성균관대학교 법학연구원 『성균관법학』 10.17008/skklr.2021.33.2.018
제33권 제2호 (2021. 6.)
Sungkyunkwan Law Review Vol.33. No.2 (June 2021) pp 589 - 622

헌법상 투명성원칙에 기반한 인공지능 규제의 방향

백수원*

【목 차】

Ⅰ. 서론
Ⅱ. 헌법상 투명성 원칙과 인공지능 규제
Ⅲ. 인공지능 관련 정책 기준의 한계와 법·규범화 가능성 검토
Ⅳ. 결론

【국 문 요 약】

과학 기술과 인간 중 누가 멀리 내다볼 수 있는, 소위 넓고 긴 안목의 진화에서 그리 현저한 성취에 앞서고 있는 것일까. 과학 기술을 창조한 것은 인간이지만 인간의 사고체계 도전할 수 있을 정도의 과학적 진보는 인간에게 두려움까지 안겨주고 있다. 인간은 구체성을 갖추지 않은 미래에 대하여 예측할 수 있는 능력이 있다. 즉 생명이 걸린 위험이 다가오기 아주 훨씬 전에도 그리고 아주 먼 거리에서 아무런 상관성도 없는 듯이 보이는 작은 조짐으로부터도 인간은 예측 능력을 발휘하여 불안을 느끼고 그에 대한 대응책을 마련한다. 그러나 점차 과학과 기술에 예측 능력을 양도함으로써 어떤 사태가 임박했을 때 기술을 사용하여 임기응변적 반응기제를 기민하게 동원하는 것에 익숙해지기 시작하였다. 인간은 보다 확실한 예측을 원했고, 그로 인해 스스로 판단하기보다는 기계나 과학적 기술 또는 시스템적 사고에 익숙해져 갔다.

이로 인해 인간의 직관력은 다소 약화되고 과학 의존도는 심화되어 갔다. 또한 멀리 떨어진 혹은 먼 미래에 대응하고자 하는 생각보다는 눈에

* 송원대학교 국방경찰학과 조교수/법학박사

보이는 현상에 대응하고자 하는 생각이 더 설득력을 얻었다. 현실에 충실해야 한다는 가치관이 '지금 여기'를 강조함으로써 지구와 인류 전체의 운명을 바라보는 먼 시간적 전망에 소홀하며 인공지능과 관련된 이슈에 있어서도 기술(technology) 영역의 문제로만 보고 선도적 규제보다는 발견된 문제를 해결하는 방향의 규제를 선호한다.

그러나 인공지능 기술의 탄생으로 우리는 매우 새로운 기술과 새로운 윤리기준을 세울 수 있는 혼돈의 시기를 겪고 있다. 이 혼돈의 시기의 끝이 언제인지 모를 만큼 빠르게 인공지능시대가 도래할 수 있다는 두려움도 있고, 기술이 법의 영역에서 어떻게 도입·적용·구현될 수 있을지에 대한 고민이 무르익어가고 있다. 그러나 한편으로는 현재가 곧 인공지능 규제를 올바른 방향으로 가져갈 수 있는 순간이라고 볼 수도 있다.

인공지능을 활용함에 있어 효과적인 규제를 논의하는 것은 이제부터라고 볼 수 있다. 투명성은 헌법원칙인 투명성 원칙에 근거하며, 인공지능 규제의 바람직한 방향을 제시할 수 있는 하나의 방안이 될 수 있다고 본다. 본문에서 구체적으로 인공지능 투명성 강화를 위한 정책의 5단계로 AI의 사용을 고지, 설명을 들을 권리의 인정, 정보에 대한 접근권 인정, 종합적인 AI 투명성 정책 그리고 AI 기반 결정 내역을 기록하고 이를 감독하는 전문 감독기관의 설치 등을 제시하였다.

