing." In *Proceedings of the Machine Learning Research, 1st Conference on Fairness, Accountability and Transparency* 81: 160-171.

○ 투표 부스 보호를 위한 AI 활용에 대한 방향성 제시

- 문제가 있는 투표 부스를 예측하는 데 AI를 합리적으로 사용할 수 있는 방법은 생각하기 어렵다.
  - ▷ 첫째, 투표소 캡처의 미묘한 특성으로 인해 수학적 추상화 및 통계 모델을 사용하여 예측하기 어렵다는 점이다.
  - ▷ 둘째, 인공지능이 완벽할 것이라고 기대할 수 없다는 점을 고려할 때, 인공지능이 문제 투표 부스를 선택하는 과정에서 오류가 발생하면 선거 결과에 불필요한 간접적인 영향을 미칠 수 있다.
  
- 투표소 보호를 위한 경찰력 배치 우선순위를 정하기 위해 AI를 사용하는 경우, 우선순위 결정 과정에서 대중의 신뢰와 믿음을 잃지 않도록 투명하고 원칙적인 방법으로 AI를 사용할 수 있는 방법을 강구할 필요가 있다.
  - ▷ 공간적 핫스팟 식별 작업은 통계 이론의 영역에서 데이터 마이닝 방법과 병행하여 연구되어 왔으며,<sup>44)</sup> 특히, 공간 스캔 통계<sup>45)</sup>는 통계적 원칙을 기반으로 한 방법들을 탄생시켰다.
    - 이러한 방법들은 엄격한 맞춤형 통계적 유의성 테스트를 사용하는 것을 원칙으로 하지만, 확장성과 효율성이 떨어져 핫스팟 치안 기법에서 활용이 제한적이다.
    - 그러나 이러한 방법의 통계적 엄밀성은 문제가 있는 부스를 예측하는 데 있어 오류의 범위를 줄여줄 수 있다.
  - ▷ P와 Sundaram(2023)<sup>46)</sup>의 새로운 연구는 핫스팟 탐지의 인구통계학적 공정성을 핫스팟 치안에 적용하고 있다.
    - 전반적인 아이디어는 문제가 있다고 판단되는 지역의 전체 인구가 인구통계학적으로 다양하고 더 넓은 인구를 대표할 수 있도록 하는 것으로 문제가 있는 투표소를 식별하는 기술을 개발할 때 투명성, 통계적 엄격성, 공정성 원칙을 혼합하려는 시도는 앞서 논의한 일부

44) P, Deepak. 2016. "Anomaly Detection for Data with Spatial Attributes." In Unsupervised Learning Algorithms, Cham: Springer International Publishing.

45) Kulldorff, Martin. 1997. "A Spatial Scan Statistic." Communications in Statistics - Theory and Methods 26(6): 1481-1496.

46) P, Deepak, and Sowmya Sundaram. 2023. "FiSH: Fair Spatial Hot Spots." Data Mining and Knowledge Discovery 37: 1374-1403.



위험을 완화할 수 있을 것이다.

## 라. 유권자 인증

○ 유권자 인증에서 발생할 수 있는 문제점과 AI 활용의 맥락

- 합법적인 유권자가 투표에 참여했는지 확인하는 것은 선거 절차의 무결성을 위해 가장 중요한 일이다.

▷ 전 세계 대부분의 선거 절차는 유권자가 투표소를 방문할 때 신분증을 지참하도록 하는 등 투표소에서 어떤 형태로든 인증을 요구한다.

- 대부분의 선행연구에서 유권자 사칭에 대한 증거는 거의 발견할 수 없으나,<sup>47)</sup> 유권자 사칭에 대한 비난은 종종 정당이 선거 과정과 결과에 대한 신뢰를 떨어뜨리고 언론의 관심을 끌기 위해 사용한다.<sup>48)</sup>

▷ 이러한 비난은 민주주의 기관이 유권자 인증에 대한 지속적인 관심을 통해 선거 과정에 대한 대중의 신뢰를 강화하도록 압력을 가하는 것이다.

· 온라인 투표의 맥락(예: 에스토니아)에서 유권자 인증을 위한 디지털 ID 사용은 강력하고 견고한 암호화 기술로 보호된다.

· 우편 투표의 유권자 인증은 유권자가 요청을 해야만 투표 패키지가 등록된 주소로 발송되기 때문에 복잡적이고 암묵적이다.

- 유권자 인증 부분에서는 투표소의 물리적 위치에서 AI 기반 유권자 인증 방법론을 사용하는 데 중점을 둔다.

---

47) James, Toby, and Alistair Clark. 2020. "Electoral Integrity, Voter Fraud and Voter ID in Polling Stations: Lessons from English Local Elections." *Policy Studies* 41(2-3): 190-209.; Ahlquist, John, Kenneth Mayer, and Simon Jackman. 2014. "Alien Abduction and Voter Impersonation in the 2012 U.S. General Election: Evidence from a Survey List Experiment." *Election Law Journal: Rules, Politics, and Policy* 13(4): 460-475.

48) Fogarty, Brian, et al. 2015. "News Attention to Voter Fraud in the 2008 and 2012 US Elections." *Research & Politics* 2(2): 1-8.

○ 유권자 인증을 위한 잠재적 AI 활용

- 유권자 인증에 AI를 사용한 사례는 매우 드물다. 그럼에도 주목할 만한 사례 중 하나는 2020년 인도 텔랑가나주 지방선거<sup>49)</sup>에서 얼굴 인식 기술을 시범 사용했다(<그림 2-7> 참조). 유권자 신원 확인을 위한 비디오 검증에 초점을 맞춘 학술적 프로토타입이다.<sup>50)</sup>

<그림 2-7> 인도 텔랑가나주 2020 지방선거에서 얼굴 인식 사용



출처: Simoes, Stanley and Muiris MacCarthaigh 2023.: 227.

- 그러나 이러한 드문 활용은 얼굴 인식,<sup>51)</sup> 지문 인식, 홍채 인식,<sup>52)</sup> 망막 스캔, 위조문서 감지<sup>53)</sup> 등의 기술을 이용한 수많은 개인 신원 확인 및 사기 탐지 메커니즘을 보여주는 AI 문헌과 매우 대조적이다.
- 신분증에 기반한 유권자 인증에서 생체인식 기반 인증으로의 전환을 기대하는 것이 합리적일 수 있다.

49) Allie, Feyaad. 2023. "Facial Recognition Technology and Voter Turnout." *The Journal of Politics* 85(1): 328-333.

50) Dunbar, J, et al. 2015. "Video Verification: An Alternative form of Identify Verification." *Procedia Manufacturing* 3: 4914-4920.

51) Li, Lixiang, et al. 2020. "A Review of Face Recognition Technology." *IEEE Access* 8: 139110-139120.

52) Mostofa, Moktari, et al. 2021. "Deep GAN-Based Cross-Spectral Cross-Resolution Iris Recognition." *IEEE Transactions on Biometrics, Behavior, and Identity Science* 3(4): 443-463.

53) Centeno, Albert. 2019. "Analysis of Background textures in Banknotes and Identity Documents for Counterfeit Detection." Ph.D. thesis, Universitat Autònoma de Barcelona.; Centeno, Albert, et al. 2019. "Recurrent Comparator with Attention Models to Detect Counterfeit Documents." In *Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR)*: 1332-1337.

- 핵심적인 측면은 이러한 기술을 인구 규모에 맞게 배포할 수 있는 새로운 데이터 인프라가 존재하는가이다.
- 대부분 국가는 유권자 등록의 하나로 사진을 수집하기 때문에 국가 차원의 시민 사진 데이터베이스를 보유하고 있다. 이러한 데이터만으로는 얼굴 인식을 구현하고 배포할 만하지 않을 수 있으며, 이러한 맥락에서 더 심층적인 형태의 국가 차원의 자료수집이 중요해진다.
- ▷ 인구 기준 세계 최대 민주주의 국가인 인도의 사례를 보면, 여러 서비스에 액세스하기 위한 게이트웨이로 Aadhar라는 생체인식 기반 데이터베이스가 등장했다.<sup>54)</sup> Aadhar는 모든 시민의 지문, 홍채 스캔, 사진을 기록하며, 전체 인구의 95%를 포괄한다. 지문 스캐너를 통해 공공 서비스에 액세스하며 Aadhar는 일상적으로 사용된다.
  - 아드하르는 유권자의 신분증과는 별개이지만, 최근 유권자 카드를 아드하르와 연결하려는 이니셔티브가 있어 선거에서 생체 인증을 사용할 수 있는 준비가 되어 있다.
  - 학계에서는 선거에서 Aadhar 기반 생체인증 및 디지털(예: 일회용 비밀번호) 기반 인증 메커니즘을 설계하는 데 많은 열정을 쏟고 있다.<sup>55)</sup>
- 최근에는 다른 생체 인식 기반 인구 데이터베이스가 등장했으며, 그 중 인도네시아의 e-ktip<sup>56)</sup> 시스템이 그 예이다. 이러한 사례는 유권자 인증을 위한 AI 사용이 가까운 미래에 널리 운영될 수 있음을 시사한다.
- 물론 서방에서는 주민등록증에 대한 상당한 저항이 있으며, 시위로 인해 대규모 생체인식 기반 신분증 프로젝트가 보류되거나 연기되었다는 점도 언급할 필요가 있다.<sup>57)</sup>

54) Rao, Ursula. and Vijayanka Nair. 2019. "Aadhaar: Governing with Biometrics." South Asia: Journal of South Asian Studies 42(3): 469-481.

55) Roopak, T. and R. Sumathi. 2020. "Electronic Voting based on Virtual ID of Aadhar using Blockchain Technology." In Proceedings of the 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA): 71-75.; Jain, Ankit, et al. 2023. "Blockchain-Based Secure E-voting System Using Aadhaar Authentication." In Predictive Data Security Using AI: Insights and Issues of Blockchain, IoT, and DevOps: 89-103.

56) Darwis, Kevindra., Charles Lim. 2011. "Design and Implementation of e-KTP (Indonesian Electronic Identity Card) Key Management System." In Proceedings of the 2011 International Conference on Advanced Computer Science and Information Systems: 143-146.

57) Ramakumar, R. 2010. "The Unique ID Project in India: A Skeptical Note." In Ethics and Policy of Biometrics, Berlin, Heidelberg: Springer.

○ 얼굴 인식을 위한 잠재적 AI 활용의 과급 효과

- 유권자 신원 확인을 위해 AI 기반 인증 메커니즘을 사용할 때 발생할 수 있는 네 가지 위험에 대해 언급한다.

▷ 첫째, 얼굴 분석 소프트웨어의 인증 및 성별 편향성 문제는 2018년에 많은 언론의 주목을 받았다.<sup>58)</sup>

- 흑인 여성과 같은 다인종 그룹은 최대 34%에 달하는 매우 높은 오류율을 보이지만, 밝은 피부의 남성은 99% 이상의 정확도를 기록했다.
- 이러한 점을 고려할 때, 상당수의 여성, 소수자, 다인종 집단이 AI 기반 인증 메커니즘을 통해 부당하게 입장이 거부되어 (수동 인증 절차로 되돌아가야 하는) 불편을 겪고 선거 과정에서 소외될 수 있다.

▷ 둘째, 얼굴 분석 외에도 지문, 홍채, 망막 기반 생체 인식 시스템은 행동 편향 가능성에 대한 실험적 분석이 매우 제한적으로 이루어졌다. 위의 기술 중 가장 널리 사용되는 지문 시스템은 어느 정도 인구통계학적 편향성을 보이는 것으로 관찰되었다.<sup>59)</sup>

- 소규모 조사에서 Aadhar 시스템에서 육체노동자들 사이에서 지문 기반 인증 오류가 높은 것으로 보고된 바 있는데,<sup>60)</sup> 이것이 우연한 결과인지 아니면 육체노동자의 경우처럼 지문이 오차범위 이상으로 희미해진 노동자에 대한 체계적인 편견인지 확인하기 위해서는 대규모 연구가 필요해 보인다.

▷ 셋째, 향상된 기술의 사용이 특히 소외된 지역사회의 투표율을 떨어뜨릴 위험이 있다.

- 2020년 인도 텔랑가나주에서 실시된 얼굴 인식 시범 사업에서 투표율이 6% 포인트 이상 통계적으로 유의미하게 감소한 것으로 나타났습니다.
- 이를 분석한 연구<sup>61)</sup>에서는 물류 문제, 사기 행위의 변화, 소외된 중

58) Buolamwini, Joy, Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In Proceedings of the Machine Learning Research, 1st Conference on Fairness, Accountability and Transparency. 81: 77-91.

59) Marasco, Emanuela. 2019. "Biases in Fingerprint Recognition Systems: Where Are We At?" In Proceedings of the 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS): 1-5.

60) Bhatti, Bharat. 2012. "Aadhaar-Enabled Payments for NREGA Workers." Economic and Political Weekly 47(49): 16-19.

교 공동체(이 경우 무슬림)의 소외감 증가 등 세 가지 잠재적 원인을 설명했다.

- 이러한 기술에 내재한 잠재력이 선거를 조작하는 국가기관의 힘을 강화할 수 있다는 점을 지적한다.

▷ 넷째, Gelb와 Diofasi(2019)<sup>62)</sup>가 지적한 바와 같이 AI 기반 생체 인증을 사용하려는 욕구에 내재된 모든 문제를 기술주의적 수단으로 해결할 수 있다는 믿음에 근거한 기술 해결주의에 대한 광범위한 우려가 있다.

▷ 생체인식 선거가 반드시 선거의 신뢰성과 공정성을 높이는 것은 아니며, 때에 따라서는 오히려 선거의 신뢰성과 공정성을 훼손할 수도 있다.

▷ 또한, 선거에 사용되는 생체인증을 최신 상태로 유지하기 위해 투입되는 강화된 리소스는 개인의 자유와 정치적 권리를 훼손하기 위한 감시를 강화하는 데 사용될 수 있다.<sup>63)</sup>

### ○ 얼굴 인식 AI 활용에 대한 방향성 제시

- 투표소 보호의 경우와 마찬가지로, 특히 생체인식 기반 인증의 경우 유권자 인증에 AI를 합리적으로 사용하기 어렵다.

- 그러나 유권자의 생체 인증에 AI를 사용할 때 발생할 수 있는 위험을 완화할 수 있는 가능한 경로에 대해 고민해 볼 수 있다.

▷ 첫째, 우리는 생체 인증이 유권자(일부)를 소외시키지 않도록 다른 병행 옵션을 도입하여 비기술적인 행정적 안전장치를 마련해야 한다. Bhatti<sup>64)</sup>가 권고한 것처럼, 생체인증이 다른 인증 옵션

61) Allie, Feyaad. 2023. "Facial Recognition Technology and Voter Turnout." The Journal of Politics 85(1): 328-333.

62) Gelb, Alan and Anna Diofasi. 2019. "Biometric Elections in Poor Countries: Wasteful or a Worthwhile Investment?" Review of Policy Research 36(3): 318-340.

63) Hosein, Gus and Carly Nyst. 2013. "Aiding Surveillance: An Exploration of How Development and Humanitarian Aid Initiatives are Enabling Surveillance in Developing Countries." Report, Privacy International.

을 밀어내고 사실상 강제적인 대안이 되어서는 안 된다.

- ▷ 둘째, 신중하게 선별된 벤치마킹(예: FRVT<sup>65</sup>)을 통해 생체인증에 대한 광범위한 감사를 시행하여 편향성을 밝혀내 이를 완화하기 위한 정책적, 기술적 노력을 유도할 수 있다.
- ▷ 셋째, 이러한 기술이 널리 배포되지 않았기 때문에 AI 기반 생체인증에 내재된 편견의 특성은 잘 알 수 없다. 따라서 포괄적인 인증 메커니즘을 개발하기 위해 상당한 주의를 기울여야 한다.
  - 설계에서 일반적인 사용자를 목표로 AI의 상업적으로 개발된 기술에 근거한다는 점에서 그 속에 내재된 편견이 작동해서는 안된다는 점에서 AI 개발은 종종 일반적인 사용자의 특성화를 포함하는 인기있는 제품 설계 도구인 페르소나<sup>66</sup>)를 사용한다.
  - 페르소나를 사용하면 상용 제품의 중요한 기준인 일반적인 사용자에게 기술이 잘 작동하는지 확인할 수 있지만, 선거 중심의 인증 AI에서는 소외되는 사람이 없도록 하는 것이 핵심 관심사이며, 이는 실패하기 쉬운 ‘엣지’ 사례를 식별하고 기술이 이들에게도 작동하도록 하는 데 초점을 맞춰 운영할 수도 있다.
  - 이러한 엣지 케이스 사고<sup>67</sup>)는 AI 기반 인증에서 더 높은 수준의 포용성과 공정성을 촉진하는 데 도움이 될 수 있다.

## 마. 영상 모니터링

### ○ 영상 모니터링에서 발생할 수 있는 문제점과 AI 활용의 맥락

- 선거가 올바르게 진행되는지 모니터링하는 것은 자유롭고 공정한 선거를 보장하기 위한 중요한 조치이다.
- 선거 모니터링을 통해 선거 후 선거 부정 혐의를 검증하거나 조사할 방법이 확보된다.<sup>68</sup>)

64) Bhatti, Bharat. 2012.

65) Grother, P., Mei Ngan, and Kayee Hanaoka. 2019. “Face Recognition Vendor Test Part 3: Demographic Effects.” Report, NIST.

66) Pruitt, John, and Tamara Adlin. 2006. The Persona Lifecycle: Keeping People in Mind Throughout Product Design. Amsterdam: Morgan Kaufmann.

67) Wachter-Boettcher, Sara. 2017. Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech. New York: W. W. Norton & Company.

68) Hyde, Susan. and Nikolay Marinov. 2014. “Information and Self-Enforcing Democracy: The Role of International Election Observation.” International Organization 68(2): 329-359.

- 가장 일반적인 모니터링은 독립적인 국제 읍서버에 의한 것이지만, 파키스탄의 자유롭고 공정한 선거 네트워크(FAFEN25)와 같이 국내 선거를 모니터링하는 기관도 존재한다.
- 선거 참관인의 존재 여부와 무관하게, 선거 기관은 감시 시스템에서 전 세계적으로 그 역할이 증가하고 있는 CCTV/감시 카메라와 같은 최신 감시 기술을 통해 선거 투명성을 높이려고 시도할 수 있다.
  - ▷ CCTV 시스템은 선거 사기를 억제하는 역할을 하면서도 인간 참관인에 의한 모니터링보다 더 효과적이라는 주장 또한 존재한다.<sup>69)</sup>
  - ▷ 참관인에 의한 선거 모니터링과 달리, CCTV 기반 모니터링은 데이터 중심의 고유한 특성을 고려할 때 선거 모니터링에서 AI가 수행할 수 있는 역할을 강화한다.
  - ▷ 실시간으로 이벤트를 식별할 수 있는 확장 가능한 AI의 출현으로 기관 보안,<sup>70)</sup> 가정 경보 시스템,<sup>71)</sup> 노인 건강 모니터링<sup>72)</sup>과 같은 다른 맥락에서 점점 더 많이 사용되고 있으며, 이는 비디오 기반 선거 모니터링 AI 사용의 가장 적절한 영역이 될 가능성을 열어준다.

#### ○ 영상 모니터링의 잠재적 AI 활용

- 러시아<sup>73)</sup>와 인도 등 전 세계적으로 투표소에 CCTV 기반 인프라가 많이 구축되었지만, 주로 선거의 투명성을 높이고 선거 부정을 방지하거나 시민이 선거를 감시하는 데 사용되어 왔다. 실시간 또는 선거 후 분석을 위해 CCTV 데이터 스트림에서 AI가 어떻게 사용되었는지에 대한 공개 정보는 매우 제한적이다.

69) Obeta, Sebastine, et al. 2021. "Credible Elections in Nigeria: The Role of Information Communication Technology." *The International Journal of Science & Technoledge* 9(10): 35-39.

70) Kakadiya, Rutvik, et al. 2019. "AI Based Automatic Robbery/Theft Detection using Smart Surveillance in Banks." In *Proceedings of the 2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*: 201-204.

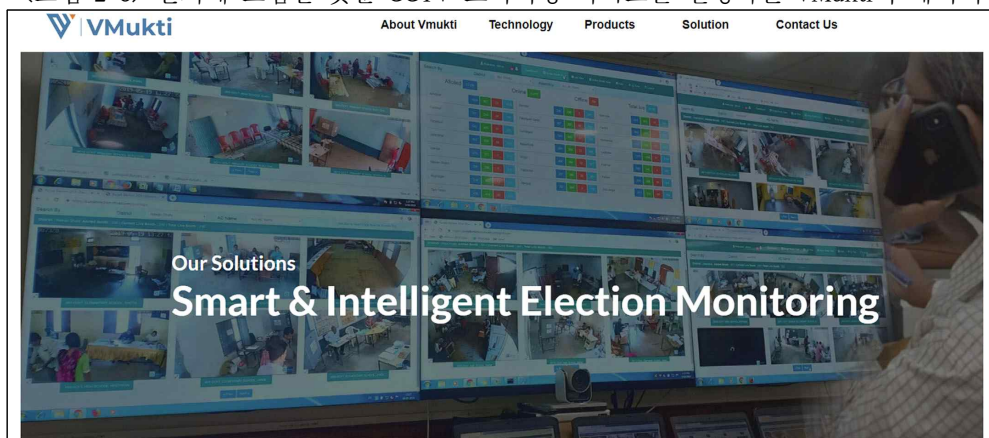
71) Liang, Chua, et al. 2021. "Smart Home Security System Based on Zigbee." In *Advances in Smart System Technologies*. Singapore: Springer.

72) Huang, Zhanyuan, et al. 2018. "Video-Based Fall Detection for Seniors with Human Pose Estimation." In *Proceedings of the 2018 4th International Conference on Universal Village (UV)*: 1-4.

73) Asmolov, Gregory. 2014. "The Kremlin's Cameras and Virtual Potemkin Villages: ICT and the Construction of Statehood." In *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. Oxford: Oxford University Press.

- 인도 비하르 주 선거에서 CCTV의 데이터를 분석하여 개표 수를 확인하기 위해 AI와 비디오 분석을 사용한 사례가 있다.<sup>74)</sup>
  - ▷ 실제 선거 후 개표가 진행되기 훨씬 전에 전자 투표기의 비디오 캡처를 수행하고 OCR 기술을 사용하여 분석하여 기술 기반 병렬 개표를 수행했다는 비하르 주 투표 과정에서의 AI 활용에 대해 자세한 설명이 있다.<sup>75)</sup>
  - ▷ 또한 인공지능 기반 분석 기술을 실시간으로 활용하여 개표 결과의 불일치를 파악하여 투표 관리자에게 경고한다.
  - ▷ 코드명 JARVIS27인 이 독점 기술은 투표 분석만을 위한 맞춤형 기술 스택이 아니라 다양한 애플리케이션에 사용되는 기술이다.
  - ▷ 이는 기존의 AI 기반 비디오 분석 기술이 투표소에서 어떻게 활용될 수 있는지를 보여주는 사례로, 잠재적인 AI 활용에 대한 논의에 초점을 맞추고 있다.
  
- 또한 <그림 2-8> VMukti28과 같이 선거를 겨냥한 비디오 분석 솔루션이 새롭게 보급되고 있다는 점도 주목할 만하다.

<그림 2-8> 선거에 초점을 맞춘 CCTV 모니터링 서비스를 설명하는 VMukti의 페이지



출처: Stanley Simoes and Muiris MacCarthaigh 2023: 230.

74) Gupta, Arvind. 2021. "Indian Elections: Scale, Use of Technology and SOTA Outcomes." In Invited talk 1 at Workshop on AI for Credible Elections: A Call to Action.

75) Ahaskar, Abhijit. 2021. "Officials Use AI Analytics to Keep Bihar Panchayat Polls Free and Fair." Story, Mint. (October 29, 2021) <https://www.livemint.com/technology/tech-news/officials-use-ai-analytics-to-keepbihar-panchayat-polls-free-and-fair-11635446470554.html> (검색일: 2023. 07. 25).



- AI 기반 비디오 분석 기술은 공공 안전 분야에서 가장 발달되어 있으며, 주로 공공장소에 설치된 감시 카메라의 데이터 스트림을 통해 사용하기 위해 개발되었다.
- 2016년에 영국 주민 한 명이 매일 평균 70대의 카메라에 찍힌다는 보고가 있었는데, 이는 현대 사회에서 비디오 모니터링이 얼마나 널리 사용되고 있는지를 보여준다.
- 비디오 기반 선거 모니터링의 맥락에서 잠재적 활용에 우선순위를 둔 몇 가지 대표적인 사례가 있다.
  - ▷ 공공 안전의 주요 활용 사례 중 하나는 흥미로운 이벤트와 각 이벤트의 내용, 시간, 장소를 식별하는 것이다.<sup>76)</sup> 폭력<sup>77)</sup>과 같은 특정 유형의 사건은 투표소의 비디오 스트림을 통해 직접적으로 유용하게 사용될 수 있다.
  - ▷ 비디오에서 이상 징후를 탐지하기 위한 더욱 일반적인 AI<sup>78)</sup>는 투표소에서 그대로 활용하여 비디오 스트림을 필터링하여 다운 스트림 수동 분석에 활용할 수 있다.
  - ▷ 공공 안전을 위해 개발된 지도 머신러닝 모델은 투표소의 통제되고 예측 가능한 환경을 고려할 때 투표소 설정에 훨씬 더 높은 정확도를 제공할 수 있을 것이다.

---

76) Zhang, Qingyang, et al. 2019. "Edge Video Analytics for Public Safety: A Review." Proceedings of the IEEE 107(8): 1675-1696.

77) Ramzan, Muhammad, et al. 2019. "A Review on State-of-the-Art Violence Detection Techniques." IEEE Access 7: 107560-107575.

78) Sultani, Waqas, Chen Chen, and Mubarak Shah. 2018. "Real-World Anomaly Detection in Surveillance Videos." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 6479-6488.

<그림 2-9> 두 대의 카메라로 이루어지는 개인 재식별:  
(싱가포르국립대, NTU) 래피드 리치 객체 검색 연구실



출처: Simoes, Stanley and Muiris MacCarthaigh. 2023. 230.

▷ 최근에는 <그림 2-9>에서 볼 수 있듯이 여러 비디오 스트림에서 동일한 사람을 추적하는 작업(예: 복수의 감시 카메라에 포착된 범죄자의 도시 내 이동 경로 추적)인 개인 재식별에 관한 관심이 급증하고 있다.<sup>79)</sup>

· 개인 재식별은 동일인이 여러 번 투표하여 동등한 대표성 원칙을 위반하는 선거 사기로 잘 알려진 중복 투표에 대한 기술적 솔루션을 개발하는 데 적합하다.

### ○ 영상 모니터링의 잠재적 AI 활용의 파급 효과

- AI 기반 선거 비디오 모니터링의 활용은 다양한 결과를 가져올 수 있다.

▷ 세밀한 선거 모니터링과 후보자별 개표를 병행할 수 있도록 OCR과 비디오 프로세싱을 통합하는 JARVIS와 같은 기술을 사용하는 것은 선거 완결성과 관련된 심각한 문제를 제기할 수 있다.

▷ 유권자의 신원도 파악된다는 점을 고려할 때, 이는 유엔 인권선언(제21.3조)에 명시된 투표 시스템의 핵심 원칙인 비밀 투표 명백하게 침해하는 경로로 간주할 수 있다.

▷ 선거에 대한 세밀한 비디오 모니터링의 근본적인 위험 중 하나는 비밀 선거권 및 개표의 투명성과 같은 선거의 공정성 원칙에

79) Ye, Mang, et al. 2022. "Deep Learning for Person Re-Identification: A Survey and Outlook." IEEE Transactions on Pattern Analysis and Machine Intelligence 44(6): 2872-2893.

대한 위협이다.

- ▷ 또한 감시 기술로 인한 유권자 집단의 소외감이 심화될 수 있다.
- ▷ 반면에 주요 이점은 선거 폭력과 조작을 억제하는 동시에 수동으로 주의를 기울여야 하는 비정상적인 활동을 효과적으로 탐지하는 것이다.
- 이러한 맥락을 고려할 때, 기술적 선거 감시를 도입해야 하는 경우 이러한 솔루션을 추구해야 한다면 선거의 공정성과 양립할 수 있으면서도 선거 폭력 및 기타 비정상적인 활동을 탐지하는 이점을 실현할 수 있는 경로를 모색해야 한다.
- ▷ 선거 완결성 원칙을 준수하기 위한 한 가지 방법은 본격적인 비디오 모니터링 대신 가벼운 모니터링만 도입하는 것이다.
  - 이 방법에는 오디오 모니터링, 세밀한 동작 감지 센서 또는 매우 낮은 해상도의 카메라(흐릿한 이미지만 캡처할 수 있어 움직임은 식별할 수 있지만 개인이나 신원은 식별할 수 없음)를 사용하거나 위의 일부 또는 전부를 포함하는 감시 설정이 포함될 수 있다.
  - 이러한 가벼운 모니터링은 흐릿한 비디오 스트림에서는 작동하지 않아 선거의 완결성에 대한 위협을 완화할 수 없으므로 OCR(리콜렉트 서비스) 또는 얼굴 인식과 같은 기술을 솔루션에 패키징할 수 없다.
- ▷ 반면에 이러한 스트림에서 효과적인 폭력 및 비정상적인 이벤트 감지를 위해서는 맞춤형 컴퓨터 비전 및 패턴인식 기술이 개발되어야 하며, 이는 AI 연구의 새로운 방향을 제시하게 될 것이다.
- ▷ 여러 감지 방식(예: 오디오, 모션, 흐릿한 비디오)이 관련된 경우, 시스템의 이점을 합리적으로 실현하기 위해 이러한 멀티모달 스트림에서 증거를 상호 연관시키는 기술의 개발이 필요하다.
- 이러한 방향을 가능하게 하거나 시작하는 방법은 다음과 같다.
  - ▷ 첫째, 입법 또는 규제, 즉 현재 비디오 모니터링이 불법이 아닌 관할권에서 본격적인 비디오 모니터링의 사용을 허용하는 법적 규정을 마련해야 한다.
  - ▷ 둘째, 가벼운 모니터링에는 이러한 모니터링이 개인을 식별할 수 없다는 사실을 시민들이 이해하고 인정하도록 해야 하는 과제가 있다. 이러한 신뢰 문제는 비기술적인 문제이기는 하지만

시민 그룹과 소외된 커뮤니티가 소외될 위험을 완화하기 위해서는 만족스럽게 해결해야 한다.

▷ 마지막 해결책은 오픈 데이터 이니셔티브를 통해 모든 모니터링 데이터를 (잠재적으로 실시간) 웹 스트림을 통해 공개하여 모니터링이 얼마나 가벼운지를 보여주는 것이다.

- 이는 핵티비즘 기반의 실험을 장려하여 문제를 부각시키고, 가벼운 모니터링 솔루션에 대한 지속적인 감시와 피드백을 기반으로 개선할 수 있는 길을 열어줄 수 있다.
- 오픈 데이터는 또한 선거에 대한 시민 감사를 장려할 수 있다. 러시아에서 시민이 CCTV 영상을 분석하여 투표용지 조작 사례를 밝혀낸 주목할 만한 사례가 있다.<sup>80)</sup>

#### ○ 선거 과정에서 인공지능 사용의 가시성

- 유권자와 일반 대중에게 나타나는 인공지능 사용의 가시성은 예를 들어, 기술 유권자 인증(예: 얼굴 인식 사용)의 사용은 명시적으로 볼 수 있는 반면, 유권자 명부 정리에 인공지능을 사용하는 것은 처리 문서를 공개 도메인에서 사용할 수 있을 때만 드러난다.
- 가시성은 공공의 감시를 보장하므로 가시성이 낮은 AI 사용은 더 위험할 수 있으므로 가시성은 중요한 고려 사항이다.
  - ▷ 다시 말해, 가시성이 매우 낮은 중간 위험도의 AI 사용은 고위험도의 높은 가시성의 AI 사용보다 장기적으로 더 나쁜 결과를 초래할 수 있다.
- (현존하거나 잠재적인) AI 사용의 가시성을 분석하는 관점에서 5가지 분야를 간략하게 검토한다.
  - ▷ 유권자 명부 유지 관리는 당연히 선거 공무원이 선거 사이에 수행하는 백 오피스 작업으로, 선거를 앞두고 선거 준비를 위해 일반적으로 활동이 급증한다. 그러나 이 업무에서 AI를 사용하는 것은 대중에게 직접적으로 드러나지 않는다.

80) HUFFPOST. 2012. "Russia Election 2012: Ballot Box Stuffing Purportedly Caught On Tape (VIDEO)." (March 5) [https://www.huffpost.com/entry/russia-election-2012-ballot-box-stuffing\\_n\\_1321379](https://www.huffpost.com/entry/russia-election-2012-ballot-box-stuffing_n_1321379) (검색일: 2023. 08. 25).

- 사후 설명 가능한 AI의 방대한 문헌을 고려할 때, 확인을 요청하는 유권자에게 이유를 제공해야 하는 규제 제약으로 인해 AI 사용이 노출되지 않을 수 있다.
  - LIME<sup>81)</sup>과 같은 사후 설명 가능한 AI 방법은 모든 결정(이 맥락에서는 분류자가 특정 유권자의 확인을 요청하기로 한 결정)에 대해 자동화된 설명을 제공할 수 있으므로, AI 사용을 노출하지 않고도 설명에 대한 규제 제약을 쉽게 확인할 수 있다.
  - 의사결정에 AI를 사용하는 것이 투명해야 한다는 규정이 없는 한(또는 다음과 같은 경우에 해당하지 않는 한) 정보에 대한 권리 규정을 고려할 때, 유권자 명부 관리에서 AI 사용의 가시성은 매우 제한적일 수 있다.
- 유권자 명부 관리의 경우처럼 투표소 위치를 결정하고 문제 부스를 식별하는데 AI를 사용하는 것도 매우 제한된 가시성 하에서 이루어질 수 있다.
  - 투표소 결정과 문제 부스 식별에는 훨씬 적은 수의 의사결정 사례(투표소 수 순서)가 필요하고 투표소 물류 배치 또는 추가 순찰 배치에 추가 수작업 검사가 필요하다는 점을 고려하면 의사결정 파이프라인에서 상당한 인적 요소가 포함될 가능성이 높다. 따라서 이 과정에 AI를 투입하는 것은 상당히 불투명하다.
  - 위에서 설명한 방법과는 달리, 유권자 인증은 AI 사용이 필연적으로 높은 가시성을 가질 수 있는 분야이다.
    - ▷ 유권자 인증을 위해 생체인식을 사용하려면 투표 관리자가 장비를 조달하고 유권자에게 이를 효과적으로 배포할 수 있도록 알려야 한다.
    - ▷ 안면인식의 경우, 대부분의 현재 알고리즘은 안정적인 사진 캡처가 필요하다. 얼굴 인식 기반 유권자 인증에 CCTV 영상을 사용할 수 있을 정도로 기술이 발전하면 AI 사용의 가시성이 줄어들 수 있지만, 현재로서는 유권자 인증 업무에 AI를 사용하는 것은 유권자에게 높은 가시성을 수반한다.
  - 영상 모니터링은 카메라와 연결 인프라를 활용하기 때문에 관찰력이 있는 유권자가 AI의 사용을 쉽게 발견할 수 있어 가시성이 높은 AI 사용 사례로 꼽을 수 있다.
    - 눈에 보이지 않는 소형 스파이 카메라의 출현은 위험을 초래할 수

81) Ribeiro, Marco, S. Sameer Singh, Carlos Guestrin. 2016. "Why Should I Trust You?: Explaining the Predictions of Any Classifier." In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 1135-1144.

있지만, 선거 기관이 가시성을 줄이기 위해 이러한 ‘트릭’을 사용하는 데 열중할 가능성은 낮다.

- 지금까지의 논의를 요약하여 간략한 개요를 제공한다. <표 2-3>은 선거 과정에서 활용할 수 있는 AI의 기술 준비도, 위험 수준 및 AI 사용의 가시성 측면에서 고려되는 5가지 분야별 각각에 대한 예비 평가의 개요를 나타낸다. 이는 객관적인 평가는 아니지만 현재 상황에 대한 정보 평가를 바탕으로 한 것이므로 비판적 접근이 필요하다.

<표 2-3> 핵심 평가에 대한 전반적인 요약

분야(Avenue)	기술 준비성 (Technology readiness)	위험 수준 (Risk level)	AI 활용 가시성 (Visibility of AI usage)
유권자 명부 관리	높음	중간	낮음
투표소 위치	중간	중간	매우 낮음
문제 부스 예측	높음	높음	매우 낮음
유권자 인증	매우 높음	높음	매우 높음
영상 모니터링	매우 높음	매우 높음	높음

출처: Simoes, Stanley and Muiris MacCarthaigh. 2023. p. 232.

### ○ 평가 내용

- 투표소 위치 결정과 관련된 AI 기술은 여러 스트림에서 사용할 수 있지만, 이러한 기술을 작업에 사용하기 위해서는 약간의 조정이 필요할 것으로 판단하여 기술 준비도를 중간으로 평가했다.
- 반면, 핫스팟 치안은 문제 투표소를 식별하는 작업에 어느 정도 쉽게 사용할 수 있으므로 기술 준비도를 높음을 평가했다.
- 유권자 명부관리 및 투표 부스 결정에 AI를 사용하는 위험 수준은 어느 정도 전문가의 감독을 보장하는 human-in-the-loop 방식으로만 사용하는 것이 사실상 불가피하므로 중간으로 평가한다.
- 반면, 비디오 모니터링의 사용은 선거의 공정성 원칙과 침해한 긴장 관계에 놓일 수 있다는 점에서 매우 높은 위험도로 평가한다.

### 3. 주요국의 생성형 AI의 선거 활용 문제 분석

#### 3.1 미국

##### 가. 생성형 AI의 선거 위해 및 방해 사례

- 미국에서는 2019년 행정명령 제13859호에서 “미국 인공지능 구상”을 공표한 이후, AI 관련 규제 입법을 추진해 왔으나, 「미국인공지능진흥법」 등의 지원책만 결의되었을 뿐 생성형 AI 기술을 규제하는 안을 의결하지 못하고 있다. 다만 미국 정부는 정치 광고에 AI를 사용했으면 출처를 표기하는 법안을 발의하였을 뿐이다. 미 정부의 AI 관련 입법 추진에도 불구하고, 생성형 AI 기술 규제 영역에 대한 준비가 미진한 상황인 것이다. 2024년 대선을 앞둔 미국 내, AI 활용 딥페이크 콘텐츠 문제 주요 사례는 아래와 같다.
- 2023년 2월 조 바이든 대통령이 트랜스젠더 혐오 발언하는 가짜 영상 유포

<그림 3-1> 조 바이든 대통령이 트랜스젠더 혐오 발언하는 가짜 영상



출처: Hudnall, Hannah. 2023. “Fact check: Video altered to show Joe Biden making transphobic remarks.” USA TODAY (February 9) <https://www.usatoday.com/story/news/factcheck/2023/02/09/fact-check-video-edited-show-joe-biden-making-transphobic-remarks/11211453002/> (검색일: 2023. 11. 01).

- 2023년 2월 3일 인스타그램에 게시된 영상에는 조 바이든 대통령이 트랜스젠더 여성을 모욕하는 음성으로 보이는 연설 모습이 담겨 있다. 내용은 동성애자를 차별하고 조롱하는 메시지이다. 해당 영상은 5일 만에 ‘좋아요’ 12,000개 이상을 기록했으며, 인스타그램에도 비슷한 게시물이 공유됐다.<sup>82)</sup> 나아가 다양한 언론매체를 통해 해당 영상은 확산되었다. 해당 영상이 마치 진짜인 것처럼 소셜미디어에 공유되었으나, 그런 발언을 했다는 증거는 없으며 전문가들은 로이터(Reuters)에 해당 영상이 딥페이크인 것으로 보인다고 전하였다.<sup>83)</sup>
- 실제 해당 영상은 AI가 생성한 음성과 미국의 우크라이나 탱크 파견을 알리는 2023년 1월 25일 CNN 생방송에서 가져온 대통령의 실제 영상을 결합한 것으로, 영상에서는 바이든의 입이 오디오와 일치하도록 조작됐다.<sup>84)</sup> 대부분의 미국 시민들은 바이든 대통령이 해당 내용을 말할 가능성이 없다는 점을 인식했지만, 콘텐츠 영상이 얼마나 현실적으로 나타나는지를 인식하며 사회적 충격을 일으킨 사례라 하겠다.

○ 2023년 3월 도널드 트럼프 전 대통령이 체포되어 경찰에 연행되는 가짜 영상 유포

<그림 3-2> 도널드 트럼프 전 대통령이 체포되는 가짜 영상



출처: Devlin, Kayleen and Joshua Cheetham. 2023. “Fake Trump arrest photos: How to spot an AI-generated image.” BBC. (March 24) <https://www.bbc.com/news/world-us-canada-65069316> (검색일: 2023. 11. 01).

- 2023년 3월 해당 영상에는 트럼프 전 대통령이 체포에 순순히 응하지

82) Hannah Hudnall. 2023.

83) Reuters. 2023a. “Video does not show Joe Biden making transphobic remarks.” (February 11) <https://www.reuters.com/article/factcheck-biden-transphobic-remarks/fact-check-video-does-not-show-joe-biden-making-transphobic-remarks-idUSL1N34Q1IW/> (검색일: 2023. 11. 01).

84) Lajika, Arijeta. 2023. “Artificial intelligence makes voice cloning easy and ‘the monster is already on the loose’ .” Fortune (February 12) <https://fortune.com/2023/02/11/artificial-intelligence-make-s-voice-cloning-easy-and-the-monster-is-already-on-the-loose/> (검색일: 2023. 11. 01).



않고 도망가거나, 경찰에 완력으로 제압되는 모습이 담겨 있다. 하지만 해당 시기의 동영상은 비전문가라도 자세히 살펴보았을 때, 영상의 일부가 뭉개져 있거나, 오류가 발견되는 등 가짜 영상임을 식별할 수 있었다.

- 해당 영상은 오픈소스 조사 매체 벨링캣(Bellingcat)의 창립자인 엘리엇 히긴스(Eliot Higgins)가 2023년 3월 경 트럼프 전 대통령의 검찰 기소가 예상되는 상황에서 해당 사실을 시각화하기로 한 데에서 비롯되었다. 그는 단순히 생성형 AI에 “도널드 트럼프가 체포되는 동안 쓰러졌다”와 같은 간단한 프롬프트를 제시함으로써 해당 영상을 제작했다고 설명했다. 히긴스는 인터뷰에서 “나는 단지 농담을 하고 있었다. 아마 5명 정도가 리트윗할 거라고 생각했다.”라고 언급했으나, 이를 후, 해당 게시물은 거의 500만 번 조회되었으며, AI 생성 이미지의 정교함 증가, 배포 용이성 및 혼란을 야기할 수 있는 잠재력에 대한 연구 사례 대상이 되었다.<sup>85)</sup>
- 해당 영상 제작에 대해 창작자인 히긴스는 회사 내에서 제재를 받은 것으로 알려졌으나, 뚜렷한 법적 위반 조치의 대상이 아니다. 히긴스는 해당 영상 제작 시 가짜뉴스(Fake News) 혹은 AI 제작 영상임을 알리는 표식을 하지 않았다. 이후 언론이 해당 영상의 이미지를 재사용할 때, 혼란 방지를 위해 가짜뉴스(Fake News)의 표식이 첨부되었다.

○ 2023년 4월 공화당, 중국 전투기가 대만 폭격하는 가상 영상 발표

- 2023년 4월 25일 공화당은 AI가 생성한 중국의 대만 공격 이미지와 경제, 사회적 재난 장면의 영상 광고를 공개하며 차기 대선에서 조 바이든 대통령에게 투표하지 말 것을 경고했다. 뉴스 음성 해설에서는 “오늘 아침 대담해진 중국이 대만을 침공했습니다.”라는 멘트와 함께 폐쇄된 은행, 미국 남부 국경으로 몰려드는 이민자 집단 행렬, 범죄 급증과 펜타닐 위기로 인해 폐쇄된 샌프란시스코가 등장한다. Forbes는 AI 기술로만 생성된 이미지를 최초로 사용한 광고로서 “이 광고는 정치 캠페인의 새로운 이정표가 될 수 있다”고 언급했다.<sup>86)</sup>

85) The Washinton Post. 2023. “Fake images of Trump arrest show ‘giant step’ for AI’s disruptive power.” (March 22) <https://www.washingtonpost.com/politics/2023/03/22/trump-arrest-deepfakes/> (검색일: 2023. 11. 01).

86) Charlier, Phillip. 2023. “Republican Party Biden attack ad features attack on Taiwan.” Taiwan E

<그림 3-3> ‘Beat Biden’ 이라는 제목의 영상 내 중국이 대만을 공격하는  
가짜 영상



출처: Charlier, Phillip. 2023.

- 해당 영상은 공화당 전국위원회(Republican National Committee)가 제작한 것으로, 바이든 대통령이 재선 출마할 것이라 발표한 직후, 유포되었다. 동 위원회는 바이든이 4년 더 집권하면 국가에 좋지 않을 것임을 동영상 통해 알리기 위해 해당 영상을 제작하였다고 기획 의도를 알렸다.
- 해당 영상에는 왼쪽 상단에 ‘완전히 AI 이미지로 제작됐다(built entirely with AI imagery.)’는 문구와 함께 워터마크가 포함되어 있어, 법적 제재의 대상이 아니다. 공화당 전국위원회는 해당 영상이 어떤 기술로 생성되었는지에 관한 언론의 질문에 답하지 않았다. 따라서 향후 이와 같은 형태의 정치적 광고는 더욱 보편화할 것으로 판단된다.
- UC Berkley의 디지털 포렌식 전문가이자 교수인 Hany Farid는 CNN과의 인터뷰를 통해, “이 광고가 명시적으로 기만적인 것은 아니라고 주장할 수 있지만, 생성형 AI가 정치적 반대자를 공격하는 데 어떻게 사용될 수 있는지 보는 것은 어렵지 않다.” 라고 지적했다.<sup>87)</sup>
- 민주당 전국위원회(DNC)는 성명을 통해 “공화당은 바이든 대통령의 결과에 반박할 수 없기 때문에 이미지를 만들어내야 했다” 라며, 해

English News. (April 26) <https://taiwanenglishnews.com/republican-party-biden-attack-ad-features-attack-on-taiwan/> (검색일: 2023. 11. 01).

87) CNN. 2023. “AI-generated GOP campaign video may be harbinger of disturbing new trend for 2024, expert says.” (April 25) <https://edition.cnn.com/politics/live-news/biden-2024-presidential-election/index.html> (검색일: 2023. 11. 01).

당 광고에 대해 평가절하했으나, 법이나 제도를 통해 이를 규제해야 한다는 의견은 나오지 않고 있다.

○ 2023년 6월 론 디샌티스(Ron DeSantis) 플로리다 주지사, 트럼프가 앤서니 파우치 전 소장과 포옹하는 허위 사진 발표

- 2023년 6월 디샌티스 플로리다 주지사는 트럼프 전 대통령이 앤서니 파우치 박사를 껴안고 있는 모습의 생성형 AI 가짜 영상을 소셜미디어에 공개했다. 주지사가 해당 영상에 대해 주지사 캠프는 도널드 트럼프를 공격하는 데 사용한 이미지를 트럼프 대통령과 그의 측근들이 유포한 밈(meme)과 비교하며, 제작 의도를 설명했다.
- 플로리다 주지사의 캠페인 팀이 게시한 해당 이미지는 2024년 공화당 백악관 경쟁자들이 어떻게 정치적 언쟁을 AI 기반 소셜미디어 분야로 끌어올려 사실과 허구를 혼합하는지 보여주는 사례로 언급되고 있다. 해당 영상에는 기자회견과 인터뷰에서 트럼프가 실제 모습을 드러내는 모습이 포함돼 있다. 그러나 25초가 지나면서 트럼프와 파우치가 포옹하거나 키스하는 장면 3장을 포함해 6장의 이미지가 등장한다. 주지사의 캠페인 팀은 해당 영상이 가짜인지 또는 AI가 이미지를 만드는 데 사용되었는지에 대해 응답하지 않았다고 한다.<sup>88)</sup>
- 해당 영상 확산과 관련한 혼란을 통해, 미국 내에서 생성형 AI를 활용하여 설득력 있는 딥페이크를 저렴하고 쉽게 만들 수 있게 되면서, 동 기술이 2024년 대선에 어떠한 영향력을 미칠지에 관해 관심이 집중되고 있다.

---

88) Ulmer, Alexandra., Tong, Anna. 2023. "With apparently fake photos, DeSantis raises AI ante". Reuters (June 9) <https://www.reuters.com/world/us/is-trump-kissing-fauci-with-apparently-fake-photos-desantis-raises-ai-ante-2023-06-08/> (검색일: 2023. 11. 01).

〈그림 3-4〉 트럼프가 앤서니 파우치 전 소장과 포옹하는 허위 사진



왼쪽부터 론 드샌티스, 도널드 트럼프, 앤서니 파우치



출처: Contorno, Steve and Donie O’Sullivan. 2023. “DeSantis campaign posts fake images of Trump hugging Fauci in social media video.” CNN (June 8) <https://edition.cnn.com/2023/06/08/politics/desantis-campaign-video-fake-ai-image/index.html> (검색일: 2023. 11. 01).; Ulmer, Alexandra., Tong, Anna. 2023.

## 나. 생성형 AI에 의한 안보 위협 사례

### ○ 2023년 5월 미 국방부 청사 폭발 가짜 이미지 유포

- 2023년 5월 22일 Bloomberg Feed라는 검증된 트위터 계정에서 ‘워싱턴 DC 국방부 건물 근처에서 대규모 폭발이 발생: 초기 보도(Large explosion near the Pentagon complex in Washington DC.: initial report)’ 라는 문구 아래, 국방부와 유사한 건물 옆에 거대한 검은 연기 기둥이 솟아오르는 이미지가 표시됐다. 면밀히 관찰하면 해당 이미지는 인공지능에 의해 생성된 것으로 추정되는 가짜임을 알 수 있으나, 당시 다우존스 주식지수가 4분 만에 85포인트 하락했다가 빠르게

게 반등하는 해프닝을 겪었다.<sup>89)</sup>

<그림 3-5> 미 국방부 건물 폭발에 관한 가짜 이미지 트윗



출처: Oremus, Will, Drew Harwell, and Teo Armus. 2023.

