

【특집1: 과학기술과 인권, 그리고 민주주의】

인공지능이 인권과 민주주의에 미치는 영향과 인공지능법안의 쟁점* **

유승익

한동대학교, 헌법학

ysi@handong.edu

<국문초록>

인공지능은 인권과 민주주의의 관점에서 긍정적 측면과 부정적 측면을 모두 갖는다. 인공지능은 정치적·경제적·사회적 구조를 획기적으로 전환할 뿐만 아니라, 새로운 헌법적·법률적 문제를 제기하고 있다. 이 글은 정치적·법적 문제의식에 한정하여 인공지능이 인권과 민주주의에 미치는 영향을 살펴보고, 최근 국회에서 논의되고 있는 인공지능 법안의 쟁점을 짚어본다.

인공지능 시스템은 거의 모든 인권과 기본권에 부정적 영향을 미칠 수 있다. 특히 표현의 자유와 개인정보자기결정권에 대한 영향은 민주적 의사형성과정에 중대한 장애를 초래할 수 있다. 특히 ‘콘텐츠 관리’와 ‘사용자 프로파일링’은 인공지능 시스템이 기본권과 민주주의에 미치는 위험성을 예시하는 대표적 사례이다. 같은 맥락에서 온라인상의 허위 정보에 대한 규제와 관리 문제가 제기된다. 통제되지 않는 인공지능 시스템은 개인의 인권을 취약하게 할 뿐만 아니라 정치·적 공동체의 운명을 민간기업의 처분에 맡기게 되는 위험도 안고 있다. 또한 감시사회가 고도화되고 분화되어 전통적 감시 권력이 빅데이터 감시 권력으로 진화하면서, 감시권력을 인권 주체의 대립적 객체가 아닌 사회적 인프라 구조처럼 자리매김하고, 감시 자본주의하에 개인은 데이터 조각으로 비인격화, 분할, 무화

* 심사위원: 오길영, 엄순영, 오동석

투고일자: 2023. 6. 7 심사개시: 2023. 6. 8. 게재확정: 2023. 6. 25.

** 이 글은 민주주의법학연구회 2023년 봄 정기학술대회 “과학기술과 인권, 그리고 민주주의”(2023. 5. 19.)에서 발표한 “인공지능이 인권과 민주주의에 미치는 영향과 규제거버넌스의 필요성”을 수정한 것이다. 사회를 맡아주신 최정학 교수님(한국방송통신대), 토론해 주신 윤현식 박사님(노동·정치·사람 정책위원)과 객민곤 교수님(서강대)께 감사드립니다.

된다. 자유는 ‘동의’에 의한 의사 결정의 위임과 인공지능 시스템에 의한 정보원의 독과점의 방식으로 포획되고, 개인과 집단의 정체성은 알고리즘에 의해 유도되면서 집단양극화된다. 자유는 외재적으로 제약되는 것이 아니라 내재적으로 왜곡된다. 인공지능사회는 빅테크, 디지털 플랫폼 등 사적 행위자가 국가나 정부의 역할을 초월하면서 사적 권력을 행사한다. 기업 권력은 정치 권력과 병존하면서 헌법 전통에 도전하고 있다.

세계 여러 나라는 인공지능 기술의 잠재적 위험에 대처하기 위한 입법을 준비하고 있다. 우리 국회에서도 인공지능법안에 대한 입법절차를 진행 중이다. 하지만 이 법안이 채택한 “우선허용·사후규제 원칙”은 인공지능 기술에 잠재한 위험에 대처하기에 부적합하며, 산업육성에 반드시 유리한 방식도 아니다. 고위험 영역 인공지능에 관한 분류는 유럽연합 등의 세분화된 분류방식과 대조적으로 금지되는 인공지능을 규정하지도 않으며, 생명·안전·기본권에 미치는 중대한 위험성이 예상되는 영역을 별다른 근거 없이 제외하고 있다. 또한 인공지능 정책의 주관 기관을 기술관료 중심의 과기정통부로 설정함으로써 산업편향적인 인공지능 정책 거버넌스에 치중하고 있다. 따라서 현재 계류 중인 법안은 전면적으로 재검토되어야 한다.

인공지능의 잠재적 위험성에 대비하고 인권과 민주주의가 기술결정론으로 대체되는 것을 방지하기 위해서는 인공지능 시스템의 한계와 오용 가능성을 직시하고 빅테크 기업에 대한 규제 거버넌스를 구축하여야 한다.

주제어: 인공지능, 알고리즘, 데이터, 인권, 민주주의, 감시

< 차례 >

- I. 문제설정
- II. 인공지능이 인권과 민주주의에 미치는 영향
- III. 우리나라 인공지능 법안의 내용과 문제점
- IV. 나오며

I. 문제설정

인공지능 기술의 기하급수적 발전과 함께 그 잠재적 위험성을 경고하는 목소리도 높아지고 있다. 대표적인 예가 최근 제프리 힌튼(Geoffrey

Hinton) 교수의 행보다.¹⁾ 그는 자신이 일생을 바친 연구를 후회하며 인공지능의 잠재적 위험성을 경고한다.

그가 지적하는 위험성은 다음과 같다. 통제되지 않는 빅테크에 의해 개발되는 인공지능 기술로 인해 가짜 사진, 동영상, 텍스트가 인터넷에 넘쳐나 일반적인 사람들은 더 이상 무엇이 진실인지 알 수 없게 될 것이며, 고용시장에서 반복적 업무를 처리하는 사람들(법률보조원, 개인비서, 번역가 등)은 기계에 의해 대체된다는 것이다. 또한 인공지능이 분석하는 방대한 양의 데이터에서 예상치 못한 행동을 학습하는 경우가 많다는 점, 인공지능 시스템이 자체 컴퓨터 코드를 생성하는 것을 넘어 스스로 그러한 코드를 실행하게 될 수 있다는 점, 킬러 로봇과 같은 자율살상무기가 현실화될 수 있다는 점 등도 지적하고 있다.

하지만 디지털 기술은 한때 ‘해방의 기술’(Liberation Technology)로 여겨지기도 했다.²⁾ 튀니지, 이집트, 레바논 등 아랍의 봄에서 트위터, 페이스북 등 소셜 미디어는 일종의 공론장이었고 봉기의 매개였다.³⁾ 디지털

1) New York Times, 2023. 5. 1., “‘The godfather of AI’ Leaves Google and warns of danger ahead”, <<https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html>>, 검색일: 2023. 5. 3. 힌튼 교수는 인공지능경망을 개발한 인공지능의 개척자 중 하나로 알려진 인물이지만, 최근 10년 이상 근무했던 구글에서 퇴사하여 화제가 되었다. 한편, 최근 미국의 생명미래연구소는 공개서한을 통해 인공지능 개발을 6개월 이상 중단해야 한다고 주장하기도 했다. 인공지능을 개발하는 연구자들조차 통제할 수 없는 인공지능 기술이 경쟁적으로 개발되고 있다는 것이다. 규제 없는 인공지능 개발이 계속 되면 통제 불가능한 상황에 직면할 수 있다는 경고이다. 이 서한에는 일론 머스크, 요슈아 벤지오, 유발 하라리 등 유명 인사들이 참여하여 화제가 되기도 했다. 힌튼 교수는 이 서명에 참여하지 않았지만, 이제 인공지능의 위험성을 비판하는 흐름에 합류하게 되었다.

2) Larry Diamond, “Liberation Technology”, *Journal of Democracy*, vol. 21, no. 3(2010), 69-83쪽.

3) 인남식, “아랍 민주화 운동과 미국의 대중동정책 변화 연구”, 서정민/인남식 엮음, *중동 민주화의 대내외 정치역학(대외경제정책연구원, 2011)*, 110쪽. “페이스북, 트위터, 유튜브 등 소위 소셜네트워크서비스(SNS)가 확산되면서 아랍 대중들끼리 소통이 가능해졌고, 시위로 이어지는 촉매제가 된 것이다. SNS의 파급효과는 상당히 컸다. 특히 튀니지의 채소 행상 모하메드 부아지지의 분신 장면이 유튜브에 실리고 페이스북으로 옮겨가면서 아랍권 대부분의 젊은이들을 자극하게 된다. 실업 상태의 확산과 곡물가의 상승 그리고 정치적 패배주의 등

기술은 대중들의 소통을 촉진하여 정치적 변혁을 이끌 수 있는 원동력이 되기도 했다.⁴⁾

기술 발전은 긍정적으로든 또는 부정적으로든 사회적 가치와 원칙에 영향을 미치며 새로운 국면의 변화를 여는 전환점이 되어 왔다.⁵⁾ 최근 혁신을 거듭하고 있는 인공지능 기술도 이러한 새로운 국면을 여는 사회적 변곡점이 될 것이라는 예상이 지배적이다.⁶⁾ 인공지능으로 통칭되는 알고리즘과 데이터 기술은 방대한 양의 데이터를 처리하는 새로운 방식이다.⁷⁾ 정치권에서도 ‘디지털 시대의 쌀은 데이터’라며 산업적 활용방안을 앞다투어 제시하고 있다. 이른바 디지털 경제에서 대량의 데이터와 정

구조적 요인이 이러한 촉진 요인을 통해 사회를 숙성시키고 결국 변혁의 단계에 이르게 된다.”