한계가 예상되는 점도 있으나 헌법의 투명성원칙에 기초한 규제가 인공지능 산업 발전적 측면에서나 규제의 효율적 측면에서 가장 주요한 가치임을 공감하고 이를 추진해야 한다. 로봇에 대한 규제는 규제 영역에 있어 새로운 패러다임을 요구한다는 점에서 큰 도전 과제이다. 그러나 기술과 규제 사이의 간극을 해소하는 일이 인간을 위하여 법이 할 수 있는 최선의 배려일 것이다.

I. 서론

검색 엔진에 흑인 소녀를 검색했을 때 '죽이는(hot)','달콤한(sugary)'이 나오는 결과가 나왔다면 이에 어떤 조치가 필요할까. 기업의 사원 채용시 불거지는 공정성 시비로 인해 인공지능채용시스템을 활용하기로 하고, 인공지능은 내재된 알고리즘으로 자가 학습한 결과 여성보다 남성을 채용하는 것이 더 이익이라고 보아 차별적 채용을 감행했다면 이는 누가 책임질 일인가. 2018년 아마존의 인공지능채용시스템으로 인한 논란이다. 당시 아마존은 인공지능이 여성을 차별하였다는 이유로 인공지능 폐기를 감행한 바 있다. 채용 알고리즘이 인간의 편견까지 같이 습득하고 있었는데, 이를 수정·보완할 수 있는 알고리즘적 조치를 확신할 수 없었기 때문에 폐기를 선택할 수밖에 없었다.[1] 이렇듯 인공지능 학습데이터에 인간의 편견이 반영되는 사례가 속속 공개되고 있는데, 인공지능의 활용 범주가 다양해진 만큼 발생하는 부작용과 문제점도 빈번해지고 있다. 윤리적 문제를 넘어 이러한 문제에 대해 어떻게 대응하는 것이 가장 적절하고 또 효과적인 것인지 법·제도적 관심과 대응이 필요한 시점이다.

인공지능에 대한 관심만큼 인공지능과 관련된 많은 법적 논의가 제기되고 있다. 넓게는 인공지능의 인격성에 관한 주제부터 구체적 입법안에 대한 제언까지 다양한 영역에서 새로운 시대에 대응하기 위한 인간의 노력을 엿볼 수 있다. 특히 법적으로 주요 이슈가 될 수 있는 영역은 헌법 단계에서부터 심도있는 고민이 필요하다고 할 수 있는데, 아직 특정된 형태의 존재성이 없는 현실에서 가상의 상황과 존재를 규율하려는 시도로 인해 그 내용이 다소 추상적일 수밖에 없다는 한계 또한 분명히 인식할 필요가 있다.

이러한 정책의 추상성과 불확실성이라는 한계가 있겠지만 꼭 필요한 인공지능 정책을 마련하기 위해서는 원칙과 원리를 세우고 이러한 기준에 기초한 방향 제시가 일차적으로 논의될 필요가 있다. 역사적으로 볼

1) BBC뉴스, "성차별:아마존, '여성차별' 논란 인공지능 채용 프로그램 폐기", BBC뉴스(2021. 4. 1.) https://www.bbc.com/korean/news-45820560

때 중요한 역사적 사건은 그것이 반드시 일어난 다음에 정책이나 헌장으로 고착된 것이 아니라, 선언 또는 정책적 틀이 선행된 다음에 현실화된 경우도 적지 않다. 인공지능과 관련한 현재의 우려는 다소 과열되거나 막연한 측면이 없지 않아 있지만, 일정한 틀을 세우고 기초적 단계에서나마 기준을 제시하는 일이 계속될 때 구체적이고 실질적인 정책도 나올 수 있을 것으로 생각한다.[2)]

이러한 관점에서 인공지능 정책 기준에 투명성 강화를 위한 기준이 마련될 필요가 있다는 제언 또한 주지할만한 일이다.[3)] 다양한 정보 속 국가 의사 결정에 영향을 줄 수 있는 투명성 원칙은 이미 헌법에 명시되지 않은 헌법원칙으로서의 기능을 한다는 점을 고려할 때 인공지능과 관련된 원칙을 세우는 데 있어 중요한 정책 결정 방향을 제시할 수 있을 것으로 기대된다.