- 사회 분석 회사인 Cyabra는 온라인 대화를 분석한 결과 약 3,785개의 계정에서 허위 사실을 언급했으며, 당시 사태로 미국 내에서 AI 제작 가짜 영상에 대한 심각성이 사회적으로 확산되었다.<sup>90)</sup>
- 해당 이미지는 최초 유포 시, <그림 3-5> 내 우측 사진과 같이 실제 이미지를 촬영한 것과 같이 전송되었으나 이후 언론보도에서 혼란 방지를 위해 좌측의 사진과 같이 붉은 엑스표를 첨부하였다.
- 해당 이미지에 관한 문제는 단순히 생성형 AI에 의한 이미지 창작이 아니라, 인증된 트위터(twitter) 계정에 의해 유포되었다는 점에 집중되었다. 최초 유포 트위터는 계정 정지되었으며, 트위터는 논평 요청에 응하지 않았다. 폭발에 대한 허위 보도는 인도의 주요 텔레비전 네트워크(리퍼블릭 TV)에서 송출되었으며, 러시아 통신 매체 RT의 보도를

89) Oremus, Will, Drew Harwell and Teo Armus. 2023. "A tweet about a Pentagon explosion was fake. It still went viral." The Washington Post (May 22) <https://www.washingtonpost.com/technology/2023/05/22/pentagon-explosion-ai-image-hoax/> (검색일: 2023. 11. 01).

90) Saliba, Emmanuelle. 2023. "How verified accounts helped make fake images of a Pentagon explosion go viral." abc News (May 24) <https://abcnews.go.com/US/verified-accounts-helped-make-fake-images-pentagon-explosion/story?id=99541361> (검색일: 2023. 11. 01).

인용해 가짜 영상을 방송하며 폭발이 일어났다고 보도했다.<sup>91)</sup>

- 2022년 3월 블로디미르 젤렌스키 우크라이나 대통령의 항복 선언 허위 영상

<그림 3-6> 젤렌스키 우크라이나 대통령의 항복 선언 허위 영상

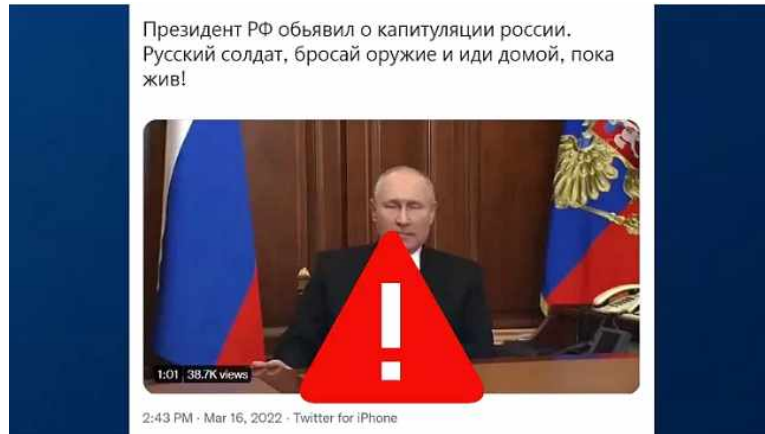


출처: The Telegraph. 2022. “Deepfake video of Volodymyr Zelensky surrendering surfaces on social media.” Youtube (March 17) <https://www.youtube.com/watch?v=X17yrEV5sl4> (검색일: 2023. 11. 01).

- 생성형 AI에 의해 제작된 젤렌스키 대통령의 항복 영상은 우크라이나-러시아 전쟁에서 의도적으로 사용된 최초 영상이다. 영상의 내용은 젤렌스키 대통령이 시민들에게 러시아에 항복하라고 촉구하는 내용과 시민들에게 ‘무기를 내려놓도록’ 독려하는 대(對) 국민 연설로 구성되었다. 해당 영상은 트위터에서 120,000회 이상 조회되며 화제가 되었으나, 정교한 딥페이크 영상이 아니기 때문에 실질적인 효과는 없었다.
- 관련 영상이 유포된 플랫폼 기업인 메타(Meta)의 보안 정책 책임자는 성명을 통해 젤렌스키 대통령의 허위 영상을 삭제했다고 확인했다.

91) Sullivan, Donie, and Jon Passantino. “‘Verified’ Twitter accounts share fake image of ‘explosion’ near Pentagon, causing confusion.” CNN. (May 23) <https://edition.cnn.com/2023/05/22/tech/twitter-fake-image-pentagon-explosion/index.html> (검색일: 2023. 11. 01).

<그림 3-7> 푸틴 러시아 대통령의 항복 선언 허위 영상



출처: Holroyd, Matthew and Fola Olorunselu. 2022. “Deepfake Zelenskyy surrender video is the first intentionally used’ in Ukraine war.” euronews. (March 16) <https://www.euronews.com/my-europe/2022/03/16/deepfake-zelenskyy-surrender-video-is-the-first-intentionally-used-in-ukraine-war> (검색일: 2023. 11. 01).

- 현재 우크라이나와 러시아는 전쟁 중이기 때문에, 이와 같은 전쟁 관련 내용물의 허위 영상은 당사국뿐 아니라 주변국에 악의적 영향력을 미칠 수 있다. 우크라이나 대통령은 자신의 인스타그램에 허위 영상 내 내용을 부인하는 동영상을 신속하게 공개하였다. 이어 우크라이나 육군은 시민들에게 항복을 촉구하는 명백한 내용을 담은 조작된 영상이 있을 수 있다는 점을 주의하라고 시민들에게 경고하며, “목표는 혼란을 야기하고, 공포를 심고, 시민들을 불신하게 하고, 우리 군대가 후퇴하도록 선동하는 것입니다. 안심하십시오. 우크라이나는 항복하지 않을 것” 이라 설명하였다.<sup>92)</sup>

○ 2023년 이스라엘-하마스 간 전쟁 중 유포된 생성형 AI 제작 전쟁 이미지

- 유럽디지털미디어관측소(European Digital Media Observatory)의 토마소 카네타(Tomasso Canetta)는 “이스라엘과 하마스 간의 갈등과 관련 허위 정보는 몇몇 사례가 있으나 AI 생성 이미지가 대량으로 사용되는 것은 아니다” 라고 일축하였다.<sup>93)</sup>

92) Holroyd, Matthew and Fola Olorunselu. 2022.

93) Eisele, Ines. 2023. “Fact check: AI fakes in Israel’s war against Hamas.” DW (October 11) <https://www.dw.com/en/fact-check-ai-fakes-in-israels-war-against-hamas/a-67367744> (검색일: 2023. 11. 24).

- 다만 AI 생성 이미지가 서사를 담아 전쟁이 진행되는 과정에서 국제 사회의 관심과 지지를 유도할 수 있도록 하는 역할을 수행할 수 있다. <그림 3-8> 내 이미지는 이의 대표적인 사례이다.

<그림 3-8> 생성형 AI로 제작된 이스라엘-하마스 전쟁 이미지



출처: Eisele, Ines. 2023.

- 언론인과 미디어 전문가에게 AI 이미지를 사용할 때 특히 가짜 전쟁과 같은 실제 사건과 관련하여 AI 이미지 사용에 대해 주의할 필요가 있음을 지적하고 있다. 실제 <그림 3-5> <그림 3-7> 내 이미지는 최초 유포 시, AI 제작에 관한 표시가 붙지 않았으나, 유포되는 과정에서 이를 감지한 사용자들이 AI 제작에 관한 표시(×자, 느낌표)를 추가하였다.

## 3.2 중국

### 가. 생성형 AI에 의한 정치 위해 사례

- 중국에서의 생성형 AI 정치 위해 사례는 중국 정부가 해외 정치 문제에 AI를 통해 개입하고 있다는 점에서 주목받고 있다.
  - 2024년 미국 선거가 다가옴에 따라 중국 공무원들이 허위 정보를 퍼뜨리고 분열적인 정치 문제에 대한 토론을 자극하기 위해 온라인에서 미국 유권자를 모방하기 위해 인공지능으로 만든 이미지를 사용했다며, 지난 2023년 9월 마이크로소프트 분석가들이 경고했다.
  - 2024년 다수의 국가에서 주요 선거가 예정되어 있는 상황에서, 미국은 중국 내 공무원 등의 활동에 의해 사용된 생성된 AI가 선거에 위



해를 가할 것에 대하여 우려하고 있다.

<그림 3-9> 마이크로소프트 위기정보 보고서(Microsoft Threat Intelligence report)에서 게재된 중국 IO 자산이 게시한 AI 생성 이미지



출처: CNN. 2023. “Suspected Chinese operatives using AI generated images to spread disinformation among US voters, Microsoft says.” (September 7) <https://edition.cnn.com/2023/09/07/politics/chinese-operatives-ai-images-social-media/index.html> (검색일: 2023. 12. 01).

## 나. 생성형 AI에 의한 저작권 분쟁 사례

- 2023년 12월 중국 베이징 인터넷 법원은 중국에서 최초로 인공지능 생성 콘텐츠(AIGC)에 대한 저작권을 인정했다.
  - 판결에서 법원은 미국 스타트업 StabilityAI의 텍스트-이미지 변환 소프트웨어 Stable Diffusion을 통해 생성된 사진(<그림 3-9> 참조)은 독창성과 지적 입력을 기반으로 저작권법의 보호를 받는 예술 작품으로 간주되어야 한다고 밝혔다.
  - 지적재산권(IP) 침해 소송은 Stable Diffusion을 사용하여 젊은 아시아 여성의 이미지를 만들어 중국의 인스타그램 같은 플랫폼 Xiaohongshu에 게시한 원고 Li 씨에 의해 2023년 5월 진행되었다. 베이징 인터넷 법원은 AI가 생성한 이미지가 예술작품이라고 판결하고 피고인 Liu에게 공개 사과를 하고 원고에게 손해 배상금 500위안 (US\$70.43)과 법원 수수료 50위안을 지불하라고 명령했다.<sup>94)</sup>

<그림 3-10> 중국 내 지식재산권 분쟁 사례였던 AI 생성 이미지



출처: Tech Trends, 2023.

- 중국이 생성형 AI에 의한 창작물 중 일부의 저작권을 인정함에 따라 다양한 논쟁이 이어지고 있다.
  - 일부 AI 생성 콘텐츠에 대한 저작권을 허용하는 것은 AI 사용에 대한 인센티브를 직접적으로 강화할 뿐만 아니라 AI 제품 및 서비스의 상업적 가치를 높이는 효과가 있다는 것이다. 반대로 AI 생성 콘텐츠에 대한 저작권을 거부하면 디지털 아티스트가 창작물의 출처를 허위로 표현하려는 유혹을 받아 의도치 않게 기만적인 관행이 조장될 수 있을 것이다.<sup>95)</sup>

94) Tech Trends, 2023. “Beijing court’s ruling that AI-generated content can be covered by copyright eschews US stand, with far-reaching implications on tech’s use.” (December 1) <https://www.scmp.com/tech/tech-trends/article/3243570/beijing-courts-ruling-ai-generated-content-can-be-covered-copyright-eschews-us-stand-far-reaching> (검색일: 2023. 12. 01).

95) Project Syndicate, 2023. “China’s Short-Sighted AI Regulation.” (December 8) <https://www.project-syndicate.org/commentary/risks-of-beijing-internet-court-ruling-allowing-copyright-of-ai-generated-content-by-angela-huyue-zhang-2023-12> (검색일: 2023. 12. 01).

### 3.3 영국

#### 가. 생성형 AI의 선거활용 문제점

##### □ 케임브리지 애널리티카 스캔들<sup>96)</sup>

○ 영국에서 허위 정보에 대한 논의는 미국 대선 당시 영국의 데이터 분석 회사인 케임브리지 애널리티카(Cambridge Analytica)가 정보 유통에 개입한 사건에서 시작되었다.

- 2015년 10월 미국 상원의원 테드 크루즈(Ted Cruz)는 당시 공화당 대통령 후보 경선에서 가장 낮은 지지율을 보이고 있었다. 그러나 케임브리지 애널리티카와 함께 당사자들의 동의 없이 잠재적 유권자 수백만 명의 페이스북 계정의 정보를 활용해 경선을 치렀다.

<그림 3-11> 미국 상원의원 테드 크루즈의 페이스북 정보 활용 스캔들



출처: Davies, Harry. 2015.

96) Davies, Harry. 2015. "Ted Cruz using firm that harvested data on millions of unwitting Facebook users." The Guardian (December 11) <https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data> (검색일: 2023. 11. 10).

- 이후 케임브리지 애널리티카가 SNS 계정을 통해 트럼프의 승리를 위해 클린턴에 대한 확인되지 않은 부정적인 뉴스를 유포하였다는 주장이 지속적으로 제기되었다.
- 2018년 3월 가디언과 옵저버, 채널4 및 미국의 뉴욕타임즈 등 주요 언론사들은 케임브리지 애널리티카의 전 직원인 크리스토퍼 와일리의 내부 고발 기사를 보도했다. 케임브리지 애널리티카가 페이스북 계정을 통해 수집한 데이터가 미국 대선, 영국의 브렉시트 국민투표, 인도, 브라질과 이탈리아 등 많은 나라의 선거 과정에 영향을 주었다는 사실이 알려졌고, 페이스북의 주가는 폭락하였다.
- 이 사건을 계기로 페이스북은 모든 지역에서 유럽연합 일반 데이터 보호 규정의 시행을 결정했다.
- 이에 대해 2018년 7월, 영국 정보위원회(Information Commissioner's Office)는 페이스북에 50만 파운드의 벌금을 부과하겠다고 밝혔다.
- 케임브리지 애널리티카 사건은 영국 의회에서 온라인 미디어에 대한 규제의 필요성에 대해 본격적으로 논의를 시작하는 계기가 되었다.

#### □ 브렉시트 국민투표

- 2016년 영국의 브렉시트 국민투표 과정에서 허위 정보 문제는 의문과 논란을 불러일으켰다. EU 분담금과 이민자 수용 분담에 대한 불만은 브렉시트에 대한 국민투표를 시행하는 계기가 되었다. 당시 영국에서는 다음과 같은 가짜뉴스들이 주요 언론에 보도되고 SNS를 통해 급속히 확산되었다.<sup>97)</sup>
  - 시리아 난민이 (영국으로) 밀려오고 있으며, 터키는 조만간 EU 회원국이 될 것이다.
  - EU에 퍼주고 있는 3억 5,000만 파운드(약 5,224억 원)의 분담금을 영국의 무상의료에 쓰겠다.
  - 코빈, 브렉시트 비용으로 노동당이 집권하면 920억 파운드를 EU에 납부할 의사가 있음을 확인하였다.

97) 전훈 · 최우정 · 김재선. 2019. “허위정보 유통과 민주주의.” 『한국법제연구원』 (9월 15일).

- 이민자 자녀 70만 명을 교육하는데 연 32억 파운드(약 4조 7천억 원)의 자금이 필요하다.

○ 브렉시트 관련 허위 정보 문제에서 영국을 뒤흔들었던 또 다른 하나의 문제는 러시아가 SNS 계정을 통 브렉시트 국민투표에 개입했다는 점이다.

<그림 3-12> 러시아 SNS 계정 15만여 개가 브렉시트 투표에 개입



출처: 김효진. 2017.

○ 러시아 SNS 계정 15만여 개가 국민투표 당시 영어로 유럽연합에 대한 비난을 쏟아내며, 브렉시트에 찬성하는 게시글을 집중적으로 올렸다는 보도가 연이어 터져 나왔다. 이에 따라 영국 의회 정보보안위원회는 본격적으로 조사를 시작했다.<sup>98)</sup>

98) 김효진. 2017. “러시아 계정 15만개 브렉시트 투표 개입 시도.” 『한겨레』 (11월 16일) <https://www.hani.co.kr/arti/international/europe/819395.html> (검색일: 2023. 11. 12).

- 영국 의회 정보보안위원회는 2019년 3월 보고서를 작성하였고, 2020년 7월이 되어서야 이 보고서를 공개했다. 이 보고서에 따르면 러시아는 영국을 서방국가 중 최우선 첩보활동 목표로 정하고 브렉시트 국민투표에 영향을 미치려 했다는 것이다. 그러나 영국 정부는 이러한 활동에 대한 증거를 수집하거나 평가하지 않았다고 언급하고 있다.
- 러시아 계정의 트위터에서의 봇과 트롤(SNS에서 여론을 호도하고 조작하는 계정 또는 메시지) 활용은 러시아가 브렉시트에 영향력을 행사하려고 했다는 증거라고 주장하였으나 영국 정부의 대응은 미온적이었다.<sup>99)</sup>
- SNS상에서 유통 확산되는 허위 정보 문제는 그 영향력이 국내의 문제로 국한되는 것이 아니라 국경을 넘어 다른 나라의 정치를 결정하고 사회를 혼란스럽게 만드는 데까지 이르고 있다는 점에서 그 심각성이 크다고 할 수 있다.

#### □ 2017년 보궐선거

- 영국의 2017년 보궐선거에서도 허위 정보가 생성, 전파되었는데, 특히 트위터 사용량의 3분의 1 이상이 사람이 아닌 봇에 의하여 사용된 것으로 나타났다.<sup>100)</sup>
- 예를 들어 2017년 6월 런던에서 차량 돌진 테러가 발생하자 런던 테러에 연루된 사람이 미국 공화당 소속 상원의원인 존 매케인(John McCain)이라는 허위 정보가 돌기도 했다.
- 이에 당시 영국 총리 테리사 메이(Theresa May)는 디지털 공간에서 테러 계획이 이루어지거나 극단주의 세력의 과격한 주장이 확산되는 것을 사전에 방지해야 한다고 주장했다. 이를 위해 소셜네트워크서비스 기업들과의 협력의 중요성을 설파했다.<sup>101)</sup>

99) 이율. 2020. “런던은 내 앞마당?...러, 전방위 내정개입에 영, 뒷북 대응.” 『연합뉴스』 (7월 22일) <https://www.yna.co.kr/view/AKR20200722086651009?section=search> (검색일: 2023. 11. 13).

100) 전훈·최우정·김재선. 2019.

101) 강상원. 2019. “소셜미디어 가짜뉴스 규제에 나선 영국 정부”. 국무조정실 규제혁신.

## □ 2019년 영국 총선

- 영국 총선 기간 정체불명의 소셜네트워크 서비스 계정과 홈페이지, 조작된 동영상 등 해외 세력의 개입 정황이 드러났다.<sup>102)</sup>
  - 정체가 모호한 집단이나 러시아에서 조작한 것으로 보이는 허위 정보들이 주를 이루었으나, 온라인상의 많은 허위 정보의 상당수가 기존 정당이나 후보자들이 생성해 낸 정보들이 드러나면서 새로운 문제가 등장했다.
  - 노동당 브렉시트 정책 담당 의원이 브렉시트 관련 질문에 대해 답변하지 못하는 것처럼 조작된 동영상이 유포되었다. 이 동영상을 유포한 보수당은 사과했다. 또한 팩트체크 계정으로 트위터 계정을 조작해서 트위터로부터 경고받았다.
  - 노동당 또한 출처가 불분명한 문건을 인용해 만약 보수당이 정권을 잡으면 브렉시트 후 미국과 무역 협정을 체결할 것이고 그 결과는 국민건강서비스가 약화로 이어질 것이라고 주장하였다. 그러나 이 문건은 러시아의 허위 정보 유포공작과 관련된 문건인 것으로 밝혀졌다.
  - 무슬림의 흥기난동 사건 또한 보수당의 지지도를 높이기 위한 공작이었다는 허위 정보가 노동당 지지자들의 페이스북에 유포되는 등 언론사를 사칭한 계정들은 수없이 많은 가짜뉴스를 생산했다.
- 이러한 일련의 사건들로 인해 영국에서는 온라인 유해물과 허위 정보에 대한 규제 논의가 본격화되었다.
- 이와 함께 2017년 발생한 14세 소녀 몰리 러셀이 소셜네트워크서비스의 허위 정보의 영향으로 자살하는 사건이 발생했다. 그 사건 이후 인스타그램에 관련 이미지와 게시물들이 연속적으로 게시되는 현상이 일어났고 소셜네트워크서비스가 만들어 내는 온라인 피해에 대한 규제 논의가 영국의 시민 단체들을 중심으로 제기 확대되었다.<sup>103)</sup>

102) Satariano, Adam and Amie Tsang. 2019. "Who's Spreading Disinformation in U.K. Election? You Might Be Surprised." New York Times. <https://www.nytimes.com/2019/12/10/world/europe/elections-disinformation-social-media.html> (검색일: 2023. 11. 12).

103) BBC. 2022. "Molly Russell: Coroner's report urges social media changes." (October 14) <https://www.bbc.com/news/uk-england-london-63254635> (검색일: 2023. 12. 01).

## 나. 정치 캠페인과 생성형 AI

### □ 2023년 영국 특별선거<sup>104)</sup>

- 무소속 앤드류 그레이는 2023년 7월 20일 국회의원 특별선거에 출마하며 인공지능을 활용해 주민들이 원하는 것을 반영한 선거 공약을 제시했다.
  - 자신만의 정책은 없다고 주장하며 유권자들의 정서를 크라우드 소싱하거나, 머신러닝을 사용하여 정치공약을 만들었다.
  - 그는 이 기술을 통해 정치인들이 더 빠르고 공정하게 유권자의 의견을 폭넓게 반영할 수 있다고 주장한다.
  
- 완전히 새로운 방식으로 유권자들과 소통할 수 있다고 주장하며 그렇다고 해서 대표자의 역할이 반드시 바뀌는 것은 아니라며 단지 우리가 무슨 일이 일어나고 있는지 훨씬 더 빨리 파악하고 더 공정하게 대변할 수 있다는 의미라고 밝혔다.
  
- 보수당 의원인 나이젤 아담스의 갑작스러운 사임으로 영국 북부의 도농 복합 지역구인 셀비와 에인스티에서 목요일 보궐선거가 실시되었다. 집권 보수당, 야당 노동당, 떠오르는 자유민주당 간의 치열한 접전이 예상된다. 군소 정당과 무소속 후보도 다수 출마했다.
  
- 정치 캠페인에서는 선거 자료 제작, 공격 광고, 기부금 요청에 AI를 활용하고 있다. AI는 이미지를 대량생산하고 중복 작업을 빠르게 수행하는데 도움이 될 수 있다.
  
- 하지만 전문가들은 이 도구가 “허위 정보의 민주화”에 이용될 것을 우려하고 있다.

104) Volpcelli, Gian. 2023 “Meet Britain’s first AI-powered candidate.” POLITICO (July 18) <https://www.politico.eu/article/britain-uk-andrew-gray-ai-artificial-intelligence-manifesto-polis-mp-nigel-adams-selby-ainsty-election/> (검색일: 2023. 12. 01).



## 3.4 독일

### 가. 독일의 AI 선거 활용 문제 개요

- 독일의 선거 과정에서 AI를 이용하여 문제가 된 구체적 사례는 아직 나타나지 않고 있다.
- 지금까지의 상황은 선거에서 AI 활용보다 가짜뉴스를 둘러싼 논의가 주요 쟁점이 되고 있다.
- 독일의 ‘새로운 책임 재단(Stiftung Neue Verantwortung)’의 보고서와 연방선거관리위원회(Bundeswahlleiter)의 홈페이지를 참조하여 가짜뉴스에 관한 독일의 입장을 살펴본다.
- 정치활동 관련 AI의 활용 사례로 2023년 11월 27일 딥페이크 기술을 활용한 가짜 비디오 사건을 살펴본다.

### 나. ‘새로운 책임 재단’ 보고서의 주요 내용

#### □ 개요

- ‘새로운 책임 재단(Stiftung Neue Verantwortung)’은 기술적으로 급속하게 진화하고 있는 사회를 제대로 이해하기 위해 2008년 베를린에 설립된 연구재단(싱크탱크)이다.
- 이 재단은 2017년 연방총선 실시 후 “왜곡된 현실 - 미국과 독일 연방총선에 드리운 가짜뉴스의 그림자”<sup>105)</sup>라는 보고서를 작성했다. 여기에 그 주요 내용을 소개한다.
  - 독일에서는 2017년 연방총선을 앞두고 가짜뉴스가 선거에 어떤 영향을 미칠지 걱정이 많았다. 그래서 선거 후 독일의 유권자들이 가짜뉴

105) Sangerlaub, Alexander. 2017. “Verezzerte Realitaten - Die Wahrnehmung von “Fake News” im Schatten der USA und der Bundestagswahl.” Stiftung Neue Verantwortung (October) [https://www.stiftung-nv.de/sites/default/files/fake\\_news\\_im\\_schatten\\_der\\_usa\\_und\\_der\\_bundestagswahl.pdf](https://www.stiftung-nv.de/sites/default/files/fake_news_im_schatten_der_usa_und_der_bundestagswahl.pdf) (검색일: 2023. 12. 01).

스 현상을 어떻게 평가하고 있었는지를 조사하였다.

- 사실확인기구들(Fact-Checking-Institutionen)이 연방총선 전에 발표한 보도에 따르면, 일반적인 우려와는 달리 독일에서 가짜뉴스가 그렇게 많지는 않았다고 한다.

## □ 가짜뉴스(Fake News)의 영향력에 관한 여론조사 결과

### ○ 개별 독일인의 미디어 신뢰도

- 전체 응답자: 1,037명 [녹색당 지지자: 87명, 독일대안당(AfD) 지지자: 69명, 비투표자: 95명 포함]
- 전체 응답자의 63%가, 녹색당 지지자의 78%가 미디어를 신뢰했지만, 전체 응답자의 33%, 녹색당 지지자의 21%는 미디어를 신뢰하지 않았다.
- 반면에 독일대안당 지지자의 26%, 비투표자의 47%만이 미디어를 신뢰했지만, 독일대안당 지지자의 70%, 비투표자의 51%는 미디어를 신뢰하지 않았다.

### ○ 선거운동에서 정보취득을 위한 주요 수단

- 전체 응답자: 1,037명 [젊은 층(18~34세): 238명, 독일대안당 지지자: 69명 포함]
- 공영 TV: 전체 응답자의 37%, 젊은 층의 20%, 독일대안당 지지자의 27%
- 온라인 뉴스: 전체 응답자의 21%, 젊은 층의 38%, 독일대안당 지지자의 18%
- 일간지: 전체 응답자의 11%, 젊은 층의 2%, 독일대안당 지지자의 4%
- 미디어를 통한 정보취득 없었음: 전체 응답자의 8%, 젊은 층의 13%, 독일대안당 지지자의 17%
- 개인 TV(Privatfernsehen): 전체 응답자의 7%, 젊은 층의 2%, 독일대안당 지지자의 3%
- SNS: 전체 응답자의 6%, 젊은 층의 13%, 독일대안당 지지자의 16%

- 공영라디오: 전체 응답자의 3%, 젊은 층의 2%, 독일대안당 지지자의 5%

○ 선거운동에서 가짜뉴스의 확산에 대한 인식 정도

- 전체 응답자: 1,037명 [미디어 비판층: 337명, 젊은 층(18~34세): 238명 포함]
- 전체 응답자의 61%는 가짜뉴스가 많았다고 느꼈고, 34%는 많지 않았다고 느낌
- 미디어 비판층의 75%는 가짜뉴스가 많았다고 느꼈고, 20%는 많지 않았다고 느낌
- 젊은 층의 72%는 가짜뉴스가 많았다고 느꼈고, 20%는 많지 않았다고 느낌
- 누군가(미디어 비판층 또는 젊은 층) 미디어를 신뢰하지 않을 경우, 가짜뉴스 또한 많다고 인식하는 것으로 나타남

□ 개별 가짜뉴스가 선거운동을 결정하는 것은 아니다.

○ 개별 가짜뉴스가 유권자가 어떤 선택을 할지에 결정적인 역할을 하는 것은 아니다. 보다 결정적인 요소는 다양한 미디어 채널을 통한 종합적인 아젠다 셋팅(Agenda Setting)이다.

- 어떤 주제가 어떤 형태로 계속해서 초점에 놓이는지, 또 아젠다가 어떤 형식과 방법으로 대중에게 전달되는지가 관건이 된다.

□ 여론조사 결과의 결론

○ 가짜뉴스는 소통전략의 일부이다. 독일에서는 특히 극우주의자나 대중영합주의자가 왜곡된 세계상을 제시하기 위해 이용하는 수단이다.

- 독일대안당 지지자의 데이터에서 보듯이 자신들이 원하는 세계상을 믿는 것이다.

- 전통적 미디어에 대한 신뢰가 감소하는 경향이다.
  - SNS, 블로그, 뉴스매체(Epoch Times, RT Deutsch 등)가 대안적 미디어의 현실이 되고 있다.
  
- 전체 응답자의 평가에 따르면, 공공의 영역에서 어느 정도 관리가 가능함에도 불구하고 가짜뉴스는 문제가 크다.
  - 미국에서 이 현상에 대한 많은 보도와 용어의 모호한 사용이 가짜뉴스 문제의 원인이 될 수 있다.

#### 다. 연방선거관리위원회의 “가짜뉴스와 사실” 코너 신설

- 연방선거관리위원회는 홈페이지에 ‘가짜뉴스와 사실(Fakten gegen Fake News)’<sup>106)</sup>이라는 코너를 신설하여 이 문제를 지속적으로 다루고 있다.

#### □ SNS의 가짜뉴스(Desinformation in Social-Media-Kanälen)

- 2021년 연방총선과 관련하여 SNS나 챗그룹(Chatgruppen)에서 유권자에게 의도적으로 잘못 전달될 수 있는 거짓 정보나 가짜뉴스가 유통되고 있다.
  - 이러한 거짓 정보는 널리 확산되고 있는데, 그 이유는 그 거짓 정보를 옮기는 사람들이 그것이 가짜뉴스인 것을 제대로 인식하지 못하기 때문이다. 그래서 연방선거관리위원회는 알게 된 가짜뉴스를 포착하여 바로잡고 있다.

106) 독일 연방선거관리위원회(Bundeswahlleiter) 홈페이지 참조. Die Bundeswahlleiterin. “Bundestagswahl 2021.” <https://www.bundeswahlleiterin.de/bundestagswahlen/2021/fakten-fakenews.html#7b083e33-e95d-4a3f-87a3-276f5e01eb6e> (검색일: 2023. 12. 05).

## □ 2021년 연방총선 관련 SNS에 확산된 가짜뉴스 사례

- 코로나19 팬데믹 때문에 예방접종을 하지 않은 유권자는 선거일에 투표 장소에 입장할 수 없다. 조건이 충족되지 않은 유권자는 선거권을 박탈한다. (전혀 사실이 아니다)
- 유권자들은 한 선거구에서 투표용지에 애완동물의 이름이 적힌 정당에 투표할 수 있다. 무소속 후보자의 후보 추천장에 실수로 정당명 대신에 애완동물의 이름을 올렸기 때문이다. (실수로 투표용지에 애완동물의 이름이 올라가진 않는다. 다만, 무소속 후보자는 투표용지의 정당 칸에 자신을 상징하는 단어나 표어를 넣을 수 있어서 생길 수 있는 문제일 뿐이다.)
- 우편 투표는 불안하다, 쉽게 조작될 수 있다. (이는 있을 수 없는 일이다.)
- 선거를 지켜보는 일은 선거구의 투표함에서만 가능하다. 그래서 우편 투표를 개표하는 일은 보이지 않는 곳에서 조작될 수 있다. (우편 투표의 개표는 전적으로 공개적으로 진행되고, 누구나 집계과정을 지켜볼 권리가 있다.)
- 투표함은 안전하지 않다. (전혀 그렇지 않다.)
- 처음 연방총선에 투표를 하는 유권자는 투표용지에 자기 이름을 적음으로써 투표 관련 로또 당첨에 참여할 수 있다. (그런 일은 없다.)
- 투표용지의 오른쪽 위 끝에 구멍을 내거나 뜯어낸 투표지는 무효이다. (무효가 아니라 모든 투표용지는 그렇게 한다.)

## 라. 독일에서 생성형 AI를 이용한 딥페이크 비디오 사례

- 2023년 11월 27일, 하나의 정치단체인 ‘정치미학센터(Zentrum für politische Schönheit)’가 생성형 AI를 이용하여 연방총리(Olaf Scholz)가 등장하는 비디오를 만들어 유튜브, 페이스북, 트위터, 인스타그램 등에 배포하였다.

- 동영상은 4분 정도의 길이로 솔츠 연방총리가 사민-녹색-자민당 연립정부는 극우 정당인 독일대안당(AfD)을 금지하는 신청서를 낼 것이라고 선언하는 내용이다.
  - 이 단체는 독일 민주주의를 수호하기 위해 극우극단주의에 반대하는 투쟁의 중요한 발걸음이라고 자평했다.
- 이에 대해 연방정부 대변인은 당일 그 비디오는 진짜가 아니고 가짜라고 밝히고, 그와 같은 딥페이크는 하나의 장난이라고 보기 어렵고, 미디어에 대한 공공의 신뢰를 손상한다고 비판하였다.
- 이 행위가 법적으로 어떤 결과를 가져올지 아직 알 수 없고, 일단 검토할 것이라고 밝혔다.
- 11월 30일 연방언론청(Bundespresseamt)은 해당 비디오를 SNS에서 삭제토록 조치했다.
- 연방언론청은 외교부와 내무부와 함께 딥페이크와 가짜정보에 어떻게 대처할지 협의 중이다.

## 4. 주요국의 AI 및 생성형 AI 선거 규제

### 4.1 미국

#### 가. AI 관련 입법 추진 개요

##### ○ AI 로드맵

- 미국은 AI 산업영역을 선도하는 기업을 보호하고, 발전을 육성하려는 의도와 자국의 안보를 강화하기 위한 의도에서 AI 및 생성형 AI 규제를 입법화하고 있다.

〈표 4-1〉 미국의 ‘AI 로드맵’의 5가지 전략

구분	주요내용
책임감 있는 AI 사용	- CISA는 AI 지원 소프트웨어 도구를 활용하여 사이버 방어(Defense)를 강화하고 중요 인프라 임무를 지원 - 연방 조달, 프라이버시, 시민권, 시민 자유에 관한 법률을 포함하여 헌법과 모든 관련 법률 및 정책에 따라 책임감 있고 윤리적이며 안전한 AI 사용을 보장
AI 시스템 평가 및 보증	CISA는 연방 민간 정부 기관, 민간 부문 기업, 주·지방·부족·자치지역(State, Local, Tribal, Territorial, SLTT) 정부를 포함한 다양한 이해관계자의 안전한 설계를 통해 AI 기반 소프트웨어를 평가하고 지원 안전하고 탄력적인 AI 개발 및 구현을 위한 모범 사례 및 지침 개발을 통해 AI 시스템을 보장
악의적인 AI 사용으로부터 중요 인프라 보호	CISA는 AI 도구를 개발, 테스트 및 평가하는 다른 정부 기관 및 업계 파트너와 협력하여 미국의 중요 인프라가 직면한 AI 위협을 평가하고 완화를 권고할 계획 이러한 노력의 일환으로 CISA는 AI 시스템과 관련된 위협, 취약성 및 완화에 대한 집중적인 협력을 촉진하기 위해 공동 사이버 방어팀(Joint Cyber Defense Collaborative, JCDC)을 설립할 예정
기관 간, 국제 파트너, 대중과 함께 주요 AI 노력에 대해 협력 및 소통	CISA는 사이버 보안과 AI에 대한 미국 정부의 전반적인 국가 전략을 위한 정책 접근법을 개발하고, AI 기반 소프트웨어 정책 문제에 대한 DHS의 전반적인 접근법을 지원하여 부처 간 노력에 기여 글로벌 AI 보안 모범 사례 및 원칙을 발전시키기 위한 국내뿐만 아니라 국제 파트너와의 협력도 포함

직원의 AI 전문성 확대	CISA는 AI 소프트웨어 시스템과 기술에 대한 교육을 지속적으로 실시할 것이며, AI 전문성을 갖춘 인턴, 펠로우 및 미래 직원을 적극적으로 채용할 계획 직원 교육으로 AI 기반 소프트웨어 시스템의 기술적 측면뿐만 아니라 법적·윤리적·정책적 측면도 반영
---------------	---

출처: 한국지식재산연구원. 2023a. “미국 국토안보수사국, ‘AI 로드맵’ 발표.” (11월 24일)  
[https://www.kiip.re.kr/board/trend/view.do?bd\\_gb=trend&bd\\_cd=1&bd\\_item=0&po\\_item\\_gb=US&po\\_no=22454](https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=US&po_no=22454) (검색일: 2023. 11. 24).

○ 미국은 2023년 11월 14일, 국토안보수사국(HSI) 산하 사이버 보안 및 인프라 보안국(CISA: Cybersecurity and Infrastructure Security Agency)은 인공지능의 안전한 개발 및 구현을 보장하고자 최초의 ‘AI 로드맵’을 발표하였다(<표 4-1> 참조).

- 바이든 대통령은 AI 행정명령 발표를 통해, 글로벌 AI 안전표준 채택 촉진, 미국의 네트워크와 중요 인프라 보호, AI가 대량살상무기를 만드는 데 사용될 수 있는 위험 대비, AI 관련 지식재산(IP) 침해 대처, 전문 인재 양성 등을 도모하고자 하고 있다.<sup>107)</sup>
- CISA는 사이버 안보에 대한 구체적 내용을 담은 AI 로드맵을 통해 이니셔티브를 추진하고, AI에 대한 CISA의 책임감 있는 접근방식을 개괄적으로 설명하는 다섯 가지 전략을 공개 <표 4-1>과 같이 공개하였다.

○ 연방국가인 미국은 연방법제와 주법제를 상하 위계 구조로 나뉘어 분류하고 있으며, 행정명령(Executive Order), 보고서 및 지침, 법률 및 법안 등을 갖추고 있다. 따라서 미국은 연방 차원에서 AI 및 생성형 AI 규제에 관한 원칙적 방향성을 제시하며, 주정부 차원에서 차별금지, 선거규제 등과 관련된 세부 법안을 마련하고 있는 것이다.

107) 한국지식재산연구원. 2023b. “미국 상무부, 미국 최초의 인공지능 행정명령에서 핵심적인 역할 수행.” [https://www.kiip.re.kr/board/trend/view.do?bd\\_gb=trend&bd\\_cd=1&bd\\_item=0&po\\_item\\_gb=US&po\\_no=22409](https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=US&po_no=22409) (검색일: 2023. 11. 24).



〈표 4-2〉 미 연방 AI 및 생성형 AI 규제 개요

법률형태	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 연방행정기관의 AI 이용 허용 해당 AI의 설계, 취득, 개발과정에서의 미국 시민보호 원칙 제시
지침	권리보호	AI 권리장전에 대한 청사진(Blueprint for an AI Bill of Rights)	AI의 개발 및 자동화 과정에서 개인, 기업, 정책입안자가 AI 기술에 대한 책임을 인식하고, 시민권 보호, 평등 접근을 보장할 것에 대해 촉구하는 핵심 원칙
법률	권리보호	AI 훈련법(AI Training for the Acquisition Workforce Act)	AI개발, 자동화 과정에서 개인, 기업, 정책입안자가 AI기술에 대한 책임, 인식, 시민권, 보호, 평등, 접근을 보장할 것을 촉구하는 핵심 원칙 제시
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)	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 위원회는 온라인 플랫폼에 대하여 소비자 보호 등을 위해 청문회 개최, 조사, 연구지원 등의 권한을 부여받으며, 광범위한 감독 권한을 가지게 될 것으로 예상된다.

출처: AI.GOV. “Making AI Work for the American People.” <https://ai.gov/> (검색일: 2023. 11. 24). 내용 등을 토대로 연구팀 작성

## 나. 생성형 AI 선거 입법 동향 및 원칙 중심 규제 동향

○ 행정명령 13960: 2023년 10월 30일 미국은 행정명령(Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence)을 통해, AI 규제의 방향성을 제시하였다. 이 명령은 AI 개발을 위한 안전, 보안, 개인정보 보호, 공정성, 소비자 보호, 근로자 지원, 혁신, 책임 있는 정부 사용의 8가지 목표를 설정한 것이다.

- 해당 행정명령은 향후 90일 내 세부 세칙 보완 이후 적용될 예정이다. 시행에 대한 보고 의무는 최장 1년이 걸릴 것으로 보고 있다. 이를 위해 ‘AI 위원회(AI council)(위원장 브루스 리드 백악관 부실장(White House Deputy Chief of Staff Bruce Reed))’가 이행을 감 독한다.
- 행정명령과 관련한 주요 내용은 아래와 같다.
  - ① 가장 강력한 AI 시스템의 개발자가 안전 테스트 결과 및 기타 중요한 정보를 미국 정부와 공유하도록 요구,
  - ② AI 시스템이 안전하고 신뢰할 수 있는지 확인하는 데 도움이 되는 표준, 도구 및 테스트를 개발,
  - ③ 생물학적 합성 스크리닝(biological synthesis screening)을 위한 강력한 새 표준을 개발하여 AI를 사용하여 위험한 생물학적 물질을 가공할 때 발생할 수 있는 위험으로부터 보호,
  - ④ AI 생성 콘텐츠를 탐지하고 공식 콘텐츠를 인증하기 위한 표준 과 모범 사례를 확립하여 AI 기반 사기와 속임수로부터 미국인 을 보호,
  - ⑤ 바이든-해리스(Biden-Harris) 행정부의 지속적인 AI 사이버 챌린 지를 기반으로 중요한 소프트웨어의 취약점을 찾아 수정하는 AI 도구를 개발하기 위한 고급 사이버 보안 프로그램을 구축
  - ⑥ 국가안전보장회의(National Security Council)와 백악관이 개발할 AI와 보안에 대한 추가 조치를 지시하는 국가안보각서(National Security Memorandum) 개발을 명령<sup>108)</sup>

108) The White House. 2023. “FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.” (October 30) <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/> (검색일: 2023. 11. 24).

- 동 행정명령은 미 정부기관이 구체적으로 AI 오용 방지를 위한 표준 제공, 콘텐츠 인증 모범 사례를 확립하고, 고급 사이버 안보 구축하도록 하였다. 이 명령은 알고리즘 편향부터 개인정보 보호, 첨단 AI 모델의 안전 규정에 이르기까지 다양한 분야를 다루고 있다.
- ▷ 국립표준안전연구소(NIST: The National Institute of Standards and Safety): AI 프로그램 모델을 공개하기 전에 테스트하기 위한 표준 제작
- ▷ 에너지부(Department of Energy) 및 국토안보부(Department of Homeland Security): 잠재적인 인프라 위협과 사이버 안보 위협에 대한 대책 마련
- ▷ 행정부: 국가 또는 경제 안보에 영향을 미칠 수 있는 모델을 개발한 개발자에 대한 보고 요구사항을 포함하여 AI 안전 및 안보에 대한 새로운 표준을 설정함
- ▷ 상무부: 콘텐츠가 AI로 생성되었는지 여부를 확인하기 위해 연방기관을 위한 워터마킹 시스템 개발
- ▷ 이와 같은 일련의 정부 행동에 따라, OpenAI, Alphabet, Meta Platforms와 같은 AI 기업은 안전성을 높이기 위해 AI 생성 콘텐츠에 자발적으로 워터마킹을 삽입하기로 합의하였다.<sup>109)</sup>
- ▷ 행정명령과 관련한 주요 사항은 다음의 <표 4-3>과 같다.

---

109) Reuters. 2023b. “OpenAI, Google, Meta pledge to watermark AI content for safety.” cybernews (November 15) <https://cybernews.com/tech/openai-google-meta-watermark-ai-content-for-safety/> (검색일: 2023. 11. 24).

〈표 4-3〉 AI 행정명령 내 주요 사항

핵심 사항	주요 내용
AI 모델 등록 의무	행정명령 내 AI 모델 등록에 관한 언급이 있으나, AI모델 구축한 기존 모델 사용 회사의 등록 의무는 없음.
안전 평가 보고 요건	행정명령에는 AI 개발자가 심각한 보안 위험을 초래할 수 있는 대규모 AI 훈련 실행을 보고하도록 규정하고 있음. 안전 평가, 즉 “레드팀 구성” 결과를 보고하도록 의무화하고 있음.
이중용도 AI 모델에 대한 정의	행정명령 내 명확히 정의되지 않음. 이중용도 AI 모델은 악의적인 목적으로 사용될 수 있는 AI 모델을 포착하려는 의도 정도로 포괄적인 의미를 부여하여 다소 모호한 상황임.
오픈소스 소프트웨어에 대한 규정	행정명령 내 오픈소스 모델과 오픈소스 AI에 대한 명확한 지침 없음. 전반적으로, 행정 명령의 일반 지침은 오픈 소스 AI와 오픈 소스로 개발 및 출시되는 모든 모델에 적용될 가능성이 높음.
빅테크 모델에 대한 규제	행정명령 내에는 빅테크 모델에 대한 구체적 설명이 명시되어 있음. 행정명령은 관련 조항과 규정이 아직 출시되지 않은 모델에 적용된다는 점을 구체적으로 명시하고 있음. 따라서 Llama 2, Falcon 70B 또는 Mistral과 같은 기존의 대규모 오픈 소스 모델은 서로 다른 수준의 조사에 직면하게 될 것임. 이러한 모델을 구축하고 수정하는 경우 팀이 최소한 규제 감독 및 보고의 하위 집합을 준수해야 할 가능성이 높음.
AI 안전 관련 사항	행정명령은 AI 시스템의 안전과 보안이 필요함을 강조함. AI 시스템에 대한 강력하고 신뢰할 수 있으며 반복 가능하고 표준화된 평가는 물론 이러한 시스템을 사용하기 전에 이러한 시스템의 위험을 테스트, 이해 및 완화하기 위한 정책 및 메커니즘이 필요함.
시민의 자유, 개인 정보 보호 및 알고리즘 편견 관련 사항	행정명령은 AI가 계속 발전함에 따라 미국인의 개인정보와 시민의 자유를 보호하도록 요구함. 요점은 AI 애플리케이션과 모델이 AI를 사용하지 않는 레거시 애플리케이션과 동일한 보호 장치, 투명성, 출처로 처리되어야 한다는 것임. 특히 숨겨진 편견이 포함되거나 올바른 프롬프트로 추출할 수 있는 익명화된 개인 식별 정보(PII)가 포함된 데이터 세트에 AI 시스템을 교육하는 것
AI 생성 코드 관련 사항	행정명령에는 AI 생성 코드에 대한 특정 조항이나 언급이 포함되어 있지 않음. AI 생성 코드는 애플리케이션 및 코드 보안, 더 구체적으로 공급망 보안과 관련된 더 광범위한 규정을 준수하도록 하는 세부안이 추가될 것으로 예측됨.
사이버안보 및 AI 도구에 관한 조항	AI를 사용하고 AI 모델, AIOps, AI를 활용하는 애플리케이션 개발 및 배포를 모니터링하고 보호하는 등 사이버 보안을 위한 새로운 도구와 새로운 연구를 육성하는 것과 관련된 조항

출처: Lim, Liqian. 2023. “What does Biden’s Executive Order on AI safety measures mean for businesses?.” Snyk Blog (November 2) [https://snyk.io/blog/biden-ai-safety-executive-order/?utm\\_medium=paid-search&utm\\_source=google&utm\\_campaign=ba\\_ai-awareness&utm\\_content=blog-executive&utm\\_term=ai%20executive%20order&gclid=CjwKCAiAslGrBhAAEiwAEzMICzbcwEaCKNIOvfCSuljECUWbecodGTCJ2ORkT0EM9WRYnGD4ZSQIvhoCc28QAvD\\_BwE](https://snyk.io/blog/biden-ai-safety-executive-order/?utm_medium=paid-search&utm_source=google&utm_campaign=ba_ai-awareness&utm_content=blog-executive&utm_term=ai%20executive%20order&gclid=CjwKCAiAslGrBhAAEiwAEzMICzbcwEaCKNIOvfCSuljECUWbecodGTCJ2ORkT0EM9WRYnGD4ZSQIvhoCc28QAvD_BwE) (검색일: 2023. 11. 24).

- 행정명령은 AI 모델 등록 의무, 안전 평가 보고 요건, 이중용도 AI 모델에 대한 정의, 오픈소스 소프트웨어에 대한 규정, 빅테크 모델에 대한 규제, AI 안전 관련 사항, 시민의 자유, 개인정보 보호 및 알고리즘 편견 관련 사항, AI 생성 코드 관련 사항, 사이버안보 및 AI 도구에 관한 조항 등을 다루고 있다.
- 행정명령이 강조하는 법 지침의 방향성은 개인권리 강화 및 알고리즘에 의한 편견을 방지하기 위함이다. 특히 이와 관련하여 행정명령이 구체적인 지침을 제시하지는 않지만, 실제로 주 수준(예: 캘리포니아) 또는 국가 수준(예: 유럽연합)의 다른 법률에서 이미 다루기 시작한 내용이 참조될 것이다.

○ ‘AI 권리장전에 관한 청사진(Blueprint for an AI Bill of Rights)’

- 2022년 10월 미 정부는 ‘AI 권리장전에 관한 청사진(Blueprint for an AI Bill of Rights)’을 공표하였다. 위 공표안에서 미 정부는 오늘날 민주주의에 가해지는 가장 큰 과제 중 하나로 미국 대중의 권리를 위협하는 방식으로 기술, 데이터 및 자동화 시스템을 사용하는 것을 지목하였다. 그 예로 확인되지 않은 소셜미디어 데이터 수집은 종종 사람들이 알지 못하거나 동의하지 않은 채 사람들의 기회를 위협하고, 개인정보를 침해하거나, 활동을 광범위하게 추적하는 데 사용되고 있으며, 채용 및 신용 결정에 사용되는 알고리즘은 기존의 원치 않는 불평등을 반영 및 재생산하거나 새로운 유해한 편견과 차별을 삽입하는 것이라 지적되었다.<sup>110)</sup>
- 이에 백악관 과학기술정책실은 바이든 대통령의 비전을 발전시키기 위해 인공지능 시대에 미국 대중을 보호하기 위해 자동화 시스템의 설계, 사용 및 배포를 안내해야 하는 5가지 원칙을 확인하였다(〈표 4-1〉 참조).
- 권리장전은 AI 관련 민권 침해 사례를 조사하는 내용을 포함하고 있다. 또한 개인인권보호 기술 개발을 우선시 하며, AI 기반 의료 행위를 위한 안전프로그램 구축, 교육자가 AI 기반 교육 도구를 사용할

110) The White House. 2022. “Blueprint for an AI Bill of Rights.” <https://www.whitehouse.gov/ostp/ai-bill-of-rights/> (검색일: 2023. 11. 24).

수 있도록 지원하는 내용을 포함하였다.

○ AI 훈련법(AI Training for the Acquisition Workforce Act)<sup>111)</sup>

- 법 제정 목표: 2022년 10월 제정된 ‘AI 훈련법’은 AI에 대한 연방 인력의 이해를 향상하고 AI의 윤리적이고 안전한 사용을 보장하기 위한 인력 확보를 목표로 제정되었다.
- 개요: 2021년 7월 미시간주 케리 피터스 상원의원 및 오하이오주 롭 포트먼 상원의원이 법안 도입, 2021년 12월 상원 통과, 2022년, 9월 29일 하원 통과, 2022년 10월 17일 바이든 대통령 최종 서명
- 주요 내용: AI 기술 구매를 담당하는 연방 직원이 AI 기술의 위험과 이점을 더 잘 이해할 수 있도록 돕는 교육 프로그램 제작

○ 2023 국방수권법(National Defense Authorization Act for Fiscal Year 2023)<sup>112)</sup>

- 법 제정 목표: 미국의 자유는 미국 군인들이 자국을 안전하게 지키는데 필요한 자원과 무기를 보유하고 있는지 여부에 달려 있으므로, 국방을 제공하는 것은 미국 헌법이 의회에 부여한 가장 중요한 책임임을 확인한다. 국방수권법은 미국 군인과 국가에 대한 약속을 이행하는 데 여전히 중요한 부분으로 전방위 영역에서 현 시점의 문제를 반영한 국방수권법을 채택하였다.
- 개요: 2022년 7월 14일 하원 통과, 2022년 12월 6일 상원 내 정치적 합의, 2022년 12월 23일 바이든 대통령 최종 서명
- AI 관련 주요 내용: 2023 국방수권법에서 군의 현대화와 관련한 내용 중 AI는 다음과 같이 언급되었다. “AI, 생명공학, 양자 컴퓨팅, 자율 시스템 등 신흥 기술에 대한 전례 없는 수준의 투자가 포함된다

111) U.S. Congress. 2022. “AI Training Act: Artificial Intelligence Training for the Acquisition Workforce Act,” (October 17) <https://www.congress.gov/117/plaws/publ207/PLAW-117publ207.pdf> (검색일: 2023. 11. 24).