- 4) 2010년 이집트 혁명에서 와엘 고넴(당시 구글 중동 지역 마케팅 담당 이사)은 “만약 사회를 해방시키고 싶다면, 인터넷만 있으면 된다”며 혁명의 단초를 제공하는 활동을 하였다. 그러나 5년 이후, 정치적 양극화를 초래하며 인터넷 그 자체가 권력이 되어 가는 현상을 겪으며, “사회를 해방시키고 싶다면, 우리는 먼저 인터넷을 해방시켜야 한다”는 말을 남겼다. 박승일, 기계, 권력, 사회: 인터넷은 어떻게 권력이 되었는가(사월의책, 2021), 37쪽 이하 참조.
- 5) 제프리 삭스는 경제체제는 어느 시간, 어느 장소가 되었든 지리, 기술, 제도에라는 세 가지 조건의 상호작용에 의존한다고 하면서, “지리의 경제적 중요성은 변화하는 지식과 기술에 의해 끊임없이 다르게 규정되어왔다”라고 지적한다. 제프리 삭스, 이중인 율김, 지리, 기술, 제도(21세기북스, 2021), 49, 52쪽.
- 6) “우리[는] 디지털 기술에 힘입어 경이로운 발전을 거듭하는 시대에 살고 있다. … 한 마디로, 우리는 변곡점에, 컴퓨터 때문에 궤도가 크게 구부러지는 지점에 와 있다. 제2의 기계 시대로 진입하고 있는 것이다.” 에릭 브린올프슨, 앤드루 맥아피, 이한음 율김, 제2의 기계시대(청림출판, 2014), 15쪽. 최근 헨리 키신저 등이 펴낸 저서에도 인공지능의 근본적 특성이 지적되고 있다. “인류의 역사는 기술의 변천사이기도 하다. 그러나 기술로 인해 사회정치적 구조가 근본적으로 바뀐 사례는 거의 없었다. … 하지만 AI는 인간 경험의 모든 영역에서 변화를 예고한다. 그 변화의 중심에는 인간이 현실을 이해하는 방식, 그리고 그 안에서 자신이 맡은 역할을 이해하는 방식을 바꿔놓는 철학적 전환이 있을 것이다”. 헨리 키신저/에릭 슈밋/대니얼 허튼로키, 김고명 율김, AI 이후의 세계(월북, 2023), 53쪽.
- 7) 대표적으로 Sue Newell/Marco Marabelli, “Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of ‘Datification’”, *Journal of Strategic Information Systems*, vol. 24, no. 1(2015), 3쪽.

보는 알고리즘 등의 가공을 통해 가치를 창출할 수 있는 자산으로 인식된다.⁸⁾ 인공지능은 인간이 세계를 경험하는 방식을 근본적으로 변화시키고 있으며, 정치적·경제적·사회적 구조를 획기적으로 전환할 것이다.

인공지능은 시민들의 권리와 자유에 관한 역량을 높인다는 측면에서 긍정적 영향을 미치기도 하지만, 새로운 헌법적·법률적 문제를 제기하기도 한다. 최근의 인공지능은 전대미문의 기술이다. 대량의 정보처리, 고도의 계산과 예측, 변화하는 상황에 대한 학습과 조정적 반응, 사물 인식 및 분류와 같은 복잡한 작업을 수행할 수 있는 자율 시스템을 인류 역사상 처음으로 다루게 된 것이다. ‘알고리즘 지배’(Algocracy),⁹⁾ ‘알고리즘 사회’(Algorithmic Society)¹⁰⁾라는 조어들이 의미하는 바처럼, 인공지능은 정치, 경제, 사회, 법의 새로운 문제계이다. 다양한 문제들이 등장하고 있다. 알고리즘 의사 결정 시스템(Algorithmic Decision Systems: ADS)을 통해 아동복지, 형사사법, 학교 배정, 교사 평가, 화재 위험평가, 노숙자 주거 우선순위 지정, 건강보험, 출입국 관리 및 입국시 위험평가, 예측치안 등 광범위한 분야에서 인공지능이 의사 결정에 관여하게 될 것이다.¹¹⁾

이 글은 정치적·법적 문제의식에 한정하여, 현재 기술 수준에서 관찰되고 예측되는 범위 안에서 인공지능이 인권과 민주주의에 미치는 영향

8) 빅토르 마이어 윈버저/케니스 쿠키어, 이지연 옮김, 빅 데이터가 만드는 세상 (21세기북스, 2013), 17쪽. “사람들은 더 이상 데이터를 유통기한이 지난 고정물로 생각하지 않게 되었다. 이전에는 비행기가 착륙하고 나면(혹은 구글에서 검색어가 처리되고 나면), 수집된 데이터는 애초의 목적을 달성했으므로 그 유용성이 끝났다고 생각했다. 하지만 이제 데이터는 비즈니스의 원자재가 되었다. 빼놓을 수 없는 경제 인풋(Input)으로서 새로운 형태의 경제적 가치를 창출하는 원료가 된 것이다”.

9) John Danaher, “The Threat of Algocracy: Reality, Resistance and Accommodation”, *Philosophy & Technology*, vol. 29, no. 3(2016), 245쪽.

10) Agnieszka M. Walorska, “The Algorithmic Society”, Denise Feldner 엮음, *Redesigning Organizations Concepts for the Connected Society*(Springer, 2020), 149-160쪽.

11) Céline Castets-Renard, “Human Rights and Algorithmic Impact Assessment for Predictive Policing”, Hans-W. Micklitz 외 엮음, *Constitutional Challenges In The Algorithmic Society*(Cambridge University Press, 2022), 93쪽.

을 살펴보고, 이에 대응할 수 있는 제도적 장치를 규제 거버넌스의 관점에서 논의하고자 한다. 특히 최근 국회에서 논의되고 있는 인공지능 법안의 내용과 문제점을 살펴보고자 한다.

II. 인공지능이 인권과 민주주의에 미치는 영향

1. 인공지능의 의미와 영향

인공지능은 매우 다양하게 정의된다.¹²⁾ 공학기술의 관점에서 인공지능이란, “지능적인 기계를 만드는 과학과 공학기술”(John McCathy), “인간이 수행한다면 지능이 필요한 일을 하는 기계를 만드는 과학”(Marvin Minsky), “주어진 상황에서 가능한 최선의 행동을 취하는 지능적 에이전트를 구축하는 문제”¹³⁾라 할 수 있다. 우리 법제에서 인공지능을 적시하여 정의하고 있는 경우는 아직 없으며,¹⁴⁾ 최근 논의되고 있는 인공지능 법안에서는 인공지능을 “학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것”으로 정의하고 있다.¹⁵⁾

인공지능이 인간과 사회에 미치는 영향은 양가적이다. 긍정적 영향과 부정적 영향 모두 가지고 있을 것이다. 반복적 노동을 대체하고, 증거에 기반한 신속한 의사 결정을 지원하며, 예측 시스템을 통해 각종 사회적 위험을 예방할 수 있다는 점은 긍정적 측면일 것이다.

12) 김진우, 나의 첫 인공지능 수업(메이트북스, 2022), 18쪽.

13) 스투어트 러셀/피터 노빅, 류광 옮김, 인공지능: 현대적 접근방식, 제3판(제이펍, 2016), 2쪽 이하.

14) 다만, 행정기본법 제20조는 자동적 처분과 관련하여, 인공지능 기술을 적용한 시스템을 포함한 ‘완전히 자동화된 시스템’으로 처분을 할 수 있다고 규정하고 있다.

15) 인공지능산업 육성 및 신뢰 확보에 관한 법률안(윤두현의원 대표발의, 의안번호: 18726) 제2조 제1호. 이 법안은 “인공지능 기술”을 “인공지능을 구현하기 위하여 필요한 하드웨어 기술 또는 그것을 시스템적으로 지원하는 소프트웨어 기술 또는 그 활용 기술”로 정의하며, 그 외에도 “고위험 영역에서 활용되는 인공지능”, “인공지능 윤리”, “인공지능 산업” 등을 정의하고 있다.

하지만 인공지능 사회에서 인공지능 시스템이 갖는 부정적 영향과 잠재적 위험성은 인권과 민주주의에도 치명적이다. 이하에서는 인권과 민주주의의 측면에서 인공지능이 갖는 부정적 영향과 위험성을 살펴본다.

2. 인공지능 시스템에 의한 인권의 침해 가능성과 부정적 영향

인공지능 시스템은 인권과 기본권 목록의 거의 모든 권리에 침해가능성을 갖는다. 인공지능은 완벽한 기술이 아니다. 품질이 낮은 데이터셋을 학습한 예측 모델은 편향된 결과나 부정확한 결과를 도출하여 차별적 결과를 초래한다. 예를 들어, 챗봇 이루다의 혐오 발언 사건, 경비로봇의 유아 공격 사건, 챗GPT의 답변 오류나 환각(Hallucination) 문제 등을 떠올려 볼 수 있다.

아래 [표]는 인공지능 시스템에 따라 침해 가능성이 있는 인권 목록을 예시한 것이다.¹⁶⁾

[표] 인공지능 시스템에 의해 침해 가능성이 있는 인권

인공지능 시스템 예시	침해될 가능성이 있는 인권
노인과 장애인 등 대상자의 맥박, 혈당, 활동 등을 감지하고 말뚝, 인지기능을 지원하는 돌봄로봇	개인정보자기결정권 침해
얼굴인식에 기반한 출입국 자동화 시스템	인종, 국가 등에 따른 차별 개인정보자기결정권 침해
지역별로 범죄 발생 확률을 예측하여 순찰 인력을 배치하는 인공지능 범죄 예측 시스템	인종, 지역 등에 따른 차별
인공지능 채용(면접) 시스템	성별, 연령, 장애, 용모, 출신지역 등에 따른 차별
공공 장소에서의 행인의 얼굴을 인식하여 용의자와 대조하는 원격 얼굴인식 시스템	이동의 자유 침해, 집회 및 결사의 자유 침해, 자의적 체포
아동이 사용하는 소셜네트워크서비스에서 선정적이고 자극적인 콘텐츠가 우선 노출되도록 하는 알고리즘	아동의 권리 침해 개인정보자기결정권 침해

¹⁶⁾ 아래 [표]는 유승익 외, 인공지능 인권영향평가 도입 방안 연구(국가인권위원회, 2022), 214쪽. 이 표에서 해당 인공지능 시스템이 반드시 인권 침해적이라는 의미는 아니며, 해당 인권과의 관련성만을 나타낼 뿐이다. 또한 해당 목록은 인공지능 시스템에 의한 인권 침해 가능성의 극히 일부 사례일 뿐이다.