본질적으로 의사소통 단계와 정책 결정 과정에서의 투명성 확보는 인공지능과 같은 새로운 영역에 대한 기준점을 설정하는데 있어 가장 기본적인 것으로 판단될 수밖에 없다. 비대화된 오늘날 국가의 역할과 힘에도 불구하고 민간 기관과의 협응을 필수로 하는 인공지능 사업에서 역설적으로 상호소통을 더욱 강조하게 되며, 이는 곧 투명성 원칙이 인공지능 정책에 있어 핵심적 요소가 될 수 있음을 확인시켜 준다.

결국 바람직한 정책은 이상적인 내용을 담고 있되, 그 이상이 잘 전파될 수 있고, 잘 소통되어 국민과 국가가 상호 협력할 수 있는 단계로 나아갈 수 있는 어떤 지점에서 시작될 수밖에 없다. 정책에 대한 국민적 협력과 순응은 투명성 확보의 과정을 거치며 완성될 수 있으며, 새로운 정책에 대한 것이라면 더욱 그러할 가능성이 높다.

아래에서 인공지능 투명성 강화를 위한 정책의 5단계로 AI의 사용을 고지, 설명을 들을 권리의 인정, 정보에 대한 접근권 인정, 종합적인 AI 투명성 정책 그리고 AI 기반 결정의 내역을 기록하고 이를 감독하는 전

---

2) 사람에 필적하는 수준의 지능을 가진 인공지능로봇에 대한 우리의 기대는 현실화되기 어렵다는 전망이 많다. 따라서 그러한 환상적인 상황보다는 현실화된 결과를 해결하는 데 주목하여야 한다는 입장으로 정완, "인공지능 로봇에 관한 법적 고찰", 경희법학 제55권 제4호, 경희대학교(2020), 204면.

3) 인공지능 사회에서 투명성 원칙의 역할에 관하여서는 김일환, "지능정보사회에서 헌법의 역할과 기능", 성균관법학 제32권 제3호, 성균관대학교(2020) 참고.

문 감독기관의 설치[4] 등이 제시되고 있는데, 그 내용을 검토하기에 앞서 정책 결정을 위한 각 각의 과정속에서 적용될 헌법 원칙으로서 투명성에 관한 내용을 확인하고자 한다.

II. 헌법상 투명성 원칙과 인공지능 규제

1. 인공지능에 대한 규제와 투명성원칙

인공지능과 관련된 제도의 특징은 사회적 합의가 매우 어렵고 사회적 갈등을 유발할 여지는 크다는 점이다.[5] 규제 방향이나 규제 형식 혹은 규제 자체의 필요성에 대해서까지 일치된 사회적 합의를 보기 어려운 이념적 특성이 있고, 세계관이나 가치관에 따라 매우 다양한 의견이 제시될 수 있다, 규제의 목적이나 목표에 대한 입장 또한 각 사회 집단별 혹은 개인별로도 매우 다를 수 있다. 게다가 인간이 인공지능의 결정에 대한 완전한 예측이나 상황인식도 불가능하다는 지적까지 나와 있는 상황에서 규제를 한다고 하여도 어떻게, 어느 정도로 하는 것이 적당한 것인지 구체적인 합의를 제시하는 것은 거의 불가능에 가까운 도전으로 볼 수도 있다.

이러한 다양한 의견이 존재하고 합의가 어려운 과제를 해결할 때 법적으로 중요한 가치는 투명성이라고 보아야 한다.