112) Republicans-Armedservices. “FY23 National Defense Authorization Act.” <https://armedservices.house.gov/sites/republicans.armedservices.house.gov/files/Final%20FY23%20NDAA%20Conf%20Highlights.pdf> (검색일: 2023. 11. 24).

(Includes unprecedented levels of investment in emerging technologies, including AI, biotechnology, quantum computing, and autonomous systems).”

○ 2022년 알고리즘 책임법(Algorithmic Accountability Act of 2022)

- 법 제정 목표: 해당 법안은 영리 조직과 비영리 조직 모두 보호된 개인 특성을 기반으로 결정을 내리는 알고리즘을 사용하는 것을 금지함을 목표로 한다. 즉 해당 법안은 기업이 실제 또는 인지된 인종, 피부색, 종교, 출신 국가, 성별, 성 정체성 또는 표현을 포함하여 광범위한 개인 특성을 기반으로 하는 경우 알고리즘을 통해 결정을 내리는 것을 불법으로 규정하는 것이다.

- 주요내용<sup>113)</sup>

▷ 책임있는 기술: 알고리즘 책임법은 규제기관에 더 나은 식별 기능을 제공함으로써 이미 소외된 집단에 대해 새로운 보이지 않는 장벽을 세우는 등을 통한 편향된 블랙박스 시스템을 해결하기 위한 중요한 단계라 지적된다. 따라서 해당 입법을 통해, 기업이 불법적인 차별과 기타 사회적 해악을 자동화하는 기술을 도입하는 것을 방지한다.

▷ 2022 알고리즘 책임법안은 알고리즘 활용이 널리 확대됨에 따라 오류나 편향성 등으로 인하여 발생할 수 있는 새로운 위험을 예방하고자, 영향평가라는 사전 안전장치를 마련하는 것에 초점을 두고, 관련 내용을 비교적 상세히 규정하고 있다. 알고리즘을 개발, 배포하는 기업에 대해 자동화된 의사결정 시스템 또는 증강된 중요 의사결정 프로세스가 소비자에게 미치는 영향을 평가하도록 한다.<sup>114)</sup>

113) 알고리즘 책임법의 주요내용은 다음에서 발췌하였다. Office of U.S. Senator Ron Wyden. 2023. “Support for the Algorithmic Accountability Act of 2023.” (September 21) [https://www.wyden.senate.gov/imo/media/doc/support\\_for\\_the\\_algorithmic\\_accountability\\_act\\_of\\_2023.pdf](https://www.wyden.senate.gov/imo/media/doc/support_for_the_algorithmic_accountability_act_of_2023.pdf) (검색일: 2023. 11. 24).

114) 홍석한. 2023. “미국 “2022 알고리즘 책임법안”에 대한 고찰.” 『미국헌법연구』 34집 제1호: 73-107.

- 적용대상: 알고리즘 책임법은 다음과 같은 경우 연방거래위원회법 (Federal Trade Commission Act) 5(a)(2)항에 따라 연방거래위원회 (FTC)가 관할권을 갖는 대상에 대해 적용된다.
  - ▷ 이전 3년 과세연도 동안 연간 수익이 \$5,000,000 이상이거나 자기자본 가치가 \$25,000,000를 초과한 경우
  - ▷ 증강된 중요 의사결정 프로세스를 사용하기 위해 개발된 자동화된 의사결정 시스템을 배포한 경우
  - ▷ 지난 3년 동안 위의 사항을 충족한 경우
  
- 영향평가 주요 요구사항: 동 법률 내 영향평가에서 요구하는 사항은 프로세스 평가, 이해관계자 협의, 개인정보 보호 평가, 성능 평가, 훈련 및 교육, 가이드라인 및 제한 사항, 데이터 문서화, 권리, 투명성 및 설명 가능성, 부정적인 영향평가, 문서 및 변화단계표 그리고 리소스 식별 등에 관한 것이다(<표 4-4> 참조).

<표 4-4> 알고리즘 책임법 2022 내 영향평가 주요 사항

영역	주요내용
프로세스 평가	새롭게 강화된 중요 의사결정 프로세스의 경우, 알려진 피해나 단점은 물론 결정 강화의 의도된 이점과 목적을 포함하여 결정에 사용된 기존 프로세스를 평가해야 함
이해관계자 협의	연락처 문서화, 협의 날짜, 조건 및 프로세스에 대한 정보, 권장 사항 등 관련 이해관계자와의 협의 필요
개인 정보 보호 평가	데이터 최소화 및 정보 보안을 포함한 개인 정보 보호 위험 및 개인 정보 보호 강화 조치에 대한 지속적인 테스트 및 평가를 수행해야 함.
성능 평가	데이터 품질 조사, 시스템 성능 모니터링, 특정 하위 그룹 모집단과 관련하여 수행된 테스트 문서화를 포함하여 시스템 또는 강화된 프로세스의 현재 및 과거 성능에 대한 지속적인 테스트 및 평가가 필요함.
훈련 및 교육	ADS의 부정적인 영향과 강화된 중요한 의사결정 프로세스 및 업계 모범 사례에 대해 관련 직원, 계약자 또는 기타 사람들을 위한 지속적인 훈련 및 교육이 필요함.
가이드라인 및 제한 사항	시스템 또는 의사결정 프로세스의 특정 사용에 대한 가이드라인의 필요성 및 개발 또는 제한 사항을 평가.
데이터 문서화	데이터 소스 및 구조, 데이터 수집 방법 및 사전 동의, 대안 탐색 여부에 대한 정보를 포함하여 시스템을 개발, 테스트, 유지 관리 또는 업데이트하는 데 사용되는 데이터 문서 및 기타 입력을 유지 관리
권리, 투명성 및	통지 및 기타 투명성 및 설명 가능성 조치를 포함한 소비자의 권리를 평가



설명 가능성	
부정적인 영향 평가	소비자와 적절한 완화 전략에 대한 강화된 중요 의사결정 프로세스의 ADS가 부정적인 영향을 미칠 수 있는 가능성을 식별함.
문서 및 변화단계표	변화단계표 기록, 테스트 날짜 및 이러한 프로세스에 관련된 담당자의 연락처를 포함하여 개발 및 배포 프로세스에 대한 지속적인 문서 관리
리소스 식별	ADS, 증강된 중요한 의사결정 프로세스 또는 성능(예: 정확성, 견고성 및 신뢰성), 공정성 및 편견, 투명성 및 성능을 포함한 영향 평가를 개선하는데 필요한 모든 기능, 도구, 표준 또는 데이터 세트 및 기타 리소스를 식별. 설명 가능성, 개인 정보 보호 및 보안, 개인 및 공공의 안전, 효율성 또는 비용 등을 포함함

출처: Holistic AI. 2023. "US Algorithmic Accountability Act: Third Time Lucky?." (October 25) <https://www.holisticai.com/blog/us-algorithmic-accountability-act> (검색일: 2023. 11. 24).

#### - 주(state)별 알고리즘 책임법

- ▷ 2021년 DC Chamber of Commerce, 알고리즘에 의한 차별금지법 (Stop Discrimination by Algorithms Act B24-0558) 채택하였다. DC는 ① 직접 또는 공급업체를 통해 25,000명 이상의 지역 주민에 대한 개인정보(구매 또는 소비 내역 포함)를 보유하거나 통제하는 기업, ② 가장 최근 회계연도 이전 3년 동안 평균 연간 총 수입이 1,500만 달러 이상인 기업, ③ 개인정보의 수집, 조립, 판매, 배포, 액세스 제공 또는 유지 관리를 통해 연간 수익의 최소 50%를 창출하는 데이터 브로커 또는 기타 법인, ④ 다른 기업을 대신하여 알고리즘 적격성 결정 또는 알고리즘 정보 가용성 결정을 수행하는 공급업체 중 하나 이상에 해당된다면 알고리즘에 의한 차별금지법의 적용을 받도록 하고 있다.

<표 4-5> 캘리포니아 주, 직장기술책임법의 주요 사항

영역	주요내용
투명성	고용주는 사용하기 전에 직장 내 모든 관련 데이터 수집, 전자 모니터링 및 알고리즘에 대해 근로자에게 고지하여야 함. 고용주는 이러한 시스템이 근로자의 성과나 생산성에 대한 평가를 포함한 고용 결정에 어떤 영향을 미칠 수 있는지 설명하여야 함. 근로자는 자신의 데이터에 액세스하고 이를 수정할 권리가 있음.
책임성	고용주는 근로자가 업무를 수행하는 데 꼭 필요한 경우 또는 유효한 사업 목적을 위해서만 근로자 데이터를 수집하여야 함. 고용주는 근무 외 직원을 모니터링하지 않아야 함. 고용주는 얼굴 인식과 같이 입증되지 않았거나 위험성이 높은 기술을 사용하지 않아야 함

	고용주는 직무와 관련 없는 직원의 행동을 예측하기 위해 알고리즘을 사용하지 않아야 함 전자 모니터링 및 알고리즘은 채용, 해고, 징계, 승진 등의 결정을 자동화하는 데 사용되지 않아야 함
의무	고용주는 사용하기 전에 모든 알고리즘 및 데이터 수집에 대한 영향 평가를 실시하고 이를 주 노동청에 제출하여야 함 고용주는 평가 전반에 걸쳐 근로자와 논의함. 전자 생산성 시스템은 구현 전 주 노동청의 검토를 받아야 함 이 법은 노동청에 의해 집행되며 근로자는 개인소송권을 가짐.

출처: Kalra, Ash. 2022. "The Workplace Technology Accountability Act." <https://www.etui.org/sites/default/files/2022-05/WTAA%20summary.pdf> (검색일: 2023. 11. 24).

- ▷ 캘리포니아 주, 직장기술책임법(Worker rights: Workplace Technology Accountability Act, California AB 1651) 채택: 캘리포니아 주는 특정한 시공간에서만 전자 모니터링과 자동화된 의사결정 시스템을 사용할 수 있도록 제한하는 법안을 채택하였다.
- ▷ 2019년 일리노이 주, 인공지능 비디오 면접법(Illinois Artificial Intelligence Video Interview Act in 2019) 채택: 일리노이 주에서는 직원 채용 시 AI 화상 면접을 시행할 경우, 사전 고지와 동의 를 의무화한 AI 비디오 면접법을 채택하였다.
- ▷ 2021년 콜롬비아 주, 알고리즘에 의한 차별금지법(stop Discrimination by Algorithms Act of 2021, SDAA Bill 24-558) 채택: 컬럼비아 주에서는 고용 등 주요결정 시 인종, 성별, 장애 등의 개인 특성에 의해 차별받지 않도록 하는 내용의 차별금지법을 제정하였다.

〈표 4-6〉 DC 알고리즘에 의한 차별금지법이 요구하는 새로운 사항

영역	주요내용
서면 계약	“서비스 제공업체”(“알고리즘 적격성 결정 또는 알고리즘 정보 가용성 결정”을 수행하는 공급업체)에 의뢰하는 기업은 법을 준수하기 위해 해당 제공업체와 서면 계약을 체결해야 함.
공지 의무	기업은 영어 및 500명 이상의 DC 거주자가 사용하는 기타 언어로 된 한 페이지짜리 공지를 작성해야 함. 해당 통지는 해당 기관의 웹사이트에 공개되어야 함. 또한 기업은 알고리즘 기반 결정을 개인에게 적용하기 전에 이 통지를 제공해야 함.
외부 전문가에 의한 감사	기업은 알고리즘 사용 관행이 보호 특성을 차별하는지 확인하기 위해 최소 5년 동안 연간 감사를 실시하고 유지해야 함. 이러한 감사에 대해 일반적으로 인정되는 지침은 아직 없기 때문에 이러한 서비스는 비즈니스 자체의 비용을 들여 외부 전문가가 수행해야 함.
연례보고서 작성	기업은 알고리즘의 성능 지표, 알고리즘 사용 이유, 기업이 접수하는 알고리즘 결정 불만사항 공개를 포함하는 연례 보고서를 법무장관실에 제출해야 함.

출처: DC Chamber of Commerce. 2023. “DC Chamber of Commerce Small Business Action Alert: Stop Discrimination By Algorithms Act Of 2021.” <https://dcchamber.org/algorithms-act/> (검색일: 2023. 11. 24).

- ▷ 2023년 뉴욕시, 자동화된 고용 결정 도구 조례(Automated Employment Decision Tools Law): 뉴욕시에서는 자동화된 고용 결정도구법이 2023년 발효되었다. 편견감시법(Bias Audit Law)으로 알려진 해당 조례는 뉴욕 거주자를 고용하는 기업이 채용과정에서 자동화 알고리즘을 사용할 경우 이를 고지하여야 할 의무를 부과한 법이다. 또한 해당 알고리즘은 소프트웨어 편향성에 대해 제3자 인증을 받아 그 결과를 공개하여야 한다.
- ▷ 뉴저지 주, 고용차별 최소화를 위해 고용 결정 시, 자동화 도구 사용을 규제하는 법안(New Jersey Bill to Regulate Use of AI Tools in Hiring Decisions, A4909) 채택: 뉴저지 주는 고용 시 차별의 최소화를 위해 고용 결정 과정에서 자동화 도구의 사용을 금지하였다.

〈표 4-7〉 미국 주 선거 및 선거운동에서의 AI 사용 적용 법령

주	법령	제정연도	내용
캘리포니아	Cal. Election Code § 20010	2019	선거 60일 전 허위 사실을 포함한 미디어 게재 금지 (해당 미디어가 조작되었다는 사실이 공개되지 않는 한)
미네소타	Minn. Stat. § 609.771	2023	선거 90일 전 선거에 영향을 미치는 딥페이크 미디어 게재 금지
텍사스	Tex. Election Code Ann. § 255.004	2019	선거 30일 전 선거에 영향을 미치는 딥페이크 영상 게시 금지
워싱턴	Wash. Rev. Code § 42.62.020	2023	선거에 영향을 미치는 합성 미디어에 대한 신고
위스콘신	Wis. Stat. § 12.05	1973	선거에 영향을 미치는 후보자 또는 선거에 대한 허위 사실 게재 금지

출처: NCSL. 2023. “Artificial Intelligence (AI) in Elections and Campaigns.” (December 7)  
<https://www.ncsl.org/elections-and-campaigns/artificial-intelligence-ai-in-elections-and-campaigns>  
 (검색일: 2023. 11. 08).

〈표 4-8〉 2023년 선거 및 선거운동에서 AI 관련 법령

주	법령	진행상황	내용
애리조나	SB 1565	기각	투표 기계 및 소프트웨어에 AI 사용 금지
일리노이	SB 1742	실패	선거 30일 전 선거에 영향을 미치는 딥페이크 영상 활용 금지
인디애나	HB 1464	실패	선거에 영향을 미치는 조작된 미디어에 대한 신고
미시간	HB 5141	도입	AI가 생성한 정치 광고에 대한 신고 (HB 5142는 공개하지 않는 것을 중범죄로 규정하고 HB 5143은 AI를 정의)
미시간	HB 5144	도입	선거 90일 전 허위 사실을 포함한 미디어 게재 금지 (해당 미디어가 조작되었다는 사실이 공개되지 않는 한) (HB5145는 이를 중범죄로 규정)
미네소타	HB 1370	제정	선거 90일 전 선거에 영향을 미치는 딥페이크 영상 게시 금지
미네소타	SB 1394	실패	선거 90일 전 선거에 영향을 미치는 딥페이크 영상 게시 금지
뉴저지	AB 5510	실패	선거 90일 전 허위 사실을 포함한 미디어 게재 금지 (해당 미디어가 조작되었다는 사실이 공개되지 않는 한)
뉴저지	SB 366	실패	선거 60일 전 허위 사실을 포함한 미디어 게재 금지

			(해당 미디어가 조작되었다는 사실이 공개되지 않는 한)
뉴저지	SB 3926	실패	선거 또는 정책 토론을 허위로 묘사하는 영상 사용 금지
뉴욕	AB 6790	실패	선거 60일 전 선거에 영향을 미치는 합성 미디어 사용 금지
뉴욕	AB 7106	실패	합성 미디어가 포함된 정치 보도와 관련하여 해당 사항 고지
뉴욕	AB 7904	실패	AI가 포함된 정치 보도와 관련하여 해당 사항 고지
뉴욕	SB 6638	실패	합성 미디어가 포함된 정치 보도와 관련하여 해당 사항 고지
뉴욕	SB 6859	실패	선거에 영향을 미치는 합성 미디어에 대한 신고
텍사스	SB 1044	실패	선거 90일 전 선거에 영향을 미치는 딥페이크 영상 게시 금지
워싱턴	SB 5152	제정	선거에 영향을 미치는 합성 미디어에 대한 신고

출처: NCSL. 2023.

## 다. 소결

### ○ 미국 내 AI 관련 규제 초점

- 미국은 AI 산업을 선도하기 위해 AI 발전 전략 중심의 로드맵을 제시하였다.
- AI 기술이 비약적으로 발전함에 따라, 미국 정부는 이에 관한 규제를 도입하고 있다. 미국 정부가 관련 영역에서 주요하게 다루는 바는 개인 권리 보호, 차별금지, AI 기술을 이용한 안보 사항이다.

### ○ 미국 내 선거, 정당에 관한 생성형 AI 창작물 유포 관련 논의

- 생성형 AI 창작물 제작 및 유포와 관련하여, 전문가는 총 4개 분류에서 문제점을 예측하고 있다. 첫째, 허위 정보의 양 증가, 둘째, 허위 정보의 질 향상, 셋째, 허위 정보의 개인화 증가, 넷째, 타당하지만 허위 정보의 비자발적인 생성이다.
- 2023년 1월부터 현재까지 발생한 미국 내 선거, 정당에 관한 생성형 AI 창작물 유포 사례들은 미국 내뿐 아니라 전 세계적 관심을 끌었다. 하지만 미국의 ‘표현의 자유’에 대한 사회적 보장과 지지가 높

은 국가로 이에 관한 법적 규제에 대해 신중한 상황이다. 이미 허위 정보에 대한 제재에 사회적 합의가 있기 때문에 생성형 AI 창작물이 허위 정보에 해당된다면 이의 법 규정을 따를 수 있다.

- 온라인 상 유통하는 정보에 대한 플랫폼 사업자의 책임은 1996년 통신품위법(CDA: Communications Decency Act) 제230조를 통해 규제된다. 통신품위법 제230조는 플랫폼 사업자에 대한 면책조항을 담고 있다. 이는 결정적으로 유럽연합의 허위 정보 규제 사례와 다른 부분이다.<sup>115)</sup>
- CDA 230조 조문: “대화형 컴퓨터 서비스의 제공자 또는 사용자는 다른 정보 콘텐츠 제공자가 제공한 정보의 게시자 또는 발표자로 간주되어서는 안 된다(no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider).”
- 즉 이것은 직접 발신한 정보가 아닌 제3자가 발신한 정보를 편집, 삭제한 경우에는 면책대상이 된다는 것이라는 것으로, 소셜네트워크서비스나 전자게시판과 같은 플랫폼 사업자가 제공하는 서비스에 게재된 명예훼손에 해당하는 기사나 댓글에 대해서 플랫폼 사업자에게는 책임을 묻지 않는다는 내용이다.<sup>116)</sup>
- 미국 사회는 생성형 AI 창작물 사례에 대해 판별 가능성과 투명성의 관점에 주목하였다. 첫째, 영상의 가짜 여부 판별 가능성이다. 이는 일반인이 영상 혹은 이미지를 접했을 때 판별 가능성과 전문가가 접했을 때 판별 가능성에 관한 논의가 중점 사항이 되었다. 2023년 초기 창작물(<그림 3-2> 도널드 트럼프 전 대통령이 체포되는 가짜 영상 사례 등)에서 화면 일부 영상의 일그러짐이나 손가락의 모양 등을 통해 해당 영상이 실체가 아닌 창작임을 알 수 있었다. 하지만 2023년 중순 이후부터의 창작물(<그림 3-4> 트럼프가 앤서니 파우치 전 소장과 포옹하는 허위 사진 등)에서는 진위여부의 판별이 어려워졌다. 해당 영상 및 이미지는 콘텐츠의 내용상 사실이기 어려운 상황을 미루어 창작물임을 짐작할 뿐 기술적 분석 없이 식별할 수 없게 된 것이다. 생성형 AI의 빠른 발전으로 미루어 2024년 이후부터는 전문

115) 이에 대한 논의는 다음 절에서 다룬다.

116) 고선규·이소영·김형호·송재민. 2021. “온라인 가짜뉴스 관련 선거관리제도 개선방안에 대한 연구”. 『한국선거협회』: 65.

가의 기술 분석으로도 영상의 진위여부 판별이 어려울 수 있다.

- 따라서 미국 사회는 선거 및 정당과 관련한 생성형 AI 창작물의 투명성에 집중하고 있다. 즉 AI에 의해 제작되었음 혹은 창작물임을 알 수 있도록 하는 표식에 대한 제도화가 필요함에 공감하고 있다. 하지만 내용상 정치와 민주주의에 관한 표현의 자유와 주장을 억제하기 위해, 생성형 AI 창작을 제도적으로 제어해야 할 필요성에 관해 논의하지는 않는다.

○ 미국 군사, 안보와 관련한 생성형 AI 창작물 유포 관련 논의

- 군사, 안보와 관련한 생성형 AI 창작물은 사회에 극단적인 악의적 영향을 미칠 우려가 있다. 군사, 안보와 관련한 생성형 AI 창작물은 국제 수준에서 악의적 의도로 제작될 가능성을 배제할 수 없기 때문에, 투명성을 높이는 것만으로 대응할 수 없는 것이다.
- 이에 미국정부는 개인의 권리보호나 차별을 방지하기 위한 위험관리 이외에 국방, 군사 영역에 관한 법률을 입법하였다.

## 4.2 영국

### 가. AI 관련 입법 추진 개요

□ 명예훼손법(Defamation Act 2013)

○ 보통법상 명예훼손죄 폐지와 명예훼손법 개정

- 영국의 보통법(Common Law) 상 명예훼손죄는 크게 신성모독적 명예훼손(blasphemous libel)과 외설적 명예훼손(obscene libel), 선동적 명예훼손(seditious libel), 사인 간 명예훼손(criminal libel, 또는 defamatory libel)으로 분류할 수 있다.<sup>117)</sup>
- 가짜뉴스의 처벌 근거가 될 가능성이 있었던 구성요건은 선동적 명예훼손과 사인 간 명예훼손이라 할 수 있다. 그러나 영국에서는 2009년까지 약 400년간 명예훼손죄로 판례로 남은 사건은 119건에 불과

117) 정태호·김훈집. 2015. “영국의 명예훼손죄 폐지와 그 교훈.” 『경희법학』 50집 제1호: 5.

하다. 또한 이 법은 주로 사회적 지위가 높은 사람들이 자신들의 기득권을 유지하기 위해 이용해왔다.<sup>118)</sup>

- 명예훼손죄는 그 표현이 문서상의 형태로 지속성을 가지고 적시된 경우에만 형사처벌이 가능했고, 말이나 몸짓 등 무형으로 이루어진 명예훼손은 처벌하지 못했다. 토론과 표현의 자유를 위축시킨다는 비판이 높아지자, 2010년 1월 1일자로 모든 유형의 명예훼손죄는 폐지되었다. 명예훼손죄 폐지와 더불어 2013년에는 명예훼손법을 개정하여 표현의 자유와 언론의 자유의 보장을 강화했다.<sup>119)</sup>

### ○ 명예훼손법의 주요 내용

- 2013년 개정된 명예훼손법은 총 17개의 조문으로 구성되어 있다. 이는 명예훼손죄 폐지 취지에 맞추어 민사상 명예훼손으로 인한 손해배상의 요건과 절차 등을 규정하고 있으나 여기에서는 생성형 AI 관련 사안만을 다루고자 한다.

#### ▷ 일반적 면책사유 : 본질적 진실과 정직한 의견<sup>120)</sup>

- 명예훼손이라고 주장하는 진술에 담긴 비방이 본질적으로 진실(substantially true)에 해당할 때는 면책될 수 있으나, 그러한 비방 내용이 본질적으로 진실이라는 입증은 이를 주장하는 측에서 해야 한다(제2조 제1항). 그리고 하나 이상의 비방이 본질적으로 진실임이 입증되지 않으면 본질적으로 진실임이 입증된 비방을 참작하여, 본질적으로 진실임이 입증되지 않은 비방이 원고의 명예를 심각하게 훼손하지 않을 때도 면책사유가 적용된다(동조 제3항). 명예훼손이라고 주장되는 진술이 둘 이상의 비방 내용을 담고 있는 경우에도 같다(같은 조 제2항).
- 개정된 명예훼손법은 “본질적 진실” 외에 “정직한 의견(honest opinion)”도 면책사유로 규정하고 있다. 공익적 사항을 논평하였을 때는 면책된다는 보통법상 “공정한 논평(fair comment)의 원칙”을 다소 완화한 것으로, 반드시 공익적 사안에 대한 의견이 아니어도 면책사유가 될 수 있다. 그 요건으로는 우선 진술이 의견표명이어야 하며, 그러한 진술이 일반적 또는 구체적 용어로 의견의 근거를 나

118) Collins, Matthew. 2014. “Collins On Defamation.” : 493, 498-499(허순철. 2015. “영국의 명예 훼손법 개정과 그 의미.” 『공법학연구』 16집 제4호: 128-129. 재인용).

119) 정태호·김훈집. 2015: 14.

120) 허순철. 2015.: 135.



타내야 한다(제3조 제2항, 제3항). 그리고 정직한 사람이라면 그러한 진술이 공표될 당시에 존재했던 사실을 근거로 의견을 표명했을 것으로 인정되거나, 그러한 진술 이전에 공표된 것으로서 특권으로 보호받는 진술 내에서 인정된 사실을 근거로 의견을 표명했을 것으로 인정되어야 한다(같은 조 제4항). 정직한 의견 여부에 관한 입증할 책임은 소를 제기한 측에 있다(같은 조 제5항).

- 한편 “특권으로 보호받는 진술(privileged statement)”은 그 진술에 대해 출판 책임을 지는 사람이 공익적 사안에 대한 출판물(제4조)이거나 동료의 심사를 거친 과학저널 또는 학술저널상의 진술(제6조), 1996년도 명예훼손법 제14조 상의 절대적 특권에 의해 보호되고 있는 재판절차 보고, 같은 법 제15조 상의 기타 제한적 특권에 의해 보호되고 있는 기타 보고임을 입증하는 경우를 말하는데, 이에 대한 입증은 진술자에게 있다. 그리고 공익적 사안에 대한 출판물일 때 피고인이 그 진술이 공익에 해당한다고 합리적으로 인지하였음을 입증해야 하며(제4조 제1항 제b호), 동료의 심사를 거친 과학저널 또는 학술저널상의 진술은 과학적·학술적 대상에 대한 진술이면서, 출간 전에 그러한 진술의 과학적·학술적 가치에 대한 독자적인 심사가 학술지 편집부나 관련분야 전문가에 의해 이루어졌어야 한다(제6조 제2항, 제3항).

▷ 웹사이트 운영자에 대한 면책규정<sup>121)</sup>

- 개정된 명예훼손법은 웹사이트 운영자에 대한 면책규정을 두고 있다(제5조). 우선 웹사이트에 명예훼손에 해당하는 진술을 직접 게시하지 않은 웹사이트 운영자는 면책이 원칙이다(제2항). 다만 원고가 그 진술 게시자의 신원 확인이 불가하고 운영자에게 그러한 진술이 자신의 명예를 훼손한다는 민원을 제기하였음에도, 웹사이트 운영자가 관련 규정에 따라 그러한 민원에 회신하지 않았음을 입증하였으면 면책되지 않는다(제3항). 이때 원고가 게시자에 대해 소를 제기할 수 있을 정도로 충분한 정보를 가진 경우에만 게시자의 “신원을 확인” 할 수 있는 것으로 본다(제4항). 그리고 민원에는 민원인의 이름을 명시해야 하며, 문제 되는 진술과 그 진술이 명예를 훼손하는 이유를 적시하고, 해당 진술이 게시되어 있는 구체적인 위치를 명시해야 하며, 관련 규정에 따른 정보를 포함하고 있어야 한다(제6항). 관련 규정은 민원이 접수되면 운영자가 취해야 할 조치에 관한 규정, 그러한 조치의 기한에 관한 규정, 기한 만료 이후에 제기된 민원을

121) 허순철. 2015: 138.

기한 전에 제기된 것으로 간주하는 재량을 법원에 부여하는 규정 등으로 국무부가 만든 것을 말한다(제5항 및 제10항).

- 한편 웹사이트 운영자가 문제 진술의 게시가 악의를 갖은 행위였음을 원고가 입증하면 본 조항에 따른 면책효력은 상실된다(제11항). 다만 웹사이트 운영자가 그 웹사이트에 다른 사람이 게시한 진술을 관리한다(moderate)는 것이 사실이라는 이유만으로는 본 규정에 따른 면책사유의 적용이 불가하다고 주장할 수 없다(제12항).
- 본 규정에 대해서는 제1조의 “중대한 해악” 과 “금전적 손실” 과 더불어 면책사유가 판례를 통해 더욱 명확해야 할 필요가 있다는 비판<sup>122)</sup>이 제기된다. 특히 “관리한다 (moderate)” 가 무엇을 의미하는지 불분명하다는 것이다. 입법취지상 편집권(editorial control)을 행사하지 않고 단지 제3자가 게시한 글을 관리하는 수준에 불과한 웹사이트 운영자를 보고하려는 것으로 보이나, 웹사이트에 따라서는 자동으로 욕설이나 금칙어가 삭제되거나 차단되도록 설정해두기도 하고, 특정 관리인이 일일이 게시물의 내용을 검토하고 삭제하기도 하여, 이를 더욱 상세하게 구분하여 규정할 필요가 있다는 것이다.

#### □ 하원의 디지털문화미디어체육부 위원회

- 영국의 경우 생성형 AI 관련 허위 정보(fake news)에 특화해 대응하는 법적 제도적 조치는 사실상 부재했다. 2017년 1월 하원의 디지털문화미디어체육부 주도로 허위 정보에 대처하기 위한 조사위원회를 구성했다.<sup>123)</sup>

- 2016년 미국 대통령 선거과정에서 드러난 허위 정보 급증 현상이 민주주의 의사결정과정을 침해할 수 있다는 위기의식이 고조됨에 따라 하원의 디지털문화미디어체육부가 구성한 조사위원회는 허위 정보 관련 대책 마련을 위한 조사를 시작했다. 조사위원회가 제시한 중점 확인 사항은 다음과 같다.

▷ [허위 정보의 정의] 허위 정보란 무엇인가? 다소 편향되더라도 적법한 기사가 선전 선동 및 거짓으로 왜곡되는 지점은 어디인가?

122) 허순철. 2015: 142.

123) UK Parliament. 2019. “Disinformation and ‘fake news’: Final Report published.” (February 18) <https://committees.parliament.uk/committee/378/digital-culture-media-and-sport-committee/news/103668/fake-news-report-published-17-19/> (검색일: 2023. 11. 12).

- ▷ [허위 정보의 영향력과 전통적 매체의 역할] 허위 정보는 대중의 세계관에 어떤 영향을 미치는가? 언론의 전통적 역할에 대한 대중의 인식에 어떠한 영향을 미치나?
- ▷ [허위 정보에 대한 반응] 세대별, 성별, 사회적 배경의 차이 등에 따른 허위 정보에 대한 반응에 차이가 있나?
- ▷ [허위 정보의 수익 창출] 허위 정보를 활용해 인터넷상의 더 많은 조회수를 확보함으로써 광고주로부터 수익 추구 방식이 허위 정보의 급증에 기여했나?
- ▷ [허위 정보 근절을 위한 생성형 기술의 활용 가능성 탐색] 검색 엔진 및 소셜미디어 플랫폼의 역할은 무엇인가? 허위 정보 근절을 위한 컴퓨터 생성 알고리즘(computer-generated algorithm)의 활용 가능성은 어느 정도인가?
- ▷ [미디어 리터러시 교육] 각기 다른 정보원을 가진 뉴스에 접근하는 방식에 대한 교육은 어떻게 이루어져야 하는가? 영국 공공방송 및 신문 독자층의 전통적 역할을 고려할 때, 영국 시민과 외국 시민 간 허위 정보를 받아들이는 정도의 차이가 있는가?
- ▷ [허위 정보 대응 해외 사례] 외국 정부는 어떻게 허위 정보에 대응하고 있는가?

- 조사단 위원인 보수당 콜린스(Damian Collins) 의원은 “급증하는 허위 정보는 민주주의를 위협할 뿐 아니라 언론에 대한 신뢰를 잠식”할 수 있음을 강조했으나 2017년 총선의 실시와 함께 위원회 활동은 정지되었다.

○ 영국은 학교 차원에서 허위 정보를 구분하는 역량 향상을 위한 교육과정의 필요성이 대두되었다. 하며 이와 관련, 경제협력개발기구(OECD) 슬라이허(Andreas Schleicher) 교육국장은 “진실과 진실이 아닌 것을 구분하는 것은 중대한 판단 능력”이라며, “알고리즘은 (다른 배경을 지닌) 사람들 간 공통 기반을 찾기 위해 토론하는 장을 창출하는 역할보다는, 서로 비슷한 사람들만을 연결하는 경향이 있다”고 지적하였다. OECD는 젊은이들이 다양한 관점으로 세상을 바라보고, 다양한 아이디어를 인정하며, 다양한 문화에 개방적인 태도를 가질 수 있는 능력을 새로운 글로벌 역량 측정 지수에 추가하여 관련 역량을 위한 교육의 중

요성을 강조했다.<sup>124)</sup>

- 지역 신문들의 주도로 허위 정보에 대응하기 위해 ‘Fighting Fake News’ 캠페인을 벌이고 있다. 해당 캠페인은 영국 뉴스 출판인들을 대변하는 뉴스미디어협회(News Media Association)의 후원으로 진행되었다.<sup>125)</sup> 학계에서도 제도적 대응 방안을 마련 중이며, 왕립학회(Royal Society), 영국학술원(British Academy), 국립데이터 과학협회(National Institute for Data Science) 등이 합작해 개인 데이터를 관리하는 규범적 틀을 구성하고 있다.
- 독일, 프랑스, 네덜란드에서는 Facebook이 기존 언론사와 협력하여 문제가 있는 기사가 뉴스피드에 표시되도록 하고 있다. 또한 당시 대통령 선거를 앞둔 프랑스의 경우, 페이스북은 정치적 허위 정보를 퍼뜨리는 것으로 의심되는 3만 개의 자동화된 계정을 정지시켰다. 하지만 영국에서는 이러한 노력이 더 늦어지고 있다는 비판이 존재한다.<sup>126)</sup> 영국에서는 2019년 1월 11일에 이러한 조치가 이루어졌다. 거짓으로 신고된 콘텐츠는 Facebook의 뉴스피드에서 더 낮은 위치에 표시되며, 사용자는 공유 중인 콘텐츠가 거짓인지 아닌지를 확인할 수 있지만, 사용자가 원할 경우, 공유를 중단하지는 않는다.<sup>127)</sup>

#### □ 인터넷서비스사업자의 대응 방안

- 2016년 9월 Amazon, DeepMind, 페이스북, IBM, Microsoft, 구글 등이 합작해 ‘인류와 사회를 위한 인공지능 파트너십 (Partnership on Artificial Intelligence to Benefit People and Society)’ 이라는 윤리 기구를 설립하였다. 인터넷서비스사업자들은 특히 2017년 총선을 앞두고 허위 정보

124) Coughlan, Sean. 2017. “Schools should teach pupils how to spot fake news.” BBC (March 18) <https://www.bbc.com/news/education-39272841> (검색일 : 2023. 10. 15).

125) THE WEEK. 2017. “Could fake news impact the general election?.” (May 4) <https://theweek.com/84203/could-fake-news-impact-the-general-election>. (검색일: 2023. 10. 12).

126) THE WEEK. 2017. “Could fake news impact the general election?.” (May 4) <https://theweek.com/84203/could-fake-news-impact-the-general-election>. (검색일: 2023. 10. 12).

127) Schroeder, Stan. 2019. “Facebook tackles fake news in the UK with a new fact-checking service.” Mashable (January 11) <https://mashable.com/article/facebook-fake-news-uk> (검색일: 2023. 12. 01).

확산을 차단하는 방안을 마련하라는 압박을 받았다.

- 이에 페이스북은 <그림 4-1>과 같이 영국 유권자들이 허위 정보에 혼란을 겪지 않고 적절히 대응할 수 있도록 광고 캠페인을 전개하였다.

<그림 4-1> 페이스북이 영국 일간지에 낸 허위 정보 예방 캠페인(2017년 5월)



출처: Collins, Katie. 2017. “Facebook posts fake-news ads in newspapers ahead of UK vote.” (May 8) <https://www.cnet.com/tech/tech-industry/facebook-posts-fake-news-ads-in-newspapers-ahead-of-uk-election/?openLogin=1> (검색일: 2023. 12. 01).

- 페이스북이 영국 유권자들에게 알려준 허위 정보 구분 요령은 대부분 일반적 수준으로 다음과 같다(<그림 4-1> 참조).

- 헤드라인(제목)에 현혹되지 말 것
- 웹페이지주소(URL)를 확인할 것
- 뉴스 정보를 확인할 것
- 서식이 낯설 경우 주의
- 사진의 조작 여부 유의
- 날짜 확인
- 증거 확인

- 기타 관련 기사 확인
- 단순 가십성 기사 혹은 농담인지 확인
- 일부 뉴스는 의도적으로 조작된 것일 수 있음에 주의

#### □ 팩트체크(fact check) 정책

○ Full Fact 및 First Draft 등 비영리 매스컴 연합기관은 총선을 앞두고 허위 정보를 감시하는 역할을 하고 있다. 팩트체크 활동에 페이스북과 구글이 부분적으로 후원하였으며, 전문가 그룹(fact-checkers)에는 경제학자, 통계학자 등도 포함되었다.<sup>128)</sup>

- 팩트체크는 일종의 정보전(information war)과 다름이 없다. Full Fact는 특정 뉴스가 허위 정보인지 아닌지 최종 판단을 내리기 전 서로 다른 전문가에게 교차확인 작업을 수행한다.

- Full Fact는 페이스북에 올라온 포스트 하나를 접하는 사람의 수가 정치 시사 관련 잡지를 구독 혹은 TV의 정치 관련 프로그램을 시청하는 사람의 수만큼 많을 수 있다고 관측하였다.

○ 영국의 Facebook 사용자는 오해의 소지가 있는 콘텐츠에 플래그를 지정할 수 있다. 그러면 Full Fact의 팩트체커가 해당 콘텐츠의 정확성을 확인하여 사실, 거짓 또는 두 가지가 혼합된 것으로 평가한다.<sup>129)</sup>

- Full Fact는 위험한 암 ‘치료법’, 테러 공격 후 퍼지는 거짓 이야기, 선거를 앞두고 투표 방법에 대한 가짜 콘텐츠 등 사람들의 안전을 해치거나 민주적 절차를 훼손할 가능성이 가장 높은 잘못된 정보를 검토하고 평가하는 데 초점을 맞춘다.

- Full Fact와 Facebook의 파트너십은 다음과 같이 작동한다. Facebook은 Full Fact가 무엇을 확인하고, 팩트체커가 어떤 내용을 작성하고,

128) Booth, Robert. 2017. “Truth seekers: inside the UK election’s fake news war room.” Guardian (May 19) <https://www.theguardian.com/politics/2017/may/19/truth-seekers-inside-the-uk-elections-fake-news-war-room> (검색일: 2023. 10. 12).

129) Schroeder, Stan. 2019. “Facebook tackles fake news in the UK with a new fact-checking service.” Mashable (January 11) <https://mashable.com/article/facebook-fake-news-uk> (검색일: 2023. 12. 01).

어떤 등급을 부여하는지에 대한 통제권이 없다. 또한, Full Fact의 팩트체크는 게시되기 전 Facebook 직원의 확인이나 승인 없이 자동으로 Facebook에 업로드된다.

#### □ 정부의 허위 정보 대응 가이드라인 및 윤리기준

- 영국 정부는 온라인 허위 정보에 대응하기 위해 온라인 정보에 대한 가이드라인을 제시하고, 허위 정보 대응 캠페인을 진행하며, ‘짐승에게 먹이를 주지마라(Don’t Feed the Beast)’ 라는 사실관계 확인을 위한 사이트를 운영하고 있다.
- 이용자 지침에는 온라인 활용 관련 교육 참가, 유해 콘텐츠 게재 등 행동 멈추기, 긍정적 행위 지속 및 증진 등의 내용이 포함되어 있다.<sup>130)</sup>
  - 허위 정보에 대한 대응방안으로는 허위 정보에 대한 저항력 증가, 허위 정보에 대한 효과적인 확인을 위한 방법, 허위 정보 유통 자제, 허위 정보가 소속기관에 영향을 줄 경우의 대응능력 강화 등을 제안한다.
  - 또한, 허위 정보에 대응하기 위한 구체적인 방법으로, <그림 4-2>와 같이 허위 정보 인식, 빠른 경고를 위한 미디어 모니터링 활용, 상황에 판단, 영향 분석, 전략적 커뮤니케이션, 결과 분석 등의 프로세스를 제안하고 있다.

130) HM Government. 2020. “Online Harms White Paper: Full Government Response to the Consultation.” (December) [https://assets.publishing.service.gov.uk/media/5fd8af718fa8f54d5f67a81e/Online\\_Harms\\_White\\_Paper\\_Full\\_Government\\_Response\\_to\\_the\\_consultation\\_CP\\_354\\_CCS001\\_CCS1220695430-001\\_V2.pdf](https://assets.publishing.service.gov.uk/media/5fd8af718fa8f54d5f67a81e/Online_Harms_White_Paper_Full_Government_Response_to_the_consultation_CP_354_CCS001_CCS1220695430-001_V2.pdf) (검색일: 2023. 11. 10).

〈그림 4-2〉 영국 정부의 허위 정보 대응가이드



출처: RESIST model: a quick guide. “RESIST” <https://cyberpolicy.nask.pl/wp-content/uploads/2020/04/RESIST-Counter-disinformation-Toolkit.pdf> (검색일: 2023. 12. 01).

- 영국 정부는 또한 허위 정보 대응사이트인 Don’ t Feed the Beast를 운영하고 있다.<sup>131)</sup>
  - 공동체가 허위 정보를 인지하도록 돕는 것을 목적으로 하며, 정보를 공유하기 전에 반드시 체크해야 할 다섯 가지 항목을 제시하고 있다.
    - ▷ 이 다섯 가지 체크리스트는 출처(source), 제목(headline), 팩트 분석(analysis), 이미지나 영상의 사실관계 확인(retouched), 문법의 오류나 오타자 확인(error)이다.
- 한편, 영국의 언론에 대한 독립 규제 기구인 IPSO(Independent Press Standard Organization)는 영국 하원의 디지털문화미디어체육부(DCMS) 위원회의 요구로 언론에 대한 규제 내용을 담고 있는 자율규약을 제정하였다.<sup>132)</sup>
  - 규약은 언론보도가 허위 정보와 구분되어야 하며, 허위 정보와 언론 보도를 구별하는 기준을 제공하고 뉴스가 효율적으로 전달되는 체계

131) 고선규 외. 2021. 133.

132) 김재선. 2020. “허위정보 규제를 위한 행정법적 대응방안 -영국의 규제 논의를 중심으로.” 『법학논집』 24(3): 223-248.



를 만들어 언론의 신뢰도를 높여야 하며, 허위 정보 여부가 문제가 될 때는 IPSO가 주도하여 그 허위 여부를 판단하여야 한다는 내용을 담고 있다.

□ 인터넷 자율규제 기구<sup>133)</sup>

- 2019년의 온라인 유해물 백서(Online Harms White Paper)<sup>134)</sup> 발표 이전 영국의 인터넷 규제는 주로 민간기구인 IWF(Internet Watch Foundation)<sup>135)</sup>가 자율적으로 시행했다. 정부로부터 독립적인 자율규제 기구이나 내무부, 디지털문화미디어체육부, 법무부, 교육부와 긴밀한 협력 체계를 구축하고 있는 IWF는 아동 성 착취물 근절을 목적으로 하는 단체로, 홈페이지, 전화, 팩스 등을 활용해 신고받고 모든 신고는 익명으로 접수하는 방식을 통해 규제를 진행해 왔다.
- IWF는 시민들의 신고를 받고 해당 자료를 검토해 불법 여부를 잠정적으로 결정한다. 신고된 콘텐츠의 내용이 불법이라고 판단되면, IWF는 국내 콘텐츠의 경우에는 영국 경찰에, 해외 콘텐츠의 경우에는 해당 해외 법집행 기구에 내용을 고지하는 방식으로 자율규제를 운영하고 있다.
- IWF는 영국 내 인터넷 서비스 사업자(ISP)들에게 불법으로 결정된 콘텐츠를 서버에서 삭제하도록 요청한다. ISP가 IWF의 콘텐츠 삭제요청을 따르지 않는 경우 이에 대한 법적 책임이 존재한다. 영국에 서버를 두고 있지 않은 URL의 경우 해당 URL이 등록된 국가의 회원사에 이에 대한 삭제를 요청한다.
- 정보통신청(Office of Communication, Ofcom)은 정부 위탁 규제기관으로 온라인 플랫폼에 대한 법적 규제 및 자율규제에 해당하는 방송 통신 및 우편 산업에 대한 규제를 담당하는 기관이다. 정부 위탁 기관이나 정부

133) 전학선·정필운·심우민·윤진희. 2020. “글로벌 플랫폼사업자의 자율규제 실태 및 협력 방안 연구.” 방송통신심의위원회.

134) Julian, Vikki. 2020. “IPSO Blog: Government responds to Online Harms consultation.” ipso (February 13) <https://www.ipso.co.uk/news-press-releases/blog/ipso-blog-government-responds-to-online-harms-consultation/> (검색일: 2023. 12. 01).

135) IWF. “How AI is being abused to create child sexual abuse imagery.” <https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/> (검색일: 2023. 12. 01).

로부터의 독립성이 보장된다.

- 영국에서 온라인 콘텐츠에 대한 규제는 IWF와 같은 민간 자율규제기관이 담당했으나 영국 정부의 온라인 플랫폼에 대한 규제 강화결정과 함께 Ofcom의 중요성과 권한이 커지고 있다. 새롭게 만들어지는 「온라인 안전법」의 주무부처로서 온라인 플랫폼의 유해 콘텐츠를 직접 삭제하는 등의 권한은 없지만 사용자의 접속을 차단하는 등 강력한 권한이 부여될 예정이다.
- 인터넷 서비스 사업자 단체인 ISPA(Internet Services Providers' Association) 또한 자율규제 관련 행위자로 회원사를 대표해 규제 당국과 정기적인 접촉을 하며 규제 당국의 인터넷 규제 가이드라인에 순응하고 있다.

#### □ 플랫폼 사업자의 책임 강화

- 2019년 영국 하원 보고서에서는 SNS의 허위 정보에 대응하기 위해 플랫폼 사업자의 책임 강화 방안이 논의되었다. 이 보고서는 플랫폼 사업자가 허위 정보를 인지하고도 이용자 보호조치를 취하지 않을 때 부과되는 책임을 중심으로 플랫폼 사업자의 입증책임의 범위, 플랫폼 사업자에 대한 과세 및 기금 마련 방안, 해외 이용자에 대한 사업자의 책임 등을 논의하고 있다.<sup>136)</sup>
- 플랫폼 사업자에게 책임을 전가하기 위해서는 플랫폼 사업자가 허위 정보가 발생, 유통 및 피해의 가능성이 있다는 것을 사전에 인지해야 한다. 따라서 이것을 판단하는 것이 가능한가에 대한 의문이 제기되었다. 영국 하원 정보위원회 보고서는 SNS 검색 기능을 활용하여 SNS 사용자들의 검색 습관, 인간관계, 주로 활용하는 검색 방법 등을 찾아낼 수 있으므로 플랫폼 사업자는 현재 기술과 서비스만으로도 허위 정보의 발생, 유통 및 발생 가능성을 인지할 수 있다고 평가하고 있다.<sup>137)</sup>

---

136) 김재선. 2020.

137) 김재선. 2020.

- 그러나 플랫폼 사업자가 허위 정보를 인지했다 하더라도 그 정보의 유효성을 판단하기 어렵기 때문에 플랫폼 사업자의 책임을 구체화하기 어렵다는 문제가 있다.<sup>138)</sup>
  - 예를 들어, 2018년 영국 정부가 페이스북이 개인정보 보호조치를 취하지 않았다는 이유로 과징금을 부과하였을 때 페이스북은 그 근거가 충분하지 않다고 주장하며 불복하였다.
  - 그러나 영국 정보위원회는 페이스북이 애플리케이션 개발업자들이 정보주체의 동의 없이 개인정보를 수집할 수 있도록 하였고, 이에 대한 안전조치를 취하지 않았다는 점을 들어 과징금이 부과될 수 있다고 주장하였다.
  
- 영국은 온라인 플랫폼 사업자의 책임성을 강화하는 방안으로 온라인 사업자에게 전체 소득의 2%에 해당하는 디지털서비스세(digital service tax)를 제도화하여 이를 기금으로 형성하고 이 기금을 허위 정보 단속 및 관리에 사용하고 있다. 디지털서비스세는 2020년 4월 20일 매출부터 적용된다.<sup>139)</sup>
  - 디지털서비스세는 유럽연합 집행위원회에서 제안한 과세 제도로, 유럽연합 내에서 기업 활동을 하고 있음에도 고정사업장이 유럽 내 존재하지 않아 세금을 내지 않고 있는 기업을 대상으로 과세하는 제도이다.
  - 영국 사용자에게 소셜미디어 플랫폼, 검색 엔진 또는 온라인 마켓플레이스를 제공하고 이러한 디지털 서비스 활동으로 인해 디지털 서비스세가 발생하는 경우 디지털 서비스세 서비스에 등록해야 한다:
  - 연간 5억 파운드 이상의 글로벌 매출 연간 2,500만 파운드 이상의 영국 매출 기업에게 적용된다. 그룹에 속한 모든 비즈니스의 수익이 이러한 기준을 따른다.
  - 따라서 이 법안은 소셜미디어, 검색 엔진 또는 온라인 마켓플레이스

138) 김재선. 2020.