인공지능 시스템 예시	침해될 가능성이 있는 인권
소셜네트워크 플랫폼에서의 알고리즘 기반 콘텐츠 관리 시스템	표현의 자유 침해 정보접근권 침해
소셜네트워크 플랫폼에서 개인의 정치 성향에 기반한 정치광고 노출 시스템	자유로운 정치참여 제한 선거권 침해
고등학교의 기존 성적에 기반한 인공지능 대학입학 시스템	지역에 따른 차별 교육권 침해
사업장 내에 설치된 생체인식, 위치추적 시스템	노동자 개인정보자기결정권 및 노동3권 침해
인공지능 판결 지원 시스템	공정한 재판을 받을 권리 침해
인공지능을 통한 사회보장급여 부정수급 탐지시스템	사회보장수급권, 장애인권리 침해, 인종 및 장애 등에 따른 차별

이상에서 살펴본 바와 같이, 알고리즘의 편향되고 잘못된 결정에 따라 표현의 자유, 평등권, 정치적 기본권, 개인정보자기결정권 등이 침해될 수 있다. 이 중에서도 표현의 자유와 개인정보자기결정권은 정치적 의사형성 과정에서 중요한 기본권이라는 점에서 특별히 강조될 필요가 있다. 예를 들어 인공지능 시스템이 적용된 소셜 미디어 등에서 의견을 자유롭게 표현할 수 없거나 ‘사적 검열’과 같은 형태의 관리가 별다른 통제 없이 이루어진다면 민주적 의사형성에 장애가 된다.¹⁷⁾ 또한 개인정보자기결정권이 제약되어 개인의 데이터 처리를 규율하는 규칙이 마련되어 있지 않다면 개인은 책임성과 투명성이라는 보호장치 없이 사적 감시체제에 노출될 수밖에 없다.¹⁸⁾

이러한 의미에서 ‘콘텐츠 관리’와 ‘사용자 프로파일링’은 인공지능 시스템이 기본권과 민주주의에 미치는 위험성을 예시해 주는 사례라 할 수 있다.

콘텐츠 관리(Content moderation)는 혐오 표현이나 허위 정보, 거짓 주장들에 대응하기 위해 개별 기술 기업들이 “내부 ‘관리자(moderator)’를

17) 홍남희, “디지털 플랫폼에 의한 ‘사적 검열(private censorship)’”, 미디어와 인권 권 통권 제6호(2018), 135쪽 이하.

18) Oreste Pollicino/Giovanni De Gregorio, “Constitutional Law in the Algorithmic Society”, Hans-W. Micklitz 외 엮음, *Constitutional Challenges In The Algorithmic Society*(Cambridge University Press, 2022), 7쪽.

고용하여 개별 콘텐츠를 확인하고 플랫폼이 정한 규칙을 위반하는 게시물을 삭제”¹⁹⁾하는 관리 방식을 말한다.²⁰⁾

사용자 프로파일링(Users' profiling)이란, “다양한 방법으로 수집된 데이터를 분석하여 개인 또는 개인 그룹에 대한 새로운 특성 또는 행태 정보를 생성하고 적용하는 등의 작업 일체를 의미하는 것”을 말한다.²¹⁾ 주요 사례로 이용자 행동 정보 분석을 통해 맞춤형 온라인 광고를 제공하는 것, 검색 엔진 사업자들이 이용자의 위치, 과거 검색 기록, 다른 이용자와의 관계 등의 분석을 통해 검색 결과를 개인화(personalization)함으로써 이용자의 검색 의도에 부합하고 연관성 있는 결과를 보여주는 것, 개인 신용평가 결과를 이자율 결정 등 금융거래에 활용하는 것 등을 들 수 있다.

특히 콘텐츠 관리는 다양한 관점에서 문제를 지적할 수 있다. ① 빅테크 등 일반 사기업이 디지털 환경에서 ‘사적’ 보호 기준을 설정함으로써 기본권 보호 여부를 결정한다. 사적 ‘검열’ 권한을 인정하는 꼴이다. ② 사적 판단을 통해 콘텐츠를 관리함으로써 공적 기준과 사적 기준의 경계를 모호하게 하고, 예측가능성이나 법적 안정성의 문제를 제기한다. ③ 디지털 플랫폼의 콘텐츠 관리는 투명성과 책임성이 결여될 수밖에 없다. 표현물의 삭제 여부를 결정하는 디지털 플랫폼 기업은 정부나 공공기관처럼 헌법상 표현의 자유에 요구되는 엄격한 심사요건을 따를 의무가 없으며, 사업 목적에 따라(대개는 상업적 목적) 자유롭게 평가하여 삭제할 수 있다.²²⁾ 표현의 자유의 보장 여부가 디지털 플랫폼의 사적 기준에 의

19) Melissa Heikkilä, “사회를 오염시키는 소셜미디어, 콘텐츠 관리만으로는 해결할 수 없다”, <<https://www.technologyreview.kr/social-media-polluting-society-moderation-alone-wont-fix-the-problem/>>, 검색일: 2023. 5. 1.

20) 콘텐츠 관리와 반대성명의 한계 및 설계의 문제를 지적하는 다음 글을 참조 B its of Freedom, “‘혐오표현’에 대항하기: ‘설계’는 ‘콘텐츠 관리’로부터 우리를 구할 수 있을까?”, <<https://act.jinbo.net/wp/39075/>>, 검색일: 2023. 5. 1. 이에 따르면, 콘텐츠 관리는 디지털 플랫폼에게 과도한 검열 권한을 부여하고, 기업들로 하여금 무엇이 옳고 그른지 결정하게 하는 민영화된 법집행으로 실질적인 법 집행을 대체하여, 표현의 자유를 더욱 취약하게 할 것이다.

21) 방송통신위원회/한국인터넷진흥원, 개인정보 처리에서의 프로파일링 사례집 (2020), 3쪽.

22) 이 문제를 자세히 서술하고 있는 룽 라이히 외, 이영래 옮김, 시스템 에러: 빅

존하게 되는 것이다.

같은 맥락에서 온라인상의 허위 정보(disinformation) 문제는 법적·정치적으로 매우 민감한 문제를 제기하기도 한다. 미국 대선 당시의 ‘피자게이트’나 영국의 브렉시트 국민투표의 사례에서 볼 수 있었던 것처럼, 정치적 의사형성에서 허위 정보의 파괴적 영향력은 개인의 프라이버시 침해에 한정되지 않는다. 이와 관련하여 허위 정보에 대한 규제수단의 헌법적 한계를 확정하는 문제, 허위 정보의 한계에 관한 기준을 설정하고 이를 관리하는 데 있어서 인공지능 시스템을 활용하는 문제와 같은 새로운 쟁점들이 등장한다. 아직 이러한 쟁점에 대한 헌법적 한계가 불분명하다.

현재의 수준에서 보자면, 디지털 플랫폼 기업은 (허위)정보의 양을 사회적 기준에 의해 ‘조절’할 수 있고, 이를 통해 정치적 지형을 형성할 수도 있다. 전 세계적으로 관찰되는 포퓰리즘은 상업적 콘텐츠 관리의 ‘완화되고 조절된’ 허위 정보를 유통시키기 위해 디지털 플랫폼을 적극 활용한다. 이탈리아의 오성운동(M5S)의 성공이 대표적인 예일 것이다.²³⁾

이렇듯 통제되지 않는 인공지능 시스템은 개인의 인권을 취약하게 할 뿐만 아니라 정치적 공동체의 운명을 민간기업의 처분에 맡길 수 있는 위험도 안고 있다.

테크 시대의 윤리학(어크로스, 2022), 307쪽 이하.

- 23) 정병기, “오성운동(M5S)의 직접 의회주의와 사이버크래틱 집중주의: 포스트포폴리스트 정치 운동의 성공과 한계”, 한국정치연구 제29권 제2호(2020), 91-116쪽. 오성운동은 정치풍자로 공영TV 출연이 금지됐던 코미디언 그릴로(Beppe Grillo)가 웹 전략가인 카사레쵸(Gianroberto Casaleggio)와 함께 2009년 창당한 정당이다. 웹 2.0 기반으로 활동하면서 인터넷 민주주의를 추구하는 정당으로 성장했고, 2013년 총선에서 25%의 지지율을 얻어 제1당이 된 후, 2018년 선거에서는 집권까지 하였다. 오성운동은 그릴로의 블로그, 온라인 플랫폼 모임인 미팅(Meetup), 당원들의 최종 의사 결정을 위해 활동하는 공간인 ‘루소’ 등 디지털 플랫폼을 적극 활용한다. 이 플랫폼들은 카사레쵸의 웹 회사가 독자적으로 제작 관리하고 있는데, 그 불투명성, 당원 제한, 리더십 집중 현상 등의 문제가 지적되고 있다.

3. 감시사회의 고도화와 분화 : 전통적 감시 권력에서 빅데이터 감시 권력으로

인공지능 사회의 출현을 통해 등장하는 가장 민감하고 중요한 문제 중의 하나는 감시사회의 성격이 고도화, 심화, 확장되고 있다는 점이다.²⁴⁾ 인공지능이 보편화된 사회는 과거 ‘정보사회’로 지칭된 사회가 진화된 형태이며,²⁵⁾ 새로운 형태의 ‘대중감시’(mass-surveillance)가 가능해지는 사회이다.²⁶⁾ 과거 푸코식의 판옵티콘 규율 체계와 일방향·전방위 시각감시 중심의 감시사회에서 “유동형(liquidity-driven) 감시와 네트워크형(network-driven) 알고리즘 감시가 공모하는 형태로 진화”하고 있다는 것이다.²⁷⁾

사실 감시사회는 새로운 주제가 아니다. 한국 현대사에서 만연했던 사찰과 정보정치부터 CCTV와 전자주민증까지 다양한 감시기법과 기술의 위험성에 대한 인식은 꾸준히 높아져 왔다.²⁸⁾ 하지만 팬데믹을 거치며 사회적 거리두기나 감염병 의심자에 대한 통제정책에 관한 효과적인 도구로서 인공지능 기술은 한층 더 각광받고 있다.²⁹⁾ 이제 인공지능 기술은 자유의 적이라기보다 보건과 건강을 지키는 안전막처럼 인식되고 있

24) 이광석, “포스트-판옵티콘 시대 감시 연구, 새로운 지형”, 김동욱 외 엮음, 스마트 시대의 위험과 대응방안(나남출판, 2015), 192쪽 이하.