투명성은 의사결정과정에 이질적이고 비합리적인 요소들이 개입하는 것을 배제하고, 이해관계인들이 특히, 정보와 관련하여 대등한 지위에서 협상이나 의사결정과정에 참여할 수 있도록 함으로써 그 결정의 민주적 정당성과 신뢰를 확보하는 기능을 수행한다. 이러한 투명성의 요청은 주로 관련된 정보의 공개와 의사결정과정에 대등한 참여를 주된 내용으로 하는 것으로서 공공부문은 물론이고 사인 간의 관계에서도 의사결정의

---

4) Thomas Wischmeyer, "*Artificial Intelligence and Transparency: Opening the Black Box*", REGULATING ARTIFICIAL INTELLIGENCE, Springer Nature Switzerland AG(2020), 75면

5) 지능정보기술의 도입 및 활용 역량에 따라 많은 사회적 문제가 있어 왔고, 특히 고용이나 소득 양극화 및 정보 격차로 인해 부익부 빈익빈이 더욱 강화된 바 있다. 김민호·이규정·김현경, "지능정보사회의 규범설정 기본원칙에 대한 고찰", 성균관법학 제28권 제3호, 성균관대학교(2016), 287면 이하 참고.

정당성을 확보하기 위한 핵심적 요소로 간주되어 왔다. 투명성의 위기, 각종 정보를 광범위하게 보유하고 있는 국가권력에 대응하여 국민의 대등한 지위를 확보할 필요성 증대, 그리고 사회가 다원화·분권화됨에 따라 정의 내지 공익의 내용을 결정하는 것이 더더욱 어려워지는 현실에서 공개와 참여를 통해 결정의 정당성을 보완할 필요성이 증대하는 현상 등은 특히 국가권력의 행사와 관련하여 투명성이 새롭게 조명을 받도록 만들고 있다. 현대 헌법은 절차적 정당성의 확보를 필수적인 요소로 한다. 따라서 국가권력행사의 결과를 비롯하여 그 절차와 내용 나아가 그에 따른 책임관계에서도 투명성이 보장될 것을 요구한다.[6]

즉, 투명성을 통해 의사결정의 신뢰확보를 용이하게 한다는 점에서 헌법적 의의가 있다고 할 수 있으며, 새로운 기술에 대한 규제에 있어서는 특히 국가가 단독으로 나서서 모든 고민을 일거에 해소시킬 수 없다는 점을 감안할 때, 반드시 민간 또는 국민적 이해와 협력이 필요한 일에 있어서 투명한 정보의 소통과 교류는 가장 당연히 되는 정책 결정의 전제라고 볼 수도 있다.

구체적으로 투명성 강화를 위해서는 결과와 절차에 대한 공개, 내용의 명확성과 절차에 대한 참여 또는 의견 수렴 가능성의 보장 및 책임에 대한 확고한 기준 제시가 있어야 한다. 이는 국가기관이 보유한 정보에 대해 국민이 최대한 접근할 수 있도록 해야 한다는 결과의 투명성과 그러한 권한 행사의 절차가 공개되어야 하며 권한 행사의 과정이 열려있어 폭넓은 국민의 참여가 보장되어야 한다는 절차의 투명성, 공개되는 정보가 국민의 입장에서 이해가 가능하도록 질적인 완전성을 갖추어야 한다는 내용의 투명성 및 책임 관계를 명확하게 하고 책임추궁을 위한 제도적 장치가 구축되어야 한다는 책임의 투명성을 의미한다.[7]

극단적으로 인공지능에 대한 정책 수립 과정에서 행정적 규제로 기술의 발전을 방해해서는 안 된다는 입장과 적극적 규제로 악용될 여지를 원천적으로 차단하자는 반대되는 입장으로 대립되는 의견이 제시되는 일이 빈번히 발생한다[8]고 하더라도 투명성을 통해 충돌이나 분쟁 발생의

6) 김일환·백수원, “헌법상 투명성원칙에 관한 시론적 고찰”, 공법연구 제39집 제2호, 한국공법학회(2010), 385면 이하 참고.

7) 백수원, “헌법상 투명성 원칙에 관한 연구”, 박사학위 논문, 성균관대학교(2010), 274면 참고.

여지를 줄여줄 수 있다.