139) HM Revenue & Custom. 2020. "Digital Services Tax." <https://www.gov.uk/government/publications/introduction-of-the-digital-services-tax/digital-services-tax> (검색일: 2023. 11. 20).

의 영국 사용자로부터 발생하는 수익 일부를 디지털 서비스세의 적용 범위에 포함시킴으로써 소수의 대규모 다국적 그룹에 영향을 미칠 것으로 예상된다.

## 나. 영국의 AI 국가 전략<sup>140)</sup>

- AI는 영국의 경제 환경을 변화시키고 산업을 혁신하고 최고 수준의 공공 서비스를 제공함으로써 영국 전역에서 국민 생활을 개선할 수 있는 잠재력을 제공한다고 확신하고 있다.
- AI는 인류 역사상 중요한 혁신 중 하나일 수 있으며, 정부는 영국이 AI가 가져올 기회에 대비하고 AI 사용 증가로 인한 복잡한 문제를 해결하는 데 앞장서는 것이 경제와 국가안보 모두에 중요하다고 판단한다.
- 영국은 앨런 튜링의 초기 연구부터 최근 딥마인드의 선구적인 발견에 이르기까지 AI 분야에서 길고 뛰어난 역사가 있다. AI 스타트업과 스케일업, 민간 자본 투자, 컨퍼런스 논문 발표 등의 측면에서 영국은 전 세계 AI 국가 중 상위권에 속한다.
- 영국은 2020년 AI 기업에 대한 민간 투자에서 미국과 중국에 이어 세계 3위를 차지했다. 이는 국가 AI 전략은 영국이 현재 강점을 바탕으로 AI의 잠재력을 극대화하면 민간 및 공공 부문 전반의 회복력, 생산성, 성장 및 혁신이 증가한다는 점을 인식하고 영국에서 AI를 위한 단계적 변화의 시작된다는 것을 의미한다.
- AI 분야의 강점을 바탕으로 향후 10년 동안 모든 관련 행위자의 역량을 총동원하기 위한 통합적 접근을 활용한다. 이를 위해 경제, 안보, 보건, 복지를 최우선 과제로 설정했다.
- 영국 정부는 지역 번영에 대한 국가적 야망과 넷 제로, 건강 회복력, 환경 지속가능성과 같은 글로벌 과제를 해결하기 위해 AI 경쟁력을 갖추는 것이 필수적이며 이를 통해 AI 역량은 글로벌 과학 강국으로서 영국

140) HM Government. 2021. “National AI Strategy.” (September) [https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National\\_AI\\_Strategy\\_-\\_PDF\\_version.pdf](https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf) (검색일: 2023. 12. 01).

이 국제적 영향력을 발휘하는 데 필수적이라고 판단한다.

- 영국 국가 AI 전략은 영국의 향후 10년을 준비할 것이며, 향후 10년에 대한 세 가지 가정을 바탕으로 수립되었다.
  - AI의 발전, 발전 및 전략적 우위의 핵심 동인은 사람, 데이터, 컴퓨팅 및 금융에 대한 접근성이며, 이 모두가 거대한 글로벌 경쟁에 직면해 있다.
  - AI는 경제 부분에서 주류가 될 것이며 영국의 모든 부문과 지역이 이러한 전환의 혜택을 누릴 수 있도록 조치가 필요하다.
  - 거버넌스와 규제 체제는 급변하는 AI의 요구에 발맞추어 성장과 경쟁을 극대화하고 혁신에서 영국의 우수성을 주도하며 시민의 안전, 보안, 선택 및 권리를 보호해야 한다.

〈표 4-9〉 영국의 국가 AI 전략 개요

구분	내용				
비전	향후 10년 동안 AI와 과학 분야의 초강대국으로 성장				
영향	모든 지역 및 부문에서 AI 도입의 이점 공유	AI 연구개발 글로벌 리더로서의 지위 유지	AI 분야의 성장을 통한 GDP 성장에 기여	영국의 가치 보호 및 강화	국가안보를 위한 강력한 AI 능력 확보
결과	원칙 1 AI 산업 생태계의 장기적인 요구에 대한 투자		원칙 2 AI가 모든 분야와 지역에 이익이 되도록 보장		원칙 3 효과적인 AI 통제·관리
	영국 공급업체 기반 확대		응용 AI 다양성 증대		영국의 확실한 AI 생태계 확보
	AI기술 경쟁 완화		산업 및 지역에서 더욱 광범위한 AI확대		AI에 대한 대중의 신뢰 향상
	새로운 AI 과학의 획기적인 발전		영국 AI 수출 확대		책임감 있는 혁신 강화
	노동 다양성 확대		윤리적 모범이 되는 공공부문 AI 확보		영국의 AI 글로벌 리더 지위 유지
	새로운 활용 사례에 AI 기술 적용		더 큰 공공가치를 위한 수익 창출		
	영국 AI 기업에 대한 투자 확대				
활동	해당 전략의 정부 활동은 향후 10년간 진행				

출처: HM Government. 2021: 14.

- 따라서 영국의 국가 AI 전략의 목표는 다음과 같다.
  - 과학 및 AI 강국으로서의 리더십을 지속하기 위해 AI 생태계의 장기적 요구에 투자하고 계획하며,
  - 영국에서 혁신의 혜택을 포착하고 AI가 모든 부문과 지역에 혜택을 줄 수 있도록 AI 기반 경제로의 전환을 지원하며,
  - 영국이 혁신과 투자를 장려하고 공공과 영국의 기본 가치를 보호하기 위해 AI 기술의 국내 및 국제 거버넌스를 올바르게 확보하는 것이다.
  
- 이 문서는 AI가 빠르게 변화하고 역동적인 분야임을 인식하고 향후 10년간의 영국의 행동 방향과 영국의 전략적 의도를 제시한다(〈표 4-10〉 참조).

〈표 4-10〉 영국의 국가 AI 전략의 세 가지 원칙과 시기별 계획

	AI 생태계의 장기적 요구에 대한 투자	AI가 모든 부분과 지역에 이익이 되도록 보장	효과적인 AI 거버넌스
단기 (향후 3개월)	<ul style="list-style-type: none"> <li>· 거시경제에서 더 나은 데이터 가용성을 가능하게 하는 정부의 역할 및 체계 발표</li> <li>· 국가 사이버 물리 시스템 체계(Cyber-Physical Infrastructure)의 역할 및 기능에 대한 협의</li> <li>· 교육부 기술부트캠프(Skills Bootcamp)를 통해 AI, 데이터 과학 및 디지털 기술 개발 지원</li> </ul>	<ul style="list-style-type: none"> <li>· NHS AI Lab을 통해 보건 및 사회복지 분야의 AI 기반 기술을 위한 국가 전략 초안 작성</li> <li>· 국방부 AI 방어 전략 수립</li> <li>· IPO를 통해 AI에 대한 저작권 및 특허에 대한 협의 시작</li> </ul>	<ul style="list-style-type: none"> <li>· CDEI(데이터윤리혁신센터) AI 인증 로드맵 발표</li> <li>· 데이터에 따른 광범위한 AI 거버넌스에서 데이터 보호 역할 결정:및 새로운 방향성 협의</li> <li>· AI를 채택하고 사용할 때 국방부가 채택할 접근 방식에 대한 세부 정보 공개</li> <li>· 국제 AI 활동에 대한 범정부적 접근 방식 개발</li> </ul>
중기 (향후 6개월)	<ul style="list-style-type: none"> <li>· 노동자들이 비즈니스 환경에서 AI를 사용할 수 있게 하도록 필요한 기술에 관한 연구를 발표하고 국가 기술 제공이 이러한 요구를 충족할 방법 파악</li> <li>· AI 스케일업(Scale-up)의 민간 자금 수요 및 과제 평가</li> <li>· 학교에서 모든 AI 프로그램을 이용할 수 있도록 국립 컴퓨팅 교육 센터 지원</li> <li>· 진로 발전경로를 보장함으로써 더욱 광범위한 사람들이 AI 관련 직업에</li> </ul>	<ul style="list-style-type: none"> <li>· 경제 전반에 걸쳐 AI의 확산에 영향을 미치는 결정 요인에 관한 연구 수행</li> <li>· ‘Innovation Missions’ 수립 시 AI 잠재력이 포함되는 방식을 고려하고 양자 및 다자간 노력을 통해 야심찬 미션 기반 협력을 촉진</li> <li>· 개발도상국의 지역 기술 혁신에 대한 영국 지원 확대</li> <li>· 실제 애플리케이션을 사용하여 AI 도전 과제에 대한 개방형 저장소 구축</li> </ul>	<ul style="list-style-type: none"> <li>· AI 관리 및 규제에 대한 혁신에 대한 국가 입장에 관한 백서 출판</li> <li>· 알고리즘 투명성을 위한 정부 부처 간 표준 개발</li> <li>· AI 표준화에 대한 영국의 참여를 조정하기 위한 AI 표준 허브 시범 운영</li> <li>· 중장기 수평 스캐닝 기능을 구축하여 AI 안전에 대한 정부의 인식</li> </ul>

	<p>진입할 수 있도록 지원, AI와 함께 일하거나 발전할 기회 강조</p> <ul style="list-style-type: none"> <li>· 미국-영국 간 AI R&amp;D 협력 선언 이행</li> <li>· AI 혁신, 상용화 및 배포를 지원하는 데 필요한 영국 컴퓨팅 용량에 대한 점진적 개시</li> <li>· 세계 최고의 AI 인재를 영국에 유치하기 위해 새로운 비자 제도 도입</li> </ul>		<p>제고</p>
<p>장기 (향후 12개월 이상)</p>	<ul style="list-style-type: none"> <li>· 반도체 공급망의 국내 접근법 점검 시행</li> <li>· AI 모델에 대해 개방형 및 기계판독형 정부 데이터셋에 대한 고려</li> <li>· 국가 AI연구·혁신프로그램(National AI Research and Innovation Programme) 자금 지원 프로그램을 조정하고 더 확대된 생태계를 지원</li> <li>· 공유 R&amp;D 과제에 대해 글로벌 파트너와 협력, 해외 개발 지원을 활용한 AI 중심 세계 파트너십 설정</li> <li>· 최고 인재, 박사 과정, 데이터 과학 전환 과정 및 산업 자금 지원 석사 과정에 걸쳐 기존 개입 지속을 통한 AI</li> </ul>	<ul style="list-style-type: none"> <li>· 잠재력이 높으나 AI 성숙도가 낮은 분야에서 AI 기술의 개발 및 채택을 촉진하기 위해 AI 사무국/UKRI 공동 프로그램 착수</li> <li>· 국가 AI 연구 및 혁신 프로그램을 통해 AI 기술의 신뢰성, 채택 가능성 및 투명성에 관한 역량 개발의 지속적인 지원</li> <li>· AI 사용이 전략적 과제에 촉매 역할을 제공할 수 있는 부분을 파악하기 위해 정부 전반에 걸쳐 협력</li> </ul>	<ul style="list-style-type: none"> <li>· AI 생태계가 글로벌 AI 표준화 환경에 참여할 수 있도록 지원하기 위한 AI 기술 표준 참여 툴킷 개발을 이해관계자와 모색</li> <li>· 다자간 및 다중 이해관계자 포럼의 파트너들과 협력하고, GPAI에 투자하여 영국의 가치와 우선순위에 맞는 AI 거버넌스를 형성하고 지원</li> <li>· 앨런 튜링 연구소와 협력하여 공공 부문의 AI 윤리 및 안전에 대한 지침 개발</li> <li>· 국가안보, 국방 및 연구 부문이 협력하여 재앙적 위협의 예측 및 예방 방안 연구</li> </ul>



	<p><b>다양성 지원</b></p> <ul style="list-style-type: none"> <li>· 영국의 사업을 위해 개방되는 동안 <b>국가안보를</b> 보호하기 위해 국가안보 및 투자법에 대한 모니터링 및 활용</li> <li>· AI를 포함한 <b>새로운 기술에 무역 거래 조항 포함</b></li> </ul>		
--	---	--	--

출처: HM Government. 2021: 9.

## 다. AI 관련 주요 계획과 행위자<sup>141)</sup>

○ 국가 AI 전략은 특정 부처 단독으로 실행할 수 없다는 점에서 이 전략은 다음과 같이 상호 연결된 다른 부처와 정부 업무 간 조정을 위한 메커니즘을 확대한다.

### - 성장과 혁신 전략 계획

▷ 혁신을 촉진할 수 있는 역량을 갖춘 다양하고 포괄적인 AI 전문가 파이프라인의 개발 필요성에 대한 인식과 성장 계획 및 최근 혁신 전략

### - 통합적 검토

▷ 영국의 우수한 AI가 국내외에서 번영과 안보를 보장하고 미래의 개방적인 국제질서를 형성할 수 있는 새로운 길을 찾기 위한 통합적 검토

### - 국가 데이터 전략

▷ 2020년 9월에 발표된 국가 데이터 전략은 책임감 있는 데이터 사용의 힘을 활용하여 생산성을 높이고, 새로운 비즈니스와 일자리를 창출하고, 공공 서비스를 개선하고, 더 공정한 사회를 지원하고, 과학적 발견을 촉진하여 영국을 차세대 혁신의 선구자로 자리매김하겠다는 비전의 제시

### - 디지털 규제계획

▷ 디지털 규제계획은 번영을 촉진하고 디지털 기술 사용에 대한 신뢰를 구축하는 방식으로 디지털 기술을 규제하는 친 혁신 접근 방식을 제시

### - 국가 사이버 전략

▷ AI 개발에 보안을 구축하는 등 신기술 보안을 위한 노력을 지속하기 위한 국가 사이버 전략

---

141) HM Government. 2021.

- 디지털 전략
  - ▷ 디지털 부문에서 정부의 야망을 더욱 구체화하기 위해 DCMS의 10대 기술 우선순위를 기반으로 하는 향후 디지털 전략
- 국방 AI 센터
  - ▷ 국방 현대화의 핵심 요소인 새로운 국방 AI 센터
- 국가안보 기술 혁신 교류
  - ▷ 국가안보 이해관계자, 업계 및 학계 파트너를 한데 모아 더 나은 국가안보 역량을 구축하기 위한 데이터 과학 및 AI 공동 창작 공간인 국가 안보 기술 혁신 교류(NSTIX)
- 국가 복원력 전략(National Resilience Strategy).
  - ▷ 영국이 기술 위협에 대응하는 방법에 초점을 맞춘 국가 복원력 전략
- AI 위원회
  - ▷ AI 위원회는 정부에 전문가 자문을 제공하고 AI 생태계의 고위급 리더십을 제공하기 위해 2019년에 설립되었다.
  - ▷ AI 위원회는 산업계, 학계, 공공 부문에서 존경받는 각 분야의 리더들이 모여 AI 부문 협약의 핵심 약속을 이행하고 있다.
  - ▷ 위원들은 분기별로 모여 AI 정책의 현재 우선순위, 기회, 과제에 대해 AI 사무국과 범정부적으로 조언한다.
  - ▷ 2021년 1월, AI 위원회는 ‘AI 로드맵’을 발표하여 정부에 AI의 전략적 방향에 대한 16가지 권고안을 제시했다.
    - 이 로드맵의 핵심은 정부가 AI 섹터 딜을 통해 이루어진 투자의 성공을 바탕으로 미래의 기술 혁신에 적용할 수 있는 국가 AI 전략을 개발할 것을 촉구하는 것이었다.
  - ▷ 그 이후로 위원회는 국가 AI 전략의 개발에 정보를 제공하기 위해 광범위한 AI 커뮤니티와의 참여 프로그램을 주도하고 있다.
  - ▷ 정부는 이 전략의 형성과 이행의 방향성 설정을 위해 AI 위원회의 역할을 업데이트하고 강화하여 지속해서 정부에 대한 전문가 자문과 AI 생태계의 고위급 리더십을 제공한다.

## 4.3 유럽연합

### 가. AI 관련 입법 추진 개요

- EU 집행위원회는 지난 2021년 4월 21일 ‘AI 규제 초안’<sup>142)</sup>을 공표하였다. 이는 집행위원회의 입법안으로, 이후 각료이사회 및 유럽의회의 공동결정 과정과 회원국 승인을 거쳐야 효력을 득하게 된다.
  - 2022년 12월 6일 집행위원회는 AI 규제에 대한 공통 입장을 채택하였다.
  - 2023년 6월 유럽의회는 AI 규제에 대한 상당한 수정 사항과 함께 협상 입장(찬성 499표, 반대 28표, 기권 93표)을 채택하였다.
  - 2023년 12월 8일 관련 합의안이 도출되었으며, 2024년 초 유럽 의회의 표결을 거쳐 발효될 것으로 판단된다.
  - 2021년 ‘AI 규제 초안’이 제안된 이후, 유럽의회와 27개 회원국은 생성형 AI와 안면인식을 포함한 여러 난제에 대한 주요 차이점을 극복해야 했다. 주요 언론에 따르면 ChatGPT 및 Google Bard와 같은 챗봇 소프트웨어의 기본 모델을 규제하는 방법과 규제 여부에 대해서도 의견 차이가 있었다는 것이다.
  
- EU 또한 미국의 경우와 같이, 연합 차원에서 AI 규제를 마련하여 각 회원국 법률의 방향성을 제시하고 있으나, 차이점은 EU는 관련 규제에 관해 세밀한 규칙을 제시하여 입법화 과정을 거치고 있다는 점이다. 미국이 연방 차원 관련 규제를 원칙과 방향성 제시의 입장에서 설계하였다면, EU는 역내 공동시장의 일원화된 규칙 질서를 법제화하는데 그 목적을 두고 있다.

---

142) European Commission. 2021. “Proposals for a Regulations of the European Parliament and the European Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending certain Union Legislative Acts”

○ EU의 AI 관련 입법 진행 과정

- AI가 EU 경제에 미치는 영향 분석을 위해 특별위원회를 구성하였다.
- 2020년 10월 20일, 유럽의회는 EU가 혁신, 윤리 표준 및 기술에 대한 신뢰를 높이는 동시에 AI를 가장 효과적으로 규제할 수 있는 방법을 설명하는 다음 <표 4-11>의 세 가지 보고서를 채택하였다.

<표 4-11> 유럽의회가 채택한 AI 관련 입법 이니셔티브

제목	작성자	주요 내용	표결결과
AI 윤리 프레임워크	Iban García del Blanco (S&D, ES)	집행위원회가 EU에서 인공지능, 로봇 공학 및 관련 기술을 개발, 배포 및 사용할 때 따라야 할 윤리 원칙과 법적 의무를 설명하는 새로운 법적 프레임워크를 제시할 것을 촉구 - 인간 중심 및 인간이 만든 AI; 안전, 투명성 및 책임; 편견과 차별로부터 보호	찬성 559표, 반대 44표, 기권 88표로 가결
AI가 피해에 대한 책임	Axel Voss (EPP, DE)	미래 지향적인 민사 책임 프레임워크를 요구하여 고위험 AI를 운영하는 사람들이 결과적인 피해에 대해 엄격하게 책임을 지도록 함 - 생명, 건강, 신체적 완전성, 재산에 해를 끼치거나 손해를 끼치거나 “검증 가능한 경제적 손실” 을 초래하는 경우 중대한 무형적 피해를 초래하는 물리적 또는 가상 AI 활동에 적용되어야 함	찬성 626표, 반대 25표, 기권 40표로 가결
AI 관련 지식재산권	Stéphane Séjourné (Renew Europe, FR)	AI 분야에서 EU의 글로벌 리더십을 위해서는 혁신적인 개발자를 보호하기 위해 효과적인 지적재산권 시스템(IPR)과 EU 특허 시스템에 대한 보호 장치가 필요 - IPR의 소유권은 인간에게만 부여되어야 함. 저작권, 데이터 수집, 영업 비밀, 알고리즘 사용 및 딥 페이크 대한 재산 권리를 명시함.	찬성 612표, 반대 66표, 기권 12표로 채택

출처: European Parliament. 2020. “Parliament leads the way on first set of EU rules for Artificial Intelligence.” (October 20) <https://www.europarl.europa.eu/news/en/press-room/20201016IPR89544/parliament-leads-the-way-on-first-set-of-eu-rules-for-artificial-intelligence> (검색일: 2023. 11. 24).

- 세 개의 보고서는 첫째, AI 윤리 프레임워크, 둘째, AI가 피해에 대한 책임에 관한 이니셔티브 및 셋째, AI 관련 지식재산권에 관한 보고서로 구성되었으며, 3건 모두 채택되었다.

- 2022년 5월 유럽의회, AI에 대한 EU 로드맵 최종보고서 채택하였다. 집행위원회의 초안과 비교하여 유럽의회 입장안의 변경된 사항은 <표 4-12>와 같다.

<표 4-12> EU 집행위원회에서 마련한 초안과 비교하여  
2022년 5월 EU 의회 안에서 변경된 주요 사항

영역	주요 변경 내용
파운데이션 모델	<ul style="list-style-type: none"> <li>· 파운데이션 모델을 고위험으로 분류하지는 않았으나, 모델 위험 관리 및 기술문서 제공 등의 엄격한 요구사항 규정</li> <li>· 생성형 AI에 사용되는 파운데이션 모델의 경우, 위법한 콘텐츠 생성 방지, 학습데이터의 저작물 정보 공개의무 등 부과</li> </ul>
AI 시스템의 정의	<ul style="list-style-type: none"> <li>· AI 시스템을 ‘다양한 수준의 자율성으로 작동하도록 설계되었으며, 명시적 또는 암묵적 목표를 위해 물리적 또는 가상 환경에 영향을 미치는 예측, 추천 또는 결정과 같은 결과를 생성하는 기계 기반 시스템’으로 정의</li> </ul>
고위험 AI 시스템 범위	<ul style="list-style-type: none"> <li>· 고위험 AI 시스템 목록에 ‘대규모 온라인 플랫폼으로 지정된 소셜 미디어의 추천시스템’ 등 추가</li> <li>· 고위험 AI 시스템 목록에 해당하는 경우에도 ‘인간의 건강, 안전 또는 기본권에 상당한 위험을 야기하는 경우에만’ 고위험으로 간주</li> </ul>
금지된 위험 AI 시스템 범위	<ul style="list-style-type: none"> <li>· 공공장소에서의 실시간 원격 생체인식 시스템 사용 전면 금지</li> <li>· 금지되는 AI 시스템 목록에 ‘민감정보를 활용한 생체인식 분류 시스템’, ‘스크래핑을 통해 안면인식 DB를 구축하는 시스템’ 등 추가</li> </ul>
고위험 AI 시스템 배포자의 의무	<ul style="list-style-type: none"> <li>· 고위험 AI 시스템을 활용하여 서비스를 제공하는 자(고위험 AI 시스템 배포자)에게 기본권 영향평가 수행 등 강화된 의무 부여</li> </ul>
중소기업 불공정 계약 규제	<ul style="list-style-type: none"> <li>· 중소기업(SME) 또는 스타트업에 일방적으로 부과되는 불공정한 계약 조건의 구속력을 부정하는 조항 추가</li> </ul>
이용자 권리	<ul style="list-style-type: none"> <li>· 이용자에게 고위험 AI 시스템에 대한 설명요구권 등 권리 부여</li> </ul>
벌칙	<ul style="list-style-type: none"> <li>· 금지되는 AI 이용 행위에 대한 과징금 상한액이 전 세계 연간 매출액의 6% 또는 최대 3,000만 유로에서 전 세계 연간 매출액의 7% 또는 최대 4,000만 유로로 상향 조정</li> </ul>
유예기간	<ul style="list-style-type: none"> <li>· 법 통과 후 시행까지 유예기간이 3년에서 2년으로 단축</li> </ul>

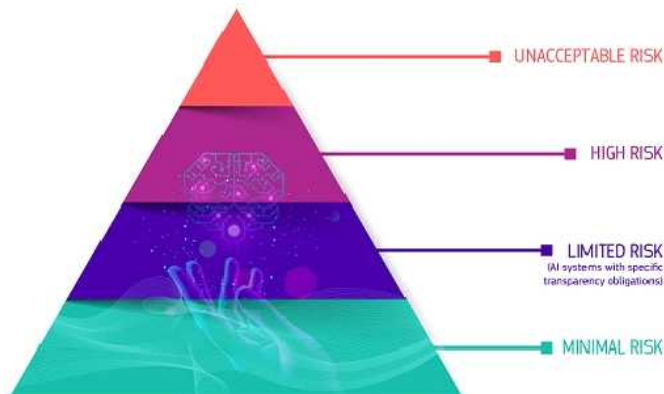
주: 파운데이션 모델(foundation model): 대규모 데이터로 사전 훈련된 AI 모델로서 출력이 범용적이어서 다른 다양한 시스템 개발의 기초로 쓰임.

출처: 김도업·마태경. 2023. “EU 의회 수정안 주요내용.” 톱슨로이터코리아(편). 『2023 국내외 AI 규제 및 정책 동향』 LAWnB Legal Essential Report. 22: 12.

## 나. 생성형 AI 선거 입법 동향 및 원칙 중심 규제 동향

- EU 차원에서 AI는 새로운 디지털 산업군에서 규제되어야 할 사항으로 여겨져 왔다. 하지만 선거에 대한 역내 이슈가 증가함으로써 주요국에서 관련한 입법이 등장하였으며, 허위 뉴스 및 가짜 뉴스가 생성형 AI에 의해 쉽게 제작, 유포될 수 있음을 인지하자, EU AI 규제(regulation) 입법 과정에서 이의 내용을 추가하게 되었다.
- EU AI 규제(regulation):
  - 동 ‘AI 규제 초안’에는 채용, 교육, 의료 또는 법 집행 분야와 같은 중요한 개인적 관심사를 다루는 생체인식시스템 또는 AI 결정과 같은 애플리케이션이 포함되며, AI에 대한 규제 프레임워크에 대한 집행위원회의 제안은 AI의 개발 및 활용에 대한 신뢰뿐만 아니라 기본적인 권리와 사용자 안전을 보호하는 것을 목표로 한다.

<그림 4-3> EU AI 위험유형 분류 피라미드



출처: European Commission. “Regulatory framework proposal on artificial intelligence.” <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> (검색일: 2023. 11. 01).

- 위험유형별 규제: EU 집행위원회는 2021년 4월 AI에 대한 최초의 EU 규제 프레임워크를 제안하였으며, 현재 입법 과정에 있다. 해당 규제 안에는 AI의 위험을 허용할 수 없는 위험(Unacceptable risk), 고위험(High risk), 제한된 위험(Limited risk), 위험이 최소화되거나 없음(Minimal or no risk)으로 분류하였다.
- ① 허용할 수 없는 위험(an unacceptable risk): 기본권 침해 등 EU 가치에 위배되는 다음과 같은 목적 하에서의 AI 시스템 활용은

금지한다. 공공기관이 AI 기술로 개인의 사회적 신용 등을 평가하는 소셜 스코어링을 방지한다. 또한 테러방지, 범죄수사 제외, 실시간 원격 생체정보기반 식별을 금지한다.

- ▷ 사람이나 특정 취약 집단에 대한 인지 행동 조작. 또한 테러방지, 범죄수사, 특정 취약 집단에 대한 인지 행동 조작: 예를 들어 어린이의 위험한 행동을 조장하는 음성 인식 장난감
- ▷ 사회적 점수: 행동, 사회 경제적 지위 또는 개인적 특성을 기준으로 사람을 분류
- ▷ 안면인식 등 실시간 및 원격 생체인식 시스템

② 고위험(a high risk): 자연인의 건강/안전/기본권에 고위험을 야기할 수 있는 AI 시스템은 요구사항 준수, 사전 적합성 평가 수행이 요구된다. 고위험군 AI는 다시 EU의 제품 안전법에 해당하는 제품에 사용되는 AI 시스템과 EU 데이터베이스에 등록해야 하는 8가지 특정 영역으로 분류된다. AI 규제 부속서2와 3에 열거된 26개 고위험군 AI 분야에 대해 집행위원회는 사전 적합성 평가, 일부 고위험 AI에 대한 독립된 인증기관의 제3자 검증 강제한다. 추후 고위험군에 새로운 유형의 AI 추가할 수 있음을 명시하였다.

- ▷ 자연인의 생체 식별 및 분류
- ▷ 중요 인프라 관리 및 운영
- ▷ 교육 및 직업훈련
- ▷ 고용, 근로자 관리 및 자영업 이용
- ▷ 필수 민간 서비스와 공공 서비스 및 혜택에 대한 접근 및 향유
- ▷ 법 집행
- ▷ 이주, 망명 및 국경 통제 관리
- ▷ 법률 해석 및 법률 적용에 대한 지원

③ 낮은 위험 혹은 최소위험(low or minimal risk): AI 시스템에 대해서는 고위험 AI 시스템에 부여된 요구사항들이 강제되지는 않으나 자발적 준수를 위한 행동강령(code of conducts)의 수립이 권



장된다. 여기에는 이미지, 오디오 또는 비디오 콘텐츠(예: 딥페이크)를 생성하거나 조작하는 AI 시스템이 포함된다.

④ 생성형 AI: 생성형 AI는 투명성 요구 사항을 준수해야 한다.

- ▷ AI가 제작한 콘텐츠를 공개
- ▷ 불법 콘텐츠 생성을 방지하도록 모델 설계
- ▷ 훈련에 사용된 저작권 데이터의 요약 게시

- 집행위원회는 고위험 AI 시스템에 관해 공급자 및 사업자의 법적 의무를 아래<표 4-13>과 같이 규정하였다.

<표 4-13> EU의 AI 시스템 공급자 및 사업자의 법적 의무

행위 자	법적 의무 내용
공급 자	(a) 요구사항의 준수 (b) 품질 관리 시스템 (규정준수를 위한 전략, 시스템 디자인 / 테스트 / 데이터 관리 관련 절차 및 방법, 위험관리 시스템, 사후(post-market) 모니터링, 책임 체계 등) 구축 (c) 고위험 AI 시스템의 기술 문서 작성 (d) 통제하에 있을 때 고위험 AI 시스템에서 자동으로 생성된 로그기록의 보관 (e) 고위험 AI 시스템의 시장 출시나 서비스 전 관련 적합성 평가 절차 수행 (f) 등록의무(제품 출시나 서비스 전 EU 데이터베이스에 AI 시스템 등록) 준수 (g) 요구사항에 위배되는 상황 발생 시 필요한 시정조치 수행 (h) AI 시스템이 사용되거나 서비스되고 있는 회원국 국가 관할 당국, 또한 해당되는 경우, 인증 기관에 비준수사항 및 수행한 시정조치를 통지 (i) 규정 준수를 알 수 있도록 고위험 AI 시스템에 CE 마크 부착 (j) 국가 관할 기관 요청 시, 고위험 AI 시스템의 요구사항 준수를 입증
사용 자	1. 고위험 AI 시스템 사용자는 시스템과 함께 제공되는 사용지침에 따라 시스템을 사용해야 함 2. 1의 의무는 연합 또는 국내법에 따른 다른 사용자 의무와 제공자가 명시한 인적 감독 조치를 구현할 목적으로 자체 자원 및 활동을 조직하는 사용자의 재량을 침해하지 않음 3. 1을 침해하지 않고 사용자가 입력 데이터에 대한 통제권을 행사하는 한, 사용자는 입력 데이터가 고위험 AI 시스템의 의도된 목적과 관련이 있는지 확인해야 함 4. 사용자는 사용지침에 따라 고위험 AI 시스템의 운영을 모니터링해야 함. 사용지침에 따른 사용이 건강/안전/기본권 위험을 초래할 수 있다고 판단되면 공급자 또는 유통업체에 이를 알리고 시스템 사용을 중지해야 함. 또한 제62조가 의미(기본권 침해)하는 심각한 사고나 오작동을 발견한 경우 공급자 또는 유통업체에게 이를 알리고 AI 시스템 사용을 중단해야 함.

	<p>5. 고위험 AI 시스템 사용자는 자신이 통제할 수 있는 범위 내에서 고위험 AI 시스템에서 자동 생성된 로그를 보관해야 함. 로그는 고위험 AI 시스템의 의도된 목적과 연합 또는 국내법에 따른 적용 가능한 법적 의무에 비추어 적절한 기간 동안 보관되어야 함</p> <p>6. 고위험 AI 시스템 사용자는, 해당되는 경우, 규정 (EU) 2016/679 제35조 또는 지침 (EU) 2016/680 제27조에 따른 데이터 보호 영향 평가를 수행할 의무를 준수하기 위해 제13조 (투명성 및 정보제공의 의무)에 의해 제공된 정보를 사용해야 함</p>
--	---

출처: 김현정. 2021. “인공지능 기반 사회에 대비한 EU의 전략과 정책 : EU의 AI 규제안을 중심으로.” 『한국과 국제사회』. 5집, 제4호: 266-267.; European Commission. 2021. “Proposals for a Regulations of the European Parliament and the European Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.” COM/2021/206 final: 13-16.

- 집행위원회는 EU의 AI 시스템 수입업체 / 유통업체에게 각각 아래 <표 4-14>와 같이 법적 의무를 부과하였다.

<표 4-14> EU의 AI 시스템 수입업체 / 유통업체별 법적 의무

행위자	법적 의무 내용
수입 업체	<ol style="list-style-type: none"> <li>1. 아래 사항의 확인할 의무               <ul style="list-style-type: none"> <li>(a) AI 시스템 공급자의 적절한 적합성 평가 절차 수행 여부</li> <li>(b) 제공자가 주어진 방식에 따른 기술 문서를 작성했는지 여부</li> <li>(c) 시스템에 요구되는 적합성 표시, 필수 문서 및 사용지침의 제공 여부</li> </ul> </li> <li>2. 고위험 AI 시스템이 규정을 준수하지 않는다고 판단되면 이를 준수할 때까지 해당 시스템을 시장에 출시해서는 안 되며, 고위험 AI 시스템이 건강/안전/기본권 위험을 나타내는 경우, AI 시스템 공급자와 시장 감시 당국에 그 영향을 알려야함</li> <li>3. 자신의 이름, 등록 상표명 또는 등록 상표 표시, 고위험 AI 시스템에 연락할 수 있는 주소 표시(가능하지 않은 경우 포장 또는 첨부 문서 (해당되는 경우)에 표시)</li> <li>4. 고위험 AI 시스템이 자신의 책임 하에 있는 동안, 적용 가능하다면, 보관 또는 운송 조건이 주어진 7가지 요구사항을 준수하는 데 위험이 되지 않도록 해야 함</li> <li>5. 합리적 요청이 있을 시, 국가 관할 당국에 쉽게 이해할 수 있는 언어로, 사용자와의 계약 또는 기타 법률에 의해 공급자의 통제를 받는 범위까지, 고위험 AI 시스템의 요구사항 준수 입증에 필요한, 자동으로 생성된 로그 액세스를 포함하여, 모든 정보와 문서를 제공. 또한 해당 시스템과 관련하여 국가 관할 당국이 취하는 조치에 대해 해당 당국에 협력</li> </ol>
유통 업체	<ol style="list-style-type: none"> <li>1. 고위험 AI 시스템을 시장에 출시하기 전 고위험 AI 시스템에 필수 CE 적합성 마크가 있는지, 필수 문서 및 사용 지침이 첨부되었는지, 적용 가능하다면, 공급자와 수입업체가 명시된 의무를 준수하였는지 확인</li> <li>2. 고위험 AI 시스템이 규정을 준수하지 않는다고 판단되면 해당 AI 시스템이 이를 준수할 때까지 해당 시스템을 시장에 출시해서는 안 되며, 고위험 AI 시스템이 건강/안전/기본권 위험을 나타내는 경우, AI 시스템 공급자, 수입업자에게 통지</li> <li>3. 고위험 AI 시스템이 자신의 책임 하에 있는 동안, 적용 가능하다면, 보관 또는 운송</li> </ol>

	<p>조건이 주어진 7가지 요구사항을 준수하는 데 위협이 되지 않도록 해야함</p> <p>4. 시장에 출시된 고위험 AI 시스템이 명시된 요구사항을 준수하지 않는다고 생각되면 유통업체는 이를 준수하기 위한 시정조치, 회수, 리콜 등을 취하거나, 공급자, 수입업자 또는 관련 운영자가 해당 시정 조치를 취하도록 해야함. 고위험 AI 시스템이 건강/안전/기본권 위협을 나타내는 경우, 즉시 회원국 국가 관할 당국에 즉시 통보하고 미준수 및 취해진 시정조치 등의 세부정보를 제공</p> <p>5. 국가 관할 당국의 합리적 요청이 있을 시, 국가 관할 당국에 고위험 AI 시스템의 요구사항 준수 입증에 필요한 모든 정보와 문서를 제공하고 국가 관할 당국이 취하는 조치에 대해 해당 당국에 협력</p>
--	---

출처: 김현정. 2021; European Commission. 2021.

- 따라서 AI 부문 對EU 수출업체는 고위험 AI에 관련한 규제사항을 준수할 의무를 지게 되며, 향후 FTA 내 고위험군 AI 규제가 포함될 가능성이 있다.

○ EU 서비스법(DSA): EU 역내 AI에 관한 규제는 결국 생성형 AI에 의해 제작된 콘텐츠 확산에 대한 초대형 플랫폼 및 검색 엔진에의 제재가 주요 관건이 될 것이다.

- 현재 DSA 적용 대상 기업은 Alibaba AliExpress(중), Google Play(미), Pinterest, YouTube(미), Snapchat, Zalando, TikTok(중), Bing, Instagram, Twitter(미), Facebook(미), LinkedIn, Wikipedia, Amazon Store(미), Google Maps(미), Apple AppStore(미), Google Shopping(미), Booking.com, Google Search(미) 이상의 19개 기업이다.
- DSA는 온라인상 불법 콘텐츠 삭제, 이용자의 권리 보호 등을 주요내용으로 하며 이를 위한 플랫폼 서비스 운영자의 의무를 규정하고 있다.<sup>143)</sup> 집행위원회는 디지털 서비스 플랫폼의 책임과 책임을 해결하기 위해 2000년 제정된 전자상거래 지침(2000 E-Commerce Directive)을 현대화하고 있다.
- 2022. 07. 05. 유럽의회 539 : 54 (기권 30)으로 승인되었다.
- DSA에는 EU에서 비즈니스를 수행하는 모든 온라인 중개 서비스에 대한 규칙이 포함되지만, 요구 사항은 회사 규모와 디지털 시장에서의

143) 한국지식재산연구원. 2021. “유럽연합 정상회의 상임의장, 탈린 디지털 서밋에서 유럽의 디지털 혁명을 강조.” (9월 28일) [https://www.kiip.re.kr/board/trend/view.do?bd\\_gb=trend&bd\\_cd=1&bd\\_item=0&po\\_item\\_gb=EU&po\\_no=20706](https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=EU&po_no=20706) (검색일: 2021. 9. 28).

역할에 따라 다르며 초안에 식별된 4개의 개별 계층으로 분류된다. 이는 △ 초대형 플랫폼(사용자 4,500만 명, 즉 EU 소비자 기반 10%), △ 온라인 플랫폼, △ 호스팅 서비스, △ 중개서비스로 분류한다.

〈표 4-15〉 DSA 내 온라인 중개 서비스 유형별 의무사항

	중개 S	호스팅 S	온라인 P	초대형 P
투명성 보고	○	○	○	○
기본권을 이유로 한 서비스 약관에 대한 요구 사항	○	○	○	○
명령에 따른 국가 당국과의 협력	○	○	○	○
연락처 및 필요한 경우 법정 대리인	○	○	○	○
이용자에 대한 정보 제공의 고지, 조치 및 의무		○	○	○
불만, 구제 메커니즘 및 법원 외 분쟁 해결			○	○
신뢰할 수 있는 신고자			○	○
이의제기 조치			○	○
제3자 공급업체(“KYBC”)의 자격 증명 심사			○	○
온라인 광고의 사용자 대면 투명성			○	○
범죄 신고			○	○
위험관리 의무 및 준법감시인				○
외부 위험 감사 및 공공 책임				○
정보 접근을 위한 추천 시스템 및 사용자 선택의 투명성				○
당국 및 연구자와 데이터 공유				○
행동 강령				○
위기대응 협력				○

출처: European Commission. “The Digital Services Act: ensuring a safe and accountable online environment.” [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en) (검색일: 2021. 09. 28).

- DSA에서 회원국 규제기관 사이에 제안된 새로운 협력 메커니즘은 집행력을 개선하고 블록 전체의 조화를 더욱 강화하는 것을 목표로 한다. 그러나 기업이 의무사항 불이행 시 아직까지 벌금은 새로운 EU 수준 기관이나 개별 회원국이 관할권에 있는 법인에 부과할 수 없음. EU 전역에서 불법 콘텐츠에 대한 정의가 없기 때문에 한 회원국이 다른 회원국에 기반을 둔 플랫폼에 해당 국가에서 콘텐츠를 제거하도록 요청하면 결과가 어떻게 될지 불분명한 상황이다.<sup>144)</sup>
- 디지털 서비스법은 서비스의 성격과 온라인 플레이어의 규모에 따라

144) Congressional Research Service. 2021. “EU Digital Policy and International Trade.” (March 25) <https://sgp.fas.org/crs/row/R46732.pdf> (검색일: 2021. 08. 25).

단계적으로 적용되는 여러 가지 조치를 규정하고 있어, 초대형 플랫폼과 검색엔진은 상대적으로 더 엄격한 요건을 적용받게 될 것이다.<sup>145)</sup>

○ EU 디지털 시장법(DMA: Digital Markets Act)

- DMA는 2022. 07. 05. 유럽의회 588 : 11 (기권 31)으로 승인되었다.
- 초대형 플랫폼 사업자인 게이트키퍼(Gatekeeper)를 정의하고 이들의 경쟁제한 및 불공정행위를 규제한다.<sup>146)</sup> 온라인 플랫폼은 사용자가 추가될수록 더 큰 시너지를 낼 수 있는 네트워크 효과의 속성을 지닌다. 따라서 집행위원회는 새로운 시장진입자 혹은 중소기업가 희생을 겪으며 해당 플랫폼의 경쟁력이 강화될 수 있다는 점에 주목한다. 불공정한 관행과 경쟁 가능성의 부족은 디지털 부문에서 더 높은 가격, 더 낮은 품질, 더 적은 선택과 혁신의 측면에서 유럽 소비자에게 피해를 주는 비효율적인 결과를 초래할 것이다. 이러한 문제를 해결하는 것은 디지털 경제의 규모(증가하는 추세와 함께 2019년 세계 GDP의 4.5%에서 15.5% 사이로 추산됨)와 사회적, 경제적 측면에서 디지털 시장에서 온라인 플랫폼의 중요한 역할을 고려할 때 가장 중요한 것이다.<sup>147)</sup> 이는 데이터 수집 및 사용이 전통적인 경쟁 지표와 함께 시장 지배력을 결정하는 데 핵심적인 역할을 하기 때문이다.
- DMA 내 게이트키퍼의 경쟁제한 및 불공정행위 분류는 다음의 <표 4-16>과 같다.

<표 4-16> DMA 내 게이트키퍼의 경쟁제한 및 불공정행위

게이트키퍼의 경쟁제한	게이트키퍼의 불공정행위
특정 상황에서 제3자가 게이트키퍼의 자체 서비스를 상호 운용할 수 있도록 허용	게이트키퍼 자체가 제공하는 서비스 및 제품을 게이트키퍼 플랫폼에서 제3자가 제공하는 유사한 서비스 또는 제품보다 순위에서 더 유리하게 취급함.
비즈니스 사용자가 게이트키퍼 플랫폼을 사용하여 생성한 데이터에 액세스할 수 있도록 함.	소비가 플랫폼 외부의 비즈니스에 연결하는 것을 방지
광고주와 출판업자가 게이트키퍼가 호스팅하는 광고에 대해 자체적으로 독립적인 검증을 수행하는 데 필요한 도구와 정보를 플랫폼에	사용자가 원할 경우 사전 설치된 소프트웨어나 앱을 제거하지 못하도록 방지

145) 최민재 · 진민정 · 이현우. 2023. “유럽의 가짜뉴스 대응정책.” 『Media 정책 리포트』 2.

146) 한국지식재산연구원. 2021.

147) European Commission. 2020a. “Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on contestable and fair markets in the digital sector (Digital Markets Act).” COM/2020/842 final.

<p>서 광고하는 회사에 제공 비즈니스 사용자가 자신의 제안을 홍보하고 게이트키퍼 플랫폼 외부에서 고객과 계약을 체결할 수 있도록 함.</p>	
---	--

출처: European Commission. 2020b. “The Digital Markets Act: ensuring fair and open digital markets.” (December 15) [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en) (검색일: 2021. 09. 28).

- 게이트키퍼가 DMA 불이행 시 제재: 대상 기업은 연간 매출의 최대 10% 벌금을 부과받는다. 일일 평균 매출의 최대 5%까지 주기적 위약금 지급을 강제한다. 집행위원회는 게이트키퍼가 DMA 의무를 체계적으로 위반한 경우 시장 조사 후 게이트키퍼에 대한 추가 구제 조치를 취할 수 있다.
- 서비스를 게이트키퍼로 지정하기 위해 여러 가지 누적 임계값이 설정되었다. 게이트키퍼는 모회사의 증권 거래소 가치가 750억 유로이고 지난 3년 동안 연간 매출액이 75억 유로를 초과했으며 EU에서 월간 활성 최종 사용자가 4,500만 명이 넘고 연간 활성 비즈니스가 10,000개를 넘는 회사이다.
- 지난 3년간 EU 내 사용자 수 기준, 게이트키퍼는 Alphabet(Google의 모회사), Amazon, Apple, Meta(Facebook 및 Instagram 모회사), Microsoft, Samsung 및 ByteDance(TikTok의 모회사)으로 예측된다.<sup>148)</sup>
- 개별국가 차원에서는 프랑스의 ‘정보조작 대처법(loi contre la manipulation de l’information)’, 독일의 ‘네트워크 집행법(NetzDG)’ 등을 통해 가짜뉴스에 대응하고 있다.

<표 4-17> 유럽 각국의 가짜뉴스 관련 법적 규제 내용

148) EURACTIV. 2023. “Digital competition law: EU Commission unveils its ‘gatekeepers’ list.” (September 7) <https://www.euractiv.com/section/law-enforcement/news/digital-competition-law-eu-commission-unveils-its-gatekeepers-list/> (검색일: 2023. 12. 01).

국가명	허위 정보 규정의 주요 내용	법적 근거	특징
프랑스	선거 전 3개월 동안 법원은 온라인 플랫폼에 허위 정보를 게시하지 못하도록 명령할 수 있음	정보조작대처법	- 선거기간에 국한 - 플랫폼에 대한 규제
독일	- 허위 정보뿐 아니라 독일 형법상 범죄가 되는 모든 종류의 게시물을 대상으로 함 - 디지털 플랫폼은 위법적인 콘텐츠나 댓글을 신고할 수 있는 시스템을 만들고, 신고가 들어오면 위법성을 판단한 후 24시간 안에 삭제해야 함	네트워크집행법	- 플랫폼에 대한 규제 - 허위 정보뿐 아니라 불법콘텐츠도 규제 대상
오스트리아	선거 또는 국민투표 기간 중 허위 뉴스의 유포를 범죄로 규정	형법 제264조	선거기간에 국한
폴란드	법원은 선거 후보자에 대한 “사실이 아닌 데이터 또는 정보”의 게시에 대해 금지 명령을 내릴 수 있음	지방선거법	선거기간에 국한
몰타	공중을 놀라게 하거나 공공의 선량한 질서 또는 평화를 교란하거나 대중 또는 특정 계층의 대중 사이에서 소동을 일으킬 가능성이 있는 허위 뉴스를 악의적으로 유포하는 행위를 범죄로 규정	형법 제82조	상시 적용
리투아니아	의도적으로 허위 정보를 유포하는 행위를 금지	일반 대중에 대한 정보 제공에 관한 법률	상시 적용
루마니아	통신 규제 당국이 코로나19 보호 및 예방 조치와 관련하여 “허위 뉴스를 조장”하는 온라인 콘텐츠의 삭제 및 접속 차단을 명령할 수 있도록 허용	대통령령	코로나19 대응조치의 일환
헝가리	팬데믹에 대한 정부의 대응을 방해하는 잘못된 정보를	가짜뉴스 대응법	코로나19 대응조치의 일환이었으나 이

	고의로 퍼뜨려 정부 대응에 대한 비판이나 두려움을 불러일으키는 경우, 최대 5년의 징역형을 선고		후 공식적인 규제로 전환됨
그리스	시민들에게 우려나 공포를 유발하거나 국가 경제, 국방 능력 또는 공중 보건에 대한 대중의 신뢰를 저해하는 정보를 대중에게 공공연하게 게시하거나 인터넷을 통해 유포한 경우, 3개월에서 5년 사이의 징역형에 처할 수 있음	형법 제191조	- 기존의 허위 정보 처벌 규정을 강화 - 언론 소유주 및 경영진도 처벌
스페인	언론매체와 소셜 네트워크의 허위 정보에 대응하기 위해 초법적 상설 기구를 설립	장관령	정부 각료들이 허위 정보 여부를 판단

출처: 최민재 · 진민정 · 이현우. 2023: 10.

## 4.4 미국과 EU의 협력과 경쟁

### 가. EU-미국 인공지능을 위한 공동로드맵

#### ○ EU-미국 무역기술위원회(TTC: Trade and Technology Council)

- EU와 미국은 2021년 6월 15일 EU-미국 무역기술협의회(TTC)를 설립하였다. TTC는 글로벌 무역, 경제 및 기술 문제에 대한 양측 간 접근 방식을 조정하고 이러한 공유 가치를 기반으로 범대서양 무역 및 경제 관계를 심화시키는 포럼 역할을 담당한다.<sup>149)</sup>
- TTC는 양측의 민감한 경제 및 통상 조치에 관해 협력 방안을 논의하는데 집중해 왔다. 스웨덴에서 개최된 제4차 장관급 회의(2023. 5. 30 ~ 31)에서는 특히 양측간 디지털 기술표준 및 지속가능한 무역 중점 합의 및 향후 협력 강화 전망, 그리고 핵심광물협정(미합의, 법적 구속력 강화 위해 향후 이사회 승인 필요 가능성) 등에 관해 논의하였다.

149) European Commission. "EU-US Trade and Technology Council." [https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council\\_en](https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council_en) (검색일: 2023. 12. 01).