25) 프랭크 웹스터, 조동기 옮김, 현대 정보사회 이론(나남출판, 2016).

26) Neil M. Richards, “The Dangers of Surveillance”, *Harvard Law Review*, vol. 126, no. 7(2013), 1934쪽.

27) 이광석, “포스트-판옵티콘 시대 감시 연구, 새로운 지형”, 192쪽. 그리고 이 글에서 주로 참고하고 있는 지그문트 바우만/데이비드 라이언, 한길석 옮김, 친에 하는 빅브라더(오월의봄, 2014)도 참조.

28) 대표적으로 한홍구 외, 감시사회(칠수와영희, 2021), 11쪽 이하.

29) 코로나 확진자 동선 파악은 카드 사용내역, 스마트폰 GPS, 주변 CCTV를 통해 이루어졌다. 이 중에서 CCTV는 단순히 영상녹화 및 저장기능에 머무르지 않고, 얼굴인식, 자동차 번호판 인식, 영상분석 추적 등이 가능한 ‘지능형 CCTV’로 발전했다. 다른 CCTV에 촬영된 영상 속 사람들과 비교해 동선을 파악할 수 있다. ICT 표준에 의한 CCTV호환이 가능해졌고, 딥러닝, 빅데이터 기술이 이를 뒷받침한다. CCTV 관제센터, 경찰서, 소방서가 연결되어 통합·관리된다. 한국정보통신기술협회, 스마트 시티 정보의 통합 관리 및 운영을 위한 플랫폼 소프트웨어 요구사항(TTAK.KO-10.1118) 참조.

다.³⁰⁾ 감시기술과 감시문화는 공공과 민간을 가로지르며 별다른 저항 없이 수용되고 있다. 이를 “감시 자본주의”라 부르기도 한다.³¹⁾

그런데 이렇게 고도화된 감시사회는 인권의 관점에서 미묘한 쟁점을 제기한다. 팬데믹 시대를 거치며 확인한 것처럼 고도화된 감시기술이 사회에 순기능을 하며 시민들의 삶을 개선하는 측면도 분명 존재한다는 것이다. 순응과 자발적 참여를 통해 감시 권력이 인권 주체의 대립적 객체로 인식되기보다 생명, 건강, 재산, 자유의 인프라 구조처럼 자리매김한다. 안전, 보호, 편의성, 효율성 등의 혜택을 위해 개인 데이터가 기록, 저장, 복구, 이동, 교차, 교환되는 것에 동의하는 것을 당연하게 받아들인다. “숨길 이유가 뭔가”(“nothing to hide”)하는 감시 이데올로기 자체가 뉴노멀로 되고, 감시에 의문을 제기할 역량 자체가 무력화된다.³²⁾

시민들의 통제에서 벗어난 기술을 통해 구현되는 이러한 새로운 감시사회에서 감시질서는 푸코적 의미에서 정상화된다.³³⁾ ‘정상적’ 감시질서 하에서 장기적으로 시민들은 새로운 형태의 감시를 준수하고 순응하며 당연시하게 된다. 이때 개인이 갖는 프라이버시권과 개인정보자기결정권은 “데이터 조각”으로 파편화되어 비인격화되고,³⁴⁾ 개인(Individual)으로서 법적 권능은 저하하고 분할되어 무화된다.

이러한 새로운 감시사회에서 정부와 민간 기업 모두에 의해 대규모 감시가 이루어지게 된다. 경우에 따라서 개인 상호 간 감시도 일상화된다. 이 모든 시나리오를 종합할 때, 개인의 사생활은 지속적이고 침습적인 감시의 대상으로 격하되고, 데이터에 대한 권리주체의 통제력은 상실된다.

30) Natalie Ram/David Gray, “Mass Surveillance in the Age of COVID-19”, *Journal of Law and the Biosciences*, vol. 7, no. 1(2020), 1쪽.

31) 쇼샤나 주보프, 김보영 옮김, 감시 자본주의 시대(문학사상, 2021).

32) Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*(W. W. Norton & Company, 2015).

33) David Lyon, *Surveillance After September 11*(Polity, 2003), 150쪽.

34) 이광석, “포스트-판옵티콘 시대 감시 연구, 새로운 지형”, 197쪽. “개인데이터의 경우에는, 가치가 추출될 수 있도록 인터넷 이용자들이 뒤에 남기는 무수한 클릭과 네트상의 동선과 흔적들, ‘데이터 배출’(data exhaust)이 빅데이터의 핵심이 된다. 이는 이용자들이 남긴 데이터 부스러기, 즉 ‘데이터 조각’으로 불리기도 한다.”

4. 자유의 의미변화: 축소와 왜곡

인공지능 사회는 자유의 핵심적 내용도 변형시킨다. 인공지능은 특정 개인이나 집단을 표적으로 자유를 직접적으로 실현하거나 제한하는 것이 아니라 대량의 데이터를 모집단으로 패턴을 인식하여 자율적으로 결정을 내릴 뿐이다. 개인이나 집단이 설정한 목적이나 내용, 개인정보의 개별적 의미나 인과적 관계(causal link)는 중요성을 상실하고 상관관계(correlation)와 예측성으로 대체된다. 기계는 인간을 위해 또는 인간을 대신하여 데이터의 상관관계에 기초하여 예측적으로 결정을 내린다. 여기에서 인간의 자유는 침해된다고 말할 수도 있지만, 엄밀히 말해 자유 그 자체가 기계적으로 환원되고 포획된다는 표현이 더 정확할 것이다.

인간 자유를 포획(capture)하는 방식은 두 가지이다. 하나는 인공지능 시스템으로 하여금 인간을 대신해서 직접 결정하도록 위임하는 경우이다. 마치 대의제하에서 자유위임의 원리와 같이 인간은 의사 결정 시스템이 자율적으로 결정할 수 있도록 위임한다. 인간 에이전트로서 이러한 위임 결정의 모멘트에만 자유를 행사하고 이후 자기 구속된다.

다른 하나는 인간이 특정한 행위나 결정을 할 때 필요한 정보를 인공지능 시스템이 제공하도록 하는 경우이다. 배경지식을 자동화된 기계가 독점 또는 과점하도록 하는 방식이다. 인간은 지식과 정보를 통해 자유롭게 결정하고 행동한다. 정보원은 인간의 결정에 영향을 미치는 핵심 요소 중의 하나이다. 어떤 정보를 취득하느냐에 따라 어떤 결정과 자유를 행사하는지 달라진다. 예를 들어, 국회의원 선거에서 특정 후보 선택은 해당 후보에 대한 이력, 정책, 여론, 평가 등의 정보에 따라 달라진다. 오늘날 직접 경험 외에 정보를 얻기 위해 이용하는 채널은 압도적으로 스마트폰이다.³⁵⁾

또한 자동화된 기술 시스템을 통해 대량으로 유통되는 정보를 일괄적으로 소비하게 됨으로써 개인이나 집단이 자신의 정체성을 과거 대면공동체나 소속 집단의 공유된 정보에 의해 구획적으로 형성하던 방식에서

35) 2022년 기준 신문구독 이용매체 현황에 따르면, 스마트폰 애플리케이션 80.2%, 스마트폰 웹브라우저 11.2%로 91.5%가 스마트폰으로 신문을 구독하고 있다. 그 외에 PC 4.8%, 종이신문 3.8% 순이다. 미디어통계포털(Kisdistat), <<https://stat.kisdi.re.kr/>>, 검색일: 2023. 5. 1.

탈피하게 되었다. 대규모로 집적된 데이터에서 알고리즘에 의해 상관관계적으로 큐레이트된 지식의 흐름에 따라 개인이나 집단의 정체성이 유동적으로 “유도”된다.³⁶⁾

이러한 인공지능 사회에서 전통적인 의미에서의 자유의 의미는 축소되고 변형된다. 첫째, 인공지능 시스템으로 하여금 대신 결정하도록 ‘동의’한 경우, 시스템이 인간의 자유를 제한하는 결정을 내렸을 때 그것을 권리 주체의 자유에 대한 제한으로 평가할 수 없게 된다. 이미 인간 에이전트는 해당 시스템이 인간의 결정보다 더 증거에 기반하고 있으며 과학적, 효과적, 중립적인 것이라는 가정하에 결정을 위임했다. 명시적으로 또는 묵시적으로 인공지능 시스템에게 의사 결정을 위임했다는 점에서, 인간의 의지나 자기 결정에 반하여 자유가 침해되었다고 말하기 힘들어진다.³⁷⁾ 인공지능 시스템을 활용한 통치는 시스템 도입 단계에서 사용자가 이용약관을 동의함으로써 저항 없이 완성된다.

둘째, 인공지능 기술이 제공하는 정보를 통해 결정을 내릴 때, 자유에 대한 제약은 더 이상 외재적인 것이 아니게 된다. 인간은 스스로 인공지능에 의해 제공된 정보를 신뢰하고 이용할 뿐이다. 여기에 외부적 강제나 폭력이 개입되지 않는다. 자신이 가진 스마트폰이나 PC를 통해 자유롭게 뉴스 플랫폼을 선택하여 이용할 뿐이다. 이러한 정보들을 믿을 만하고 유효한 근거로 스스로 받아들일 뿐이다. 물론 이러한 정보 수용은 인터넷 플랫폼에서 제공하는 웹상의 정보들이 과학적으로 정확하거나 신뢰할 수 있다는 가정에 기초한다.