생각건대, 규제가 어렵거나 새로운 영역이라는 이유로 정책 도입에 주저하는 것은 해당 산업을 약화시켜 국제적 경쟁에서 뒤처지기만 할 뿐, 사회 전체에 이익이 되지 못하는 경우가 많았다. 인공지능에 대한 각국의 경쟁은 어제오늘의 일이 아니기 때문에 더 이상 규제의 어려움을 핑계로 좌시할 수 없는 상황이다. 따라서 규제의 어려움에도 불구하고 투명성에 기초한 규제책을 마련할 필요가 있고, 다양한 의견을 수렴하는 과정에서 바람직한 제도가 탄생할 수 있다는 것을 주지하여야 한다.

2. 인공지능 규제를 위한 이론적 전개

(1) 인본주의와 탈인간중심주의 그리고 인공지능 규제

인공지능을 어떻게 규제하여야 할 것인가를 고민하기에 앞서 탈인간중심주의적으로 접근할 것인가, 인간중심으로 검토할 것인가를 먼저 논할 필요가 있다. 이를 위해 인간은 어떤 존재이고, 어떤 이유로 오늘날의 규제 형식을 갖추게 되었는지에 대한 검토 또한 필수적이다. 인공지능은 인간을 모방하며 탄생하였기 때문에, 인간과 유사한 의식의 흐름을 상정하고 규제를 마련하는 것이 가장 본질적이면서도 기초적인 규제 설정 방향이 될 수밖에 없다[9]는 관점이 전통적 관점이다. 반면, 인간 중심적 사고에서 벗어나 완전히 새로운 존재임을 인정하고 새로운 인격체로 해석하여야 하기 때문에 규제방식도 인간적 사고를 넘어서야 한다는 입장도 있다.

전통적인 헌법적 관점에서 인간의 생물학적·법적 특징은 매우 중요하다. 헌법 이론은 모든 범주에 있어 인간의 특성과 인간 중심의 사고로 규율되어 있기 때문이다. 법적 지위를 새롭게 부여하는 일에 있어서도

---

8) Karen Yeung & Martin Lodge, “Algorithmic Regulation: An Introduction”, ALGORITHMIC REGULATION, Oxford University Press(2019), 7면

9) 전통적인 헌법학이 내재하는 기본 가치는 인본주의라고 할 수 있다. 결국 이 사회는 인간으로 구성되어 있으며, 인간의 본질이 사회를 통해 발현되고 또 그것이 마땅히 그래야 한다는 것이 규범적으로 녹아있는 것이 헌법이라고 본다. 이러한 관점에서 인공지능에 관한 규제 또한 시작된다는 점이 새로운 기술에 대한 우리 사회의 기준이자, 한계라고 본다. 인간중심주의를 벗어나 새로운 과학 현상을 검토해야 한다는 의견으로 김연식, “과학 기술의 발달에 따른 탈인간중심주의 헌법학 모색”, 헌법학연구 제26권 제4호, 한국헌법학회(2020), 290면 이하 참고.

인간의 범주를 넓혀서 그것을 해결하거나 적어도 인간과 관련지어 해석하는 형태가 주류를 이룰 수밖에 없었다. 헌법상 인간이 아닌 존재는 인간과 구분 짓기 위해, 또는 인간과의 유사점을 찾기 위해 규율되는 매우 예외적으로 다루어질 수밖에 없는 대상이라고 할 수 있다. 태아나 뇌사자와 같이 법적으로 인간의 지위를 향유하기 어려운 경우에도 인간 경계를 확장함으로써 인간에 준하는 지위를 부여하는 방식이 바로 우리 법체계의 규율 방식이다.[10]