〈표 4-18〉 TTC 10개 작업그룹

작업그룹	세부내용
1. 기술 표준	인공지능, 사물인터넷, 생명공학, 제약제품, 의료기기, 로봇틱스, 블록체인 등
2. 기후 및 청정 기술	무역과 기술을 포함하는 기후, 에너지, 환경 이니셔티브
3. 공급망 안보	반도체, 배터리, 중요 미네랄, 활성 의약품 성분
4. ICTS 보안 및 경쟁력	데이터 보안 표준, 안전하고 탄력적이며 다양한 통신 및 ICT 인프라 공급망, 5G/6G
5. 데이터 거버넌스 및 기술 플랫폼	기술 플랫폼, 콘텐츠 규제, 목적 광고 및 빅데이터 화용에 대한 책임 설정
6. 보안 및 인권을 위협하는 기술의 오용	인권을 침해하는데 사용되는 사이버 위협 및 기술에 대응, 잘못된 정보 작업 수행자 대응
7. 이중용도 품목의 수출 통제에 관한 협력	수출통제 조정, 정보공유 개선, 인권에 영향을 미치는 감시 기술 등 민감하고 새로운 기술의 위험 평가
8. 투자 심사 협력	외국인 투자심사를 위한 정보 공유 개선
9. 중소기업의 디지털 기술 접근 및 사용 촉진	중소기업에 더 많은 고객이 다가가고 디지털 기술이 소외된 지역의 사회에 혜택을 줄 수 있도록 지원
10. 글로벌 무역 과제	비시장 경제정책 및 관행에 초점 신기술 제품 및 서비스에 대한 불필요한 장벽을 피하고 노동권과 양질의 일자리 촉진

출처: European Commission. “Digital in the EU-US Trade and Technology Council.”  
<https://futurium.ec.europa.eu/en/EU-US-TTC> (검색일: 2022. 09. 15).

### ○ TTC 내 AI에 관한 논의

- 설립 이후, TTC는 파트너십을 위한 장관급 회의를 개최해 왔으며, 제 4차 장관급 회의에서 ① 믿을 수 있는 인공지능, ② 신기술(충전 인프라, 양자기술, 6G 등), ③ 정보 조작 방지, ④ 의약품, 해양 장비 영역 등에서의 양측 무역 증가, ⑤ 지속가능한 무역 등에 관한 논의를 진전하였다.
- 동 회의에서 양측이 동의한 AI 관련 협력 사안은 아래와 같다.
  - ▷ AI 관련 용어, 분류, 표준 및 새로운 위험을 다루는 3개의 전문가 그룹 출범
- 생성형 AI 시스템 관련 협력에 대해 중점 논의
  - ▷ 기후 변화, 자연재해, 의료, 에너지 및 농업 분야의 글로벌 과제를 해결하기 위해 공익을 위한 인공지능에 관한 행정 협정에 따라 특정 영역의 우선순위를 지정함.<sup>150)</sup>

○ TTC ‘인공지능을 위한 공동로드맵’ 주요 사항

- 2022년 12월 1일 TTC는 ‘인공지능을 위한 공동로드맵(TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management)’ 을 공표하였다.
- TTC의 AI 로드맵에서는 TTC의 AI 관련 활동이 다음의 국제기구와의 정보교환을 통해 개념과 용어의 공유된 이해와 일관된 적용을 진행할 것이라 적시하였다.
  - ▷ 국제표준개발기구(ISO: International Organization for Standardization)와의 연계
  - ▷ OECD AI 거버넌스 작업반(AIGO) 및 OECD AI 전문가 네트워크(ONE.AI) 내에서 지속적인 노력
  - ▷ AI 위험관리프레임워크와 관련 가이드 및 도구를 개발하려는 NIST의 노력
  - ▷ 국가 AI 구상법 및 AI 권리장전 청사진
  - ▷ EU AI법
  - ▷ 유럽표준기구(European standards organizations)에서 개발한 작업
  - ▷ 신뢰할 수 있는 AI에 대한 ALTAI 평가 목록 등 EU 고위 전문가 그룹의 결과물
- 위 로드맵은 총 11장의 문건으로 구성되었다. 이의 서문에는 ‘기술 정책에 관한 유럽연합과 미국 간의 합의가 기술표준의 공동 개발 이상의 진전을 이루지는 않을 것이며 기술표준의 사용에 있어서는 근본적으로 다르다는 점’ 을 적시하고 있다(<표 4-12> 참조).

---

150) ibid.

<표 4-19> TTC ‘인공지능을 위한 공동로드맵’ 내 협력의 접근방식에 관한 제시

양측은 신뢰할 수 있는 AI를 발전시키기 위해 사회적, 기술적 요인(사회 기술적 관점)의 조합을 고려하는 위험 기반 접근 방식을 적용한다. EU 사례는 제안된 EU AI 법과 AI에 관한 고위 전문가 그룹(HLEG)의 작업에 나타나 있다. 미국의 사례는 미국 국립표준기술연구소(NIST)의 AI 위험관리 프레임워크 초안과 백악관 과학기술정책국(OSTP)의 AI 권리 장전 청사진에서 제시되어 있다. EU와 미국은 위험 평가에 대한 책임 할당, 시설에 대한 가능한 법적 책임을 포함하여 규제 접근 방식에 대해 서로 다른 견해를 가질 수 있다.

Both sides apply risk-based approaches that consider the combination of societal and technical factors (socio-technical perspective) to advance trustworthy AI. EU examples are represented in the proposed EU AI Act and the work of the High-Level Expert Group (HLEG) on AI. United States examples can be seen in the National Institute of Standards and Technology (NIST) draft AI Risk Management Framework as well as the White House Office of Science and Technology Policy (OSTP) Blueprint for an AI Bill of Rights. (See Appendix for summaries of these examples.) While the EU and United States may have different views on regulatory approaches – including allocation of responsibility for risk assessment, possible legal responsibility for the establishment.

출처: Trade and Technology Council. 2022. “TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management.” (December 1) <https://digital-strategy.ec.europa.eu/en/library/ttc-joint-roadmap-trustworthy-ai-and-risk-management> (검색일: 2023. 12. 01).

- 단기목표

▷ 포용적 협력 채널 구축:

- 1) AI 용어 및 분류, 2) 신뢰할 수 있는 AI 및 위험 관리를 위한 AI 표준 및 도구, 3) 기존 및 신규 AI 위험 모니터링 및 측정에 관한 3개의 전문 실무 그룹을 구성
- 세 전문가 그룹 각각에 대한 작업 계획을 개발
- 전문가 워크숍 등 이해관계자 및 전문가 협의 계획을 수립

▷ 공유 용어 및 분류의 발전:

- 위험, 위험 관리, 위험 허용 범위, 위험 인식 및 신뢰할 수 있는 AI의 사회 기술적 특성을 포함하되 이에 국한되지 않는 주요 EU 및 미국 문서와 국제 표준의 용어 및 분류를 매핑

▷ AI 표준:

- EU 또는 미국이 관심을 갖고 있는 국제 표준에 대한 현황 분석을

수행하고, 국제 표준 개발에 대한 각 당사자의 참여 및 기여 수준을 평가

- 협력을 위한 관심 국제 표준을 식별
- 각 국제회의에 전문가 및 관련 이해관계자의 참여를 촉진
- 표준화 기관.

▷ 도구 개발:

- 도구 선택, 포함 및 개정을 위한 프로세스를 확립
- 신뢰할 수 있는 AI 도구를 평가하기 위한 기준을 설정

▷ 기존 및 새로운 AI 위험 모니터링 및 측정:

- 분류에 대한 파일럿 시도를 기반으로 할 수 있는 사용 사례 및 사고 보고를 기반으로 기존 AI 위험을 추적하기 위한 목표와 방법론을 설정
- 새로운 AI 위험에 대한 테스트 및 평가를 위한 연구 방법론을 식별

- 장기목표

▷ 전문가 워크숍

- 진행 상황을 검토 및 평가하고 필요한 경우 로드맵을 업데이트
- 로드맵 결과와 학습 내용을 협력하고 공유할 수 있는 기회를 식별

▷ 공유 용어 및 분류의 발전:

- 용어 및 분류에 대한 공유된 이해를 개발하거나 수정

▷ AI 표준:

- 확인된 특정 항목과 관련하여 국제 표준화 포럼에서 가능한 협력을 조직
- 표준 개발 또는 배포에 있어 전문가와 협력하고 지원
- 상호 이익

▷ 도구 개발:

- 공유 허브/저장소에 추가할 측정항목과 방법론을 식별
- 공유 허브/저장소를 업데이트하고 유지함

▷ 기존 및 새로운 AI 위험 모니터링 및 측정:

- AI 사고에 대한 실증적 연구를 통해 정보를 얻을 수 있는 AI 위험에 대한 벤치마크 및 평가를 작성
- 새로운 위험과 미래 위험에 대해 이론적 정보를 바탕으로 분석적으로 예측

- 로드맵 어디에도 법적 격차를 메우는 데 사용되는 표준과 이를 추적하고 평가하는 데 사용되는 방법에 관한 것 외에는 법률의 조화 또는 실제로 법률의 유사성에 대한 언급이 없으므로, 미국과 EU AI 안전 규정의 상호 운용성을 위한 여지를 결정하려면 기존 도구에 대한 법적 분석이 필요하다.<sup>151)</sup>
- 로드맵은 규제에 대한 위험 기반 접근 방식에 대한 공동 지원의 내용이 중심이며, <표 4-11>에서 제시한 바와 같이, 양측이 위험에 대비하기 위한 접근 방식의 측면에서 AI에 관한 협력을 진전한다면, AI 규제에 관한 양측의 균열이 존재하는 것인지에 관해 논할 필요가 있다. TTC AI 로드맵 내에는 “과학, 국제 표준, 공유 용어, 검증된 지표 및 방법론에 의해 지원되는 EU 및 미국 접근 방식의 중요성” 만을 강조하며, 법적 조화와 양측 AI 규제에 관한 정합성 여부에 관한 내용이 없다.

## 나. EU-미국 AI 규제 격차

### ○ TTC ‘인공지능을 위한 공동로드맵’ 내 과학기반 언어 사용에 관한 동의 사항

- TTC AI 로드맵은 다음 세 가지 구성 요소로 표준을 공식화하기 위한 합의를 제공하였다.
  - ▷ ‘신뢰할 수 있는(trustworthy)’, ‘위험(risk)’, ‘해로움(harm)’, ‘위험 임계값(“risk threshold)’, ‘편향(bias)’, ‘강건함(robustness)’, ‘안전(safety)’, ‘해석 가능성(interpretability)’ 및 ‘보안(security)’ 과 같은 AI 관련 주요 용어에 대한 공유 용어 및 분류법을 발전시키기 위해 ISO, OECD, IEEE 등 국제적으로 이미 논의가 진행되고 있는 사항을 중용하기로 합의했다.
  - ▷ 신뢰할 수 있는 AI 및 위험 관리를 위한 국제 기술 표준 및 도구에 대한 리더십과 협력을 도모한다. 후자에는 측정항목 및 방법론의 공유 허브/저장소, 공통 표준 기반 구축, 기존 솔루션 환

151) Vranckaert, Koen. 2023. “How Cautious is too Cautious? The US and EU Artificial Intelligence Roadmap.” (May 9) <https://www.law.kuleuven.be/citip/blog/how-cautious-is-too-cautious-the-us-and-eu-artificial-intelligence-roadmap-part-1/> (검색일: 2023. 12. 01).

경에 대한 연구뿐만 아니라 모든 이용 가능한 이해관계자(비즈니스, 시민 사회, 학계, 정부 등)의 추가 연구 지원이 포함된다.

- ▷ 기존 및 새로운 위험 추적기와 AI 위험의 상호 운용 가능한 테스트 및 평가로 구성된 기존 및 새로운 AI 위험을 모니터링하고 측정한다.<sup>152)</sup>

## ○ EU-미국 간 AI 규제 격차

### - 자발성 및 강제성

- ▷ EU의 AI 법안은 AI 안전과 시장규제를 강력한 규제를 통해 제어하고자 하는 경성법, 즉 규칙 기반 질서 확립에 있다. EU AI법은 회원국이 결정하는 집행 권한을 국가 당국에 부여하도록 규정하고 있다. 집행위원회는 3천만 유로 또는 이전 회계연도 전 세계 연간 총 매출액의 6%(금지된 AI 관행 사용의 경우), 2천만 유로 또는 4%(기타 AI 위반의 경우)의 벌금을 부과한다.<sup>153)</sup>
- ▷ 미국의 접근 방식은 연성법과 행정명령으로 제한된 원칙 기반 질서이다. 미국은 주요 AI 빅테크 기업의 성장을 통해 시장 주도를 목표로하고 있다. 미국 AI 권리 장전은 시행되도록 설정되지 않은 일련의 원칙일 뿐 AI 개발자에게 지침을 제공함에 그치고 있다.

### - 제재의 대상

- ▷ EU AI법은 위험관리 위반에 관한 제재의 대상으로 플랫폼을 규제하고 있으나, 미국의 경우는 그렇지 않다. EU의 플랫폼 규제는 △ 초대형 플랫폼(사용자 4,500만 명, 즉 EU 소비자 기반 10%), △ 온라인 플랫폼, △ 호스팅 서비스, △ 중개서비스로 분류하여 이행된다(<표 4-9> 참조).

152) Trade and Technology Council. 2022.

153) Vranckaert, Koen. 2023.

- 위험관리 방식

- ▷ EU AI법은 4개 위험 계층을 분류하여, ‘용인하지 않는 AI’ 영역을 지정하고 있으나, 미국의 AI 시스템은 사전 금지하지 않고 있다. EU AI법은 ① 사람에게 해를 끼치는 특정 조작 모드(취약한 특성의 잠재의식 및 오용)에 대한 AI 사용, ② 초기 데이터 수집과 관련이 없거나 불균형하다고 간주되는 맥락에서 사회적 점수를 매기기 위해 AI를 사용함 등을 제한한다. 심지어 중국도 실시간 생체 인식의 사용 일부를 제한하나, 미국은 주별 알고리즘 사용에 대한 제한을 둘 뿐, 연방 차원에서 제한하는 영역이 없다.
- ▷ 일부 AI 시스템은 제품 안전에 관한 새로운 입법 체계에 포함되어 특정 필수 요구 사항을 준수하도록 요구하고 준수 추정을 제공하는 표준의 역할을 제공한다(EU AI act 제안 제8조~15조 및 제40조). EU AI법은 AI 기반 챗봇, 생체 인식 분류 시스템, 감정 감지 시스템, 딥페이크 생성기 등 기타 AI 시스템은 자신이 AI 시스템임을 알려주기만 하면 되도록 제시하고 있다.
- ▷ 반면 AI 권리장전에 대한 청사진은 인공지능 시대에 미국 대중의 권리를 보호하기 위해 자동화 시스템의 설계, 사용 및 배포를 안내하는 5가지 원칙과 관련 관행을 제공한다. 이런 의미에서 ‘AI 권리장전에 대한 청사진’은 EU AI 고위 전문가 그룹의 신뢰할 수 있는 AI에 대한 윤리 지침과 더 유사하다 할 것이다.

<표 4-20> EU AI법과 미국 AI 권리장전 내 위험관리 규제 격차 비교

EU AI법	미국 AI 권리장전
지속적이고 정기적으로 업데이트되는 위험 관리 시스템	안전하고 효과적인 시스템: 자동화 시스템은 다중 이해관계자 협의를 거쳐 개발되어야 하며, 안전을 보장하기 위해 개발 전 테스트를 거쳐야 함. 해를 끼치도록 설계되어서는 안됨.
알고리즘 편견을 피하기 위해 훈련, 검증 및 테스트 데이터 세트가 AI 시스템의 대상 모집단과 일치하는지 확인하는 시스템	알고리즘 차별 보호: AI 시스템이 불법 차별에 가담하지 않도록 AI 시스템을 설계하고 조치를 취해야 함.
감사자가 AI 시스템의 결함을 이해할 수 있도록 하기 위한 기술 문서화 및 자동 이벤트 로깅	데이터 개인정보 보호: 내장된 보호 기능을 통해 귀하는 내장된 보호 기능을 통해 악의적인 데이터 관행으로부터 보호되어야
사용자가 AI 시스템의 기능을 이해할 수 있도록 투명한 정보 제공	
이해 가능성을 높이는 도구 제공부터 기각	

<p>메커니즘 및 킥 스위치에 이르기까지 효과적인 인간 감독을 제공하기 위한 요구 사항</p> <p>외부 공격으로부터 AI 시스템을 보호하기 위한 정확성, 견고성 및 사이버 보안 메커니즘.</p>	<p>하며 귀하에 대한 데이터가 사용되는 방식에 대한 권한을 가져야 함. 이 원칙은 EU의 일반 데이터 보호 규정에 있는 많은 원칙을 다시 적용함.</p> <p>공지 및 설명: 사용자는 자동화된 시스템이 사용되고 있다는 사실을 항상 알고 있어야 하며 해당 시스템이 귀하에게 영향을 미치는 결과에 어떻게, 왜 기여했는지 이해해야 함.</p> <p>인간의 대안, 고려 및 대체: AI 시스템 사용을 거부하고 발생할 수 있는 문제를 신속하게 고려하고 해결할 수 있는 사람에게 접근할 수 있어야 함</p>
---	--

출처: Vranckaert, Koen. 2023.

- 생성형 AI 창작물의 저작권 인정 여부

- ▷ EU AI법과 미국 모두 생성형 AI 창작물에 대한 저작권을 인정하지 않는다. 반면 중국은 생성형 AI 창작물의 저작권 인정하는 방안을 발표하여 이들과 대조를 이룬다.

## 4.5 중국

### 가. 생성형 AI에 대한 규제 도입 배경

○ 사회 가치를 위한 접근

- 중국은 2023년 8월 세계 최초로 생성형 AI 법규제(生成式人工智能服务管理暂行办法)를 도입, 시행하는 국가이다. 중국의 생성형 AI 법규제는 콘텐츠 관리가 중점으로, ‘사회주의적 가치를 구현해야 한다’는 중심 원칙 아래에 당국의 허가를 받아야 생성형 AI 서비스를 제공할 수 있다는 ‘라이선스 제도’를 도입하였다.<sup>154)</sup>
- 중국은 이전 시기 공산당을 비판하는 내용을 검열하기 위해 사회관계망 서비스나 해외 플랫폼에 대한 제어했던 것과 마찬가지로 생성형 AI에 관한 검열을 시작하였다.

154) 김다운. 2023. “각국의 생성 AI 규제 및 정책 동향.” 『투이컨설팅』 (8월 16일) <https://www.2eco.kr/news/articleView.html?idxno=302827> (검색일: 2023. 12. 01).



- 중국 사이버 관리국이 생성형 AI 규제를 만든 이유 또한 아래와 같이 밝히고 있다. “중국 IT기업은 데이터셋에 사회주의 핵심 가치를 반드시 반영해야 한다” 고 강조하며, 해당 규정을 위반하는 기업은 벌금을 내야 한다. 벌금은 1만 위안(약 190만원)에서 최대 10만 위안(약 1천900만원) 부과와 형사 조사도 받는다.<sup>155)</sup>

## 나. 생성형 AI에 대한 규제

### ○ 중국의 AI 관련 규제

- 중국은 다음과 같은 규정을 제안하고 시행함으로써 AI 규제에 있어 다른 국가보다 강하게 나가고 있다.
  - ▷ 2022년 3월 1일 발효. 인터넷 정보 서비스에 대한 알고리즘 권장에 관한 관리 규정(알고리즘 권장 규정)
  - ▷ 2023년 1월 10일 발효. 인터넷 정보 서비스의 심층 합성 관리에 관한 규정(심층 합성 규정)
  - ▷ 2023년 7월 13일 발표되어 2023년 8월 15일 발효. 생성 인공지능 서비스 관리에 관한 잠정 조항(생성형 AI 규정)
  - ▷ 공개 협의를 위해 2023년 4월 14일에 발표된 과학 및 기술 활동의 윤리적 검토를 위한 시험 조치(초안 윤리적 검토 조치)

### ○ 생성형 AI 서비스 규제 지침

- 2023. 8. 15. 시행 중국 생성형 AI 규제: 중국 국가인터넷정보판공실(国家互联网信息办公室)은 최근 홈페이지를 통해 국가발전개혁위원회(国家发展和改革委员会), 교육부(教育部), 과학기술부(科学技术部), 공업정보화부(工业和信息化部),公安部(公安部), 국가방송총국(国家广播电视总局) 등 6개 부처와 함께 ‘생성형 AI 서비스 관리 잠정 방법(生成式人工智能服务管理暂行办法)’을 발표했다.<sup>156)</sup> 해당 규정은 총

155) 김미정. 2023. “중국, 생성 AI 검열 시작 . . . “공산당 비판하면 벌금” .” 『ZDNET Korea』 (4월 14일) <https://zdnet.co.kr/view/?no=20230414074132> (검색일: 2023. 12. 01).

156) KOTRA. 2023. “중국, 생성형 AI 서비스 관리에 나서.” (7월 31일) [https://dream.kotra.or.kr/kotra-news/cms/news/actionKotraBoardDetail.do?SITE\\_NO=3&MENU\\_ID=80&CONTENTS\\_NO=2&bbsGbn=242&bbsSn=242&pNttSn=204600](https://dream.kotra.or.kr/kotra-news/cms/news/actionKotraBoardDetail.do?SITE_NO=3&MENU_ID=80&CONTENTS_NO=2&bbsGbn=242&bbsSn=242&pNttSn=204600) (검색일: 2023. 12. 01).

24개 조항으로 구성되었다.

<표 4-21> 중국 생성형 AI 서비스 관리 잠정 방법 주요 내용

적용 범위	<p>(제2조) 생성형 AI 기술을 이용하여 중국 주민을 대상으로 텍스트·이미지·음성·동영상 등 내용을 생성하는 서비스를 제공하는 경우 이 방법이 적용됨</p> <p>(제2조) 생성형 AI 기술을 개발하고 적용하지만, 대중에 제공하지 않는 산업단체·기업·연구기관 등은 이 규정의 적용 대상이 되지 않음</p>
기술개발 촉진	<p>(제5조) 다양한 산업분야에서 생성형 AI 기술의 활용 및 긍정적이고 건강한 고품질 콘텐츠의 생성을 장려하고, 적용 시나리오의 탐색·최적화 및 활용 시스템의 구축을 추진함. 생성형 AI 분야 기술 혁신·데이터 자원 구축·기술 전환/활용, 리스크 예방 등 관련 산업단체·기업·교육/연구기관·공공문화기구 등의 협력을 지원함</p> <p>(제6조) 생성형 AI 알고리즘·프레임워크·칩 및 지원 소프트웨어 플랫폼 등 기본기술의 자립적 혁신을 장려하고, 평등하고 서로 이익이 되는 국제 교류·협력을 추진하며, 생성형 AI 관련 국제규칙 제정에 참여함. 생성형 AI 인프라 및 공공 훈련 데이터 자원 플랫폼 구축을 촉진하고, 컴퓨팅 자원의 공유 및 효율적 이용을 추진함. 공공데이터의 유형별·등급별 개방을 촉진하고 고품질 공공 훈련 데이터를 확충함. 안전하고 신뢰도 높은 칩·소프트웨어·도구·컴퓨팅 및 데이터 자원의 사용을 장려함.</p>
준수의무	<p>보안 평가</p> <p>(제17조) 여론 속성 또는 사회적 동원 능력을 갖춘 생성형 AI 서비스를 제공하는 경우 관련 국가 규정에 따라 안전 평가를 수행하고 ‘인터넷 정보 서비스 알고리즘의 권장 관리 규정’ (互联网信息服务算法推荐管理规定)에 따라 알고리즘 제출·변경 및 등록 취소 절차를 수행해야 함</p>
	<p>데이터 훈련·라벨링</p> <p>(제7조) 서비스 제공자는 법에 따라 사전 훈련, 최적화 훈련 등 훈련 데이터 처리 활동을 수행하고 다음 규정을 준수해야 함</p> <ol style="list-style-type: none"> <li>(1) 합법적인 출처를 가진 데이터 및 기본 모델을 사용</li> <li>(2) 타인의 지적재산권 침해 불가</li> <li>(3) 개인정보와 관련된 경우 개인의 동의를 얻거나 법률 및 행정법규의 규정에 부합해야 함</li> <li>(4) 훈련 데이터의 품질을 향상시키고 훈련 데이터의 진정성·정확성·객관성 및 다양성을 향상시키기 위한 효과적인 조치를 취해야 함</li> <li>(5) ‘사이버보안법’, ‘데이터 보안법’, ‘개인정보보호법’ 및 기타 법률, 행정법규의 관련 규정 및 관련 주관 부서의 관련 감독 요구 사항</li> </ol> <p>- (제8조) 생성형 AI 기술 연구·개발 과정에서 데이터 라벨링을 수행하는 경우 서비스 제공자는 이 방법의 요구사항에 따라 △명확하고 구체적이며 실행 가능한 라벨링 규칙을 제정하고, △데이터 라벨링 품질 평가를 수행하며, △라벨링 내용의 정확성을 확인하기 위해 샘플링을 수행해야 함. 또한 △라벨링 담당자에게 필요한 교육을 제공하고 준법 의식을 향상시키며, △라벨링 담당자의 표준화된 작업 수행을 감독·지도해야 함</p>

	콘텐츠 관리	<p>(제4조) 인공지능 서비스의 제공 및 사용은 법률·행정법규를 준수하고 사회의 공중도덕과 윤리를 존중하며 다음 규정을 준수해야 함</p> <p>(1) 사회주의 핵심가치를 반영하여야 하며 △국가 권력 전복, △사회주의 체제 전복, △국가 분열 선동, △민족단결 저해, △테러 조장, △극단주의, △인종 증오와 차별, △폭력, △음란 및 허위 정보, △경제 및 사회 질서 교란 위험이 있는 내용을 포함해서는 안 됨</p> <p>(2) 알고리즘 설계·훈련 데이터 선택·모델 생성 및 업데이트·서비스 제공 등의 과정에서 인종·민족·신앙·국적·지역·성별·연령·직업 등에 대한 차별을 사전에 방지해야 함</p> <p>(3) 지적재산권을 존중하고, 상업 윤리와 상업 기밀을 유지해야 하며 불공정 경쟁 행위에 연루해서는 안 됨</p> <p>(4) 타인의 합법적인 이익을 존중하고 타인의 심신 건강에 상해, 초상권·명예권·개인의 프라이버시·개인정보를 침해해서는 안 됨</p> <p>(5) 서비스 유형의 특성을 기반으로 생성형 AI 서비스의 투명성을 개선하고 생성된 콘텐츠의 정확성과 신뢰성을 향상시키기 위한 효과적인 조치를 취해야 함</p>
	콘텐츠 표기	<p>(제12조) 서비스 제공자는 ‘인터넷 정보 서비스의 심층 합성 관리에 관한 규정’ (互联网信息服务深度合成管理规定)에 따라 이미지·동영상 등 생성된 콘텐츠에 대해 표기를 해야 함</p>
	개인 정보 보호	<p>(제11조) 사용자의 입력 정보·사용기록에 대해 보호의무를 이행해야 하며 △불필요한 개인정보 수집, △사용자 신원을 식별할 수 있는 입력정보·사용기록의 불법 보관, △타인에게 사용자 입력정보·사용기록의 불법 제공은 안 됨</p> <p>(제11조) 서비스 제공자는 법률에 따라 개인정보 열람·복사·정정·보완 및 삭제에 대한 개인의 요청을 적시에 수락·처리해야 함</p>
	운영 상 규제	<p>(제13조) 서비스 제공자는 사용자의 정상적인 사용을 보장하기 위해 서비스 과정에서 안전하고 안정적이며 지속적인 서비스를 제공해야 함</p> <p>(제10조) 서비스 제공자는 서비스의 적용 대상자·장소·용도를 공개하고, 생성형 AI 기술에 대한 사용자의 과학적·이성적 이해와 합법적 사용을 유도하며, 생성형 인공지능에 대한 미성년자의 과도한 의존 및 중독을 방지하기 위해 효과적인 조치를 취해야 함</p> <p>(제15조) 서비스 제공자는 건전한 민원신고 메커니즘을 구축하고, 쉽게 접근할 수 있는 신고경로를 제공하며, 처리 절차·기한을 제시하고, 민원 신고를 적시에 접수·처리하며 결과를 피드백해야 함</p>
	위반 시 조치	<p>(제14조) 서비스 제공자는 불법 내용을 발견한 경우 생성 중지·전송 중지·내용 삭제 등 조치를 적시에 취하고 모델 최적화 교육을 수행하며 주무기관에 보고를 해야 함</p> <p>(제14조) 서비스 제공자는 사용자의 불법 행위를 발견한 경우 경고·기능 제한·서비스 중단/종료 등 조치를 취하고 관련 기록을 보관하며 주무부서에 보고해야 함</p>

처벌	<p>(제21조) 서비스 제공자가 이 규정을 위반한 경우 주무부서는 ‘인터넷보안법’, ‘데이터안전법’, ‘개인정보보호법’, ‘과학기술진보법’ 등 법률·행정법규에 따라 처벌해야 함. 법률·행정법규에 관련 규정이 없는 경우 주무부서는 직책에 따라 경고·통보·시정명령 등 조치를 취하며, 주무부서의 시정명령을 거부하거나 관련 상황이 심각한 경우 해당 서비스를 중지함</p> <p>치안관리처벌법 위반에 해당하는 경우 치안관리 처벌을 부과하며, 형법 위반에 해당하는 경우 형사책임을 추궁함</p>
----	--

출처: KOTRA. 2023.

- 24항으로 구성된 이 규정은 생성형 AI 서비스 제공·이용 관련 관리 감독 체계에 대한 내용은 물론 기술개발 촉진, 데이터 처리 활동 및 데이터 라벨링 교육에 대한 요구사항이 포함되어 있으며, 또한 인종·민족·성별 등에 대한 차별을 방지하고 개인정보 및 미성년자 보호 관련 내용도 포함됐다(<표 4-21> 참조).

## 4.6 독일

### 가. AI와 민주주의에 대한 독일의 주요 논의

#### □ 개요

○ 이 논문<sup>157)</sup>은 인공지능(Kunstliche Intelligenz)의 발전과 확산이 민주주의에 미치는 영향에 대해 최근 독일의 시각을 잘 보여주고 있다. 여기서 다루고 있는 주요 주제는 아래의 세 가지이다.

- 인공지능의 확산이 민주적 공공의 구조와 역학에 미치는 영향
- 인공지능의 확산이 민주적 통치에 미치는 기회와 위험
- 인공지능의 확산이 시민들의 정치참여 변화에 미치는 영향

157) Thiel, Thorsten. 2023. “KI und Demokratie: Entwicklungspfade, Aus Politik und Zeitgeschichte, Bundeszentralen für politische Bildung.” (October 13) <https://www.bpb.de/shop/zeitschriften/apuz/kuenstliche-intelligenz-2023/541497/ki-und-demokratie-entwicklungspfade/> (검색일: 2023. 12. 01).

○ 이 논문의 주요 내용을 요약하여 살펴보면 다음과 같다.

- 민주적 공공의 구조와 역학에 대한 영향

▷ 다양성과 투명성의 증가: AI 기술의 도입은 정보 접근성과 다양성을 증가시킬 수 있다. 풍부한 데이터를 기반으로 한 AI 알고리즘은 다양한 의견을 포함하고 다양한 주제에 대한 정보에 접근할 수 있게 한다.

▷ 편향과 소수 의견의 부상: 반면에 AI 시스템은 훈련 데이터의 편향을 반영할 수 있다. 편향된 데이터를 기반으로 한 알고리즘은 소수 의견에 민감해질 우려가 있다.

- 민주적 통치에 대한 기회와 위험

▷ 정보 분석과 정책 향상: AI는 막대한 양의 데이터를 신속하게 분석하여 정책 결정에 활용할 수 있다. 이는 효율성과 정확성을 향상할 수 있지만, 의사결정의 투명성과 책임성 문제를 동반할 수 있다.

▷ 알고리즘 투명성과 판단의 어려움: AI 알고리즘은 종종 블랙박스로 여겨진다. 이로 인해 정책이나 결정의 이해가 어려워질 수 있고, 알고리즘의 판단 기준에 대한 의문이 제기될 수 있다.

- 시민 정치참여의 변화

▷ 디지털 의사소통 강화: AI를 통한 의사결정 프로세스의 디지털화는 시민들에게 더 많은 참여 기회를 제공할 수 있다. 온라인 의견 조사 및 피드백 메커니즘이 강화될 수 있다.

▷ 기술 접근성 부족과 디지털 격차: 그러나 기술 접근성이 부족한 경우 일부 시민은 이러한 기회에 참여하지 못할 수 있으며, 이는 디지털 격차를 증가시킬 수 있다.

## □ 디지털 기술의 발전

- 급격한 기술의 발전은 정보처리 과정을 개선했을 뿐만 아니라 사회적 행위의 목적과 가능성을 보는 우리의 시각을 혁명적으로 바꿔놓았다.
  - 인터넷, SNS, 스마트폰은 지식을 어떻게 조직하는지 또는 경제가 어떻게 작동하는지와 같은 사례에서 그 방식에 큰 변화를 촉진하고 있다.
  - 정치와 민주주의의 영역도 디지털 기술에 기반한 구조변화를 외면할 수 없다.
  - 1990년대부터 공공 영역으로 확대된 인터넷의 지속적인 디지털화와 긴박함은 민주주의에 대한 위기와 기회가 되고 있다.
  
- 2022년 말 챗GPT(ChatGPT)가 공식적으로 출시되면서 AI 사용이 새로운 차원에 이르렀고, 또 민주주의에 대한 영향력으로 생성형 인공지능을 둘러싼 논란이 화두가 되고 있다.
  - 특히 민주주의와 관련해서는 가짜뉴스와 정보의 조작 위험성이 논란의 초점이 된다.
  - 인공지능이 우리의 일상에 광범위하게 들어옴으로써 다양한 방면에서 민주주의에 대한 도전이 되고 있다.
  
- 20세기 중반 이후 인공지능의 발전과정은 대략 두 단계로 나누어 볼 수 있다.
  - 단순히 정보를 분류하고 논리적 규칙에 따라 데이터를 가공하는 단계(소위 상징적 인공지능)
  - 기계적 학습을 통해 문제를 분석하고 데이터와 가능성을 결정하여 새로운 해법을 찾아내는 단계
  
- 2010년부터 본격화된 인공지능은 이후 급속하게 발전하고 있다.
  - 그러한 발전의 근저에는 인공지능을 적용할 수 있는 컴퓨터 계산 용

량의 무한한 확대와 방대한 데이터양의 확보가 있다.

- 이러한 인공지능의 성공에는 사진인식과 번역 분야의 기술 향상이 중요했으며, 이런 점들이 경제적 수익과 연결되었고, 이는 또다시 기술의 발전을 가져왔다.
- 텍스트 또는 비디오 생성과 같은 새로운 응용 분야를 개척한 것이 최근의 성공 사례이다.

#### □ 인공지능의 확산이 민주적 공공 영역에 미치는 영향

○ 2022년 이전에는 인공지능에 의한 민주주의 변형에 관한 논란은 주로 SNS 콘텐츠의 내용을 조절하는데 인공지능을 사용하는 것이었다.

- 여기에는 플랫폼의 구조적 및 조직적 권력이 더 강화될 수 있다는 우려가 있었다.

○ 이후 많은 사람이 생성형 인공지능을 이용하게 됨에 따라 새로운 단계에 이르게 되었다. 그것은 소통에 따른 결과물을 생성하거나 변경하는데 드는 비용이 현저하게 감소했다는 점이다.

- 사진, 비디오 또는 목소리 등이 모두 실시간으로 만들어지거나 변형되어 전달되고, 어떤 주제나 어떤 문체로도 생산이 가능하다.
- 생성형 인공지능은 개인적 행동반경을 뛰어넘어 아예 정보환경 자체를 바꾸어 놓았다.

○ 이러한 것들이 공공의 영역이나 민주주의에 어떤 영향력을 행사할 것인가? 첫째, 거짓 정보 또는 가짜뉴스의 의도적 전염병적 확산에 대한 우려

- 생성형 AI는 얼굴 또는 목소리를 위조하여 가짜 동영상이나 오디오를 만들어내는 소위 딥페이크(deepfake) 기술을 가능케 한다. 이미 누구나 낮은 단계의 디지털 기술을 이용하여 사진이나 비디오를 손보고 있다는 사실을 알고 있다.
- 과거에도 정치적 소통단계에서 조작과 모방이 부분적으로 존재했지

만, 생성형 AI는 이를 훨씬 더 중요한 문제로 만들었다. 정보 조작의 기술이 향상되어 만들어진 질 높은 생성물을 독자들이 가짜라고 판단하기 어렵게 되었기 때문이다.

- 이러한 가짜뉴스는 동질적 집단별로 메신저 서비스의 단체대화방에서 유통되기 때문에 발견하기 어려울 뿐만 아니라 없애기도 쉽지 않다. 전략적 행위자는 이러한 극단화 과정을 심화하기 위해 생성형 AI를 이용할 수 있다.

○ 둘째, 의도적 가짜뉴스 이외에 오인되거나 부정확한 정보 (Misinformation)의 문제

- 부정확한 정보는 AI가 연습데이터의 기계적 해석을 통해 새로운 내용을 만들어내는 내재적 불확실성으로부터 도출된다. 즉 생성형 AI가 만들어내는 정보는 검증이나 증거 없이 맥락을 만들어내기 때문에 ‘환각적’ 일 수 있다.
- 이런 방식이 설득력이 있는 이유는 AI에 의해 만들어진 정보가 인간적 기대감이나 편견에 편승하고 있기 때문이다.
- 이런 현상은 미디어의 높은 권능을 전제로 하는데, 이는 정보의 권위는 크고, 정보의 검증가능성은 작기 때문이다.

○ 한 사회에서 가짜뉴스나 잘못된 정보가 증가하면, 이는 민주적 소통에 커다란 도전이 된다.

- 사회에서 정보를 교환하거나 타협할 수 있는 기반이 약화하고, 사회가 파편화된다.
- 독일 사회의 민주적 소통과정은 복수의 미디어 체계, 폭넓은 민주시민교육, 다양한 대안적 소통 채널 등으로 매우 복잡하다. 바로 이런 점들이 디지털 방식으로 제작되는 많은 생성물이 조작될 가능성을 꾸준히 제약하고 있다.
- 공적 영역의 소통에 대한 위협은 단기적 조작보다도 민주적 여론의 구조적 인프라 조건에 대한 장기적인 잠식이 중요하다.



- 생성형 AI는 조작과 관련하여 이미 많은 비난을 받고 있어서 미디어와 공공의 소통에 대한 신뢰를 떨어뜨리는 역할을 할 수 있다.
- 생성형 AI는 조작하거나 위조하여 만들어 낸 제작물을 통해 미디어와 저널리즘의 수익에도 영향을 미칠 수 있다.

#### □ 인공지능의 확산이 민주적 통치에 미치는 영향

○ 인공지능 기술의 확산은 생성형 인공지능과 공적 소통을 넘어서서 새로운 도전에 직면하고 있다.

- 특히 주목할 만한 분야 중 하나는 보조 또는 심지어 자동화된 의사결정 시스템에 의한 민주 정부 실천의 변화이다.
- 이러한 시스템은 유럽뿐만 아니라 전 세계적으로 더 많이 보급되고 있으며, 현재까지 사용된 많은 시스템이 기대에 못 미치고 있지만 계속해서 확산하고 있다.

○ 이 지속적인 성장의 원인은 주로 패턴인식에 특화된 인공지능 기술이 갖는 분석 능력에 기인한다.

- 이러한 능력은 자료수집과 응용프로그램 개발을 촉진하며 데이터 중심 및 효율적인 정책에 대한 기대를 높인다.
- 더 정확한 분석 및 예측 및 시뮬레이션의 사용은 제한된 시간 범위, 동적 상황에서의 응답 부재 및 제어 도구의 부정확한 적용과 같은 대표적 민주주의 절차의 약점을 보완하기 위한 것이다.
- 인공지능 기반의 절차는 정치를 더 나아지게 하고 이로써 민주주의 체제의 정당성을 향상할 것으로 기대된다.

○ 민주주의 국가에서 공권력의 행사는 특히 높은 수준의 통제를 해야 한다.

- 공권력은 기본권과 시민 권리에 구속되어야 하고, 평등과 정의의 기준을 충족하며, 국민은 자신이 어떻게 통치되고 있는지 이해할 수 있어야 하며, 항상 남용과 잘못된 전개에 효과적으로 반대할 수 있어야 한다.

- AI 기술의 향상은 이러한 과정에서 다양한 도전을 유발하고 있다.
- 기본권 및 평등 문제의 관점에서 어려움은 과거 패턴을 계속 확장하는 논리에 있다.
    - 여기에서 패턴은 ‘예측 치안(predictive policing)’에서 건강서비스 관리에 이르기까지 사례에서 보듯이 권력 및 지위의 비대칭성으로 나타난다.
    - 인공지능이 필요한 곳은 행정비용을 줄이거나 비용축소의 압박이 있는 곳으로, 특히 국경 감시와 같은 곳이다. 이러한 영역에서는 종종 취약한 그룹이 의사결정의 불리한 대상이 된다.
  - 통제와 책임에 대한 민주적 규범과 관련하여 추가적인 문제가 발생한다.
    - 기계 학습을 기반으로 하는 AI 프로세스의 불투명한 기술적 구조가 문제이고, 또 정치적 기관이 그 개발 및 운영을 위해 민간 부문 행위자의 인프라 및 서비스에 강하게 의존한다는 점이 문제이다.
    - 통제 및 책임이 실제로 가능함을 보장하는 명확한 규정이 없으면, AI 시스템은 민주적 절차의 보장을 약화한다.
  - 이는 데이터 기반의 의사결정이 자주 시민의 민주적 자기 결정과 긴장 관계에 놓인다는 사실이다.
    - 여기서 비판점은 민주주의가 AI 프로세스에 의존함으로써 데이터를 물신화하고 개인의 참여적 자기 결정의 가능성을 부정한다는 것이다.
    - 아이러니하게도 그러한 입장 자체는 사회적 행동과 민주적 정치화의 능력을 처음부터 과소평가하고, 강력한 기술적 명령에 기반을 두고 있다는 점이다.

□ 인공지능의 확산이 시민들의 정치참여 변화에 미치는 영향

○ AI 기술의 사회적 수용이 시민들의 정치참여에 어떤 영향을 미칠 수 있을까, 또 AI는 정치참여의 가능성과 방식을 변경할 수 있을까?

- 이에 대한 답은 앞서 살펴본 두 분야보다 더 불투명하다. 왜냐하면 이는 적용방식뿐만 아니라 정치체제에서의 통합과 수용에도 달려있기 때문이다. 여기에는 상향식과 하향식의 두 가지 방향이 존재한다.

○ 먼저 하향식(top-down) 강화란 정치인이 AI 프로그램을 활용해 시민과의 교류를 구체적으로 확대한다는 뜻이다.

- 이는 전반적으로 의사소통 중심의 정책을 확대하려는 추세와 일치하며, 디지털화가 직접적이고 즉각적인 의사소통을 촉진한다는 기대에 부응하는 것이다.

- 이미 행정 분야에서는 챗봇을 이용한 소통 노력이 있는 것처럼 AI 응용프로그램을 사용하여 정치를 더 반응하고 설명할 수 있게 만들 수 있다.

- 이러한 노력은 정치계와 일반 시민 사이의 정기적인 의사소통 프로세스를 지원할 수 있다.

- AI 응용프로그램은 참여환경을 구조화하고, 복잡한 토론을 하나로 모아 시민 참여가 효과적으로 이루어지는 것을 명확하게 보여줄 수 있다.

- 여기에서 민주주의의 잠재력을 느낄 수 있지만, 새로운 참여 기회가 누구에 의해 어느 정도로 어떻게 사용되는지, 또 그것이 ‘진정한’ 참여인지 아니면 정치 마케팅의 한 형태에 불과한 것인지는 아직 제대로 알 수 없다.

○ 반면에 상향식(Bottom-up) 관점에서는 AI 프로그램이 어떻게 새로운 형태의 시민사회 자체조직을 가능하게 하는지, 또 새로운 형태의 시위와 행동을 창출하는 데 도움이 되는지를 살펴본다.

- 디지털화가 정치적 행동과 관련하여 변혁적인 영향을 미칠 수 있다는 것은 잘 연구되어 있다.

- 그러나 정치적 행동에 맞춰진 사용 방식이 아직 많이 개발되지 않았고, 주로 정치 교육이나 정보에 중점을 두고 있기 때문에 현재 AI 응용프로그램에 특화된 동력은 없다.
- 정치참여의 대리인으로 봇(Bots)의 사용을 제안하는 등 정치 시스템을 재구성하기 위한 대규모 제안은 드물다.
  - 하향식 방향과 마찬가지로 상향식 구조 개발과 관련하여 더 많은 참여 기회가 어느 정도 동등한 참여를 촉진하는지 또는 이것이 기존 권력 및 자원 비대칭을 강화하는 경향이 있는지 묻는 것도 중요하다.
- 경험적 증거에 따르면, 미래에는 민주 정치가 명시적인 정치적 참여보다는 측정 가능한 시민의 행동에 더 많은 기반을 두게 될 것이다.
  - 이러한 경향은 데이터가 풍부한 사회에서 정치가 시민 태도의 분석을 통해 더욱 합법화된다는 사실에 기인한다.
  - 그러한 발전에 유리한 점은 복잡한 사회에서 고비용의 참여 형태가 종종 비효과적이고 불공정한 것으로 간주되는 반면, 정기적 선거와 같은 단순한 형태의 참여는 참여에 대한 높은 기대를 충족시키지 못하기 때문이다.
  - 따라서 간접적인 측정보다 관찰에 의존할 때 시민의 폭넓은 견해와 관심사에 대한 중립적이고 포괄적인 고려가 정확하게 이루어진다.
  - AI는 시민에게 적절하고 항상 최신의 ‘가독성’을 제공해야 하지만, 그것이 기술주의적 탈정치화를 촉진하는 것은 아닌지 살펴야 한다.

## □ 전망

- AI 프로세스가 점점 더 사회적으로 활용된다는 관점에서 민주적 실천을 위해 제시된 잠재적 발전 경로는 일관성이 없고 모호한 미래상을 보여주고 있다.
  - 인공지능은 대중적 비판이 제시하는 것처럼 단순히 민주적 과정에

대한 경쟁자가 아니다. 개발의 변혁적 효과는 효과적인 조작 가능성의 증가에만 국한되지 않는다.

- 따라서 개인, 국가 및 민간 부문 행위자의 행동을 더 잘 이해하기 위해서는 AI 기술과 민주적 발전의 연계를 더 자세히, 그리고 민주적 삶의 구체적인 측면에서 이해하는 것이 필요하다.

○ 민주적 관점에서 볼 때 디지털 기술의 발전은 좀 더 적극적인 정치적 발전을 가능하게 한다.

- 기술 혁신만 자유롭게 허용되던 디지털화의 초기 단계와는 대조적으로 현재 시점에서는 민주적으로 민감한 규제가 필요하다.
- AI 프로세스의 현재 단계는 주로 응용 시나리오(예: 안면 인식 또는 딥페이크)와 AI의 데이터 기반 문제(데이터 보호 및 저작권)에 중점을 두고 있다.
- 민주적 관점에서 볼 때 AI의 개발, 적용 및 사용과 관련하여 사회적 공동결정 가능성에 더 큰 관심을 기울여야 한다.
- 이는 단순히 투명성에 대한 요구에 국한되지 않고, 기술적 가능성과 경제적 유인에도 불구하고 공익과 시민적, 민주적 권리를 강화할 수 있는 개입이 가능한 제도의 설립을 의미한다.
- 이는 간단한 투명성 요구로 그치지 않고, 기술적 가능성 및 경제적 인센티브에 맞서 공공 이익 및 시민 및 민주적 권리를 강조하는 효과적인 제도의 수립을 의미한다.

## 나. 연방정치교육원의 “디지털 가짜정보” 위험 분석<sup>158)</sup>

□ 디지털 가짜정보 또는 가짜뉴스(digitale Desinformation)

- 우리는 정보의 홍수 시대에 살고 있으며, 누구나 전 세계와 순식간에 소통할 수 있다. 동시에 보고, 읽고, 들을 수 있는 수많은 뉴스, 비디오,

158) Kettemann, Matthias. 2021. “Ordnungsbildung in Online-Kommunikationsräumen.” Bundeszentralen für politische Bildung (October 4) <https://www.bpb.de/themen/medien-journalismus/digitale-desinformation/341326/ordnungsbildung-in-online-kommunikationsraeumen/> (검색일: 2023. 12. 01).

포스트 등이 넘쳐나고 있다.

- 문제는 누가 그런 정보를 우리에게 오도록 결정하는지, 어떻게 좋은 정보와 나쁜 정보를 구분할 수 있는지, 과연 옳고 그른 정보를 정확하게 구분할 수 있는지, 만약 그렇다면 누구에게 그런 권한이 있는지 등이다.
- 그러한 권한이 플랫폼에 있는가, 아니면 국가에 있는가?

○ 디지털 가짜정보로부터 민주적 공공의 영역을 보호하기 위해 개입해야 한다는 목소리가 날로 커지고 있다.

- 가짜정보를 정의하는 일도 쉬운 문제가 아니다. 예를 들어 풍자는 일반적으로 가짜정보라고 하지 않기 때문이다.
- 미디어에 대한 신뢰를 떨어뜨리려는 전략적 목적을 가지고 거짓 정보를 의도적으로 확산시킬 때 비로소 가짜정보라는 말은 의미가 있다.

○ 가짜정보의 방지 문제와 관련하여 법률, 공동체 기준(Gesellschaftsstandard), 플랫폼디자인(Plattformdesign) 등에 논의의 초점이 모이고 있다.

□ 디지털 가짜정보가 위험한 이유

- 간단히 쉽게 전파할 수 있다.
- 최소비용으로 다수에게 전파가 가능하다.
- 일반적으로 추가적인 정보원이 없이는 그 거짓 정보를 구별하는 것이 어렵다.
- 편견을 갖고 있거나 부족한 지식을 보유한 사람들은 온라인상의 가짜정보를 쉽게 받아들일 수 있다.
- 익명성과 부족한 대면 접촉은 협력적 인간관계를 어렵게 하고, 온라인

상의 공격적 언행을 쉽게 받아들이게 만든다.

□ 가짜정보가 사회질서를 위협할 것인가?

○ 아니다.

- 거짓말은 항상 존재하고, 가짜정보는 온라인상의 거짓말로 하나의 새로운 현상이지만, 독일 사회에 근본적인 가치변화를 가져올 수는 없다고 본다.
- 누군가 가짜정보를 연구한다면, 개별 가짜정보의 구체적 효과를 측정하기가 쉽지 않다는 사실을 알게 될 것이다.
- 사람은 자신의 지식을 수많은 다양한 채널을 통해서 얻게 된다. 인터넷을 통해 그러한 채널은 점점 더 넓어지고 있다.
- 하지만 우리가 뉴스를 습득하는데 있어서 전통적 미디어는 학생이나 젊은 층에 이르기까지 여전히 중요한 역할을 하고 있다.

○ 일반적으로 누군가 편견을 갖고 있거나 모자란 지식을 가진 경우, 거짓정보나 가짜정보에 취약하게 된다.