그러나 이러한 정보에 대한 신뢰는 인과성에 기초한 과학이나 객관적 검증과는 전혀 관계가 없다. 플랫폼은 다만 기존 파라미터에서 신호를 추출함으로써 행동 패턴을 학습하는 알고리즘을 사용하여 개인과 집단이 갖는 기존의 신념을 반복적으로 강화할 뿐이다. 알고리즘이 객관성이나

³⁶⁾ Holger Pötzsch, “Archives and Identity in the Context of Social Media and Algorithmic Analytics: Towards an Understanding of iArchive and Predictive Retention”, *New Media & Society*, vol. 20, no. 9(2018), 3304쪽.

³⁷⁾ Andrea Simoncini/Erik Longo, “Fundamental Rights and the Rule of Law in the Algorithmic Society”, in Hans-W. Micklitz 외 엮음, *Constitutional Challenges In The Algorithmic Society*(Cambridge University Press, 2022), 40쪽.

정확성을 보증한다는 것은 확증편향에 의해 생성되는 기존의 신념을 확률적으로 반복함을 의미할 뿐이다. 사용자는 관심 있는 커뮤니티별로 모여 해당 커뮤니티의 기존 선호도를 학습한 알고리즘에 의해 추출된 정보를 ‘객관적’ 정보로 습득하게 된다. 전체적으로 보면, 집단양극화와 분리, 통약불가능성을 초래한다. 결국 자유는 외재적으로 제약되는 것이 아니라 내재적으로 왜곡된다.

5. 사적 권력의 문제

인공지능 사회에서 빅테크 기업이나 디지털 플랫폼 기업 등 사적 행위자의 지위와 역할이 국가나 정부의 역할을 초월하면서 공사의 경계가 희석된다. 빅테크 기업은 대중들의 데이터를 통해 막대한 수익을 얻고 있지만, 별다른 대안 없이 빅테크가 구축한 디지털 플랫폼 서비스에 더 의존하게 된다(Lock-in 효과). 이러한 플랫폼은 애초 귀찮은 일에서 해방해주고 보다 창의적인 활동을 할 수 있도록 약속했지만, 역설적으로 디지털 플랫폼 자체가 불가결의 존재로 되어 가고 있다.

개인들은 점점 더 자신들의 의사 결정에 영향을 미치는 기술 시스템에 둘러싸이게 되지만, 이러한 현상을 이해하거나 통제할 가능성은 점점 더 줄어들게 된다. 결과적으로 민주적 의사 결정에 의식적으로 참여할 가능성도 축소될 수밖에 없다.³⁸⁾ 이러한 현상은 알고리즘 자체가 불투명성을 가지고 있는 것이기도 하지만, 다른 한편으로는 인공지능 시스템이 사적으로 개발되기 때문이기도 하다.

인공지능 사회에서 혁신의 주도세력은 기하급수적 성장을 거듭하면서 인터넷 관련 상품 및 서비스 시장을 주도하는 거대 빅테크 기업이다.³⁹⁾ 감시 자본주의 또는 “플랫폼 자본주의”는 이러한 빅테크 기업들이 압도

38) Oreste Pollicino/Giovanni De Gregorio, “Constitutional Law in the Algorithmic Society”, 5쪽.

39) 예를 들어, 스마트 기기(Apple, Samsung, Huawei, Xiaomi), 웹 검색 엔진(Google), 소셜 미디어 기업(Facebook, Instagram, Twitter), 클라우드 서비스 제공업체(Amazon, Microsoft, Google), 전자상거래 기업(Amazon, Netflix), 소셜 플랫폼(Zoom, Cisco Webex) 등이다. Andrea Simoncini/Erik Longo, “Fundamental Rights and the Rule of Law in the Algorithmic Society”, 32쪽.

적 기술과 자본으로 사회적 권력을 행사하는 새로운 시스템이다.⁴⁰⁾

시가총액의 약 25%를 차지하는 5대 빅테크 기업(애플, 아마존, 마이크로소프트, 페이스북, 구글(알파벳))을 ‘몰리고폴리’(moligopoly)라는 합성어로 정의하기도 한다.⁴¹⁾ 이러한 몰리고폴리 기업들은 사용자들의 거래를 통해 발생하는 네트워크 효과로부터 수익을 얻을 뿐만 아니라, 사실상의 정치적 영향력까지도 획득하고 있다. 그 가장 극단적인 예는 케임브리지 애널리티카(Cambridge Analytica)가 유권자 타겟팅을 통해 2016년 미국 대선과 영국의 브렉시트 국민투표에 개입한 사건일 것이다.⁴²⁾

인공지능 사회는 사적·사회적 기업 권력이 정치 권력과 병존하고 교차하는 사회이다. 이러한 새로운 권력의 출현은 헌법 전통에 도전하고 있으며, 이에 대응하여 헌법은 어떻게 진화해야 하는지에 관해 새롭고도 중대한 문제를 제기한다.

Ⅲ. 우리나라 인공지능 법안의 내용과 문제점

인공지능 사회는 인권과 민주주의의 가치에 다양한 위협요인이 되지만, 이를 예방하거나 완화한 제도적 장치는 충분히 마련되지 않은 실정이다.

인공지능 산업은 아직 혁신이 진행 중인 민간 주도 분야라는 점, 인공

40) 쇼사나 주보프, 김보영 옮김, 감시 자본주의 시대, 420쪽. “역사적으로 다른 어떤 시기에도 지금처럼 전례 없는 부와 권력을 지닌 민간 기업들이 돈으로 살 수 있는 모든 첨단 과학의 노하우로 전 지구적인 유비쿼터스 컴퓨팅 지식 및 통제 아키텍처를 구축하고 유지하며, 이를 토대로 행위의 경제에 입각한 사업 활동을 이처럼 자유롭게 영위한 적은 없었다.” 또한 주보프는 페이스북 등을 감시 자본가로 지칭하며, “사회적 권력과 통제권을 위한 투쟁이 상대해야 할 눈에 보이지 않는 적은 이제 계급이나 생산관계가 아니라 자동화된 행동 수정이다”라고 강한 어조로 비판한다.

41) Nicolas Petit, *Big Tech and the Digital Economy: The Moligopoly Scenario* (Oxford University Press, 2020); Oreste Pollicino/Giovanni De Gregorio, “Constitutional Law in the Algorithmic Society”, 32쪽 이하.

42) 이에 대하여는 시바 바이디야나단, 홍권희 옮김, 페이스북은 어떻게 우리를 단절시키고 민주주의를 훼손하는가(아라크네, 2020), 207쪽 이하. 또한 브리태니커, 고영태 옮김, 타겟티드(한빛비즈, 2020); 로저 맥나미, 김상현 옮김, 마크 저커버그의 배신(에이콘출판, 2020) 참조.

지능 기술에 대한 무정부적 탈규제는 인권과 민주주의의 본질적 내용과 근본원칙을 방기할 수 있다는 점, 새로운 인공지능 기술이 등장했을 때 법제도가 개입하면 너무 늦은 경우가 많다는 점 등이 종합적으로 고려되어야 한다.

따라서 사후적 책임 분배와 구제수단을 강구하는 사법적 조치만으로는 인공지능 시스템의 위험성을 예방하거나 완화할 수 없다. 인공지능 기술의 안전성을 확보하고 위험성을 완화하기 위해서는 기술의 형성과정을 역추적하여 적절한 시점에 개입하는 것이 중요하다. 따라서 보다 두터운 보호를 위해서는 규제를 통한 보호, ‘설계’에 의한 보호, ‘기본값’(default setting)에 의한 보호까지도 필요하다.⁴³⁾

세계 여러 나라는 인공지능 기술에 잠재하는 위험에 대처하기 위한 입법적 준비에 한창이다. 신뢰가능한 인공지능 기술의 발전을 위해 공정성, 윤리성, 투명성과 설명가능성, 안전성 등의 차원에서 통제와 규율을 강화하고 있다.

우리 국회도 인공지능 기술에 대응하는 입법에 속도를 내고 있다. 지난 2월 14일 국회 과학기술정보방송통신위원회의 법안소위는 ‘인공지능산업 육성 및 신뢰 기반 조성에 관한 법률안’을 통과시켰다. 이하에서는 이 법안을 둘러싼 주요 쟁점과 문제점을 검토하고 앞으로의 과제를 살펴본다.

1. 법안의 주요 내용

해당 인공지능법안은 2020년 7월부터 과방위에 발의된 인공지능 관련 법안 7개를 통합한 법안이다(대표발의 의원은 이상민, 양향자, 민형배, 정필모, 이용빈, 윤영찬, 윤두현 의원이다. 통합된 수정안은 아직 공개되어 있지 않다).⁴⁴⁾

⁴³⁾ Andrea Simoncini/Erik Longo, “Fundamental Rights and the Rule of Law in the Algorithmic Society”, 41쪽. 여기에서 더 나아가 설명가능성과 의사 결정에 대한 인간의 개입(Human-in-the-loop for decisions) 원칙 등의 구체적인 의미를 전달한다는 의미에서 “교육에 의한 보호”까지도 요구된다.

⁴⁴⁾ 심사의견 중, 해당 법안의 구체적인 조문은 적시하고 있지 않다는 의견이 있었다. 해당 법안은 2023. 2. 14. 과방위 법안심사소위를 통과하였다고 알려졌지만, 아직 병합·수정된 이른바 ‘소위원장 대안’을 공개하지 않고 있다(2023. 6. 기

이 법안은 인공지능산업을 진흥하고 인공지능사회의 신뢰기반 조성에 필요한 사항을 규정함으로써 국민의 권익과 존엄성을 보호하고 국민의 삶의 질 향상과 국가경쟁력을 강화하는 데 이바지함을 목적으로 한다고 밝히고 있다(윤두현의원안 제1조). 인공지능산업 육성 도모와 신뢰기반 조성으로 요약된다.