- 선입견에 대한 인간적 약점은 독자로 하여금 자신의 생각에 합치하는 정보를 선호하게 하고, 자신의 세계관에 반하는 정보를 꺼리게 한다. 가짜정보 유포자는 바로 이점을 이용한다. 점점 더 많은 정보제공자가 인터넷에서 이런 점을 전략적으로 이용하여 정보를 유포하고 있다.
- 사람들이 흔히 자신이 어떤 뉴스를 판단할 때 뉴스 자체가 아니라 자신이 이용하는 매체에 대한 신뢰도에 따라 거기에 나온 뉴스의 가치를 판단하는 경향이 있다.
- 사람이 언제나 모든 필요한 정보를 스스로 알아낼 수는 없기 때문에 그러한 신뢰는 어쩔 수 없는 측면이 있다.
- 전통적인 기자의 기사를 신뢰해야 한다. 그러나 그러한 신뢰는 온라인 세상에서는 새로이 정의되어야 한다.

## □ 가짜정보 논의의 법률적 측면

○ 국가는 디지털 세계에서 인권과 자유를 보호하기 위해 일차적인 책임이 있다.

- 국가의 책임은 단순히 키보드나 스마트폰에서 끝나지 않는다. 개인의 자유를 보호하는 일은 다소 모순적인 측면이 존재한다. 인권을 침해하지 않기 위해서 국가는 소극적 의무(negative Verpflichtung)를 다해야 하지만, 또 인권을 보호하고 실제로 실행하기 위해서는 적극적 의무(positive Verpflichtung)가 필요하기 때문이다.

- 국가의 모든 개입은 항상 어떤 맥락 속에서만 평가할 수 있다. 국가는 인터넷상에서 법률이나 플랫폼의 규칙에 대한 통제를 통해서 그런 의무를 수행하려고 노력하고 있다.

○ 온라인 소통 공간은 국제적 법률, 지역적 기준, 국내 법률, 플랫폼 관련 민법 등 다양한 차원에서 규제받고 있다.

- 대부분의 가짜정보는 국가에 의해 간단히 금지될 수 없다. 대다수 가짜뉴스는 (현행법에 따르면) 처벌 대상이 아니기 때문이다.

- 개인에 대한 가짜정보가 단순한 의견이 아니라 거짓 주장을 담고 있다면, 금지할 수 있다. 중상모략은 처벌할 수 있다.

- 그러나 거짓 정보를 확산하려는 권리를 제한하려는 새로운 규정은 사전에 반드시 의사 표현의 자유와 대립하지 않는지 따져보아야 한다.

## □ 가짜뉴스 회피방안

○ 가짜정보에 대한 제재는 개인 생명이나 신체에 대한 직접적 위협의 정도가 매우 높을 때에만 헌법적으로 가능하다.

- 선거 이전이나 도중에 정보의 홍수 속에서 가짜정보로부터 유권자를 보호하는 일은 매우 중요하다. 거짓 주장을 적시에 교정하기는 어렵기 때문이다.<sup>159)</sup>

159) 선거와 관련하여 모든 정보를 일일이 검증하는 것은 불가능하다. 성복숙, 2022. “20대 대선 관련



- 가짜뉴스 확산을 막기 위해서는 국가의 개입뿐만 아니라 자발적 팩트체크도 중요하다. 특히 정보에 대한 시스템적 개입이 효과적이다.
- 가짜뉴스는 전적으로 법적인 문제라기보다는 하나의 사회적 해결 과제이다.

#### 다. 독일의 AI 및 생성형 AI 선거 규제 개요

- 생성형 AI가 급속하게 발전하고 있음에 따라 향후 독일의 선거 과정에서도 생성형 AI가 영향을 미칠 것으로 예상되는바, 그에 대한 독일의 대책을 살펴본다.
- 가짜뉴스나 (생성형) AI가 선거에 미칠 영향과 관련하여 독일 연방선거관리위원회, 연방정치교육원(BpB), 독일의 최신 논문 등을 통해 그에 대한 대책을 알아본다.
- 독일의 (생성형 AI에 의한) 가짜정보에 대한 규제는 현재 상황에서 법적 접근보다는 자율규제 쪽에 더 가깝다고 할 수 있다.

#### 라. 연방선거관리위원회

##### □ 2021년 연방총선 대비 연방선거관리위원장 인터뷰

- 2021년 9월 26일 연방의회 선거를 앞두고 연방선거관리위원장 (Bundeswahlleiter, Dr. Georg Thiel)은 선거에 대한 가능한 사이버공격, 가짜뉴스 선거운동, 정보 조작 시도 위험 등의 내용을 언급하며, “어떻게 연방총선을 안정되게 치를 것인가, 누가 선거의 안전을 돌볼 것인가, 어떤 대책들이 가능한가?”에 대해 인터뷰했다. 여기에 인공지능과 관련한 주요 내용을 소개한다.<sup>160)</sup>

가짜뉴스의 양상과 그 사회적 영향.” 『인문사회 21』, 13집 3호.

160) Bundeswahlleiter. 2021. “Erkennen und Bekämpfen von Desinformation.” Die Bundeswahlleiterin. <https://www.bundeswahlleiterin.de/bundestagswahlen/2021/fakten-fakenews.html#7b083e33-e95d-4a3f-87a3-276f5e01eb6e> (검색일: 2023. 11. 24).

□ 연방총선의 안전에 대한 책임

○ 연방총선을 규정에 맞추어 안전하게 치르는 것은 잘 작동하고 있는 우리 민주주의의 기본조건이다.

- 연방선거관리위원장은 주 및 지역선거관리위원장, 그 밖의 다수 기관과 협력하여 선거를 준비하여 치르고, 그 결과를 발표할 책임과 권한이 있다.

○ 선거의 안전과 규정에 따른 선거 진행은 다양한 측면을 포함한다.

- 먼저 정보와 소통 기술의 작동 측면인데, 반드시 정보의 누락이나 조작을 방지해야 한다.

- 다른 측면은 우편 선거나 선거함의 조작 시도를 방지하는 것이다. 이를 위해 명확한 규정이 필요하며, 우리는 이의 준비와 실행을 위해 수많은 대책을 준비하고 있다.

□ 선거에 정보처리기술을 적용하는 문제

○ 연방총선의 선거 당일, 선거 결과를 예측하고 임시결과를 발표하기 위해 정보처리기술(Informationstechnik)을 적용하고 있다.

- 원칙적으로 독일의 선거는 분권적으로(dezentral) 조직되어 있다.

- 이는 각각의 선거관리기구가 정보처리기술을 사용하는 것에 권한을 갖는다는 의미이다.

○ 투표함의 투표용지를 세는 일은 각 투표소(Wahllokal)에서 수작업으로 이루어진다.

- 게마인데는 각 투표소의 선거 결과를 취합하여 지역구선거관리위원회(Kreiswahlleitung)에 전달한다.

- 지역구선거관리위원회는 지역구의 선거 결과를 집계하고, 그 결과를 16개 주(州) 선거관리위원회(Landeswahlleitung)에 보낸다.

- 각각의 주 선거관리위원회는 주의 선거 결과를 집계하여 그 결과를

연방선거관리위원회에 전달한다.

○ 게마인데 차원에서부터 정보처리기술이 적용될 수 있다.

- 어떤 경우에든지 주 선거관리위원회는 최종 선거 결과를 늦어도 선거 당일 밤까지는 연방선거관리위원회에 이메일 등의 방식으로 (elektronisch) 전달한다.
- 연방선거관리위원회는 이 결과들을 가지고 임시 공식 선거 결과를 계산하여 선거 당일 밤 또는 다음 날 아침에 이 결과를 발표한다.

#### □ 연방총선의 안전을 위한 노력

○ 연방총선에서 정보처리기술의 보안성 확보는 최우선 과제이며, 이를 위해 다수의 기관이 일하고 있다.

- ‘연방정보처리기술보안청(Das Bundesamt für Sicherheit in der Informationstechnik, BSI)’ 은 특히 연방과 주 선거관리위원회가 정보 보안의 위협을 감지하고 대응할 수 있도록 지원하고 있다.
- 연방선거관리위원회는 선거에서 위 보안청과 다른 보안관청과 함께 개표과정을 관찰하고, 이미 준비된 보안예방책을 지속적으로 점검하고 이를 적용하고 있다.
- IT 기술의 가능한 취약점을 찾아내고 그것을 보완하기 위해 사이버공격과 같은 다양한 침입전략에 대해 준비하고 있으며, 여러 시나리오에 대비하고 있다.

#### □ 선거에 대한 하이브리드 위협

○ 독일에서 ‘하이브리드 위협(hybride Bedrohungen)’ 이란 개념은 공공의 의사결정에 영향을 미치려는 특정한 시도라고 본다. 이러한 영향력은 다양한 방법과 도구를 통해 생겨난다.

- 예를 들어 사이버공격이나 선전, 가짜뉴스 등 거짓이나 왜곡된 정보를 언론이나 SNS에 특정한 의도를 가지고 확산하는 것이다.

- 그 주체는 한편으로 국내의 행위자일 수도 있고, 다른 한편으로 그런 수단을 가진 다른 나라에서도 가능하다.

○ 연방내무부(BMI)는 그와 같은 하이브리드 위협은 연방총선에서도 가능하다고 본다.

- 그런 행위는 독일에서 주권을 가진 시민의 정치적 의사 형성에 반하는 일이다.
- 연방내무부는 산하 정보기관과 연방정부의 다양한 기관들과 함께 위협상황을 분석하고 있다.
- 연방내무부는 연방총선에서 제3국을 통한 불법적 영향력 행사를 방지하기 위해 부처 간 대책을 조율하고 있다.

#### □ 하이브리드 위협에 대한 연방선거관리위원회의 역할

○ 유권자에게 객관적 자료와 사실에 근거한 정보를 제공하는 일은 매우 중요하다. 동시에 선거 결과에 대한 정확한 집계도 중요한 일이다.

- 가짜정보 캠페인은 거짓이나 왜곡된 정보로 선거의 진행 과정, 보안과 유효성을 훼손하는 것을 목적으로 한다. 그들은 유권자에게 영향을 주고 선거 과정과 결과에 대한 불신을 유도하는 것을 목표로 한다. 이것을 통해 우리의 민주주의에 대한 의혹을 불러오려는 것이다. 우리는 모두 이에 맞서 싸워야 한다.

○ 연방선거관리위원회는 선거와 관련하여 정파를 초월한 공식적이고 신뢰할 만한 정보원이다.

- 위원회는 다양한 채널을 통해 종합적인 정보를 적극적으로 제공하고 있다.
- 홈페이지([www.bundeswahlleiter.de](http://www.bundeswahlleiter.de))는 배경정보와 최신 뉴스를 제공한다. 트위터와 인스타그램(@Wahlleiter\_Bund)은 선거와 관련된 모든 주요 정보를 올리고 있다.

- 또한 위원회의 부서는 가짜뉴스를 감지하고 찾아내기 위해 전통적 양식의 언론과 최신 SNS의 상황을 관찰하고 있다.
- 최근 홈페이지에 “가짜뉴스와 사실(Fakten gegen Fake News)” 이라는 코너를 신설하였다. 여기서는 SNS에서 선거에 대해 잘못되거나 왜곡된 주장을 찾아내 알려주고 있다. 언론이나 유권자는 이곳에 문의할 수 있다.
- 그밖에 연방선거관리위원회는 선거와 관련한 주요 주제에 대해 연방 정치교육원(Bundeszentrale für politische Bildung)과 협업하고 있다.
- 우리는 정확하고 종합적인 공식적 정보가 유권자에게 무언가 영향력을 행사하려는 시도를 막는 중요한 수단이라고 생각한다.

○ 언론은 어떤 경우에도 객관적이고 신뢰할만한 정보를 위해 애써야 하고, 이를 통해 굉장히 가치 있는 일을 하게 된다.

- 독일은 다양하고 비판적인 언론환경을 가지고 있으며, 언론은 객관적이고 중립적인 보도를 하고 있다.
- 일부 언론은 잘못된 정보나 소문을 찾아내고 바로잡는 특화된 사실 확인자(Faktenprüfer/factchecker)이다. 언론은 선거에 관한 올바른 정보를 확산시킬 수 있는 중요한 바탕이다.

## □ 가짜뉴스 대책

○ 유권자는 스스로 가짜뉴스를 가려낼 책임이 있다.

- 유권자는 다양한 채널을 통해 종합적인 정보를 얻을 수 있다. SNS에 막 올라온 정보에 대해 언제나 비판적으로 살펴보고 신뢰성을 따져보아야 한다.

○ 연방총선은 선거관리기구와 기관, 선거관리 보조자, 정당, 언론, 시민 등 우리 모두에게 하나의 거대한 도전이다.

- 그러나 선거는 우리 민주주의의 초석이다. 모두가 협력하여 규정에 따

라 안전한 선거를 치르는 것은 무슨 일이 있어도 대단히 값진 일이다.

#### 마. 연방선거관리위원회의 “가짜뉴스와 사실(Faken gegen Fake News)” 코너

##### □ 가짜뉴스의 인지와 극복

○ 가짜뉴스란 여론에 영향을 미치거나 속이기 위해 (증명이 가능한) 거짓되거나 왜곡된 정보를 확산하는 것을 말한다.

- 가짜뉴스를 감지하고 바로잡을 수 있는 하나의 중심기관이 독일에는 아직 없다.

○ 선거나 선거 과정과 관련된 가짜뉴스를 인지하고 바로잡는 일은 연방선거관리위원회의 몫이다.

- SNS상 콘텐츠 내용의 사실 여부에 대해서는 SNS 사업자가 책임을 진다. 이를 감독하는 권한은 주 언론기구들(Landesmedienanstalten)에 있다.

##### □ 가짜뉴스의 인지(Erkennen von Desinformationen)

○ 가짜뉴스를 이용한 선거운동을 감지하는 일은 특히 정보기관들(Nachrichtendienste)과 기타 연방 소속의 보안관청, 예를 들어 연방범죄수사청(Bundeskriminalamt), 연방헌법수호청(das Bundesamt für Verfassungsschutz)에 있다.

- 그 밖에도 가짜뉴스를 이용한 선거운동을 감지하는 일에 대한 특별한 책임은 언론사, 방송국과 기타 미디어에 있다.

##### □ 예방과 바로잡기(Vorbeugung und Richtigstellung)

○ 가짜뉴스를 바로잡기 위한 가장 중요한 수단은 객관적이고 신뢰할 만한 정보이다. 예방적 홍보작업과 거짓 정보를 바로잡는 일은 특히 아래 기

관의 과제이다.

- 선거 과정을 위한 연방과 16개 주 선거관리위원회(Bundes- und Landeswahlleiter)
- 연방총선이나 정당을 위해 일반 정보를 제공하는 연방과 주 정치교육원(Bundeszentrale und Landeszentralen für politische Bildung)
- 연방정부의 대책, 법안, 전략, 정책 배경 관련 연방언론청(Bundespresseamt)
- 정당과 후보자의 정치공약과 정책내용
- 그밖에 정보나 홍보 관련 중요한 역할을 하는 곳은 언론사, 방송국과 기타 미디어이다. 그들은 객관적 보도 자세와 사실 체크를 통해 거짓 정보를 가려내어 그 확산을 억제하고 있다.

○ 시민의 미디어와 디지털 역량 강화는 어떤 경우에도 가짜뉴스의 확산을 억제하는 데 기여하고 있다.

- 정보를 비판적 시각으로 마주할 수 있는 것은 누가 어떻게 뉴스를 만들었는지를 이해하는 것을 전제로 한다.
- 또한 디지털 기술이 어떻게 작동하는지, 예를 들어 어떻게 뉴스가 SNS에 퍼지는지에 대한 기본적인 기술적 이해가 필요하다.

□ 선거관리기구의 중립화, 선거 과정의 투명한 통제

○ 선거관리위원회에 대한 시민의 신뢰와 선거 결과의 수용을 위한 기본조건은 법적으로 의무화된 선거관리위원회의 중립성과 헌법으로 보장된 선거의 공공성 원칙이다.

- 선거의 모든 중요한 단계는 공개적으로 검증할 수 있다.

○ 투표의 정확한 결과산정을 보장하기 위해서는 특히 다음 사항들이 중요하다.

- 언제든지 다시 셀 수 있는 유권자의 선택을 담은 물리적 투표용지의

## 제출

- 서명을 할 수 있는 선거관리지도부와 선거관리위원회의 공개적 논의와 회의
- 복수의 선거관리 지도부와 선거관리위원회: 다수의 감시원칙 (Mehraugenprinzip), 상호견제 가능
- 선거 관련 착오나 문제가 있을 때 연방의회와 연방헌법재판소에서의 재검표 가능성

## □ 사이버 공격으로부터 선거 보호

- 연방정보처리기술보안청(BSI), 16개 주 선거관리위원회와 연방선거관리위원회는 잠정적 선거 결과를 보장하는 요구사항을 충족해야 한다.
- 그 요구사항은 선거 당일 저녁 잠정 결과의 전달에서 기술적으로 투명하고 높은 안전수준을 보장하는 것이다.
- 2009년 연방헌법재판소의 결정에 따라 연방총선거 유럽의회선거에서 선거용 기계를 사용하지 않고 있다.
  - 모든 투표행위는 반드시 투표용지를 사용하는 종이 형태로 이루어진다. 따라서 투표행위는 IT 안전성에 의존하지 않아도 되고, 조작 위험성이 없다.

## □ 잠정 선거 결과(Vorläufiges Wahlergebnis)

- 잠정 선거 결과의 집계에서 부분적으로 IT의 지원을 받는다.
  - 어떤 기기나 소프트웨어를 사용할 것인지는 권한을 가진 게마인데와 선거위원회가 결정한다.
  - 그들은 IT의 안전성에 스스로 책임을 진다. 연방정보처리기술보안청(BSI)은 그들을 지원한다.



○ 잠정 선거 결과를 산출하고 전달하는 과정에서 IT 인프라 구조에 대한 사이버공격은 잠정 결과의 발표를 지연시킬 수 있다.

- 그러나 새로이 연방의회를 구성하는 데 결정적인 역할을 하는 것은 선거일로부터 약 3주 후에 나오는 최종 선거 결과이다.

#### □ 최종 선거 결과(Endgültiges Wahlergebnis)

○ 최종 선거 결과는 물리적 서류 형태, 즉 선거기록문서에 따라 산출된다. 그래서 그 결과는 사이버 공격으로 조작될 수 없다.

- 투표용지와 선거관리위원회에 의해 만들어진 결과를 산출한 문서는 언제든지 검증할 수 있다.

- 이 최종 문서는 원칙적으로 다음 선거의 60일 이전까지는 보존되어야 한다.

#### 바. 독일의 AI 규제 방식

○ 독일의 (생성형 AI에 의한) 가짜정보에 대한 규제방식은 아직은 법제적 접근보다는 자율규제 쪽에 더 가깝다고 할 수 있다.

○ 물론 2017년에 제정된 ‘네트워크 법집행개선법 (Netzwerkdurchsetzungsgesetz, 흔히 Facebook-Gesetz라고도 함)이 있어서 일견 법률에 의한 규제를 시도하고 있는 것처럼 보인다.

- 하지만 이 법의 주요 내용은 SNS 사업자에게 특정의 위법정보(가짜 정보)에 대한 민원 처리의 의무화, 처리결과에 대한 보고서 작성 및 공표의 의무화, 위반자에 대한 고액의 과태료 부과이다.

- 이 법이 대응을 의무화하고 있는 위법 콘텐츠는 반드시 형법상의 범죄의 구성요건에 해당하는 것이라고 명시하고 있다. 형법상 위법에 해당하지 않는 정보는 이 법의 대상에서 제외된다.

○ 즉, 이 법은 가짜정보나 AI 활용에 대한 새로운 규제방안을 내놓은 것이 아니라, 원래 형법에 처벌하게 되어 있는 내용을 SNS상에서도 적용하라는 의미에 불과하다.

- 실제로 이 법안의 입법 당시 청문회 과정에서도 논란이 많았으며, 그 주요 쟁점은 이 법안이 표현의 자유를 제약할 위험이 있다는 것이었다.

○ 따라서 독일에서는 가짜정보와 관련하여 법률에 의한 규제보다는 아직 자율규제에 중점을 두고 있는 것으로 보인다.

#### 아. 생성형 AI 관련 규제의 갈림길에 선 독일<sup>161)</sup>

□ 독일 연방의회 디지털 위원회는 2023년 5월 24일 전문가들을 초청하여 청문회를 실시했다. 여기에 그 주요 내용을 소개한다.

○ 생성형 인공지능(AI) 분야에서는 효과적인 규제와 혁신적인 비즈니스 모델 촉진 사이에 적절한 균형이 필요하다.

- 청문회에서 전문가들은 직장과 사회 세계에 대한 생성형 AI의 잠재력과 위험을 평가하고 시간이 핵심이라는 점을 한목소리로 강조했다.

- 이는 AI 기반 텍스트에 라벨을 표시하는 방법과 같은 투명성 문제와 관련하여 적용되며, 유럽의회가 6월 중순에 투표할 예정인 AI 법안에 대한 유럽의 접근방식과 관련하여도 적용된다.

○ 고객과 경쟁업체에서의 불확실성

- 하이델베르그 스타트업 Aleph Alpha의 요나스 안드루리스(Jonas Andrulis)는 규제 수준을 가능한 한 낮게 유지하고 속도를 높일 것을 주문했다.

- 그는 현재의 규제 상태에 대해 ‘고려할 가치가 있다고

161) Deutscher Bundestag. “Generative Künstliche Intelligenz: Regulierung steht am Scheideweg.” <https://www.bundestag.de/dokumente/textarchiv/2023/kw21-pa-digitales-ki-947976> (검색일: 2023. 12. 01).

(bedenkenswert)' 고 하면서 초안에는 모호한 법률 용어와 영향 평가가 많이 포함되어 있다고 지적했다. 따라서 고객과 경쟁업체에는 불확실성이 높다.

- AI는 미국의 거대 기술 기업들이 현재 '투명하지 않은 제안'으로 다시 분위기를 조성하고 있는 '신산업 혁명'의 기본 기술이다. 그러나 독일과 유럽에는 실질적, 경제적 주권이 매우 중요하다.
- Jonas Andrulis는 AI 분야의 중소기업과 관련하여 “유럽에는 마이크로소프트나 구글이 없다” 라고 언급했다.

#### ○ AI 애플리케이션을 위한 일종의 TÜV

- 인공지능 연방협회의 라스무스 로테(Rasmus Rothe)도 경쟁력을 유지하기 위해 중소기업을 배려하고 좀 더 실용적이어야 한다고 주장했다.
- 그는 생성형 AI가 인간보다 많은 작업을 더 잘 수행할 것이며, 그런 기술이 경제성장으로 이어질 수 있다고 지적했다.
- 우리는 항상 AI의 구체적인 용도를 살펴야 하며, 앞으로는 AI 애플리케이션을 위한 TÜV 방향에 대해 더 많이 생각해야 한다고 그는 제안했다.
- 프랑크푸르트(오데르) 비아드리나 대학의 필립 해커(Philipp Hacker)는 생성형 AI가 우리가 소통하고, 생활하고 일하는 방식을 크게 변화시킬 것이라고 강조했다.
- 규제와 관련하여 우리는 현재 코스를 올바르게 설정해야 하는 교차로에 있다. 유럽의회의 최근 초안은 ‘올바른 방향’ 으로 가고 있지만, 여전히 심각한 결함이 있다고 그는 지적했다.
- 무엇보다도 그는 사회적으로 유용한 애플리케이션이 느려지거나 차단될 위험이 있다고 생각한다. 그는 EU에서는 사용할 수 없는 구글이 개발한 AI 기반 챗봇 ‘바드(Bard)’ 를 예로 들었다.
- 또 다른 문제는 중소기업이 감당하기 힘든 컴플라이언스(Compliance) 비용이다.

- 해커는 교육 데이터 등 특정 최소 표준과 의료 분야 등 고위험 생성형 AI 애플리케이션이 사용되는 영역에 대한 애플리케이션 관련 규정을 지지한다고 밝혔다.

#### ○ 개발자와 사용자 간 협력

- 암스테르담 대학교의 나탈리 헬베르거(Natali Helberger)도 기본 모델에 관한 규제의 진전을 높이 평가했다.
- 이러한 모델은 다양한 목적으로 사용되므로 개발자에게 시스템 보안에 대해 생각할 동기를 부여하는 것이 중요하다. 그렇지 않으면 예측할 수 없을 정도로 많은 다운스트림 사용으로 인해 훈련 데이터의 적법성 또는 품질 문제가 발생할 수 있다.
- 예를 들어 디지털 서비스법(Digital Services Act, DSA)의 거대 온라인 플랫폼에 대한 시스템적 위험을 평가하는 조항에서 배울 수 있다고 그녀는 언급했다.
- 그녀는 위험 모니터링을 강화하기 위해 연구를 더 많이 활용하는 데 찬성하고, 책임과 관련하여 개발자와 사용자 간의 협력이 필요하다는 점을 지적했다.
- ‘Oxford Internet Institute’의 산드라 와처(Sandra Wachter)도 이러한 측면을 강조했다. 설계를 결정하는 사람이 위험에 영향을 미친다. 예를 들어 그는 편견에 대응하기 위해 새로운 텍스트를 포함해야 하는지를 결정할 수 있다고 그녀는 설명했다.

#### ○ 직업 세계에 미치는 영향

- 베를린 응용과학대학교의 크리스티안 켈러만(Christian Kellermann)은 AI가 업무와 고용에 미치는 영향을 설명했다.
- 그는 현재 연구 상황은 좋지만, 설명력은 여전히 제한적이라고 지적했다. 이는 작업 대체 정도에 대한 신뢰할 수 있는 진술과 관련이 있다.
- 생성형 AI는 번역, 코드 생성 또는 창의적 글쓰기에서 잘 작동한다고

가정할 수 있으므로 번역가, 여론 조사원 또는 PR 전문가 등이 특히 이에 노출될 수 있고, 특히 조립 또는 공장 근로자의 경우에는 상황이 다를 것이라고 그는 언급했다.

- ALLAI의 카텔리진 물러(Catelijne Muller)는 유럽의 규제 노력이 보조를 맞출 수 없는 경우는 아니라고 강조했다.
- 대규모 시스템은 이제 규제될 수 있다: OpenAI와 같은 회사도 허위 정보, 가짜 뉴스 또는 시스템의 ‘환각’ 과 같은 문제를 고려하여 스스로 이를 요구하고 있다고 언급했다.

#### ○ 우선순위, 속도, 관할

- 킬(Kiel) 전문대학의 도리스 베셀스(Doris Weßels)는 현재 많은 세부 사항이 지배하고 있다고 말했다. 세 가지 기본 문제가 있는데, 첫째, 우선순위의 부족, 둘째, 불충분한 속도, 셋째, 관할의 분산 필요하다.
- Weßels는 다른 자금의 조달 옵션이 필요하다는 관점에서 프로세스와 구조가 느려지고 있지만, 높은 역동성이 필요하다고 주장했다.
- 그녀는 AI 법의 분명한 근거와 표시 의무에는 찬성했지만, 과잉 규제가 있어서는 안 된다고 강조했다.

## 4.7 스위스

### 가. 최근 생성형 AI 선거규제 관련 논의<sup>162)</sup>

- 스위스 5개 정당이 2023년 10월 22일 연방 선거 캠페인에서 인공지능(AI)의 사용을 제한하기로 합의했다.
- 공동 성명서에 서명한 사회민주당(the Social Democratic Party), 중앙당(the Centre Party), 녹색당(the Green Party), 자유녹색당(Liberal Green Party), 개신교당(the Protestant Party)은 각 주(州) 지부에 녹화 광고, 포

162) Ajansi, Anadolu. 2023. “5 Swiss political parties agree to limit use of AI ahead of elections.” (September 25) <https://www.aa.com.tr/en/europe/5-swiss-political-parties-agree-to-limit-use-of-ai-ahead-of-elections/3000856> (검색일: 2023. 09. 26).

스터, 광고 등 선거운동 과정에서 AI를 사용할 때 이를 명시적으로 신고해야 한다는 행동강령을 준수하도록 해야 한다고 주장했다.

- 또한 상대 후보를 공격하는 이른바 ‘네거티브’ 캠페인에 AI를 사용하는 것을 금지하는 조항 또한 포함된다.
- AI는 새로운 형태를 의사소통할 수 있게 하는 기회이나, 사실을 왜곡하거나 특정 정치 행위자가 말하지 않았을 내용을 특정 정치 행위자에게 들릴 수 있는 위험도 있다고 강조했다.
- 급진자유당은 딥페이크 등을 예로 들며 AI의 악용을 금지하기 위한 노력이 필요함을 강조했지만 성명서에 서명하지는 않았다.
- 스위스 최대 정당 스위스 인민당은 네거티브 캠페인의 경계가 모호하다고 주장하며 AI의 선거 사용 금지에 대해 유보 의견을 밝혔다.
- 스위스 정부는 여전히 새로운 정책이나 법안의 형성에 관한 관심보다는 기존의 접근에 의존하고 있다.

#### 나. AI 규제에 대한 의회의 논의: AI와 민주주의의 위기와 기회 - 연방의회 의원의 질의와 정부의 답변<sup>163)</sup>

##### □ 연방의회 사민당 의원의 질의 2023.06

- 인공지능(AI) 분야의 급속한 발전은 민주주의에 대한 도전이다.
  - AI 시스템을 사용하면 허위와 선전을 통한 조작의 위험이 증가하여 여론 형성과 의지 형성은 물론, 우리의 투표 및 투표 행동에 직접적인 영향을 미칠 것이라는 우려가 있다.
  - 투표권을 가진 사람들이 어떤 정보를 신뢰할 수 있는지, 그 정보가 어떻게 제공되는지를 더 이상 알지 못하게 되면 기관에 대한 신뢰도가 떨어진다. 이는 민주주의에 대한 위협이다.

163) Celine, Widmer. 2023. “Künstliche Intelligenz. Gefahren und Potenziale für die Demokratie.” Das Schweizer Parlament (June 15) <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233812> (검색일: 2023. 12. 02).

○ 그러나 AI는 다른 많은 분야와 마찬가지로 정치 분야에서도 위험뿐만 아니라 기회도 제공한다.

- 알고리즘은 예를 들어 정치 참여나 정보 분야 등 민주주의 강화에 기여할 수 있다고 생각된다.

○ 이러한 배경에서 연방정부에 제기한 질문은 다음과 같다.

1. 연방정부는 특히 투표 및 선거와 관련하여 민주주의의 위험과 기회를 인식하고 있는가, 그것은 무엇인가?
2. AI가 민주주의에 위협이 되는 것을 막기 위해서는 어떤 규제가 필요한가?
3. 이러한 위험과 기회가 알려지지 않을 경우, 연방정부는 이와 관련하여 가능한 한 빨리 이를 명확히 하기 위해 무엇을 하고 있는가?
4. 연방정부는 사람들이 진짜 정보와 가짜 정보를 구별하도록 하는데 어떤 대책이 효과적이라고 생각하는가?
5. 연방정부는 신뢰할 수 있는 정보에 대한 접근을 어떻게 강화하고자 하는가?

□ 연방정부의 답변 2023.08

○ 최근 몇 년 동안 인공지능(AI)의 개발과 응용이 더욱 진전되었다.

- 2019년 12월 “인공지능의 과제” 보고서에서 연방정부는 스위스가 인공지능의 과제를 해결하기에 전반적으로 좋은 위치에 있음을 보여주었다.

- 의견 및 의사결정 영역에서 AI의 사용과 관련하여 연방정부는 대중매체의 해당 발전을 모니터링하고자 했다.

- 소셜미디어와 검색 엔진의 경우, 연방정부는 DETEC와 연방총리실에 대중 커뮤니케이션을 위한 AI의 위험성을 식별하는 보고서 작성을 의뢰했다.

- 보고서 “중개자와 커뮤니케이션 플랫폼”은 2021년 11월에 출판되었다. 이에 따라 연방정부는 2023년 4월 5일 초대형 통신 플랫폼을 규제하기로 했다.

○ 1번부터 3번까지의 질문에 대해:

- “중개자 및 커뮤니케이션 플랫폼” 보고서는 커뮤니케이션 플랫폼과 대중 커뮤니케이션에 사용하는 알고리즘의 긍정적인 잠재력과 부정적인 잠재력을 모두 식별한다.
- 긍정적인 잠재력에는 무엇보다도 정보 찾기(정보의 자유)가 더 쉬워지고, 대중 커뮤니케이션에 더 쉽게 참여하고, 모든 의견이 대중에게 공개될 기회가 향상되고(평등), 공개 의견의 다양성이 증가하는 것이 포함된다.
- 부정적인 가능성에는 무엇보다도 증오심 표현의 확산(그 결과 공격을 받은 사람과 그들의 입장이 공개 담론에서 제외됨), 잘못된 정보, 사용자 권리 침해(예: 계정 삭제) 및 정보 부족과 불투명성이 포함된다.
- 부정적인 가능성을 고려하여 2023년 4월 5일 연방정부는 스위스에서 많이 사용되는 통신 플랫폼에 대한 규제를 모색하기로 했다.
- 또한 인공지능에 대한 최초의 구속력 있는 국제협약을 위해 (스위스 출신 의장 주도하에) 유럽평의회(Europarat)에서 협상이 진행 중이다.
- 이는 인권, 민주주의, 법치주의를 보호하는 조건에서 AI 시스템을 사용할 수 있도록 보장하기 위한 것이다. 연방정부는 현재 유럽평의회 협상이 2024년 봄까지는 마무리될 수 있다고 보고 있다.
- AI와 관련하여 진행 중인 유럽평의회 협정에 따라 부담해야 하는 새로운 의무에 대해 연방정부는 2024년 말까지 대책을 마련할 예정이다.

○ 질문 4에 관하여:

- 연방정부는 정보에 대한 평가(“진짜/가짜” 또는 “참/거짓”)가 무엇보다도 미디어 역량의 문제라고 생각한다.
- 교육 주권은 주(州)에 있다. 하지만 연방정부는 법적 가능성의 범위 내에서 디지털 기술(“디지털 활용 능력”) 증진에 기여할 것이다.



○ 다섯 번째 질문에 관해:

- 커뮤니케이션 플랫폼과 달리 언론매체는 언론 편집 기준에 따라 콘텐츠를 선택하고 출판 전에 정보의 진실성을 확인하는 경우가 많다.
- 따라서 스위스의 강력한 미디어 센터는 국민에게 신뢰할 수 있는 정보를 제공하는 데 중요하다.
- 연방정부는 보고서의 일부를 인용한 “미래 지향적인 미디어 후원 전략을 지금 시작하라” 라는 가설에 따른 정책 아이디어를 제시할 것이다.

#### 다. 스위스의 AI 역량 예측<sup>164)</sup>

□ 연방의회 녹색당 의원의 질의 2022.03

- 스위스 연방정부(Bundesrat)는 AI 발전을 모니터링하고 예측하기 위해 AI 모니터링 및 조기 탐지 프로그램 역량을 제공하는 과제를 맡고 있다.
- AI(인공지능/자동화된 의사결정 시스템, 특히 기계적 학습 능력을 갖춘)의 성능과 사회에서의 잠재적 응용이 급속도로 증가하고 있다.
  - 스위스는 AI 발전을 모니터링하는 능력을 강화하지 않으면 뒤처질 위험이 있다.
- 연방정부는 AI의 발전을 예측하고 모니터링할 수 있는 자체 역량을 개발해야 한다.
  - 기술평가재단 ‘TA Swiss’의 2020년 연구에서 정기적이고 통합적인 모니터링의 필요성이 강조되었다.

164) Die Bundesversammlung. 2022. “Vorausschauende KI-Kompetenz der Schweiz.” (May 17) <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20223298> (검색일: 2023. 12. 01).

- 모니터링 프로젝트는 AI의 기본 조건을 개발하고 AI의 사회적 영향을 관리하는 데 기여한다.
  - 또한 중요한 영역(예: 알고리즘 오작동으로 인한 정전)에 널리 배포된 AI 애플리케이션의 문제를 줄일 것이다.
  
- 이러한 이니셔티브를 통해 스위스는 연방정부의 과거 긍정적인 이니셔티브(예를 들어 UVEK가 검토 중인 연방 행정당국을 위한 AI 지침)를 기반으로 하면서도 최상의 구조 조건을 예측하고 설정하는 데 사용할 수 있는 새로운 자원을 제공할 것이다.
  
- 모니터링 중 일부는 외부 기관에 위탁할 수 있지만, 연방정부는 강력한 리더십 역할을 수행해야 한다. 그래야 아래의 사항이 보장될 수 있다.
  1. 정치적 의사결정에 맞춤화되어 사용할 수 있다.
  2. 모니터링 외부에서도 긍정적인 영향을 미치는 연방정부 내 역량을 구축한다(예를 들어 연구/산업과의 긴밀한 교류, 연방정부의 자체 AI 사용에 대한 더 많은 전문 지식 등).
  3. 외부 행위자와의 종속 관계가 발생하지 않는다.
  
- 다양한 연방행위자 또는 연방정부와 가까운 사람들은 이미 AI(예: CNAI)를 다루거나 기술의 결과(예: TA Swiss)를 평가하고 있다.
  - 그러나 위에 설명된 바와 같이 그 필요성에도 불구하고 AI 모니터링이 아직 존재하지 않는다.
  
- 스위스는 새로운 기술에 대해 성급하고 광범위한 규제를 하지 않는다.
  - 이로 인해 AI의 기회와 위험을 식별하고 혁신적인 비즈니스 입지를 강화하기 위해서는 강력한 조기 탐지 프로그램이 더욱 중요하다.
  - 이것이 Predictive AI Competence Switzerland 프로그램에 연간 수십만 프랑을 들이는 이유이다.

□ 연방정부의 답변 2022.05

- 연방총리실(BK, 스위스의 연방총리실은 Bundesrat의 스텝 기관임)이 담당하는 연방의 디지털화 전략은 신기술에 대한 모니터링 도입을 전제한다.
  - CNAI(Competence Network for Artificial Intelligence)를 마련함으로써 하위 영역인 인공지능(AI)에 대한 모니터링이 이루어진다.
  
- 또한 CNAI가 소속된 연방통계청(BFS) 산하 데이터과학역량센터(DSCC)는 무엇보다도 연방정부 내 AI 분야 역량을 개발하는 임무를 맡고 있다.
  - DSCC(Data Science Competence Center)는 연방기술연구소(Federal Institutes of Technology, 특히 EPFL Extension School, [www.extensionschool.ch](http://www.extensionschool.ch))와 긴밀히 협력하여 2023년까지 체계화할 데이터 과학 및 AI 교육 프로그램을 설계했다.
  
- 연방 행정부의 AI 활용과 관련하여 연방정부는 2020년에 주요 지침을 채택했다.
  - 연방정부는 연방 환경, 교통, 에너지 및 통신부(DETEC)에 관련 부서와 협력하여 이러한 지침의 적용과 개발을 정기적으로 검토하도록 의뢰했다.
  - DETEC는 올해 CNAI와 협력하여 첫 번째 평가를 수행할 예정이다.
  
- 중개자 및 통신 플랫폼의 AI 사용과 관련하여 연방의회는 DETEC에 2022년 말까지 통신 플랫폼을 규제해야 하는지 여부와 방법을 명확히 하도록 의뢰했다.
  
- 연방 행정부의 AI 활용과 관련하여 연방정부는 2020년에 주요 지침을 채택했다. 연방정부는 AI 사용과 관련된 과제를 인식하기 위해 업계, 시민 사회, 학계 및 기술계와의 다자간 교류와 상호작용이 근본적으로 중요하다라는 견해를 가지고 있다.

- 이것이 바로 연방정부가 2019년에 연방통신국(OFCOM)이 이끄는 다중 이해관계자 플랫폼을 초대하여 AI 주제에 대한 교류를 진행한 이유다.

- 이러한 모든 조치는 연방정부가 AI 사용을 모니터링하도록 보장하므로 추가 명령을 내릴 필요는 없다.

## 라. 스위스의 생성형 AI 규제

- 스위스의 법률 체계는 허위 정보와 관련된 엄격한 규정을 규정하고 있지는 않지만, 일반적인 원칙과 강력한 법리 체계가 선거 판사의 법률 적용에 지침이 되고 있다.

- 예를 들어, 연방 대법원 판사들은 허위 정보에 대한 좁은 해석을 채택하여 허위 정보 사건에 대한 법원의 개입 의무는 사적 행위자의 영향력이 유권자의 의견 형성 과정을 심각하게 방해하거나 막는 경우에만 발생할 수 있다고 판단했다. 이러한 행위는 투표 무효로 이어질 수 있다.

- 스위스는 원칙에 기반한 기술 중립적인 규제 방식을 채택하고 있으므로 AI 시스템 사용으로 인해 발생하는 위험은 기존 법률에 대한 목표 추가 및 조정을 통해 대응할 수 있다. 또한 유럽 표준이 스위스에서도 모범 사례로 자리 잡을 것으로 기대할 수 있으므로 AI 법안 마련은 빠른 시일 내에 기대하기 어렵다.<sup>165)</sup>

---

165) KPMG. 2023. “EU KI-Verordnung: «Brüssel-Effekt» in der Schweiz?.” (May 24) <https://kpmg.com/ch/de/blogs/home/posts/2023/05/eu-ki-verordnung-bruessel-effekt-in-der-schweiz.html> (검색일: 2023. 08. 24).

○ 2020년 11월, 연방의회는 연방을 위한 인공지능 가이드라인을 채택했다.<sup>166)</sup>

- 총 7개의 가이드라인은 연방 행정부와 연방의 행정업무를 위탁받은 기관에 일반적인 참조 기준을 제공하고 일관된 인공지능 정책을 보장하는 것을 목표로 한다.
- 이 가이드라인의 적용과 관련성은 2022년에 처음으로 검토되었다. 연방통신부(OFCOM)는 연방정부 내에서 설문조사를 실시하여 관련 기관에 가이드라인 적용 현황을 보고하고 개정이 필요한 부분을 파악하도록 요청했다.
- 이 가이드라인은 연방정부에서 AI 사용에 대한 일반적인 기준으로서 유용하며 현재로서는 개정하거나 업데이트할 필요가 없다고 본다.
- 그러나 연방정부 내에서 가이드라인과 그 구체적인 적용에 대해 보다 심도 있게 논의할 필요가 있음을 인정하고 있다.
- OFCOM은 이를 위해 ‘플랫폼 3자 협의체’의 행정위원회를 활용하여 연맹의 인공지능 역량 네트워크(CNAI)도 가이드라인의 활용도를 높일 수 있다.
  - ▷ 법률 문제 전문가 그룹으로서 ‘플랫폼 삼자 협의체’와 CNAI를 지원하는 전문 법무 그룹(‘Knotenpunkt Recht’)도 인공지능 프로젝트의 법적 요건과 관련하여 가이드라인의 적용과 이행을 도울 것이다.
  - ▷ 기술 발전의 빠른 속도를 고려할 때, 가이드라인은 최신의 상태로 업데이트될 필요가 있다. 연방 위원회의 위임에 따라 OFCOM은 관련 연방 기관과 협력하여 2024년에 가이드라인에 대한 평가를 재 실시할 예정이다.
  - ▷ 그때까지 OFCOM은 AI의 개발과 적용을 모니터링하고, 연방 법무부 및 FDFA의 국제법국과 함께 국내 및 국제 수준의 AI 규제를 지속해서 감시하게 될 것이다.
  - ▷ 또한 OFCOM은 다른 관련 연방 기관과 협력하여 AI에 대한 적

166) Federal Office of Communications OFCOM. 2022. “Monitoring of the “Artificial Intelligence” Guidelines for the Confederation: Evaluation of the application and relevance of the guidelines.” (December 9) <https://cna1.swiss/wp-content/uploads/2022/12/Monitoring-der-Leitlinien-KI-EN.pdf> (검색일: 2022. 12. 15).

절한 국제 프레임워크를 개발의 책임을 진다.

- 연방의회는 인공지능(AI)의 잠재력을 활용하면서 사회에 미치는 위험을 최소화하고자 한다. 이를 위해 2023년 11월 22일 회의에서 인공지능에 대한 가능한 규제 접근 방식에 대한 개요를 작성하도록 DE TEC에 의뢰했다. 이 개요는 2024년 말까지 마련될 예정이다.<sup>167)</sup>
  - 연방의회는 AI의 발전, 기회, 도전 과제에 대해 논의했다. 2024년 말까지 AI 규제에 대한 가능한 접근 방식을 파악하고 관련 법률 분야를 담당하는 모든 연방 기관을 참여시키도록 DE TEC에 지시했다.
  - 이 분석은 기존 스위스 법률을 기반으로 하며, 유럽연합의 AI 규정 (“AI 법”) 및 유럽평의회 AI 협약과 호환되는 스위스의 가능한 규제 접근법을 모색할 것이다. 두 국제 규정 모두 스위스와 관련이 있다. 특히 기본권 준수에 중점을 두고 규제의 필요성을 검토할 것이다. 다양한 규제 접근법의 기술 표준과 재정 및 제도적 영향도 고려될 것이다.
  - 이 분석에는 신중한 법률, 경제 및 유럽 정책 설명이 포함되며 모든 부서에 걸쳐 협력이 필요하다. 이 작업은 EU 디지털 정책에 관한 부서 간 조정 그룹의 프레임워크 내에서 수행된다. 연방통신부와 유럽자유연맹(FDFA)의 유럽 부서가 전담 부서로, 플랫폼 삼자 협의체와 그 관리 위원회, 연방 AI 역량 네트워크와 그 법률 허브, 연방 행정부 내 AI 실무 그룹 등이 참여하게 된다.
  - 연방의회는 이 분석을 통해 2025년에 AI 규제 제안에 대한 구체적인 명령을 내리고 책임을 규제하기 위한 근거를 마련할 계획이다.

---

167) Der Bundesrat. 2023. “Bundesrat prüft Regulierungsansätze für Künstliche Intelligenz,” (November 22) <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-98791.html> (검색일: 2023. 12. 01).

## 4.8 핀란드

### 가. 허위 정보로부터 가장 자유로운 국가로 선정된 핀란드의 대응

#### ○ 2023 유럽 미디어 리터러시 인덱스<sup>168)</sup>

- 2023년 유럽 미디어 리터러시 지수 결과에 따르면, 총 41개국 중 핀란드가 74점으로 1위를 차지했다. 그 뒤를 이어 덴마크가 73점으로 2위, 노르웨이가 72점으로 3위, 에스토니아와 스웨덴이 각각 71점으로 비슷한 점수로 4위와 5위, 아일랜드가 70점으로 6위를 차지했다.
- 확장된 미디어 리터러시 지수 결과에 따르면 추가국 중 캐나다는 68점으로 47개국 중 7위를 차지했고, 호주는 63점으로 10위를 차지했다. 17위인 한국과 18위인 미국의 점수는 60점으로 같으며, 일본은 77점으로 22위를 차지했다. 캐나다는 주로 북유럽 국가들로 구성된 최고 성과자 1그룹에 합류했다. 호주, 한국, 미국, 일본은 주로 서유럽과 중부 유럽 국가들로 구성된 2그룹에 속했다. 이스라엘은 42점으로 32위, 남유럽과 중유럽 국가들이 속한 3번째 클러스터에 속했다.

---

168) Lessenski, Marin. 2023. "The Media Literacy Index 2023: Measuring Vulnerability of Societies to Disinformation." <https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf> (검색일: 2023. 12. 01):.

<그림 4-4> 2023년 미디어 리터러시 순위



출처: Lessenski, Marin. 2023: 19.

<그림 4-5> 핀란드의 우수 사회 지표

Press Freedom	Happiness	Transparency	Social justice	Gender equality
1. Norway	1. Finland	1. Denmark	1. Denmark	1. Iceland
2. Finland	2. Denmark	2. New Zealand	2. Sweden	2. Norway
3. Sweden	3. Norway	3. Finland	3. Finland	3. Sweden
4. Netherlands	4. Iceland	4. Singapore	4. Czech Republic	4. Finland
5. Denmark	5. Netherlands	5. Sweden	5. Slovenia	5. Nicaragua

출처: Mackintosh, Eliza. 2019. "Finland is winning the war on fake news. What it's learned may be crucial to Western democracy." CNN (May 17) <https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/> (검색일: 2023. 12. 01).

○ 핀란드는 행복, 언론 자유, 성평등, 사회 정의, 투명성, 교육 등 거의 모든 지표에서 꾸준히 상위권 또는 그에 근접한 순위를 차지하고 있어 외부로부터 오는 충격에 강하며 사회 균열을 찾아내어 악용하기 어렵다.

○ 핀란드는 또한 550만 명의 국민이 연간 6,800만 권에 가까운 책을 대출



하고, “헬싱키의 거실” 이라고 불리는 최첨단 도서관에 1억 1,000만 달러를 투자하는 등 독서를 장려하는 오랜 전통을 가지고 있다. 핀란드는 국제학업성취도평가(PISA)에서 EU 국가 중 가장 높은 점수를 받았다.

- 전 세계 다른 지역에서 언론에 대한 신뢰도가 하락하는 가운데 핀란드는 지역 언론과 공영방송이 강세를 유지하고 있다. 로이터 연구소의 디지털 뉴스 보고서 2018에 따르면 핀란드는 미디어 신뢰도 부문에서 1위를 차지했으며, 이는 국민들이 뉴스를 얻기 위해 다른 소스로 눈을 돌릴 가능성이 적다는 것을 의미한다.
- 그러나 일부에서는 단순히 미디어 리터러시와 비판적 사고를 가르치는 것만으로는 충분하지 않으며, 허위 정보의 확산을 막기 위해 소셜미디어 회사 측에서 더 큰 노력을 기울여야 한다는 주장이 존재한다.
- 2016년 미국 대선에 러시아가 개입한 정황이 드러나기 훨씬 전부터 러시아의 선거 개입을 조사했다는 이유로 수많은 공격을 받아온 핀란드 공영방송 YLE의 저널리스트 제시카 아로는 “페이스북, 트위터, 구글/유튜브는 러시아 트롤의 조력자이기 때문에 반드시 규제되어야 한다” 고 주장했다.
- 2014년 제시카 아로는 러시아와 연계된 허위 정보 캠페인이 핀란드 국민에게 어떤 영향을 미쳤는지 조사했다.
  - 많은 핀란드인이 러시아의 이러한 활동을 목격했지만, 이는 핀란드인들이 너무나도 잘 알고 있던 옛 소련 선전이 새로운 디지털 기술로 탈바꿈한 것일 뿐이라 핀란드인들은 트롤을 피할 수 있었다고 주장했다.
  - 이 기사로 인해 아로는 CIA 요원, 나토의 비밀 조력자, 마약상, 광적인 러시아 혐오자라는 비난을 받으며 끊임없는 명예훼손 캠페인의 표적이 되었다.
  - 지난해 헬싱키 지방법원이 명예훼손 혐의로 친푸틴 활동가 두 명, 즉 반이민 친러시아 웹사이트 MV-레티를 운영한 러시아계 핀란드인 일야 야니츠킨과 자칭 ‘인권 운동가’ 이자 러시아 국영 뉴스 매체 RT에 자주 출연하는 요한 백만에게 중형을 선고했다.<sup>169)</sup>

169) BBC. 2018. “Jessikka Aro: Finn jailed over pro-Russia hate campaign against journalist.” (October 18) <https://www.bbc.com/news/world-europe-45902496> (검색일: 2023. 11. 20).