법안의 주요 내용은 다음과 같다.

① 인공지능산업 육성 및 신뢰 확보를 위한 추진체계로, 과학기술정보통신부(이하, ‘과기정통부’라 함) 장관은 3년마다 인공지능기술 및 인공지능산업의 진흥과 국가경쟁력 강화를 위하여 인공지능 기본계획을 수립·시행, 국무총리 소속으로 심의·의결기관인 인공지능위원회와 그 산하에 전문위원회(인공지능 신뢰성 전문위원회 포함) 설치, 지능정보사회진흥원 산하 국가인공지능센터 설치가 포함된다.

② 인공지능 기술개발 및 산업 육성을 위하여, 인공지능기술 및 알고리즘의 연구·개발 및 인공지능제품 및 서비스 출시 등에 대한 ‘우선허용·사후규제 원칙’, 인공지능기술 개발 및 안전한 이용 지원 사업 실시, 인공지능기술의 표준화, 인공지능 학습용데이터 관련 시책의 수립, 기업의 인공지능기술 도입·활용 지원, 창업 활성화, 인공지능 융합의 촉진, 과기정통부장관의 법령정비 등 제도개선 노력의무, 전문인력의 확보 시책 추진, 국제협력 및 해외시장 진출의 지원, 인공지능집적단지 지정, 대한인공지능협회의 설립 등이 포함된다.

③ 인공지능윤리 및 신뢰성 확보를 위하여, 정부는 인공지능사업자 및 이용자가 인공지능의 개발·이용과정에서 지켜야 할 인공지능 윤리원칙 제정·공표, 신뢰할 수 있는 인공지능 기반조성을 위한 시책 마련(과기정통부장관), 인공지능 신뢰성 검·인증 지원 사업 추진(과기정통부장관), 고위험 영역 인공지능의 확인, (제품 또는 서비스 제공자의 이용자에 대한) 고위험 영역 인공지능 고지 의무 등을 규정하였다.⁴⁵⁾

준). 국회는 본회의 상정 직전까지 이를 공개하지 않는데, 이 법안의 사례처럼 상당한 기간 입법정보의 공백을 방치하고 있는 꼴이다. 제도적 개선이 필요한 부분이다.

45) 기타 고위험 영역 인공지능과 관련한 사업자의 책무, 민간자율인공지능윤리위원

④ 기타 보칙으로 인공지능산업의 진흥을 위한 재원의 확충 방안 마련, 실태조사, 통계 및 지표의 작성·관리 및 공표, 비밀누설 등에 대한 벌칙 조항, 국가인공지능센터 또는 이와 유사한 명칭 사용한 자에 대한 과태료 부과 조항을 규정하고 있다.

2. 법안의 문제점

이 법안을 둘러싼 논의의 주요 쟁점은 이 법안이 과연 신뢰가능한 인공지능을 위한 제도 설계로 적합한 것인지, 인공지능 기술에 내재되어 있는 위험성을 통제하거나 규제할 수 있는 원칙과 수단을 제공하고 있는지에 모아진다.

첫째, 법안이 인공지능 기술 규제방식으로 채택하고 있는 “우선허용·사후규제 원칙”의 문제이다. 이른바 ‘포괄적 네거티브 규제방식’이다.⁴⁶⁾ 인공지능은 신생 기술이므로 산업기반 조성을 위해 규제의 강도를 완화하자는 취지로 이해된다. 이 법안은 인공지능기술, 제품 또는 서비스가 ‘국민의 생명·안전·권익에 위해가 되거나 공공의 안전 보장, 질서 유지 및 복리 증진을 현저히 저해할 우려가 있는 경우’가 아니면 제한할 수 없도록 함으로써 사전규제를 엄격히 제한한다. 현저한 위험이 사전에 확인되지 않는다면, 사전규제는 사실상 불가능하다.

그러나 이러한 “우선허용·사후규제 원칙”은 모든 산업에 적용할 수 있는 것이 아니다. 특히, 국민의 안전·생명·건강에 위해가 되거나 환경에 중대한 위협을 초래할 우려가 있는 경우, 개인정보의 안전한 보호 및 처리 여부가 문제되는 경우 네거티브 규제방식의 도입은 부적합하다.⁴⁷⁾ 인공지능 기술은 기본적 인권에 직접적이면서도 광범위하게 부정적 영향을 미칠 수 있다. 개인정보자기결정권을 비롯하여 표현의 자유, 프라이버시

회의 설치 등도 논의되고 있는 것으로 알려져 있다.

46) 행정규제기본법 제5조의2에 따르면, 국가나 지방자치단체가 신기술을 활용한 새로운 서비스 또는 제품과 관련된 규제를 법령등이나 조례·규칙에 규정할 때, 네거티브리스트, 포괄적 개념정의, 유연한 분류체계, 그리고 사후규제를 규정한다. 포괄적 네거티브 규제방식에 대하여는 정관선/박균성, “네거티브 규제의 재검토”, 법제 제699권(2022), 189쪽 이하.

47) 행정규제기본법 제19조의3 참조.

권, 평등권, 생명권, 안전권 등 거의 모든 인권 목록이 문제된다.⁴⁸⁾ 챗 GPT와 같은 인공지능 기술이 막대한 전력을 소모하여 기후환경을 파괴한다는 지적도 있다.⁴⁹⁾ 인공지능 기술에 잠재한 위험성을 고려한다면 이러한 사후규제방식은 적절한 규제수단이 아니다.

특기할 만한 점은 이러한 규제방식이 기존 ICT 법제의 규제 프레임에 따르고 있다는 것이다. 인공지능법안의 해당 규정은 정보통신융합법이나 지능정보화기본법 등의 내용을 거의 그대로 답습하고 있다.⁵⁰⁾ 인공지능 정책은 민간기업 간 경쟁 촉진보다는 공정성, 윤리성, 투명성, 설명가능성, 견고성, 안전성과 같은 인공지능의 신뢰성을 확보하기 위한 인프라 구축에 중점을 두어야 하며,⁵¹⁾ 이를 위해서는 기술의 전 형성과정에서 규제 거버넌스가 요구된다. 사후규제만으로는 새로운 데이터 인프라 구조 구축이나 인공지능 기술개발 단계 등에서 제기되는 윤리문제 및 위험성에 적절히 대처할 수 없다. 인공지능 정책에는 사전규제에 의한 보호, 설계에 의한 보호, 기본값(default setting)에 보호, 인공지능 리터러시 교육 등 다양한 정책패키지가 각 단계에 합리적으로 배치되어야 한다. 사후규제는 단기적으로 신기술의 시장진출을 용이하게 할 수는 있으나, 규제에 대한 예측가능성이 없어 산업육성에 반드시 유리한 것도 아니다.⁵²⁾

둘째, 고위험 영역 인공지능의 정의와 분류 문제이다. 법안은 “사람의

48) 자율주행차의 보행자 사망 사고, 경비로봇의 유아 공격, 인공지능 스피커의 오주문, 챗봇 이루다의 혐오 발언, 법무부 출입국 인공지능 식별추적 시스템의 얼굴정보 무단 이용, 그리고 챗GPT의 개인정보 및 보안 이슈 등 인공지능의 위험성은 최근까지도 수차례 보고되고 있다.

49) 중앙일보, 2023. 2. 14., “‘AI 더러운 비밀’…구글보다 챗GPT가 지구에 더 나쁜 이유”, <<https://www.joongang.co.kr/article/25140421>>, 검색일: 2023. 5. 1.

50) 정보통신융합법 “제3조의2(우선허용·사후규제 원칙) ① 누구든지 신규 정보통신 융합등 기술·서비스를 활용하여 사업을 할 수 있으며, 국가와 지방자치단체는 신규 정보통신융합등 기술·서비스를 활용하는 과정에서 국민의 생명과 안전을 저해하는 경우에 이를 제한할 수 있다.” 지능정보화기본법도 유사한 규정을 두고 있다. “제31조(규제 개선 등) ① 누구든지 지능정보기술, 지능정보서비스 및 지능정보기술 제품을 개발·제공·활용할 수 있으며, 정부는 지능정보기술, 지능정보서비스 및 지능정보기술 제품을 개발·제공·활용하는 과정에서 사람의 생명과 안전을 저해하는 경우 등에 한정하여 이를 제한할 수 있다.”

51) 한상기, 신뢰할 수 있는 인공지능(클라우드나인, 2021) 참조

52) 정관선/박균성, “네거티브 규제의 재검토”, 208쪽.

생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미칠 우려가 있는 영역에서 활용되는 인공지능”을 고위험 영역 인공지능으로 정의한다. 고위험 외에 다른 분류규정은 없다. 그러나 이는 국제적 표준으로 인식되는 유럽연합의 「인공지능법(AI Act)」(안)에서 인공지능이 생성할 수 있는 위험 수준을 네 단계로 구분하고 각 수준에 따라 공급자와 사용자에게 다른 의무를 설정하고 있는 것과 대조를 이루고 있다.

유럽연합의 법안은 인공지능의 위험을 수용불가한 위험, 고위험, 저위험, 최소위험으로 분류하면서, 특히 수용불가한 위험도를 가진 인공지능에 대해서는 그 개발 및 활용을 엄격히 금지한다. 여기에는 잠재의식기술을 사용하는 인공지능 시스템, 사람의 취약성을 공격하는 시스템, 사회적 점수평가(social scoring)에 활용되는 시스템, 공개장소에서 실시간 원격 생체 인식 시스템, 성별, 인종 등 민감한 특성을 사용하는 생체 인식 분류 시스템, 예측 치안 시스템 등이 포함된다. 반면, 우리 법안은 금지되는 인공지능 시스템에 관한 규정 자체가 없다.