- 핀란드는 유럽연합 국가 중 허위 정보 캠페인의 책임자에게 유죄 판결을 내린 첫 번째 국가로 극단적인 증오 발언과 언론의 자유라는 허울 사이에 선을 그은 것이다.

#### 나. AI 및 빅데이터에 적용되는 법률 및 규정<sup>170)</sup>

- 민간 부문에서의 AI 사용에 대한 명시적인 판례는 존재하지 않지만, 헌법상의 권리가 민간 AI 사용의 배경이 되는 것은 분명하다. 그러나 명시적인 프레임워크가 존재하지 않기 때문에 AI와 ADM은 일반적인 부문별 규칙에 따라 다루어진다.
- 헌법 외에도 차별금지 및 평등 대우법과 데이터 보호에 관한 법률이 가장 관련성이 높은 입법 규칙이다. 판례는 드물지만, 핀란드 국가차별금지 및 평등재판소 본회의 216/2017 판결은 사적 행위자가 성별, 연령, 언어, 거주지 등의 차별적 속성을 알고리즘에 통합하는 경우 자동화된 통계 모델을 사용할 수 없음을 명확히 했다.
- 핀란드 AI 센터는 “세계의 실제 사람들을 위한 실제 AI” 를 중심으로 활동하며 윤리적 측면을 운영에 포함했다. 기본권과의 관계에서 윤리적이고 책임감 있는 AI와 관련된 다양한 미래 발전 방향이 존재한다. 초기의 AI 핀란드 프로그램과 기업들의 적극적인 자율규제 움직임은 좋은 출발점이었기 때문에 핀란드가 이러한 아이디어를 실천에 옮기는 데 적극적으로 나서야 한다는 분위기가 지배적이었다.
- 지난 선거 이후 오로라 AI 프로젝트(공표)가 진행되지 않으면서 정부 측에서는 상대적으로 침묵을 지키고 있다. 윤리적 AI에 대한 정부의 약속과 선언은 전략에 머물러 있지만, 실질적인 의미는 정착되어 가고 있다.
- AI에 대해서는 정부의 한 가지 방향성보다는 여러 부처, 지방 당국과 단체, 연구기관이 더 광범위하게 책임을 분담할 수 있다. 핀란드 법무부는

170) Suomeksi, Tiivistelmä. 2023. “Comparison of AI policies and programmes in Finland, the Netherlands and Sweden - case Finland.” White paper (Oktober 2) <https://www.vttresearch.com/sites/default/files/2023-02/Comparison-of-AI-policies-case-Finland-10022023.pdf> (검색일: 2023. 12. 01).

기본권과 입법권의 관점에서 적극적으로 활동하는 기관이다. 핀란드 데이터 보호 옴부즈맨은 스웨덴의 한 학교에서의 얼굴 인식 사용과 관련된 사건처럼 AI와 관련된 눈에 띄는 결정은 없지만, 텔레마케팅 목적의 자동화된 로봇 발신기 사용에 대해 언급했다.

○ 법적-윤리적 관점은 ETAIROS, 알고리즘 에이전시, 법률 및 법률 기술 연구소 등 다양한 연구기관과 프로젝트에서 연구하고 있다. 투르쿠 대학교와 탐페레 대학교는 로봇과 복지 서비스의 미래, IHAN 윤리(IHAN Etiikka), 스마트 시티를 위한 인간 중심 AI 솔루션과 같은 프로젝트를 통해 윤리 연구에 적극적으로 참여하고 있다.

○ 법무부의 주도로 2018년부터 선거 개입에 대응하기 위한 준비를 시작했다. 대비는 법적 의무의 일부로 우리 주변의 세계가 변화하고 새로운 유형의 위협과 위협이 확인됨에 따라 당국은 이러한 변화가 핀란드와 핀란드 사회에 미치는 영향평가를 시작했다.

- 대비책에는 선거 간섭에 대응하고 당국, 정당 및 기타 기관에 대응하는 방법에 대한 절차 및 협력 개발과 교육 제공이 포함된다.

- 이를 위해 관련 기관은 소셜미디어 기업과 직접 접촉하여 신뢰와 이를 유지하는 방법을 주제로 대비했다.<sup>171)</sup>

○ 핀란드의 경우 2014년부터 AI 관련 정책 프로그램을 가동하였으며 규제 보다는 역량 강화의 측면에서 접근하고 있다.<sup>172)</sup>

- 러시아가 미국 대선에 개입하기 2년 전인 2014년에 핀란드 정부가 시작한 가짜뉴스 방지 이니셔티브의 일환으로, 주민, 학생, 언론인, 정치인에게 분열을 조장하는 허위 정보에 대응하는 방법을 교육하는 프로그램을 시작했다.

- 허위 뉴스 관련 교육이 평생 교육 프로그램으로 지역 대학을 중심으

171) Prime Minister's Office. 2019. "Finland has the best elections in the world. And why is that." <https://vnk.fi/en/-/suomessa-on-maailman-parhaat-vaalit-mieti-miksi-https://vnk.fi/en/-/suomessa-on-maailman-parhaat-vaalit-mieti-miksi> (검색일: 2023. 12. 01).

172) Mackintosh, Eliza. "Finland is winning the war on fake news. What it's learned may be crucial to Western democracy." CNN <https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/> (검색일: 2023. 12. 11).

로 진행되고 있다.

- ▷ 예를 들어 에스포 성인 교육 센터에서는 “러시아 트롤 군대에 당한 적이 있나요?” 라는 제목의 슬라이드에 이미지 및 동영상 조작, 반쪽짜리 진실, 헐박, 허위 프로필 등 소셜미디어에서 독자를 속이기 위해 사용되는 방법에 대한 교육을 진행한다.
- ▷ 트위터 프로필 페이지의 다이어그램이 포함된 또 다른 슬라이드에서는 봇을 식별하는 방법(스톡 사진 찾기, 하루 게시물의 양 평가, 일관성 없는 번역 및 개인정보 부족 확인 등)을 설명했다.
- ▷ 이 강의는 버락 오바마 대통령의 ‘딥페이크’ 로 마무리하여 앞으로의 정보 전쟁의 과제를 강조하고 있다.

- 최근 헬싱키에 소재한 프랑스-핀란드 학교는 최근 핀란드의 팩트체크 기관인 Faktabaari(팩트바)와 협력하여 초등학생부터 고등학생까지 EU 선거에 대해 학습할 수 있는 디지털 리터러시 ‘툴킷’ 을 개발했다. 이 툴킷은 블록의 미디어 리터러시 전문가 그룹에 제출되어 회원국 간에 공유되었다.

▷ 연습 문제에는 YouTube 동영상과 소셜미디어 게시물에서 발견되는 주장 조사, 다양한 ‘클릭베이트’ 기사의 미디어 편향성 비교, 잘못된 정보가 독자의 감정에 미치는 영향 조사, 심지어 학생들이 직접 가짜 뉴스 기사를 작성해 보도록 하는 것까지 포함된다.

▷ 학생들은 소셜미디어에서 좋아요를 누르거나 공유하기 전에 ‘누가 이 글을 썼을까? 어디에 게시된 것인가? 다른 출처에서 동일한 정보를 찾을 수 있는가?’ 라는 질문을 던지도록 교육받는다.

○ 선거에서 시민들이 투표에 참여하도록 독려하고 미디어를 비판적으로 읽는 것이 왜 중요한지 상기시키기 위한 캠페인을 활용한다.

- 커뮤니케이션 캠페인의 목적은 선거에 관한 신뢰할 수 있는 공식 정보를 어디서 찾을 수 있는지, 선거 개입의 의미는 무엇인지, 선거 개입을 인식하고 이에 대비하는 방법은 무엇인지에 대한 대중의 인식을 제고하는 것이다. 이 캠페인은 정보 환경의 변화와 더 일반적으로

정보 운영에 대한 인식을 높여 민주주의 교육을 지원하고자 한다.

- 핀란드가 가짜 뉴스와의 전쟁에서 승리하고 있다는 가장 큰 신호는 다른 국가들이 핀란드의 청사진을 본받으려 한다는 사실이다. 싱가포르를 비롯한 여러 유럽연합 국가의 대표들이 이 문제에 대한 핀란드의 접근 방식을 배우고 있다.
- AI 강국이라고 할 수 있는 핀란드는 선거과정 관리에서는 여전히 전통적인 방식을 고수하고 있다. 110년 이상 선거를 안정적으로 운영해 왔으며, 2023년 봄, 핀란드 의회 선거와 유럽의회 선거에서도 이 강력한 전통이 계속되었다. 선거 조직과 관리는 여전히 펜, 종이, 기표소라는 검증된 방식에 의존하고 있다.

## 5. 결론

### 가. 유형화에 따른 국가별 비교분석

- 생성형 AI 선거 규제 관련 선행연구는 대부분 허위 정보를 다룬다. 본 연구는 생성형 AI 규제를 원칙-규칙 기반 규제로 유형화했다. 그러나 선거 분야에서 생성형 AI 규제 방식에는 국가별로 차이가 존재한다.
- 그러나 이러한 각 유형을 결정하는 주요 요인은 다음과 같다(〈그림 5-1〉 참조). 생성형 AI 분야는 기술 전환이 시기 전 세계 국가들의 주요 경제안보 영역으로 이 분야 산업경쟁력의 측면이 규제를 결정하는 주요한 요인이 된다. 따라서 이러한 경향은 이 분야 기술에 앞선 미국이 원칙 기반 규제를 이 분야 기술을 따라잡아야 하는 유럽연합의 경우 규칙 기반 규제가 이루어졌다.
- 그러나 중국을 제외한 사례 국가 모두 민주주의 가치를 수호한다는 점에서 언론의 자유를 포함한 주요한 민주주의 가치를 훼손할 가능성이 높은 강력한 입법화를 선호할 가능성은 작을 것이다. 특정 이념에 근거한 국가 혹은 권위주의 국가의 경우 강력한 법제화를 선호할 것이다. 최근 신생민주주의 국가에서 생성형 AI 선거 규제의 법제화 경향이 두드러진다. 이는 법에 대한 국가적, 문화적 이해에 따른 것일 수도 있다.
- 본 연구가 사례로 선택한 민주주의 국가들의 경우 다양한 새로운 문제가 발생하자 엄격한 법제화 추구 움직임을 보였으나 대부분 실패하거나 국내 모든 행위자가 법제화에 적극적이지는 않다.
- 미국은 전체적인 AI 기술 및 산업 발전, 개인 권리 보호의 원칙 기반 아래, 관련 행정부 혹은 지방정부의 폭넓은 해석에 의한 규제 적용이 진행되고 있다. 가치적 목적보다는 시장중심적 AI 규제의 원칙 기반 접근이 진행되고 있다.

<그림 5-1> 생성형 AI 규제 유형화

		규제원칙		
		원칙 기반	규칙 기반	유형화 요인
입법 목적	경제 목적	미국	유럽연합	AI 관련 기술과 기업 경쟁력의 차이
	정치, 사회적 가치 목적	중국	유럽연합 회원국 스위스, 영국	가치의 수호 그러나 그것을 지키는 방법의 차이

○ 생성형 AI로 인한 문제에 대한 선거규제 관련해서 규제 방식은 법제형과 자율규제형으로 나누어 볼 수 있다.

- 주요 국가의 사례를 통해 논란이 되는 생성형 AI의 규제에 대한 강력한 입법을 통해 대처할 것인지(규칙 기반) 아니면 신기술에 대한 국가 경쟁력과 시장 중심의 규제(원칙 기반)를 할 것인지에 대한 논란이 있다.

- 선거 캠페인 기간 다양한 문제가 발생하고 있다는 점에서 국가의 정치 세력에 따라 강력한 입법을 요구하는 목소리가 있으나 여전히 대부분의 사례 국가가 자율규제에 의존하고 있다.

▷ 과도기를 포함해 AI법이 시행되기까지는 몇 년이 걸릴 것이기 때문에 EU는 그때까지 기업들의 자발적인 자율규제로 공백을 메우고자 한다. 이를 위해 주요 기술 기업과 주요 AI 개발자의 참여를 유도할 필요가 있다고 본다. 이것은 입법 이후에도 많은 문제를 내포하고 있다는 점에서 우어줄라 폰 데어 로이엔 EU 집행위원장은 AI의 이점과 위험을 평가하기 위해 유엔 기후변화에 관한 정부 간 패널(IPCC)을 모델로 한 국제기구 설립을 제안하고 있다.

- 선거운동과 선거 과정에서 문제 발생이 빈번해짐에 따라 법제화 규제 방식의 필요성이 높아지고 있다.

- 그러나 생성형 AI 규제에서 강력한 법제화 방식을 추구하고 있는 EU

의 경우 산업경쟁력 확보를 동기로 한다는 점에서 선거 과정에 대한 생성형 AI 규제에서 회원국들의 규제 완화에 대한 요구가 빈번함에 따라 특정 기간, 즉 선거기간에만 법제화가 주요 규제 방식으로 자리잡을 가능성이 있다. 그러나 대부분 국가에서 다양한 생성형 AI 선거법 제정이 실패하고 있다는 점에서 단기적으로는 국가 대부분에서 자율규제를 유지할 것으로 보인다.

- 중국에서 가장 먼저 AI 인터넷 당국이 인공지능 규제 초안을 발표했다. 목표는 AI 기술을 촉진하는 원칙 기반 아래, 그러니 중국의 근본적인 사회주의 가치에 부합해야 한다는 점에서 정치, 사회적 가치 목적을 지향한다.

## □ 미국

### ○ [생성형 AI로 인한 선거 관련 문제]

- 생성형 AI를 활용한 정치인에 대한 허위영상, 가짜뉴스 생성 및 유포는 다음의 사항들이 있다.
  - ▷ 정치인의 선거공약, 신념, 이념과 배치되는 행위를 하는 허위 영상, 가짜뉴스 생성 및 유포
  - ▷ 정치인의 명예가 훼손될 수 있는 영상(예, 체포) 제작 유포
  - ▷ 정치적 적대 상대와 화해, 포용 등의 행위를 하는 허위영상, 가짜뉴스 생성 및 유포
- 정당이 생성형 AI를 활용하여 정치광고 생성에 관한 사례가 있다.
  - ▷ 정당이 공식 선거운동을 위한 정치광고에서 발생하지 않은 일(예, 중국의 대만 침공)을 제작 유포
- 안보에 위협이 될 수 있는 영상 생성 및 유포에 관한 사례가 있다.
  - ▷ 국내 안보 위기 발생에 관한 허위 영상, 가짜뉴스 유포
  - ▷ 국외 안보 상황의 급진적 변화 혹은 위기 발생에 관한 허위 영상, 가짜뉴스 유포



○ [생성형 AI 규제] 원칙기반의 규제 방향성

- 생성형 AI에 대해 개인의 권리, 차별금지 등 원칙에 기반한 규제 방향성을 제시하였다.
- ▷ 행정명령 13960을 통해, AI 개발을 위한 안전, 보안, 개인정보 보호, 공정성, 소비자 보호, 근로자 지원, 혁신, 책임 있는 정부 사용의 8가지 목표를 원칙적으로 제시할 뿐이다.
- ▷ 또한 ‘AI 권리장전에 관한 청사진’을 통해 AI 기술에 대한 개인권리 보호에 초점을 둔 세부지침을 두고 있다.

○ [생성형 AI 선거규제] 자율규제 방식

- 생성형 AI를 활용한 정치인에 대한 허위 영상, 가짜뉴스 제작 및 유포, 정당이 생성형 AI를 활용하여 정치광고 생성 행위에 관해 투명성 강화(AI 생성물임을 표기) 이외 강제성을 띤 법제화를 진행하지 않는다.
- 미국 정부는 투명성 보장, 즉 생성형 AI에 의해 제작된 창작물임을 표시하게 함으로써 표현의 자유를 보장한다.
- 플랫폼에 대한 제어 또한 원칙 제시 뿐, 유포에 대한 기업의 책임은 면책대상 조항으로 다루고 있다.
- ▷ CDA 230조 내 조문에 의해 기업은 허위 영상, 가짜뉴스 유통의 통로로 플랫폼이 사용된다 하더라도 면책대상이 된다.
- ▷ 다만 플랫폼 기업들이 생성형 AI 창작물에 의한 선거 위해 사례 발생을 방지하기 위한 자발적인 노력을 정부가 독려하고 있다.

○ [규제 이외 관련 문제 해결 대안]

- 생성형 AI 창작물에 의한 문제점을 방지하기 위한 기업의 자발적 참여를 독려하고 있다.
- ▷ 구글, 마이크로소프트, 오픈 AI, 엔트로픽 등 미국 AI 관련 기업들이 자발적으로 구성한 ‘프런티어 모델 포럼’은 공동으로 생성형 AI의 안전하고 책임있는 개발 및 안전장치를 마련하기 위

해 노력하고 있다.

- ▷ 이들은 안전장치로써 워터마크 기술 등을 공동개발하기 위한 논의를 이어 나가고 있다.

#### ○ [한국적 함의]

- 생성형 AI에 관한 투명성 관련 규제 도입이 필요하다.
  - ▷ 표준 워터마크에 관한 규제가 필요하다.
- 허위사실 유포와 관련한 기존의 법제가 존재하므로, 정치활동과 관련하여 생성형 AI에 대한 규제는 제한될 필요가 있다.
- 국내 플랫폼 기업의 자발적 공동 노력의 대응이 필요하다.

#### □ 영국

##### ○ [생성형 AI로 인한 선거 관련 문제]

- 케임브리지 에털리티카의 Facebook 계정을 통해 수집한 데이터를 영국의 브렉시트 국민투표에 영향을 주었다는 사실은 영국의회의 온라인 미디어 규제의 필요성을 깨닫는 계기가 되었다.
- 뒤이어 2017년 보궐선거와 2019년 영국 총선에서도 봇이 생성한 다양한 허위 정보가 선거과정을 어지럽혔다.

##### ○ [생성형 AI 규제] 규칙 기반의 규제 방향성

- 유럽연합의 광범위한 입법 움직임에 조용하여 데이터 정보법과 디지털 서비스세를 활용해 플랫폼 기업에게 책임성과 투명성을 요구하고 있다.
- 플랫폼 기업에 대해서는 유포에 대한 기업의 책임은 면책 대상으로 다루고 있다.
  - ▷ 일반적 면책 사유와 웹사이트 운영자에 대한 면책 규정이 적용된다.

○ [생성형 AI 선거규제] 자율규제 방식

- 영국도 미국과 마찬가지로 생성형 AI를 활용한 정치인에 대한 허위 영상, 가짜뉴스 제작 및 유포, 정당이 생성형 AI를 활용하여 정치광고 생성 행위에 관해 투명성 강화(AI 생성물임을 표기) 이외 강제성을 띤 법제화를 진행하고 있지 않다.
- 영국은 문제가 있는 AI 창작물의 경우 Facebook이 기존 언론사와 협력하여 뉴스피드에 표시되도록 하고 있다. 허위로 신고된 정보는 뉴스피드에서 더 낮은 위치에 표시되며 사용자가 진위를 판별할 수 있다.
- 이러한 자율규제는 민간 자율규제 기관인 IWF가 담당하고 있었으나 정부의 위탁기관인 정보통신청(Ofcom)이 주무기관으로 유해 콘텐츠를 직접삭제할 권한은 없으나 사용자의 접속 차단 권한은 가지고 있는 점차 권한의 강화가 예상된다.

○ [규제 이외 관련 문제 해결 대안]

- 팩트체크는 비영리 매스컴 연합기관들이 수행하고 있으며 대표적으로 Full Fact는 Facebook과의 파트너십을 통해 활동한다. 즉 영국의 Facebook 사용자가 오해가 소지가 있는 내용에 플래그를 지정하면 Full Fact가 정보를 검토하고 평가한다.

○ [한국적 함의]

- 한국에서 플랫폼 사업자에 대한 규제 어렵다는 점에서 영국의 Full Fact와 Facebook의 파트너십을 통한 문제해결 방식에 대한 고려가 필요하다.

□ 유럽연합

○ [생성형 AI로 인한 선거 관련 문제]

- EU는 유일하게 유럽의회에 관한 선거를 진행하고 있으나, 유럽의회 의 정치적 역할이 유럽 내에서 그다지 크지 않은 탓에 생성형 AI에

의한 선거 관련 문제는 현재까지 일어나지 않고 있다.

- ▷ 유럽의회 선거가 5년마다 치러지고 있으며, 최근 선거가 2019년이었다. 이에 생성형 AI의 영향이 있었던 시기 선거가 없었다.

#### ○ [생성형 AI 규제] 규칙질서 기반의 규제 방향성

- 유럽연합은 가장 광범위한 AI 규제를 입법화하고 있는 지역이다.
  - ▷ AI의 위험을 허용할 수 없는 위험(Unacceptable risk), 고위험(High risk), 제한된 위험(Limited risk), 위험이 최소화되거나 없음(Minimal or no risk)으로 분류하며, 구체적인 유형 내용을 적시하여 법적 제어를 진행하고 있다.
  - ▷ 생성형 AI와 관련하여, 첫째, AI가 제작한 콘텐츠를 공개, 둘째, 불법 콘텐츠 생성을 방지하도록 모델 설계, 셋째, 훈련에 사용된 저작권 데이터의 요약 게시의 규칙질서 규제를 진행하고 있다.
- 유럽연합은 생성형 AI의 허위 영상, 가짜뉴스가 생성되어 유포되는 과정에서 거대 플랫폼이 이에 대한 책임있는 관리를 할 것을 법적으로 제어한다.
  - ▷ EU는 디지털서비스법(DSA)에 의해 거대 플랫폼 기업의 책임을 엄격하게 제어하고 있다. 이에 초대형 플랫폼은 위험관리 의무, 범죄 신고의 의무 등이 있으며, 플랫폼 내 유포에 대한 책임을 기업이 지게 되는 것이다.

#### ○ [생성형 AI 선거규제] 자율규제 방식

- 유럽연합 내 2019년 이후 현재까지 선거가 치러진 바 없다. 향후 2024년 유럽의회 선거가 예정되어 있어, 해당 시기 선거와 관련한 유럽연합의 입장이 있을 것으로 판단된다.

○ [한국적 함의]

- 생성형 AI에 관한 투명성 관련 규제 도입이 필요함.
  - ▷ 표준 워터마크 도입, 생성형 AI 창작물 내 인용 자료에 대한 출처 표기 관련 규제가 필요하다.
- AI 활용에서 반드시 제어되어야 할 고위험군 영역에 대한 식별과 법적 제어가 필요하다.
- 국내에서 플랫폼 기업을 법적으로 제어하는 것은 현실적으로 어려운 상황이다. 이에 기업들이 자발적 공동 노력을 할 수 있도록 정부가 위원회 방식의 대응을 할 필요가 있다.

□ 독일

- [생성형 AI로 인한 선거 관련 문제] 독일의 선거 과정에서 AI를 이용하여 문제가 발생한 실제 사례는 많지 않은 편이다.
  - 생성형 AI의 활용에 따른 문제보다 가짜뉴스를 둘러싼 논의가 주요 쟁점이다.
  - 2023년 11월 생성형 AI를 활용하여 제작한 연방총리의 독일대안당 해체 주장 동영상은 독일에서도 딥페이크 기술을 활용한 가짜 비디오가 문제가 될 수 있다는 점을 보여주었다.
    - ▷ 연방정부는 이 사건에 대해 매우 비판적 견해를 밝혔으나, 처벌이나 대책에는 아직 명확한 입장을 내놓지 않고 있다.
- [생성형 AI 규제] 생성형 AI 분야에서는 효과적인 규제와 혁신적인 비즈니스 모델 촉진 사이에서 적절한 균형을 추구한다.
- [생성형 AI 선거규제] 독일에서는 생성형 AI를 활용한 제작물에 관해 법률에 따른 규제보다는 아직 자율규제에 중점을 두고 있는 것으로 보인다.

- [규제 이외 관련 문제 해결 대안] 독일 사회의 다양한 민주적 소통과정은 생성형 AI의 가짜뉴스 제작을 억제하고 있다.
  - 복수의 미디어 체계
  - 폭넓은 시민교육 강화
  - 다양한 대안적 소통 채널
  
- [한국적 함의] 정치과정에 생성형 AI를 활용하는 문제에 대해 자율적 규제를 우선시하는 독일의 사례는 우리에게도 비슷한 시사점을 준다.
  - 법률적 규제가 가져올 기본권(표현의 자유 등) 침해에 대한 논란이 우려된다.
  - 가짜뉴스에 대한 사실 여부 판단을 법률에 따르는 방식이 물리적으로 가능할지를 고려해야 한다.
  - 민주주의 관련 시민교육의 강화와 승자독식의 선거제도를 개선하는 것도 생성형 AI 시대의 시급한 해결과제이다.

## □ 스위스

- [생성형 AI로 인한 선거 관련 문제]
  - 스위스의 경우 생성형 AI 관련 선거 개입문제는 크게 대두되지 않았다. 최근 연방선거 캠페인에서 문제점을 느낀 스위스 5개 정당이 선거운동에서 AI 사용 제한을 제안하고 있기는 하지만 합의는 요원해 보인다.
  
- [생성형 AI 규제] 규칙 기반의 규제 방향성
  - 유럽연합의 AI 규정("AI 법") 및 유럽평의회 AI 협약과 호환되는 스위스의 가능한 규제 접근법을 모색하고 있다.
  - 스위스의 법률 체계는 허위 정보와 관련 법적 규정은 하고 있지 않지만, 일반적인 원칙과 강력한 법리 체계가 선거 판사의 법률 적용에

지침이 되고 있다.

- 스위스는 기본적으로 새로운 기술에 성급하고 광범위한 규제를 하지 않는다는 원칙을 가지고 있다.
- 스위스는 원칙에 근거한 기술 중립적인 규제 방식을 채택하고 있으므로 AI 시스템 사용으로 인해 발생하는 위험은 기존 법률에 대한 목표 추가 및 조정을 통해 대응할 수 있다고 본다.

○ [생성형 AI 선거규제] 자율규제 방식

- 2023년 4월 5일 연방정부는 스위스에서 많이 사용되는 통신 플랫폼에 대한 규제를 모색하기로 했다. 그러나 이외에 스위스는 생성형 AI 관련 규제에 가장 소극적인 모습을 보이고 있으며 기본적으로 자율 규제를 따른다.
- OFCOM은 AI의 개발과 적용을 모니터링하고, 연방 법무부 및 FDFA의 국제법국과 함께 국내 및 국제 수준의 AI 규제를 지속해서 감시하게 될 것이며 OFCOM은 다른 관련 연방 기관과 협력하여 AI에 대한 적절한 국제 프레임워크의 개발 책임을 진다.

○ [규제 이외 관련 문제 해결 대안]

- 이 분야의 대응과 관련해 스위스의 경우 다른 국가들과 비교하면 가장 초기 단계로 이제 가이드라인을 설정하고 이에 대한 모니터링과 긍정적 혹은 부정적 영향에 대한 식별과 평가가 진행 중이다.

○ [한국적 함의]

- 유럽연합의 회원국이 아니나 유럽연합 혹은 유럽평의회와의 협력을 통해 생성형 AI 분야에 대한 규제 보조를 맞추려는 시도가 눈에 띈다.
- 기술의 영역이라는 점에서 다양한 관련 분야의 다양한 국제적 교류를 시도해야 한다.

## □ 핀란드

### ○ [생성형 AI로 인한 선거 관련 문제]

- 핀란드의 공영방송 YLE는 2016년 미국 대선에 러시아의 개입 정황을 파악하기 이전부터 러시아의 선거 개입에 대한 조사를 수행했으며 조사를 담당한 기자는 Facebook, Twitter, Google/Youtube는 러시아 트롤의 조력자이기 때문에 반드시 규제해야 한다는 주장을 펼쳤다.
- 러시아와 국경을 마주하고 있는 핀란드의 경우 옛 소련의 정치 선동이 새로운 기술로 탈바꿈했을 뿐 본질적인 변화는 없다고 보고 있다. 따라서 이러한 문제에 핀란드인들은 잘 대처할 수 있다는 것이다.

### ○ [생성형 AI 규제] 규칙 기반의 규제 방향성

- 헬싱키 지방법원은 허위 정보로 인한 명예훼손 혐의로 친푸틴 활동가 두 명에 이례적인 중형이 선고되었다.
- 핀란드는 허위 정보를 확산시킨 책임자에게 유제 판결을 내린 첫 번째 국가로 극단적인 발언과 언론의 자유에 명확한 경계를 설정했다.
- 기본권과의 관계에서 윤리적이고 책임감 있는 AI와 관련된 다양한 미래 발전 방향이 존재한다.

### ○ [생성형 AI 선거규제] 자율규제 방식

- 핀란드의 초기 AI 프로그램과 기업들의 적극적인 자율규제 움직임은 좋은 출발점이었기 때문에 핀란드가 이러한 아이디어를 실천에 옮기는 데 적극적으로 나서야 한다는 분위기가 지배적이었다.
- 핀란드 법무부의 주도로 2018년부터 선거 개입에 대응하기 위한 준비는 영향 평가로부터 시작되었다.
  - ▷ 갈수록 늘어나는 선거 간섭에 대응하고 당국, 정당 및 기타 기관에 대응 방법에 대한 절차 및 협력 개발과 교육 제공 등을 들 수 있다.
  - ▷ 이를 위해 관련 기관은 소셜미디어 기업과 직접 접촉하여 신뢰와 이를 유지하는 방법이 주로 다루어졌다.



○ [규제 이외 관련 문제 해결 대안]

- 2014년부터 핀란드는 AI 관련 정책 프로그램을 가동하였으며 규제보다는 역량 강화의 측면에서 접근한다.
  - ▷ 생성형 AI 역량강화를 통해 정보를 식별할 수 있는 능력을 배양하기 위해 학교와 평생교육원에서 구체적으로 봇의 식별방법 등을 교육한다,
  - ▷ 학교 차원에서 팩트체크 기관인 Faktabaari와 협력을 통해 초중고학생들을 대상으로 EU선거에 대한 교육을 시행하고 있고 교제로는 미디어리터러시 툴킷을 개발하여 사용하고 있다.

○ [한국적 함의]

- 생성형 AI 양산하고 있는 다양한 문제에 규제 접근보다는 미디어리터러시를 강화하는 개인의 역량 향상의 측면에 집중하고 있다는 점에서 플랫폼 규제 등이 어려운 한국적 환경과 AI 등 인터넷 접근성이 높은 한국적 조건에서 이 문제는 개인의 역량 강화의 측면으로 접근하는 것이 더 효과적일 것이다.

**나. 생성형 AI 선거 규제 동향과 한국적 함의**

- 2023 미디어 리터러시 보고서는 자유민주주의 정부를 가진 국가에 사는 사람들이 “민주주의 제도가 없거나 제한적인 국가에 사는 사람들보다 잘못된 정보에 대해 걱정할 가능성이 더 높다” 라고 밝혔으며, 교육 수준이 높은 사람들 사이에서 허위 정보에 대한 우려가 더 크다는 것을 발견했다. 허위 정보의 영향에 가장 취약한 사회가 동시에 허위 정보의 확산과 영향에 대해 가장 신경쓰지 않는다는 것은 우려스러운 일이다.

○ 미국, 영국, 유럽연합, 독일, 핀란드 사례

- 생성형 AI의 기술 및 산업 발전을 고려하며, 정치·사회에 미치는 영향에 대해 신중한 접근을 하고 있다.
- 사례 국가는 표현의 자유라는 기본권의 원칙 아래 생성형 AI의 규제

원칙을 확립하고자 한다.

- 표현의 자유, 법의 실효성, 법의 효율성을 고려할 때 생성형 AI 투명성 강화, 개인 권리 및 알고리즘 규제, 빅테크플랫폼에 대한 권고지침 마련을 통한 자율규제가 단기적으로는 설득력 있다.
  - 한국은 AI 산업 선도 국가로서 지나친 규제는 지양할 필요가 있다. 중국의 사례처럼 전방위적인 차단과 규제는 기술 발전을 저해함과 동시에 법적 실효성도 보장할 수 없다.
  - 생성형 AI에 관한 투명성 관련 규제 도입이 필요함.
    - ▷ 표준 워터마크 도입, 생성형 AI 창작물 내 인용 자료에 대한 출처 표기 관련 규제가 필요하다.
  - AI 활용에서 반드시 제어되어야 할 고위험군 영역에 대한 식별과 법적 제어가 필요하다.
  - 국내에서 플랫폼 기업을 법적으로 제어하는 것은 현실적으로 어려운 상황이다. 이에 기업들이 자발적 공동 노력을 할 수 있도록 정부가 위원회 방식의 대응을 할 필요가 있다.
- 정치과정에 생성형 AI를 활용하는 문제에 대해 법률적 규제보다 자율규제를 우선시하는 사례는 우리에게도 같은 시사점을 준다.
- 허위 정보에 대한 사실 여부 판단을 법률에 따르는 방식이 물리적으로 가능할지를 고려해야 한다.
  - 최근 한국 페이스북에서 특정 용어의 사용에 대해 제재를 가했는데, 이는 이용자의 반발을 가져왔다.
  - 영국의 사례처럼 국가 차원 혹은 민간 차원의 독립이 보장된 기관의 설립을 통해 이러한 문제를 지식과 근거에 따라 판단하도록 할 필요가 있다.
- 가짜뉴스의 생성을 억제하는 독일과 핀란드 사회의 저력은 복수의 미디어 체계, 폭넓은 시민교육, 다양한 대안적 소통 채널 등의 민주적 소통을 통한 개인 역량의 강화에서 찾을 수 있다.

- 한국 사회도 독일처럼 다양한 미디어 체계와 소통 채널을 갖추고 있으나, 민주주의에 대한 시민교육은 매우 부족한 상황이다. 시민교육의 강화를 통해 건전한 시민의식을 갖추는 것이 가짜뉴스에 대한 대책이라고 할 수 있다.
- 핀란드는 생성형 AI에 대한 더 정확하게는 외부의 개입에 가장 강력하게 대처할 수 있는 국가로 선정되었다. 한국의 미디어리터러시는 27위로 슬로베니아나 슬로바키아 보다 낮다(<그림 5-2>와 <그림 5-3> 참조). 교육 수준이 미디어 리터러시 역량을 담보하지 못한다는 점에서 생성형 AI에 대한 미디어 리터러시 역량 향상을 위한 조치가 필수적이다. 규제를 통한 위협의 완화는 필수적이거나 가장 근본적인 해법에 대한 투자가 필요하겠다.
- 경험적 연구가 필요하겠으나 일련의 사례 연구를 통해 확인한바 대체로 승자독식의 선거제도는 극단적 경쟁을 불러오고, 이는 허위 정보를 활용할 유혹을 크게 한다. 대체로 중앙집권적인 국가보다 분권적인 국가에서 이러한 위협이 위협이 될 가능성이 작다고 하겠다. 따라서 적절한 선거제도의 개혁논의는 AI시대의 시급한 과제라고 할 수 있겠다.

〈그림 5-2〉 2023 확장된 미디어 리터러시 순위

Expanded Media Literacy Index 2023 (clusters 1 and 2)				Expanded Media Literacy Index 2023 (clusters 3,4 and 5)			
Ranking (1-47)	Country	Scores (100-0)	Cluster	Ranking (1-47)	Country	Scores (100-0)	Cluster
1	Finland	74	1	28	Slovakia	48	3
2	Denmark	73	1	29	Italy	47	3
3	Norway	72	1	30	Croatia	45	3
4	Estonia	71	1	31	Malta	45	3
5	Sweden	71	1	<b>32</b>	<b>Israel</b>	<b>42</b>	<b>3</b>
6	Ireland	70	1	33	Hungary	41	3
<b>7</b>	<b>Canada</b>	<b>68</b>	<b>1</b>	34	Cyprus	39	3
8	Switzerland	67	1	35	Greece	38	3
9	Netherlands	64	2	36	Ukraine	38	3
<b>10</b>	<b>Australia</b>	<b>63</b>	<b>2</b>	37	Serbia	33	4
11	Iceland	62	2	38	Moldova	32	4
12	Belgium	61	2	39	Montenegro	32	4
13	Germany	61	2	40	Romania	32	4
14	Portugal	60	2	41	Bulgaria	31	4
15	United Kingdom	60	2	42	Turkey	29	4
<b>16</b>	<b>South Korea</b>	<b>60</b>	<b>2</b>	43	Bosnia and Herzegovina	24	5
<b>17</b>	<b>USA</b>	<b>60</b>	<b>2</b>	44	Albania	23	5
18	Austria	59	2	45	North Macedonia	22	5
19	Czech Republic	58	2	46	Kosovo	21	5
20	Spain	58	2	47	Georgia	20	5
21	France	57	2				
<b>22</b>	<b>Japan</b>	<b>57</b>	<b>2</b>				
23	Latvia	55	2				
24	Slovenia	55	2				
25	Lithuania	54	2				
26	Luxembourg	53	2				
27	Poland	53	2				

출처: Lessenski, Marin. 2023: 7.

<그림 5-3> Media Freedom 순위

**Annex: Ranking by indicators**

**Expanded Media Literacy Index 2023:  
Media freedom ranking**

Media Freedom ranking	Media freedom score	Country	Overall ranking	Overall score	Cluster
1	32.6	Norway	3	72	1
2	30.8	Denmark	2	73	1
3	30.4	Sweden	5	71	1
4	30	Finland	1	74	1
5	29.2	Estonia	4	71	1
6	28.4	Netherlands	6	70	1
7	28.2	Portugal	14	60	2
8	27.8	Switzerland	8	67	1
9	27.2	Iceland	11	62	2
10	26.8	Belgium	12	61	2
11	26.6	Netherlands	9	64	2
12	26.4	Luxembourg	26	53	2
13	26.2	Canada	7	68	1
14	26.2	Lithuania	25	54	2
15	25.8	Germany	13	61	2
16	25	Czech Republic	19	58	2
17	23.4	Austria	18	59	2
18	23.2	United Kingdom	15	60	2
19	23.2	Latvia	23	55	2
20	23	France	21	57	2
21	23	Slovakia	28	48	3
22	22.4	Australia	10	63	2
23	22	Spain	20	58	2
24	21.8	USA	17	60	2
25	20.2	Slovenia	24	55	2
26	19.4	Cyprus	34	39	3
27	18.8	South Korea	16	60	2
28	18.2	Italy	29	47	3
29	17.8	Japan	22	57	2
30	17.8	Malta	31	45	3
31	16.4	Croatia	30	45	3
32	16.4	Poland	27	53	2
33	16.4	Romania	40	32	4
34	14.6	Israel	32	42	3
35	14.4	Montenegro	39	32	4
36	13.8	Moldova	38	32	4
37	13.6	Kosovo	46	21	5
38	12.2	Bosnia and Herzegovina	43	24	5
39	12.2	Bulgaria	41	31	4
40	12	Hungary	33	41	3
41	11.2	Serbia	37	33	4
42	10.4	Greece	35	38	3
43	10.2	Georgia	47	20	5
44	9.8	North Macedonia	45	22	5
45	9	Albania	44	23	5
46	8.2	Ukraine	36	38	3
47	0	Turkey	42	29	4

출처: Lessenski, Marin. 2023: 10.

〈표 5-1〉 생성형 AI 선거규제 국가별 비교 분석

	미국	EU	독일	핀란드	영국	스위스	중국
생성형 AI로 인한 선거 관련 문제	다양한 선거교란 사례	-	없음	없음	다양한 선거교란 사례	야당의 문제 제기	없음
생성형 AI 규제	시장중심 AI 규제 원칙기반 접근	기술경쟁력 중심 규칙 기반 접근	EU AI 법/민주주의 가치 기반	EU AI 법/민주주의 가치 기반	EU AI 법 기반/민주주의 가치 기반	EU AI 법 기반/민주주의 가치 기반	AI 기술축진을 위한 원칙 기반/ 사회주의 가치기반
생성형 AI 선거규제	다양한 입법 시도 그러나 실패 자율규제 기반	AI 법 입법 성공 / 발효 전까지 자율규제 기반	기본법 활용 자율규제 기반	기본적인 법리 적용 자율규제 기반	EU의 적극적인 법안(디지털 서비스세 등) 활용 자율규제 기반	입법에 대한 소극적 움직임 자율규제기반	최초로 인공지능 규제 법안 마련
규제 이외 관련 문제 해결 대안	관련 이해관계자와의 교류와 소통	-	미디어리터러시 역량 강화 정치교육	미디어리터러시 역량 강화 AI관련 일반인 대상 교육	팩트체크 (Facebook과 IWF의 협력)	관련 이해관계자와의 교류와 소통	-
한국적 함의	표준 워터 마크 관련 규제 도입	생성형 AI관련 투명성 관련 규제 도입	개인의 역량 강화를 위한 정치교육	개인의 역량 강화를 위한 정치교육	팩트체크를 위한 독립적인 정부 및 민간 기관과 기업 간의 협력	규제에 소극적이며 모니터링과 역량 개발 기관 설립으로 대응	-

출처: 저자 작성

## 요약

### ■ 선거 과정에서 AI 활용 사례 및 잠재적 활용 가능성

○ 선거과정에서 AI 활용 및 잠재적 활용 가능성에 대한 논의를 살펴보면 기술 준비도, 위험 수준 및 AI 사용의 가시성 측면에서 고려되는 5가지 분야(유권자 명부관리, 투표소 위치, 문제부스예측, 유권자 인증, 영상 모니터링) 각각에 대한 예비 평가라고 할 수 있다. 이는 아직까지 기술의 수준 및 법적 적용 가능성에 대한 논란이 존재한다는 점에서 객관적인 평가가 될 수 없으나 현재 상황에 대한 정보 평가를 바탕으로 이루어졌다는 점에서 선거과정에서 AI 활용에 대한 가능성에 대한 탐색의 의미가 있다.

- 투표소 위치 결정과 관련된 AI 기술은 작업에 사용하기 위해서는 어느정도 조정이 불가피하다는 점에서 기술 준비도를 중간으로 평가했다. 위험도가 높지 않다는 점에서 활용 가능성이 있다.
- 투표소에서 발생하는 폭력 문제를 해결하기 위해 문제 투표소를 식별하는 작업에 어느 정도 쉽게 사용할 수 있으므로 기술 준비도를 높음을 평가했으나 이는 안면인식과 관련한 사안으로 이 기술을 치안의 관점으로 이해하면 법적 예외에 해당하여 활용 가능성이 높은 영역이라 할 수 있다.
- 유권자 명부관리 및 투표 부스 결정에 AI를 사용하는 위험 수준은 어느 정도 전문가의 감독을 보장하는 human-in-the-loop 방식으로만 사용하는 것이 사실상 불가피하므로 중간으로 평가한다.
- 영상 모니터링의 사용은 선거의 공정성 원칙과 침해한 긴장 관계에 놓일 수 있다는 점에서 매우 높은 위험도로 평가한다.

## ■ 주요국의 생성형 AI 선거 규제

### □ 미국

- 미국은 AI 산업을 선도하기 위해 AI 발전 전략 중심의 로드맵을 제시하였다. AI 기술이 비약적으로 발전함에 따라, 미국 정부는 이에 관한 규제를 도입하고 있다. 미국 정부가 관련 영역에서 주요하게 다루는 바는 개인권리 보호, 차별금지, AI 기술을 이용한 안보 사항이다.
- 2023년 1월부터 현재까지 발생한 미국 내 선거, 정당에 관한 생성 AI 창작물 유포 사례들은 미국 내뿐 아니라 전 세계적 관심을 끌었다. 하지만 미국의 ‘표현의 자유’에 대한 사회적 보장과 지지가 높은 국가로 이에 관한 법적 규제에 대해 신중한 상황이다. 이미 허위 정보에 대한 제재에 사회적 합의가 있기 때문에 생성 AI 창작물이 허위 정보에 해당된다면 이의 법 규정을 따를 수 있다. 따라서 미국 사회는 선거 및 정당과 관련한 생성 AI 창작물의 투명성에 집중하고 있다. 즉 AI에 의해 제작되었음 혹은 창작물임을 알 수 있도록 하는 표식에 대한 제도화가 필요함에 공감하고 있다. 하지만 내용상 정치와 민주주의에 관한 표현의 자유와 주장을 억제하기 위해, 생성 AI 창작을 제도적으로 제어해야 할 필요성에 관해 논의하지는 않는다.
- 이와 대조적으로 미국 군사, 안보와 관련한 생성 AI 창작물 유포와 관련하여 미국 정부는 좀 더 강경한 입장이다. 군사, 안보와 관련한 생성 AI 창작물은 사회에 극단적인 악의적 영향을 미칠 우려가 있기 때문이다. 군사, 안보와 관련한 생성 AI 창작물은 국제 수준에서 악의적 의도로 제작될 가능성이 있으므로, 투명성을 높이는 것만으로 대응할 수 없는 것이다. 이에 미국 정부는 개인의 권리보호나 차별을 방지하기 위한 위협 관리 이외에 국방, 군사 영역에 관한 법률을 입법하였다.



## □ 영국

- 영국의 국가 AI 전략의 세 가지 목표를 설정하고 있다. 첫째, 과학 및 AI 강국으로서의 리더십을 지속하기 위해 AI 생태계의 장기적 요구에 투자하고 계획하며, 둘째, 영국에서 혁신의 혜택을 포착하고 AI가 모든 부문과 지역에 혜택을 줄 수 있도록 AI 기반 경제로의 전환을 지원하며, 마지막으로 영국이 혁신과 투자를 장려하고 공공과 영국의 기본 가치를 보호하기 위해 AI 기술의 국내 및 국제 거버넌스를 올바르게 확보하는 것이다.
- 온라인 플랫폼에 대한 법적 규제 및 자율규제에 직접 관여하고 있는 정보통신청(Office of Communication, Ofcom)은 정부 위탁 규제 기관으로 방송 통신 및 우편 산업에 대한 규제를 목적으로 하는 기관이다. 정부 위탁 기관이지만 정부로부터의 독립성이 보장된다. 새롭게 만들어지는 「온라인안전법」의 주무부처로서 온라인 플랫폼의 유해 콘텐츠를 직접 삭제하는 등의 권한은 없지만 사용자의 접속을 차단하는 등 강력한 권한이 부여될 예정이다.
- 인터넷 서비스 사업자 단체인 ISPA(Internet Services Providers' Association) 또한 자율규제의 역할을 담당하고 있다. ISPA는 회원사를 대표해 규제 당국과 정기적인 접촉을 하며 규제 당국의 인터넷 규제 가이드라인을 따르고 있다.
- 온라인 플랫폼 사업자의 책임성을 강화하는 방안으로 온라인 사업자에게 전체 소득의 2%에 해당하는 디지털서비스세(digital service tax)를 제도화하여 이를 기금으로 형성하고 이 기금을 허위 정보 단속 및 관리에 사용하고 있다. 디지털서비스세는 2020년 4월 20일 매출부터 적용된다.
- 그러나 플랫폼 사업자가 허위 정보를 인지하였다고 하더라도 그 정보의 유해성을 판단하기가 쉽지 않기 때문에 플랫폼 사업자의 책임을 계산하기가 쉽지 않다는 문제가 있다.
- 영국은 자율규제를 기본적인 규제 방향으로 설정하여 모든 관련 행위자

와의 공조를 통해 AI 파트너십이라는 윤리기구를 설립하여 활동하며 다양한 허위뉴스 캠페인을 벌이는 등 다양한 방식으로 대응하고 있다.

#### □ 유럽연합

- EU 집행위원회는 지난 2021년 4월 21일 'AI 규제 초안' 을 공표, 2023년 12월 8일 관련 합의안이 도출되었으며, 2024년 초 유럽의회의 표결을 거쳐 발효될 것으로 판단된다.
- EU가 발의한 AI 규제는 AI 사용에 관한 세계 최초의 포괄적 AI 규제법이다. EU는 디지털 전략의 일환으로 AI 산업을 규제하여 혁신적 기술 개발 및 사용을 위한 더 나은 조건을 보장하고자 포괄적 입법 과정을 거치고 있다.
- EU의 AI 규제는 해당 기술영역을 4개의 위험군으로 분류함으로써 제어된다. EU 규제 최종안에서 생성 AI는 투명성 요구사항을 준수해야 하는 사항으로 포함되었다. 해당 내용에는 첫째, AI가 제작한 콘텐츠임을 공개할 의무, 둘째, 불법 콘텐츠 생성을 방지하도록 모델 설계할 의무, 셋째, 훈련에 사용된 저작권 데이터의 요약 게시 의무가 부과되었다.
- 수정안에는 '생성 AI에 대한 의무' 항목에 Chat GPT와 같은 모델을 기반으로 하는 생성형 AI 시스템의 투명성 요구 사항과 학습에 사용된 저작권이 있는 데이터의 자세한 요약도 공개적으로 제공되어야 한다는 내용이 포함되어 있으나, 저작권에 관한 전반적인 쟁점은 포함되어 있지 않았으며, 최종 가결될 때까지 해당 사항이 포함될 가능성은 높지 않아 보인다.

#### □ 미국과 EU의 협력과 경쟁

- EU와 미국은 EU-미국 무역기술위원회(TTC: Trade and Technology Council)을 통해 'EU-미국 인공지능을 위한 공동로드맵' 을 발표하는 등 다양한 노력을 기울이고 있으나, 법제도적 합치를 이룰 것으로 예상되지 않는다. 이는 양측이 AI 기술 발전과 사회에의 영향에 관한 시각

격차가 크기 때문이다.

- 이에 반해 중국은 AI 기술을 선도, 미국 빅테크 기업을 위협하는 시장 성장을 주도하며, 사회주의 가치를 국가 내에서 수호하여야 한다는 이중적 입장에 있다. 최근 AI 관련 다양한 규제를 입법하고, 최초의 ‘생성형 AI’ 규제 또한 공포하였으나, 이의 법적 마찰에 대한 우려의 시각이 높은 상황이다.

#### □ 독일

- 독일의 선거 과정에서 AI를 이용하여 문제가 발생한 실제 사례는 많지 않은 편이나 최근 생성형 AI를 활용하여 제작한 연방총리의 독일대안당 해체 주장 동영상은 독일에서도 딥페이크 기술을 활용한 가짜 비디오가 문제가 될 수 있다는 점을 보여주었다.

- 연방정부는 이 사건에 대해 비판적 입장을 보였으나, 처벌이나 대책에는 아직 명확한 입장을 내놓지 않고 있다.

- 독일 선거제도는 다음과 같은 특성을 갖는다.

- 독일은 (한국과 같은 소선거구 단순다수제가 아니라) 권역별 연동형 비례대표제를 시행하기 때문에 네거티브 형태의 허위 정보가 결정적 역할을 하기 어려운 경향이 있다.

- 원칙적으로 독일의 선거는 분권적으로(dezentral) 조직되어 있어서 연방선거관리위원장은 주 및 지역선거관리위원장, 그 밖의 다수 기관과 협력하여 선거를 준비하여 치르고, 그 결과를 발표할 책임과 권한이 있다.

- 이는 각각의 선거관리기구가 정보처리기술을 사용하는 것에 대한 권한을 갖는다는 의미이다.