유럽연합의 법안은 건강, 안전, 기본권 또는 환경에 대한 위험성을 포함하여 고위험 영역을 설정하고 있으며, 최근 논의를 통해 선거운동 등 정치캠페인에서 유권자에 영향을 미치는 인공지능 시스템과 4,500만 명 이상의 사용자를 보유한 소셜 미디어 플랫폼에서 사용하는 추천 시스템도 고위험군 목록에 추가하였다. 우리 인공지능법안의 경우, 에너지, 먹는 물 등의 공급, 보건의료의 제공 및 의료체계, 의료기기, 핵물질과 원자력 시설 등에서 사용하는 인공지능을 고위험 영역으로 분류하고 있다.

하지만 사람의 생명, 안전, 기본권에 미치는 중대한 위험성이 충분히 예상되는 영역이 다수 고위험 분류에서 제외되었다는 비판이 제기되고 있다.⁵³⁾ 특히 생체정보의 분석·활용의 경우 범죄수사나 체포업무라는 일부 업무에 한정하여 고위험 영역으로 분류하고 있어, 이를 금지하는 유럽연합의 입법태도와 대조를 이룬다. 법률상 고위험군으로 명시되지 않았던

53) 예를 들어, 유럽연합 법안은 ① 실시간 또는 사후적으로 사람의 생체정보를 활용하여 신원확인 작업을 수행하는 인공지능, ② 지원서 선별, 후보자 평가, 승진결정, 작업 할당, 업무 성과 모니터링 등 인사 관리 업무에 사용되는 인공지능, ③ 직업 훈련 기관의 선정 및 지원 결정, 교육생 및 훈련생 평가에 사용되는 인공지능을 포함하고 있지만, 우리 인공지능법안에서는 제외되었다.

라도 대통령령을 통해 사후 고위험으로 규율할 수 있다는 반론도 가능하지만, 충분히 예상되는 위험에 대한 사전의 법률적 규율과 사후의 위입법법적 규율은 규제강도를 달리한다.

또한 유럽연합의 법안은 허용할 수 없는 위험이나 고위험 인공지능시스템이 제재사항을 준수하지 않을 경우 3천만 유로 내지 직전회계연도의 전 세계 연간 총매출액의 6% 중 높은 금액을 과태료로 부과한다. 반면 우리 법안에 따르면, 고위험 영역 인공지능에 대한 사업자의 이용자에 대한 사전고지의무, 사업자의 신뢰성 확보조치와 그에 대한 준수 ‘권고’만을 규정하여, 고위험을 완화하거나 방지할 수 있는 실효적 장치를 찾아볼 수 없다.

셋째, 인공지능 정책의 거버넌스 문제이다. 법안은 과기정통부를 인공지능 정책의 주관기관으로 설정하고 있다. 인공지능에 관한 폭넓은 정책 권한을 부여하고 있는 것이다.

인공지능 기본계획 수립, 인공지능위원회의 간사위원 참여, 국가인공지능센터 설치, 인공지능 기술 개발 활성화 사업, 학습용 데이터 관련 시책 수립, 전문인력 확보, 인공지능 윤리원칙 및 그 실천방안에 대한 권고, 고위험 영역 인공지능의 확인 등이 과기정통부의 권한에 포함된다. 그러나 이 영역 중 상당 부분은 공정거래위원회, 고용노동부, 국가인권위원회, 개인정보보호위원회, 방송통신위원회, 산업통상자원부, 행정안전부, 각 지자체 등 다른 기관의 업무와 중첩된다. 산업진흥에 본연의 가치를 두고 있는 과기정통부가 개인정보보호, 소비자보호, 인권보장, 차별금지, 제품안전 등 광범위한 영역의 가치를 두루 고려하면서 기본계획, 시책, 윤리원칙을 마련할 수 있는지 의문이다.

인공지능법안은 인공지능 등에 관하여 ‘다른 법률에 특별한 규정이 있는 경우를 제외하고는 이 법이 정하는 바에 따른다’(윤두현의원안 제4조)고 규정하는데, 이를 통해 형식적으로 다른 규제기관의 작용을 무력화시키지는 않겠지만, 해당 인공지능법안이 인공지능에 관한 최초의 법률로서 갖는 위상, 이 법에 따라 정립될 과기정통부와 인공지능위원회의 지위와 권한 및 영향력으로 볼 때, 이 인공지능법안이 통과된다면 인공지능 정책에 관한 거버넌스는 과기정통부 중심으로 형성될 것이다. 인공지능 기술의 다면적 성격과 내재적 위험성에 대한 선제적·종합적 대응을 요구하는

환경하에서 기술관료 중심의 거버넌스는 산업편향의 결과를 산출할 가능성이 크다.

3. 향후 과제

우선, 인공지능 기술, 제품, 서비스가 갖는 잠재적 위험성을 예방하고 완화할 수 있는 규제방식을 채택해야 한다. 이와 관련하여 국제기구와 세계 각국은 인공지능에 대한 영향평가제도를 도입하고 있다. 특히 인권실사(HRDD)와 그 핵심도구인 인권영향평가를 인공지능 규제정책에 포함할 필요가 있다. 인공지능 인권영향평가는 인공지능 기술의 잠재적·현실적 위험성을 예방, 완화하기 위한 제도적 대안으로 주목받고 있다. 법안이 추상적으로 선언하는 우선허용·사후규제 원칙은 영향평가제도의 도입을 봉쇄하는 조항이다.

둘째, 인공지능 사용 영역의 위험도와 심각도에 대한 사회적 합의와 이에 기초한 입법이 필요하다. 유럽연합이 제시한 금지/고위험/제한적/최소의 위험도 구분을 기준으로 할 때, 먼저 우리 사회에서 금지되어야 할 인공지능 활용영역을 규범적으로 확인해야 한다. 헌법적 기준에 따른 위험도 평가가 필요하다. 유럽연합의 법안처럼 잠재의식기술을 활용한 광고 행위도 금지되어야 한다. 고위험 영역에 대해서는 광범위한 사회적 합의가 요구된다. 예를 들어, 앞서 언급한 공공장소 원격 실시간 생체 인식(얼굴인식)의 위험도를 어떻게 분류할 것인지, 입학시험, 성적평가, 노동자 모니터링, 플랫폼 노동 등에 활용되는 인공지능의 위험도를 어떻게 평가할 것인지 등 다양한 이슈는 사회적 합의가 필요하며 입법은 이를 반영해야 한다. 공론조사방식도 고려될 수 있다.

셋째, 인공지능 규제와 정책을 주도할 독립적인 기관 설치를 고려해야 한다. 지금 법안이 상정하고 있는 것처럼 과기정통부가 인공지능 정책을 주도한다면, 인공지능 정책은 균형을 잃고 산업편향에 빠질 것이다. 세계적인 입법경향에 부합하는 인공지능 거버넌스를 구축하기에 과기정통부는 좁은 틀이다. 2022년 5월 국가인권위원회는 「인공지능 개발과 활용에 관한 인권 가이드라인」을 발표하면서 “인공지능을 독립적이고 효과적으로 감독할 수 있는 체계 수립”을 권고한 바 있다.

4. 소결

인공지능의 위험성에 대한 입법적 대응은 신약에 대한 검증 통제에서 배울 필요가 있다. ‘우선허용·사후규제 원칙’으로는 인공지능의 잠재적 위험성에 대비할 수 없다. 인공지능은 인간 정신의 깊숙한 영역까지 조정할 수 있는 침습적 기술로 활용될 수 있으므로 장단기적 안전성을 검증하기 위한 제도적 장치가 마련되어야 한다.

2020년 유엔사무총장은 인공지능 사용에 대한 책임성 보장을 위한 적절한 법률체계와 감독체계 수립, 피해구제수단 마련을 권고한 바 있고, 2021년 유엔인권최고대표도 인공지능 사용의 부정적 인권영향을 방지·완화하는 인권실사와 규제체계 도입을 권고한 바 있다. 국가인권위원회, 개보위, 금융위, 방통위도 관련 가이드라인이나 규제방안을 제시한 바 있다. 하지만 국회 과방위와 과기정통부가 주도하고 있는 인공지능법안은 이러한 국내외의 입법 경향과 동떨어진 내용을 담고 있다. 이대로 입법되면 국내 인공지능 산업은 갈라파고스화되고, 시민들은 보호장치 없이 인공지능의 위험성에 그대로 노출될 것이다.

IV. 나오며

인공지능 시스템이 제기하는 문제는 프라이버시나 데이터권에 국한되지 않는다. 인간의 고유한 능력으로 여겨왔던 활동이나 판단능력을 기계적으로 구현할 수 있는 시대가 다가온 것이다. 자동화된 의사 결정 시스템은 민주적 가치에 총체적인 영향을 미칠 수 있다는 점에 주목해야 한다. 민주적 가치를 추구하는 국가와 정치단체는 인공지능 시스템의 한계와 오용 가능성을 파악하고 이에 대비하는 제도와 전략을 준비해야 한다.

인공지능 시스템을 자율규제라는 이름으로 안전장치 없이 방임한다면, 이 시스템을 지배하는 행위자들(테크노크라트와 빅테크의 소유자)이 감시 자본주의와 디지털 자본주의의 논리에 따라 초국가적 차원에서 권리와 자유의 보호 기준을 자의적으로 결정하게 될 것이다. 인권과 민주주의는

기술결정론으로 대체될 것이다. “기술에 의존하는 세상에서는 그 기술에 대해 정통한 정치꾼이 곧 왕이다”.⁵⁴⁾

고도화된 인공지능 시스템을 지속가능한 방식으로 운영하기 위해서는 규제 거버넌스의 역할을 이해하는 것이 중요하다. 공공부문과 민간부문, 기술혁신과 위험 사이에서 균형 있는 방향잡기가 필요하다. 통제되지 않은 인공지능 기술을 활용하여 확증편향으로 점철된 차별적 편견을 조장하고, 유권자의 특정 성향을 프로파일링하여 가짜 뉴스를 세뇌될 때까지 발신할 수도 있다. 인간의 특정한 반응을 ‘유도’하는 기술도 날로 발전하고 있다. 오늘날 인공지능 기술의 지배자인 빅테크나 디지털 기업이 헌법적 한계를 초월하는 권력을 획득하도록 방치할 수는 없다. 인공지능 시대의 빅테크 기업에 대한 규제 거버넌스는 일순위 의제로 논의되어야 한다.