- 독일에서 유권자는 스스로 허위 정보를 가려낼 책임이 있다고 본다.

- 유권자는 다양한 채널을 통해 종합적인 정보를 얻을 수 있다. SNS에 막 올라온 정보에 대해 언제나 비판적으로 살펴보고 신뢰성을 따져

보아야 한다.

- 대부분의 가짜뉴스는 국가에 의해 간단히 금지될 수 없으며, (헌행법에 따르면) 처벌 대상이 아니다.
- 가짜정보에 대한 제재는 개인 생명이나 신체에 대한 직접적 위협의 정도가 매우 높을 때만 헌법적으로 가능하다.
- 가짜뉴스를 억제하는 독일 사회의 저력은 복수의 미디어 체계, 폭넓은 민주시민교육, 다양한 대안적 소통 채널 등의 민주적 소통과정이다. 바로 이런 점들이 디지털 방식으로 제작되는 많은 생성물이 조작될 가능성을 제약하고 있다.
- 가짜뉴스는 전적으로 법적인 문제라기보다는 하나의 사회적 해결 과제이다.

- 따라서 독일에서는 가짜뉴스 및 AI의 활용한 정보생성에 대해 법률에 의한 규제보다는 아직 자율규제에 중점을 두고 있는 것으로 보인다.

## □ 스위스

- 스위스 5개 정당이 2023년 10월 22일 연방 선거 캠페인에서 인공지능(AI)의 사용을 제한하기로 합의했다. 공동 성명서에 서명한 사회민주당(the Social Democratic Party), 중앙당(the Centre Party), 녹색당(the Green Party), 자유녹색당(Liberal Green Party), 개신교당(the Protestant Party)은 각 주(州) 지부에 녹화 광고, 포스터, 광고 등 선거운동 과정에서 AI를 사용할 때 이를 명시적으로 신고해야 한다는 행동 강령을 준수하도록 해야 한다고 주장했다. 또한 상대 후보를 공격하는 이른바 ‘네거티브’ 캠페인에 AI를 사용하는 것을 금지하는 조항 또한 포함된다.
- 급진자유당은 딥페이크 등을 예로 들며 AI의 악용을 금지하기 위한 노력이 필요함을 강조했지만, 성명서에 서명하지는 않았으며, 스위스 최대 정당 스위스 인민당은 네거티브 캠페인의 경계가 모호하다고 주장하며 AI의 선거 사용 금지에 대해 유보 의견을 밝혔다.

- AI에 대한 스위스의 규제 움직임은 그다지 적극적이지 않다. 그러나 규제 요구가 높아지자 기관의 설립을 통해 모니터링과 역량 개발의 업무를 수행하게 하고 있다.
- 연방 행정부의 AI 활용과 관련하여 연방정부는 2020년에 주요 지침을 채택했다. 연방정부는 AI 사용과 관련된 과제를 인식하기 위해 업계, 시민 사회, 학계 및 기술계와의 다자간 교류와 상호작용이 근본적으로 중요하다라는 견해를 가지고 있다.
- 유럽연합의 AI 법과 유럽평의회는 규제 입장을 따르고자 하며, 유럽연합의 법안이 서명에 성공함에 따라 앞으로 스위스의 행보에 주목할 필요가 있다.

#### □ 핀란드

- 2023년 유럽 미디어 리터러시 지수 결과에 따르면, 총 41개국 중 핀란드가 74점으로 1위를 차지했다. 핀란드는 행복, 언론 자유, 성평등, 사회 정의, 투명성, 교육 등 거의 모든 지표에서 꾸준히 상위권 또는 그에 근접한 순위를 차지하고 있어 외부로부터 오는 충격에 강하며 사회 균열을 찾아내어 악용하기 어렵다.
- 전 세계 다른 지역에서 언론에 대한 신뢰도가 하락하는 가운데 핀란드는 지역 언론과 공영방송이 강세를 유지하고 있다.
- 핀란드는 러시아와의 지리적 인접성으로 인해 오랫동안 러시아의 선전에 노출되어 있었다는 점에서 허위 정보 대응에 오랜 노력을 기울였고 핀란드는 유럽연합 국가 중 허위 정보 캠페인의 책임자에게 유죄 판결을 내린 첫 번째 국가이다.
- 기본권과의 관계에서 윤리적이고 책임감 있는 AI와 관련된 다양한 미래 발전 방향이 존재한다. 초기의 AI 핀란드 프로그램과 기업들의 적극적인 자율규제 움직임은 좋은 출발점이었기 때문에 핀란드가 이러한 아이디어를 실천에 옮기는 데 적극적으로 나서야 한다는 분위기가 지배적이었다.

- 법무부의 주도로 2018년부터 선거 개입에 대응하기 위한 준비를 시작했다. 대비책에는 선거 간섭에 대응하고 당국, 정당 및 기타 기관에 대응하는 방법에 대한 절차 및 협력 개발과 교육 제공이 포함된다. 이를 위해 관련 기관은 소셜미디어 기업과 직접 접촉하여 신뢰와 이를 유지하는 방법을 주제로 대비했다.
- 핀란드의 경우 2014년부터 AI 관련 정책 프로그램을 가동하였으며 규제보다는 역량 강화의 측면에서 접근하고 있다.
- 선거에서 시민들이 투표에 참여하도록 독려하고 미디어를 비판적으로 읽는 것이 왜 중요한지 상기시키기 위한 캠페인을 활용한다.
- AI 강국이라고 할 수 있는 핀란드는 선거 과정 관리에서는 여전히 전통적인 방식을 고수하고 있다. 2023년 봄, 핀란드 의회 선거와 유럽 의회 선거에서도 선거 조직과 관리는 여전히 펜, 종이, 기표소라는 검증된 방식에 의존한다.

## ■ 유형화에 따른 국가별 비교분석

- 중국에서 가장 먼저 AI 인터넷 당국이 인공지능 규제 초안을 발표했다. 목표는 AI 기술을 촉진하는 원칙 기반 아래, 그러나 중국의 근본적인 사회주의 가치에 부합해야 한다는 점에서 정치, 사회적 가치 목적을 지향한다.
- 미국은 전체적인 AI 기술 및 산업 발전, 개인 권리 보호의 원칙 기반 아래, 관련 행정부 혹은 지방정부의 폭넓은 해석에 의한 규제 적용이 진행되고 있다. 가치적 목적보다는 시장중심적 AI 규제의 원칙 기반 접근이 진행되고 있다.
- 영국은 EU를 탈퇴했으며 스위스는 EU의 회원국이 아님에도 불구하고 EU와 생성형 AI 규제에 있어 보조를 맞추고 있다. 물론 영국의 경우 플랫폼 사업자에 관한 책임성 요구를 강한 법으로 규제하고 있으나 스위스의 경우 여전히 구체적인 법안을 상정하지 않고 자율규제를 추구하고 있으나 EU의 법안과의 조율을 언급하고 있다는 점에서 앞으로의 규제

방향은 좀 더 지켜볼 필요가 있다.

- 그러나 스위스와 독일 그리고 핀란드처럼 기존법의 원칙을 통해 이 분야를 규제하는 법문화를 가진 국가들의 경우 법제화로의 선회는 크게 기대하기 어렵다.
- 영국의 경우 AI 분야 투자에서 세계 3위를 기록할 만큼 이 분야에 적극적으로이라는 점에서 연구 자금을 3억 4,400만 유로로 3배 늘릴 계획이다.
- 핀란드와 독일의 경우 자율규제의 방향성을 보이면서 개인이 문제를 스스로 해결할 수 있는 역량 강화의 측면에 집중하고 있다.
- 스위스, 핀란드, 그리고 독일과 같은 생성형 AI 문제에 덜 취약한 국가들의 공통점은 그 국가들의 제도적 특징에서 찾을 수 있다.
  - 예를 들어 독일은 (한국과 같은 소선거구 단순다수제가 아니라) 권역별 연동형 비례대표제를 시행하기 때문에 네거티브 형태의 가짜뉴스가 결정적 역할을 하기 어려운 경향이 있다.
  - 원칙적으로 독일의 선거는 분권적으로(dezentral) 조직되어 있어서 연방선거관리위원장은 주 및 지역선거관리위원장, 그 밖의 다수 기관과 협력하여 선거를 준비하여 치르고, 그 결과를 발표할 책임과 권한이 있다.
  - 이는 각각의 선거관리기구가 정보처리기술을 사용하는 것에 대한 권한을 갖는다는 의미이다.
- 서방 민주주의 국가에서 AI 분야에서 최초로 만들어진 법이기 때문에 분명 신호탄이 될 것이다. EU도 매우 중요한 시장인 만큼 모든 AI 기업은 이 새로운 표준을 따르게 될 것이다.
- 한 가지 분명한 것은 EU가 규제를 통해 경쟁 상대에게 뒤쳐진 부분을 만회하려고 하고 있지만 AI 기술은 개발의 시작 단계에 불과하며 이는 법적 규제방식에도 적용된다. 이 또한 국제적으로 자리를 잡고 더욱 발전해야 할 것이다.

## ■ 생성형 AI 선거 규제 동향과 한국적 함의

- 2023 미디어 리터러시 보고서는 자유민주주의 정부를 가진 국가에 사는 사람들이 “민주주의 제도가 없거나 제한적인 국가에 사는 사람들보다 잘못된 정보에 대해 걱정할 가능성이 더 높다”고 밝혔으며, 교육 수준이 높은 사람들 사이에서 가짜 뉴스에 대한 우려가 더 높다는 것을 발견했다. 가짜 뉴스의 영향에 가장 취약한 사회가 동시에 허위 정보의 확산과 영향에 대해 가장 우려하지 않는다는 것은 우려스러운 일이다.
- 미국, 영국, 유럽연합, 독일, 핀란드 사례
  - 생성형 AI의 기술 및 산업 발전을 고려하며, 정치·사회에 미치는 영향에 대해 신중한 접근을 하고 있다.
  - 사례국가는 표현의 자유라는 기본권의 원칙 하에 생성형 AI의 규제 원칙을 확립하고자 한다.
  - 표현의 자유, 법의 실효성, 법의 효율성을 고려할 때 생성형 AI 투명성 강화, 개인 권리 및 알고리즘 규제, 빅테크플랫폼에 대한 권고지침 마련을 통한 자율규제가 단기적으로는 설득력 있다.
  - AI 산업 선도 국가로서 지나친 규제는 지양할 필요가 있음. 중국의 사례처럼 전방위적인 차단과 규제는 기술 발전을 저해함과 동시에 법적 실효성도 보장할 수 없다.
- 정치과정에 생성형 AI를 활용하는 문제에 대해 법률적 규제보다 자율규제를 우선시하는 사례는 우리에게도 같은 시사점을 준다.
  - 허위 정보에 대한 사실 여부 판단을 법률에 따르는 방식이 물리적으로 어려움이 많다.
  - 최근 한국 페이스북에서 특정 용어의 사용에 대해 제재를 가했는데, 이는 이용자의 반발을 가져왔다.
- 허위 정보의 생성을 억제하는 독일과 핀란드 사회의 저력은 복수의 미



디어 체계, 폭넓은 시민교육, 다양한 대안적 소통 채널 등의 민주적 소통역량이라는 점에서 이러한 역량을 향상시킬 수 있는 조치가 필요하다.

## 참고 문헌

- 강상원. 2019. “소셜미디어 가짜뉴스 규제에 나선 영국 정부.” 국무조정실 규제 혁신.
- 고선규·이소영·김형호·송재민. 2021. “온라인 가짜뉴스 관련 선거관리제도 개선방안에 대한 연구.” 『한국선거협회』: 65.
- 김다운. 2023. “각국의 생성 AI 규제 및 정책 동향.” 『투이컨설팅』 (8월 16일) <https://www.2e.co.kr/news/articleView.html?idxno=302827> (검색일: 2023. 12. 01).
- 김도엽·마태경. 2023. “EU 의회 수정안 주요내용.” 톰슨로이터코리아(편). 『2023 국내외 AI 규제 및 정책 동향』 LAWnB Legal Essential Report. 22: 12.
- 김미정. 2023. “중국, 생성 AI 검열 시작 . . . “공산당 비판하면 벌금” .” 『ZDNET Korea』 (4월 14일) <https://zdnet.co.kr/view/?no=20230414074132> (검색일: 2023. 12. 01).
- 김용우. 2023. “가트너 ‘생성형 AI 기대감 최고조’ .” 『지디넷코리아』 (8월 17일) <https://zdnet.co.kr/view/?no=20230817104503> (검색일: 2023. 11. 01).
- 김재선. 2020. “허위정보 규제를 위한 행정법적 대응방안 -영국의 규제 논의를 중심으로.” 『법학논집』 24(3): 223-248.
- 김현정. 2021. “인공지능 기반 사회에 대비한 EU의 전략과 정책 : EU의 AI 규제안을 중심으로.” 『한국과 국제사회』. 5집, 제4호: 266-267.
- 김효진. 2017. “러시아 계정 15만개 블랙리스트 투표 개입 시도.” 『한겨레』 (11월 16일) <https://www.hani.co.kr/arti/international/europe/819395.html> (검색일: 2023. 11. 12).
- 성복숙. 2022. “20대 대선 관련 가짜뉴스의 양상과 그 사회적 영향.” 『인문사회 21』, 13집 3호.
- 이율. 2020. “런던은 내 앞마당?…러, 전방위 내정개입에 영, 뒷북 대응.” 『연합뉴스』 (7월 22일) <https://www.yna.co.kr/view/AKR20200722086651009?section=search> (검색일: 2023. 11. 13).
- 전윤미. 2023. “생성형 AI가 기존 AI와 결정적으로 다른 점.” 『애플경제』 (3월 17일) <https://www.apple-economy.com/news/articleView.html?idxno=71123>

- (검색일: 2023. 11. 01).
- 전학선 · 정필운 · 심우민 · 윤진희. 2020. “글로벌 플랫폼사업자의 자율규제 실태 및 협력 방안 연구.” 방송통신심의위원회.
- 전훈 · 최우정 · 김재선. 2019. “허위정보 유통과 민주주의.” 『한국법제연구원』 (9월 15일).
- 정태호 · 김훈집. 2015. “영국의 명예훼손죄 폐지와 그 교훈.” 『경희법학』 50집 제1호.
- 최민재 · 진민정 · 이현우. 2023. “유럽의 가짜뉴스 대응정책.” 『Media 정책 리포트』.
- 한국지식재산연구원. 2023a. “미국 국토안보수사국, ‘AI 로드맵’ 발표.” (11월 24일) [https://www.kiip.re.kr/board/trend/view.do?bd\\_gb=trend&bd\\_cd=1&bd\\_item=0&po\\_item\\_gb=US&po\\_no=22454](https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=US&po_no=22454) (검색일: 2023. 11. 24).
- \_\_\_\_\_. 2023b. “미국 상무부, 미국 최초의 인공지능 행정명령에서 핵심적인 역할 수행.” [https://www.kiip.re.kr/board/trend/view.do?bd\\_gb=trend&bd\\_cd=1&bd\\_item=0&po\\_item\\_gb=US&po\\_no=22409](https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=US&po_no=22409) (검색일: 2023. 11. 24).
- \_\_\_\_\_. 2021. “유럽연합 정상회의 상임의장, 탈린 디지털 서밋에서 유럽의 디지털 혁명을 강조.” (9월 28일) [https://www.kiip.re.kr/board/trend/view.do?bd\\_gb=trend&bd\\_cd=1&bd\\_item=0&po\\_item\\_gb=EU&po\\_no=20706](https://www.kiip.re.kr/board/trend/view.do?bd_gb=trend&bd_cd=1&bd_item=0&po_item_gb=EU&po_no=20706) (검색일: 2021. 09. 28).
- 허순철. 2015. “영국의 명예 훼손법 개정과 그 의미.” 『공법학연구』 16집 제4호.
- 홍석한. 2023. “미국 “2022 알고리즘 책임법안” 에 대한 고찰.” 『미국헌법연구』 34집 제1호.
- KOTRA. 2023. “중국, 생성형 AI 서비스 관리에 나서.” (7월 31일) [https://dream.kotra.or.kr/kotranews/cms/news/actionKotraBoardDetail.do?SITE\\_NO=3&MENU\\_ID=80&CONTENTS\\_NO=2&bbsGbn=242&bbsSn=242&pNttSn=204600](https://dream.kotra.or.kr/kotranews/cms/news/actionKotraBoardDetail.do?SITE_NO=3&MENU_ID=80&CONTENTS_NO=2&bbsGbn=242&bbsSn=242&pNttSn=204600) (검색일: 2023. 12. 01).
- Abbasi, Mohsen, Aditya Bhaskara, and Suresh Venkatasubramanian. 2021. “Fair clustering via equitable group representations.” Proceedings of the 2021 ACM conference on fairness, accountability, and transparency.
- Ahaskar, Abhijit. 2021. “Officials Use AI Analytics to Keep Bihar Panchayat

- Polls Free and Fair.” Story, Mint. (October 29, 2021) <https://www.livemint.com/technology/tech-news/officials-use-ai-analytics-to-keepbihar-panchayat-polls-free-and-fair-11635446470554.html> (검색일: 2023. 07. 25).
- Ahlquist, John, Kenneth Mayer, and Simon Jackman. 2014. “Alien Abduction and Voter Impersonation in the 2012 U.S. General Election: Evidence from a Survey List Experiment.” *Election Law Journal: Rules, Politics, and Policy* 13(4): 460-475.
- AI.GOV. “Making AI Work for the American People.” <https://ai.gov/> (검색일: 2023. 11. 24).
- Ajansi, Anadolu. 2023. “5 Swiss political parties agree to limit use of AI ahead of elections.” (September 25) <https://www.aa.com.tr/en/europe/5-swiss-political-parties-agree-to-limit-use-of-ai-ahead-of-elections/3000856> (검색일: 2023. 09. 26).
- Allie, Feyaad. 2023. “Facial Recognition Technology and Voter Turnout.” *The Journal of Politics* 85(1): 328-333.
- Asmolov, Gregory. 2014. “The Kremlin’s Cameras and Virtual Potemkin Villages: ICT and the Construction of Statehood.” In *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. Oxford: Oxford University Press.
- BBC. 2022. “Molly Russell: Coroner’s report urges social media changes.” (October 14) <https://www.bbc.com/news/uk-england-london-63254635> (검색일: 2023. 12. 01).
- \_\_\_\_\_. 2018. “Jessikka Aro: Finn jailed over pro-Russia hate campaign against journalist.” (October 18) <https://www.bbc.com/news/world-europe-45902496> (검색일: 2023. 11. 20).
- Berger, Jonah, et al. 2008. “Contextual priming: Where people vote affects how they vote.” *Proceedings of the National Academy of Sciences* 105(26): 8846-8849.
- Besaw, Clayon. 2021. “Election Violence Spiked Worldwide in 2020-Will this Year be Better?” *The Conversation* (February 18) <https://theconversation.com/election-violence-spikedworldwide-in-2020-will-this-year-be-better-153975>. (검색일: 2023. 08. 25).

- Bhatti, Bharat. 2012. "Aadhaar-Enabled Payments for NREGA Workers." *Economic and Political Weekly* 47(49): 16-19.
- Birhane, Abeba, et al. 2022. "The values encoded in machine learning research." *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*.
- Blumenthal, Jeremy and Terry Turnipseed. 2011. "The Polling Place Priming (PPP) Effect: Is Voting in Churches (Or Anywhere Else) Unconstitutional." *Boston University Law Review* 91(2): 561-599.
- Booth, Robert. 2017. "Truth seekers: inside the UK election's fake news war room." *Guardian* (May 19) <https://www.theguardian.com/politics/2017/may/19/truth-seekers-inside-the-uk-elections-fake-news-war-room> (검색일: 2023. 10. 12).
- Brady, Henry and John McNulty. 2011. "Turning Out to Vote: The Costs of Finding and Getting to the Polling Place." *American Political Science Review* 105(1): 115-34.
- Braga, Anthony, Rod Brunson, and Kevin Drakulich. 2019. "Race, Place, and Effective Policing." *Annual Review of Sociology* 45(1): 535-555.
- Bundeswahlleiter. 2021. "Erkennen und Bekämpfen von Desinformation." *Die Bundeswahlleiterin*. <https://www.bundeswahlleiterin.de/bundestagswahlen/2021/fakten-fakenews.html#7b083e33-e95d-4a3f-87a3-276f5e01eb6e> (검색일: 2023. 11. 24).
- Buolamwini, Joy, Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of the Machine Learning Research, 1st Conference on Fairness, Accountability and Transparency*. 81: 77-91.
- Celine, Widmer. 2023. "Künstliche Intelligenz. Gefahren und Potenziale für die Demokratie." *Das Schweizer Parlament* (June 15) <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaefft?AffairId=20233812> (검색일: 2023. 12. 02).
- Centeno, Albert. 2019. "Analysis of Background textures in Banknotes and Identity Documents for Counterfeit Detection." Ph.D. thesis, *Universitat Autònoma de Barcelona*.
- Centeno, Albert, et al. 2019. "Recurrent Comparator with Attention Models to Detect Counterfeit Documents." In *Proceedings of the 2019 International*

- Conference on Document Analysis and Recognition (ICDAR), 1332–1337.
- Charlier, Phillip. 2023. “Republican Party Biden attack ad features attack on Taiwan.” Taiwan English News. (April 26) <https://taiwanenglishnews.com/republican-party-biden-attack-ad-features-attack-on-taiwan/> (검색일: 2023. 11. 01).
- CNN. 2023. “AI-generated GOP campaign video may be harbinger of disturbing new trend for 2024, expert says.” (April 25) <https://edition.cnn.com/politics/live-news/biden-2024-presidential-election/index.html> (검색일: 2023. 11. 01).
- \_\_\_\_\_. 2023. “Suspected Chinese operatives using AI generated images to spread disinformation among US voters, Microsoft says.” (September 7) <https://edition.cnn.com/2023/09/07/politics/chinese-operatives-ai-images-social-media/index.html> (검색일: 2023. 12. 01).
- Collins, Katie. 2017. “Facebook posts fake-news ads in newspapers ahead of UK vote.” (May 8) <https://www.cnet.com/tech/tech-industry/facebook-posts-fake-news-ads-in-newspapers-ahead-of-uk-election/?openLogin=1> (검색일: 2023. 12. 01).
- Collins, Matthew. 2014. “Collins On Defamation.” : 493, 498-499.
- Congressional Research Service. 2021. “EU Digital Policy and International Trade.” (March 25) <https://sgp.fas.org/crs/row/R46732.pdf> (검색일: 2021. 08. 25).
- Conradie, Lizelle. 2019. “Rules-based approach vs Principle-based approach to regulation in the Financial Industry.” Etude Risk Management <https://www.etude.co.za/article.php?article=32> (검색일: 2023. 11. 01).
- Contorno, Steve and Donie O’Sullivan. 2023. “DeSantis campaign posts fake images of Trump hugging Fauci in social media video.” CNN (June 8) <https://edition.cnn.com/2023/06/08/politics/desantis-campaign-video-fake-ai-image/index.html> (검색일: 2023. 11. 01).
- Coughlan, Sean. 2017. “Schools should teach pupils how to spot fake news.” BBC (March 18) <https://www.bbc.com/news/education-39272841> (검색일 : 2023. 10. 15).
- Darwis, Kevindra., Charles Lim. 2011. “Design and Implementation of e-KTP (Indonesian Electronic Identity Card) Key Management System.” In

- Proceedings of the 2011 International Conference on Advanced Computer Science and Information Systems: 143-146.
- Davies, Harry. 2015. "Ted Cruz using firm that harvested data on millions of unwitting Facebook users." The Guardian (December 11) <https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data> (검색일: 2023. 11. 10).
- DC Chamber of Commerce. 2023. "DC Chamber of Commerce Small Business Action Alert: Stop Discrimination By Algorithms Act Of 2021." <https://dcchamber.org/algorithms-act/> (검색일: 2023. 11. 24).
- Deutscher Bundestag. "Generative Künstliche Intelligenz: Regulierung steht am Scheideweg." <https://www.bundestag.de/dokumente/textarchiv/2023/kw21-pa-digitales-ki-947976> (검색일: 2023. 12. 01).
- Devlin, Kayleen and Joshua Cheetham. 2023. "Fake Trump arrest photos: How to spot an AI-generated image." BBC. (March 24) <https://www.bbc.com/news/world-us-canada-65069316> (검색일: 2023. 11. 1).
- Dey, Palash, Neeldhara Misra, Swaprava Nath, Garima Shakya. 2019. "A Parameterized Perspective on Protecting Elections." In Proceedings of the Twenty Eighth International Joint Conference on Artificial Intelligence, IJCAI-19: 238-244.
- \_\_\_\_\_. 2021. "A Parameterized Perspective on Protecting Elections." Theoretical Computer Science 874: 15-31.
- Die Bundesversammlung. 2023. "Künstliche Intelligenz. Gefahren und Potenziale für die Demokratie." (June 15) <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233812> (검색일: 2023. 12. 01).
- \_\_\_\_\_. 2022. "Vorausschauende KI-Kompetenz der Schweiz." (May 17) <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20223298> (검색일: 2023. 12. 01).
- Die Bundeswahlleiterin. "Bundestagswahl 2021." <https://www.bundeswahlleiterin.de/bundestagswahlen/2021/fakten-fakenews.html#7b083e33-e95d-4a3f-87a3-276f5e01eb6e> (검색일: 2023. 12. 05).
- Dunbar, J, et al. 2015. "Video Verification: An Alternative form of Identify Verification." Procedia Manufacturing 3: 4914-4920.

- Eisele, Ines. 2023. “Fact check: AI fakes in Israel’s war against Hamas.” DW (October 11) <https://www.dw.com/en/fact-check-ai-fakes-in-israels-war-against-hamas/a-67367744> (검색일: 2023. 11. 24).
- Ensign, Danielle, et al. 2018. “Runaway Feedback Loops in Predictive Policing.” In Proceedings of the Machine Learning Research, 1st Conference on Fairness, Accountability and Transparency 81: 160-171.
- EURACTIV. 2023. “Digital competition law: EU Commission unveils its ‘gatekeepers’ list.” (September 7) <https://www.euractiv.com/section/law-enforcement/news/digital-competition-law-eu-commission-unveils-its-gatekeepers-list/> (검색일: 2023. 12. 01).
- European Commission. “Digital in the EU-US Trade and Technology Council.” <https://futurium.ec.europa.eu/en/EU-US-TTC> (검색일: 2022. 09. 15).
- \_\_\_\_\_. “Regulatory framework proposal on artificial intelligence.” <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> (검색일: 2023. 11. 01).
- \_\_\_\_\_. “The Digital Services Act: ensuring a safe and accountable online environment.” [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en) (검색일: 2021. 09. 28).
- \_\_\_\_\_. 2021. “Proposals for a Regulations of the European Parliament and the European Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.” COM/2021/206 final: 13-16.
- \_\_\_\_\_. 2020a. “Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on contestable and fair markets in the digital sector (Digital Markets Act).” COM/2020/842 final.
- \_\_\_\_\_. 2020b. “The Digital Markets Act: ensuring fair and open digital markets.” (December 15) [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en) (검색일: 2021. 09. 28).
- European Parliament. 2020. “Parliament leads the way on first set of EU rules



- for Artificial Intelligence.” (October 20)  
<https://www.europarl.europa.eu/news/en/press-room/20201016IPR89544/parliament-leads-the-way-on-first-set-of-eu-rules-for-artificial-intelligence> (검색일: 2023. 11. 24).
- Federal Office of Communications OFCOM. 2022. “Monitoring of the “Artificial Intelligence” Guidelines for the Confederation: Evaluation of the application and relevance of the guidelines.” (December 9)  
<https://cnaai.swiss/wp-content/uploads/2022/12/Monitoring-der-Leitlinien-KI-EN.pdf> (검색일: 2022. 12. 15).
- Fitzsimmons, Zack, Omer Lev. 2020. “Selecting Voting Locations for Fun and Profit.” In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20, 224(30).
- Fogarty, Brian, et al. 2015. “News Attention to Voter Fraud in the 2008 and 2012 US Elections.” *Research & Politics* 2(2): 1-8.
- Garnett, Holly and Sean Grogan. “I Came, I Saw, I Voted: Distance to Polling Locations and Voter Turnout in Ontario, Canada.” *Canadian Journal of Political Science/Revue canadienne de science politique* 54(2): 316-334.
- Gartner. “What Is Artificial Intelligence?.”  
<https://www.gartner.com/en/topics/artificial-intelligence> (검색일: 2023. 11. 01).
- Gelb, Alan and Anna Diofasi. 2019. “Biometric Elections in Poor Countries: Wasteful or a Worthwhile Investment?” *Review of Policy Research* 36(3): 318-340.
- Grother, P., Mei Ngan, and Kayee Hanaoka. 2019. “Face Recognition Vendor Test Part 3: Demographic Effects.” Report, NIST.
- Gui, Jie, Zhenan Sun, Yonggang Wen, Dacheng Tao, Jieping Ye. 2021. “A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications.” *IEEE Transactions on Knowledge and Data Engineering* 35: 3313-3332.
- Gupta, Arvind. 2021. “Indian Elections: Scale, Use of Technology and SOTA Outcomes.” In Invited talk 1 at Workshop on AI for Credible Elections: A Call to Action.
- HM Government. 2021. “National AI Strategy.” (September)

- [https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National\\_AI\\_Strategy\\_-\\_PDF\\_version.pdf](https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf) (검색일: 2023. 12. 01).
- \_\_\_\_\_. 2020. “Online Harms White Paper: Full Government Response to the Consultation.” (December)  
[https://assets.publishing.service.gov.uk/media/5fd8af718fa8f54d5f67a81e/Online\\_Harms\\_White\\_Paper\\_Full\\_Government\\_Response\\_to\\_the\\_consultation\\_CP\\_354\\_CCS001\\_CCS1220695430-001\\_\\_V2.pdf](https://assets.publishing.service.gov.uk/media/5fd8af718fa8f54d5f67a81e/Online_Harms_White_Paper_Full_Government_Response_to_the_consultation_CP_354_CCS001_CCS1220695430-001__V2.pdf) (검색일: 2023. 11. 10).
- HM Revenue & Custom. 2020. “Digital Services Tax.”  
<https://www.gov.uk/government/publications/introduction-of-the-digital-services-tax/digital-services-tax> (검색일: 2023. 11. 20).
- Holistic AI. 2023. “US Algorithmic Accountability Act: Third Time Lucky?.” (October 25) <https://www.holisticai.com/blog/us-algorithmic-accountability-act> (검색일: 2023. 11. 24).
- Holroyd, Matthew and Fola Olorunselu. 2022. “Deepfake Zelenskyy surrender video is the first intentionally used’ in Ukraine war.” euronews. (March 16)  
<https://www.euronews.com/my-europe/2022/03/16/deepfake-zelenskyy-surrender-video-is-the-first-intentionally-used-in-ukraine-war> (검색일: 2023. 11. 01).
- Hosein, Gus and Carly Nyst. 2013. “Aiding Surveillance: An Exploration of How Development and Humanitarian Aid Initiatives are Enabling Surveillance in Developing Countries.” Report, Privacy International.
- Huang, Zhanyuan, et al. 2018. “Video-Based Fall Detection for Seniors with Human Pose Estimation.” In Proceedings of the 2018 4th International Conference on Universal Village (UV): 1-4.
- Huber, Gregory, et al. 2021. “The racial burden of voter list maintenance errors: Evidence from Wisconsin’s supplemental movers poll books,” Science Advances 7(8).
- Hudnall, Hannah. 2023. “Fact check: Video altered to show Joe Biden making transphobic remarks.” USA TODAY (February 9)  
<https://www.usatoday.com/story/news/factcheck/2023/02/09/fact-check-video-edited-show-joe-biden-making-transphobic-remarks/11211453002/> (검색일: 2023. 11. 01).
- HUFFPOST. 2012. “Russia Election 2012: Ballot Box Stuffing Purportedly Caught

- On Tape (VIDEO).” (March 5)  
[https://www.huffpost.com/entry/russia-election-2012-ballot-box-stuffing\\_n\\_1321379](https://www.huffpost.com/entry/russia-election-2012-ballot-box-stuffing_n_1321379) (검색일: 2023. 08. 25).
- Hyde, Susan. and Nikolay Marinov. 2014. “Information and Self-Enforcing Democracy: The Role of International Election Observation.” *International Organization* 68(2): 329-359.
- IWF. “How AI is being abused to create child sexual abuse imagery.”  
<https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/> (검색일: 2023. 12. 01).
- Jain, Ankit, et al. 2023. “Blockchain-Based Secure E-voting System Using Aadhaar Authentication.” In *Predictive Data Security Using AI: Insights and Issues of Blockchain, IoT, and DevOps*: 89-103.
- James, Toby, and Alistair Clark. 2020. “Electoral Integrity, Voter Fraud and Voter ID in Polling Stations: Lessons from English Local Elections.” *Policy Studies* 41(2-3): 190-209.
- Julian, Vikki. 2020. “IPSO Blog: Government responds to Online Harms consultation.” ipso (February 13)  
<https://www.ipso.co.uk/news-press-releases/blog/ipso-blog-government-responds-to-online-harms-consultation/> (검색일: 2023. 12. 01).
- Jung, Christopher, Sampath Kannan and Neil Lutz. 2020. “Service in your neighborhood: Fairness in center location.” *Foundations of Responsible Computing(FORC)*.
- Kakadiya, Rutvik, et al. 2019. “AI Based Automatic Robbery/Theft Detection using Smart Surveillance in Banks.” In *Proceedings of the 2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*: 201-204.
- Kalra, Ash. 2022. “The Workplace Technology Accountability Act.”  
<https://www.etui.org/sites/default/files/2022-05/WTAA%20summary.pdf> (검색일: 2023. 11. 24).
- Kettemann, Matthias. 2021. “Ordnungsbildung in Online-Kommunikationsräumen.” *Bundeszentralen für politische Bildung* (October 4)  
<https://www.bpb.de/themen/medien-journalismus/digitale-desinformation/341326/ordnungsbildung-in-online-kommunikationsraeumen/> (검색일: 2023. 12. 01).

- Kleindessner, Matthäus, Pranjal Awasthi, and Jamie Morgenstern. 2020. “A notion of individual fairness for clustering.” arXiv preprint arXiv:2006.04960.
- KPMG. 2023. “EU KI-Verordnung: «Brüssel-Effekt» in der Schweiz?.” (May 24) <https://kpmg.com/ch/de/blogs/home/posts/2023/05/eu-ki-verordnung-bruessel-effekt-in-der-schweiz.html> (검색일: 2023. 08. 24).
- Kulldorff, Martin. 1997. “A Spatial Scan Statistic.” *Communications in Statistics - Theory and Methods* 26(6): 1481–1496.
- Lajika, Arijeta. 2023. “Artificial intelligence makes voice cloning easy and ‘the monster is already on the loose’ .” *Fortune* (February 12) <https://fortune.com/2023/02/11/artificial-intelligence-makes-voice-cloning-easy-and-the-monster-is-already-on-the-loose/> (검색일: 2023. 11. 01).
- Lessenski, Marin. 2023. “The Media Literacy Index 2023: Measuring Vulnerability of Societies to Disinformation.” <https://osis.bg/wp-content/uploads/2023/06/MLI-report-in-English-22.06.pdf> (검색일: 2023. 12. 01).
- Li, Lixiang, et al. 2020. “A Review of Face Recognition Technology.” *IEEE Access* 8: 139110–139120.
- Liang, Chua, et al. 2021. “Smart Home Security System Based on Zigbee.” In *Advances in Smart System Technologies*. Singapore: Springer.
- Lim, Liqian. 2023. “What does Biden’s Executive Order on AI safety measures mean for businesses?.” *Snyk Blog* (November 2) [https://snyk.io/blog/biden-ai-safety-executive-order/?utm\\_medium=paid-search&utm\\_source=google&utm\\_campaign=ba\\_ai-awareness&utm\\_content=blog-executive&utm\\_term=ai%20executive%20order&gclid=CjwKCAiAsIGrBhAAEiwAEzMICzbcwEaCKNIOVfCSuljECUWbecodGTCJ2ORkT0EM9WRYnGD4ZSQlvhoCc28QAvD\\_BwE](https://snyk.io/blog/biden-ai-safety-executive-order/?utm_medium=paid-search&utm_source=google&utm_campaign=ba_ai-awareness&utm_content=blog-executive&utm_term=ai%20executive%20order&gclid=CjwKCAiAsIGrBhAAEiwAEzMICzbcwEaCKNIOVfCSuljECUWbecodGTCJ2ORkT0EM9WRYnGD4ZSQlvhoCc28QAvD_BwE) (검색일: 2023. 11. 24).
- Mackintosh, Eliza. 2019. “Finland is winning the war on fake news. What it’s learned may be crucial to Western democracy.” *CNN* (May 17) <https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/> (검색일: 2023. 12. 01).
- Mahabadi, Sepideh, and Ali Vakilian. 2020. “Individual fairness for k-clustering.” *International conference on machine learning*. PMLR.

- Marasco, Emanuela. 2019. "Biases in Fingerprint Recognition Systems: Where Are We At?" In Proceedings of the 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS): 1-5.
- McKinsey & Company. 2023a. "What's the future of generative AI? An early view in 15 charts." (August 25) <https://www.mckinsey.com/featured-insights/mckinsey-explainers/whats-the-future-of-generative-ai-an-early-view-in-15-charts#/> (검색일: 2023. 11. 01).
- \_\_\_\_\_. 2023b. "What every CEO should know about generative AI." (May 12) <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/what-every-ceo-should-know-about-generative-ai#/> (검색일: 2023. 11. 01).
- McQuillan, Dan. 2022. Resisting AI: An Anti-fascist Approach to Artificial Intelligence. Bristol: Bristol University Press.
- Merivaki, Thessalia. 2020. "“Our voter rolls are cleaner than yours” : Balancing access and integrity in voter list maintenance," American Politics Research 48(5): 560-570.
- Mostofa, Moktari, et al. 2021. "Deep GAN-Based Cross-Spectral Cross-Resolution Iris Recognition." IEEE Transactions on Biometrics, Behavior, and Identity Science 3(4): 443-463.
- Mueen, Abdullah, et al. 2009. "Exact Discovery of Time Series Motifs." In Proceedings of the 2009 SIAM International Conference on Data Mining (SDM): 473-484.
- NCSL. 2023. "Artificial Intelligence (AI) in Elections and Campaigns." (December 7) <https://www.ncsl.org/elections-and-campaigns/artificial-intelligence-ai-in-elections-and-campaigns> (검색일: 2023. 11. 08).
- Ntoutsis, Eirini, et al. 2020. "Bias in data-driven artificial intelligence systems—An introductory survey." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 10(3).
- Obeta, Sebastine, et al. 2021. "Credible Elections in Nigeria: The Role of Information Communication Technology." The International Journal of Science & Technology 9(10): 35-39.
- Office of U.S. Senator Ron Wyden. 2023. "Support for the Algorithmic Accountability Act of 2023." (September 21)

- [https://www.wyden.senate.gov/imo/media/doc/support\\_for\\_the\\_algorithmic\\_accountability\\_act\\_of\\_2023.pdf](https://www.wyden.senate.gov/imo/media/doc/support_for_the_algorithmic_accountability_act_of_2023.pdf) (검색일: 2023. 11. 24).
- Oremus, Will, Drew Harwell and Teo Armus. 2023. “A tweet about a Pentagon explosion was fake. It still went viral.” *The Washington Post* (May 22) <https://www.washingtonpost.com/technology/2023/05/22/pentagon-explosion-ai-image-hoax/> (검색일: 2023. 11. 01).
- P, Deepak. 2016. “Anomaly Detection for Data with Spatial Attributes.” In *Unsupervised Learning Algorithms*, Cham: Springer International Publishing.
- P, Deepak, and Sowmya Sundaram. 2023. “FiSH: Fair Spatial Hot Spots.” *Data Mining and Knowledge Discovery* 37: 1374-1403.
- P, Deepak and Savitha Abraham. 2020. “Representativity Fairness in Clustering.” In *Proceedings of the 12th ACM Conference on Web Science 202-11*. Association for Computing Machinery
- Perri, Lori. 2023. “What’s New in the 2023 Gartner Hype Cycle for Emerging Technologies.” *Gartner* (August 23) <https://www.gartner.com/en/articles/what-s-new-in-the-2023-gartner-hype-cycle-for-emerging-technologies> (검색일: 2023. 11. 01).
- Prime Minister’s Office. 2019. “Finland has the best elections in the world. And why is that.” <https://vnk.fi/en/-/suomessa-on-maailman-parhaat-vaalit-mieti-miksi-https://vnk.fi/en/-/suomessa-on-maailman-parhaat-vaalit-mieti-miksi> (검색일: 2023. 12. 01).
- Project Syndicate. 2023. “China’s Short-Sighted AI Regulation.” (December 8) <https://www.project-syndicate.org/commentary/risks-of-beijing-internet-court-ruling-allowing-copyright-of-ai-generated-content-by-angela-huyue-zhang-2023-12> (검색일: 2023. 12. 01).
- Pruitt, John, and Tamara Adlin. 2006. *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Amsterdam: Morgan Kaufmann.
- Ramakumar, R. 2010. “The Unique ID Project in India: A Skeptical Note.” In *Ethics and Policy of Biometrics*, Berlin, Heidelberg: Springer.
- Ramzan, Muhammad, et al. 2019. “A Review on State-of-the-Art Violence Detection Techniques.” *IEEE Access* 7: 107560-107575.
- Rao, Ursula. and Vijayanka Nair. 2019. “Aadhaar: Governing with Biometrics.”

- South Asia: Journal of South Asian Studies 42(3): 469-481.
- Republicans-Armedservices. “FY23 National Defense Authorization Act.”  
<https://armedservices.house.gov/sites/republicans.armedservices.house.gov/files/Final%20FY23%20NDAA%20Conf%20Highlights.pdf> (검색일: 2023. 11. 24).
- RESIST model: a quick guide. “RESIST”  
<https://cyberpolicy.nask.pl/wp-content/uploads/2020/04/RESIST-Counter-disinformation-Toolkit.pdf> (검색일: 2023. 12. 01).
- Reuters. 2023a. “Video does not show Joe Biden making transphobic remarks.” (February 11)  
<https://www.reuters.com/article/factcheck-biden-transphobic-remarks/fact-check-video-does-not-show-joe-biden-making-transphobic-remarks-idUSL1N34Q1IW/> (검색일: 2023. 11. 01).
- \_\_\_\_\_. 2023b. “OpenAI, Google, Meta pledge to watermark AI content for safety.” cybernews (November 15)  
<https://cybernews.com/tech/openai-google-meta-watermark-ai-content-for-safety/> (검색일: 2023. 11. 24).
- Ribeiro, Marco, S. Sameer Singh, Carlos Guestrin. 2016. “Why Should I Trust You?: Explaining the Predictions of Any Classifier.” In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 1135-1144.
- Rodon, Toni, and Marc Guinjoan. 2022. “Beaten Ballots: Political Participation Dynamics Amidst Police Interventions.” Political Science Research and Methods 10(1): 153-170.
- Roopak, T. and R. Sumathi. 2020. “Electronic Voting based on Virtual ID of Aadhar using Blockchain Technology.” In Proceedings of the 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA): 71-75.
- Rutchick, Abraham. 2010. “Deus ex machina: The influence of polling place on voting behavior.” Political Psychology 31(2): 209-225.
- Saliba, Emmanuelle. 2023. “How verified accounts helped make fake images of a Pentagon explosion go viral.” abc News (May 24)  
<https://abcnews.go.com/US/verified-accounts-helped-make-fake-images-pentagon-explosion/story?id=99541361> (검색일: 2023. 11. 01).

- Sängerlaub, Alexander. 2017. "Verezzerte Realitäten – Die Wahrnehmung von "Fake News" im Schatten der USA und der Bundestagswahl." Stiftung Neue Verantwortung (October) [https://www.stiftung-nv.de/sites/default/files/fake\\_news\\_im\\_schatten\\_der\\_usa\\_und\\_der\\_bundestagswahl.pdf](https://www.stiftung-nv.de/sites/default/files/fake_news_im_schatten_der_usa_und_der_bundestagswahl.pdf) (검색일: 2023. 12. 01).
- Satariano, Adam and Amie Tsang. 2019. "Who's Spreading Disinformation in U.K. Election? You Might Be Surprised." New York Times. <https://www.nytimes.com/2019/12/10/world/europe/elections-disinformation-social-media.html> (검색일: 2023. 11. 12).
- Schmidt, Adam and Laura Albert. 2022. "Locating ballot drop boxes." arXiv preprint arXiv:2207.10707.
- Sharma, S. 1999. "Booth Capturing: Judicial Response." Journal of the Indian Law Institute 41(1): 44-55.
- Shepherd, Michael, Adriane Fresh, Nick Eubank, and Joshua Clinton. 2021. "The politics of locating polling places: race and partisanship in North Carolina Election Administration, 2008-2016." Election Law Journal: Rules, Politics, and Policy 20(2): 155-177.
- Simoës, Stanley and Muiris MacCarthaigh. 2023. "AI and Core Electoral Processes: Mapping the Horizons." AI Magazine.
- Simon, Felix, Sacha Altay, and Hugo Mercier. 2023. "Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown." Harvard Kennedy School Misinformation Review. 4(5): 2.
- Stepanov, Nicola-Paul. 2022. "Fairness in Facility Location through Clustering." Master's thesis, University of Amsterdam.
- Sullivan, Donie, and Jon Passantino. "'Verified' Twitter accounts share fake image of 'explosion' near Pentagon, causing confusion." CNN. (May 23) <https://edition.cnn.com/2023/05/22/tech/twitter-fake-image-pentagon-explosion/index.html> (검색일: 2023. 11. 01).
- Sultani, Waqas, Chen Chen, and Mubarak Shah. 2018. "Real-World Anomaly Detection in Surveillance Videos." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 6479-6488.
- Suomeksi, Tiivistelmä. 2023. "Comparison of AI policies and programmes in Finland, the Netherlands and Sweden – case Finland." White paper



- (Oktober 2)  
<https://www.vttresearch.com/sites/default/files/2023-02/Comparison-of-AI-policies-case-Finland-10022023.pdf> (검색일: 2023. 12. 01).
- Tech Trends. 2023. “Beijing court’s ruling that AI-generated content can be covered by copyright eschews US stand, with far-reaching implications on tech’s use.” (December 1)  
<https://www.scmp.com/tech/tech-trends/article/3243570/beijing-courts-ruling-a-i-generated-content-can-be-covered-copyright-eschews-us-stand-far-reaching> (검색일: 2023. 12. 01).
- The Telegraph. 2022. “Deepfake video of Volodymyr Zelensky surrendering surfaces on social media.” Youtube (March 17)  
<https://www.youtube.com/watch?v=X17yrEV5sl4> (검색일: 2023. 11. 01).
- The Washinton Post. 2023. “Fake images of Trump arrest show ‘giant step’ for AI’s disruptive power.” (March 22)  
<https://www.washingtonpost.com/politics/2023/03/22/trump-arrest-deepfakes/> (검색일: 2023. 11. 01).
- THE WEEK. 2017. “Could fake news impact the general election?.” (May 4)  
<https://theweek.com/84203/could-fake-news-impact-the-general-election>. (검색일: 2023. 10. 12).
- The White House. 2023. “FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.” (October 30)  
<https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/> (검색일: 2023. 11. 24).
- \_\_\_\_\_. 2022. “Blueprint for an AI Bill of Rights.”  
<https://www.whitehouse.gov/ostp/ai-bill-of-rights/> (검색일: 2023. 11. 24).
- Thiel, Thorsten. 2023. “KI und Demokratie: Entwicklungspfade, Aus Politik und Zeitgeschichte, Bundeszentralen für politische Bildung.” (October 13)  
<https://www.bpb.de/shop/zeitschriften/apuz/kuenstliche-intelligenz-2023/541497/ki-und-demokratie-entwicklungspfade/> (검색일: 2023. 12. 01).
- Trade and Technology Council. 2022. “TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management.” (December 1)  
<https://digital-strategy.ec.europa.eu/en/library/ttc-joint-roadmap-trustworthy-a>

- i-and-risk-management (검색일: 2023. 12. 01).
- Turkoglu, Celik and Erol Genevois. 2020. “A comparative survey of service facility location problems.” *Annals of Operations Research* 292(1): 399-468.
- U.S. Congress. 2022. “AI Training Act: Artificial Intelligence Training for the Acquisition Workforce Act,” (October 17) <https://www.congress.gov/117/plaws/publ207/PLAW-117publ207.pdf> (검색일: 2023. 11. 24).
- UK Parliament. 2019. “Disinformation and ‘fake news’: Final Report published.” (February 18) <https://committees.parliament.uk/committee/378/digital-culture-media-and-sport-committee/news/103668/fake-news-report-published-17-19/> (검색일: 2023. 11. 12).
- Ulmer, Alexandra., Tong, Anna. 2023. “With apparently fake photos, DeSantis raises AI ante” . *Reuters* (June 9) <https://www.reuters.com/world/us/is-trump-kissing-fauci-with-apparently-fake-photos-desantis-raises-ai-ante-2023-06-08/> (검색일: 2023. 11. 01).
- Verma, Arvind. 2005. “Policing Elections in India.” *India Review* 4(3-4): 354-376.
- Volpcelli, Gian. 2023 “Meet Britain’ s first AI-powered candidate.” *POLITICO* (July 18) <https://www.politico.eu/article/britain-uk-andrew-gray-ai-artificial-intelligence-manifesto-polis-mp-nigel-adams-selby-ainsty-election/> (검색일: 2023. 12. 01).
- Vranckaert, Koen. 2023. “How Cautious is too Cautious? The US and EU Artificial Intelligence Roadmap.” (May 9) <https://www.law.kuleuven.be/citip/blog/how-cautious-is-too-cautious-the-us-and-eu-artificial-intelligence-roadmap-part-1/> (검색일: 2023. 12. 01).
- Wachter-Boettcher, Sara. 2017. *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. New York: W. W. Norton & Company.
- Wang, Hongzhi, Mohamed Bah, and Mohamed Hammad. 2019. “Progress in outlier detection techniques: A survey.” *Ieee Access* 7: 107964-108000.
- Weisburd, David. 2016. “Does Hot Spots Policing Inevitably Lead to Unfair and

- Abusive Police Practices, or Can We Maximize Both Fairness and Effectiveness in the New Proactive Policing?." University of Chicago Legal Forum 2016(16): 661-689.
- Woodruff II, James. 2011. "Where the Wild Things Are: The Polling Place, Voter Intimidation, and the First Amendment." University of Louisville Law Review 50: 253-285.
- Yamin, Muhammad, Andrii Shalaginov, and Basel Katt. 2020. "Smart policing for a smart world opportunities, challenges and way forward." Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC) 1.
- Ye, Mang, et al. 2022. "Deep Learning for Person Re-Identification: A Survey and Outlook." IEEE Transactions on Pattern Analysis and Machine Intelligence 44(6): 2872-2893.
- Zhang, Qingyang, et al. 2019. "Edge Video Analytics for Public Safety: A Review." Proceedings of the IEEE 107(8): 1675-1696.