과거 산업 비즈니스 모델에 입각한 디지털 정책의 관행적 시행에서 벗어나 인공지능 기술이 생성하는 제도적 환경에 적응하는 인공지능 정책을 형성해야 한다. 현재 국회에서 계류 중인 인공지능법안은 기존 ICT 법제의 규제방식을 그대로 채용하고 있으며, 인공지능 법제가 풀어야 할 주요 쟁점인 인공지능 위험도 분류 문제를 별다른 사회적 합의나 뚜렷한 기준 없이 ‘고위험 영역’이라는 단일 척도로 제도화하고자 한다. 고위험 인공지능을 규제하고 금지할 실효성 있는 제도적 장치는 몰각한 채 과거 정통부라는 산업부처 주도로 입법절차가 진행되고 있다. 현재 해당 인공지능법안은 소관위 전체회의 의결을 앞두고 있다.

세계적으로 인공지능에 대한 규제정책은 일반 산업이나 디지털 및 인터넷 경제에 대응하는 정책적 경향과 차별화되고 있다. 인공지능 기술, 제품, 서비스가 갖는 잠재적 위험성을 경고하면서 인공지능 개발자와 기업이 스스로 사전 규제의 필요성을 주장하는 역설적 상황도 펼쳐지고 있다. 고도 기술로서 인공지능의 지속가능한 운영을 위해서는 세계적 입법 경향에 부합하는 규제 거버넌스가 구축되어야 한다.

54) 데이비드 린시먼, 최이현 옮김, 쿠데타, 대재앙, 정보권력(아날로그, 2020), 170쪽.

<참고문헌>

이광석, “포스트-관용콘 시대 감시 연구, 새로운 지형”, 김동욱 외 엮음, 스마트 시대의 위협과 대응방안, 나남출판, 2015, 189-207쪽.
인남식, “아랍 민주화 운동과 미국의 대중동정책 변화 연구”, 서정민/인남식 엮음, 중동 민주화의 대내외 정치역학, 대외경제정책연구원, 2011, 73-148쪽.
정관선/박균성, “네거티브 규제의 재검토”, 법제 제699권, 2022, 189-213쪽.
정병기, “오성운동(M5S)의 직접 의회주의와 사이버크래틱 집중주의: 포스트포폴리스트 정치 운동의 성공과 한계”, 한국정치연구 제29권 제2호, 2020, 91-116쪽.
홍남희, “디지털 플랫폼에 의한 ‘사적 검열(private censorship)’”, 미디어와 인권권 통권 제6호, 2018, 135-172쪽.

김진우, 나의 첫 인공지능 수업, 메이트북스, 2022.

박승일, 기계, 권력, 사회: 인터넷은 어떻게 권력이 되었는가, 사월의책, 2021.
방송통신위원회/한국인터넷진흥원, 개인정보 처리에서의 프로파일링 사례집, 2020.
유승익 외, 인공지능 인권영향평가 도입 방안 연구, 국가인권위원회, 2022.
한홍구 외, 감시사회, 철수와영희, 2021.

데이비드 런서먼, 최이현 옮김, 쿠데타, 대재앙, 정보권력, 아날로그, 2020.

로저 맥나미, 김상현 옮김, 마크 저커버그의 배신, 에이콘출판, 2020.

롭 라이히 외, 이영래 옮김, 시스템 에러: 빅테크 시대의 윤리학, 어크로스, 2022.

브리테니 카이저, 고영태 옮김, 타겟티드, 한빛비즈, 2020.

빅토르 마이어 쾨버거/케니스 쿠키어, 이지연 옮김, 빅 데이터가 만드는 세상, 21세기북스, 2013.

쇼샤나 주보프, 김보영 옮김, 감시 자본주의 시대, 문학사상, 2021.

스튜어트 러셀/피터 노빅, 류광 옮김, 인공지능: 현대적 접근방식, 제3판, 제이펍, 2016.

시바 바이디야나단, 홍권희 옮김, 페이스북은 어떻게 우리를 단절시키고 민주주의를 훼손하는가, 아라크네, 2020.

제프리 삭스, 이종인 옮김, 지리, 기술, 제도, 21세기북스, 2021.

지그문트 바우만/데이비드 라이언, 한길석 옮김, 친애하는 빅브라더, 오월의봄, 2014.

프랭크 웹스터, 조동기 옮김, 현대 정보사회 이론, 나남출판, 2016.

헨리 키신저/에릭 슈밋/대니얼 허튼로커, 김고명 옮김, AI 이후의 세계, 월북, 2023.

Castets-Renard, Céline, “Human Rights and Algorithmic Impact Assessment for Predictive Policing”, Hans-W. Micklitz 외 엮음, *Constitutional Challenges In The Algorithmic Society*, Cambridge University Press, 2022, 93-110쪽.

- Danaher, John, “The Threat of Algocracy: Reality, Resistance and Accommodation”, *Philosophy & Technology*, vol. 29, no. 3, 2016, 245-268쪽.
- Diamond, Larry, “Liberation Technology”, *Journal of Democracy*, vol. 21, no. 3, 2010, 69-83쪽.
- Newell, Sue/Marabelli, Marco, “Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of ‘Datification’”, *Journal of Strategic Information Systems*, vol. 24, no. 1, 2015, 3-14쪽.
- Pollicino, Oreste/De Gregorio, Giovanni, “Constitutional Law in the Algorithmic Society”, Hans-W. Micklitz 외 엮음, *Constitutional Challenges In The Algorithmic Society*, Cambridge University Press, 2022, 3-24쪽.
- Pötzsch, Holger, “Archives and Identity in the Context of Social Media and Algorithmic Analytics: Towards an Understanding of iArchive and Predictive Retention”, *New Media & Society*, vol. 20, no. 9, 2018, 3304-3322쪽.
- Ram, Natalie/Gray, David, “Mass Surveillance in the Age of COVID-19”, *Journal of Law and the Biosciences*, vol. 7, no. 1, 2020, 1-17쪽.
- Richards, Neil M., “The Dangers of Surveillance”, *Harvard Law Review*, vol. 126, no. 7, 2013, 1934-1965쪽.
- Simoncini, Andrea/Longo, Erik, “Fundamental Rights and the Rule of Law in the Algorithmic Society”, in Hans-W. Micklitz 외 엮음, *Constitutional Challenges In The Algorithmic Society*, Cambridge University Press, 2022, 27-41쪽.
- Walorska, Agnieszka M., “The Algorithmic Society”, Denise Feldner 엮음, *Redesigning Organizations Concepts for the Connected Society*, Springer, 2020, 149-160쪽.
- Lyon, David, *Surveillance After September 11*, Polity, 2003.
- Petit, Nicolas, *Big Tech and the Digital Economy: The Moligopoly Scenario*, Oxford University Press, 2020.
- Schneier, Bruce, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*, W. W. Norton & Company, 2015.

<Abstract>

The Impact of AI on Human Rights and Democracy and the Issues in the AI Bill

Yoo, Seung Ik

Research Professor, Handong University

Artificial intelligence (AI) encompasses both positive and negative aspects from the perspectives of human rights and democracy. AI not only has the potential to radically transform political, economic, and social structures but also raises new constitutional and legal issues. This article focuses on the political and legal awareness and examines the impact of AI on human rights and democracy, as well as highlights the key points of contention in recent discussions of AI legislation in the legislature.

AI systems can have detrimental effects on nearly all human rights and fundamental freedoms. Particularly, they can significantly impede democratic processes by undermining freedom of expression and the right to self-determination of personal information. “Content moderation” and “user profiling” serve as representative examples that illustrate the risks posed by AI systems to fundamental rights and democracy. In the same vein, the regulation and management of misinformation online become contentious issues. Uncontrolled AI systems not only render individuals vulnerable to human rights violations but also jeopardize the fate of political communities by placing decision-making power in the hands of private enterprises. Moreover, as surveillance societies become more sophisticated and differentiated, traditional surveillance powers evolve into big data surveillance powers, where surveillance authority solidifies as a social infrastructure rather than an antagonistic object of human rights subjects. Under surveillance capitalism, individuals are depersonalized, fragmented, and reduced to data fragments. Freedom

is captured through delegated decision-making based on “consent” and monopolistic control of information sources by AI systems, while individual and group identities are manipulated algorithmically, leading to polarization. Freedom is not externally constrained but internally distorted. The AI society enables private actors such as big tech companies and digital platforms to exercise private power beyond the role of the state or government, challenging constitutional traditions.

Many countries around the world are preparing legislation to address the potential risks of AI technologies. National Assembly of South Korea is also undergoing the legislative process regarding AI bills. However, the “permit-first-regulate-later principle” adopted by these bills is inadequate to address the inherent risks associated with AI technologies and is not necessarily favorable to industrial growth. Furthermore, unlike the detailed classification systems in the European Union and elsewhere, these bills fail to regulate prohibited AI and unreasonably exclude areas with significant risks to life, safety, and fundamental rights without sufficient justification. Additionally, by designating the Ministry of Science and ICT, a technocracy-centered institution, as the main governing body for AI policies, the current bills emphasize an industry-biased AI policy governance. Therefore, the pending bills require comprehensive reconsideration.

Key phrases: Artificial Intelligence(AI), Algorithms, Data, Human Rights, Democracy, Surveillance