인간은 동물의 일원이지만 다른 동물에서 볼 수 없는 고도의 지능을 소유하고 독특한 삶을 영위하는 고등동물이다. 현생의 인간은 호모사피엔스(homo sapiens)의 일종으로 분류되며, 직립보행(直立步行)을 하며 언어를 사용한다는 특징이 있다.[11] 그러나 법적인 의미에 있어 인간은 이러한 신체적 특징이나 언어 사용 유무만으로 구분되지 않는다. 법률적 의미에서 인간이 갖는 유의미성은 "법적 권리와 의무의 주체"라는 점이다.[12] 다만, 인본주의적 관점에서는 진화론이 갖는 한계에도 불구하고[13], 인간이라는 존재가 지구상 생물의 발전에 있어 최고의 단계에 있다는 점이 곧 법적으로도 다른 존재가 구분되는 우월적 지위를 향유할 수 있게 하는 본질적 요인이기는 하다. 따라서 이러한 우월적 지위로 말미암아 법적 지위를 갖고 있는 존재라는 관점에서는 인공지능은 인간의 필요에 의한 존재에 불과하므로 인공지능 규제 또한 인간 중심으로 이루어져야 한다는 입장을 견지한다.[14]

---

10) 김연식, 앞의 논문, 291~292면.

11) 두산백과, https://terms.naver.com/entry.naver?docId=1135955&cid=40942&categoryId=31611, 검색어: "인간", 검색일: 2021.5.30.

12) John Chipman Gray(edited by Davide Campbell&Philip Thomas), "The nature and sources of the law". Routledge(1997), 13-23면

13) 진화론적으로 인간은 여러 단계의 진화를 거친 존재로 분류되지만 현재는 이러한 진화로 동물이 인간이 되는 것을 목격할 수 없으며, 심지어 같은 영장류에서도 인간으로 진화하는 모습을 오늘날 찾을 수 없다. 결국 현재 일어날 수 없는 일이 과거에는 어떤 이유에서 어떻게 가능했는지 설명하지 못하고 있다는 점에서 한계가 있는 가설임은 분명하다.

14) 이러한 입장에서는 인간 개념의 확대를 통해 법체계 안으로 비인간을 포섭하는 것은 본질적으로 인간의 존엄성에 기초하기 때문에 인간과 연관을 지을 수 있을 때만 법적으로 의미가 있다고 본다. 인격의 확장성에 관한 논의에 관하여서는 박기주, "제4차 산업혁명 시대 인공지능(로봇)의 법적 지위-새로운 인(人) 개념의 법적 설계 가능성을 중심으로", 한국지식재산연구원(2017), 3면 이하 참고.

이와 달리 인공지능 기술은 이러한 인간 개념의 확장으로 그 변화에 대응하기 어렵다는 이른바 '탈인간중심주의적 입장'에서는 이러한 인간 중심적 사고를 넘어서야만 비로소 인공지능을 제대로 이해하고 규율할 것이라고 기대한다.[15] 현생인류인 호모사피엔스가 동물과 확연히 구분되는 사회적 진보를 성취한 것은 기술의 발명과 발전에 따른 것으로 본다. 1만 년 전의 농경사회의 출현으로 인간은 자연에 단순히 의지하는 것을 넘어, 시대 전환에 대한 융합적 성찰을 통해 그 환경과 조건을 이용하여 물질적 잉여를 생산하고 정착 생활과 인구증가로부터 사회라는 것을 만들게 되었다. 농업혁명에서 비롯된 긴 문명화의 기간 동안 인간은 다양한 언어와 사회적 제도, 종교와 관습을 만들어 냈다. 이렇듯 여러 차례의 신기술에 대한 도전과 적응 과정을 거쳐 새로운 법·규범이 변화 발전하여 오늘날의 법체계가 완성되었기 때문에 새로운 기술에 대한 새로운 규제적 접근이 필요하다고 본다. 따라서 탈인간중심적 사고에서는 사람에게 적용되는 인격화가 법인과 비(非)인간 조직에게도 적용되기 때문에 인공지능을 하나의 독립적인 사회체계 보아 인간과 다른 별도의 존재로 인식한다.[16]

사견으로 이러한 탈인간중심주의 이론이 갖는 이론적 의미는 납득하더라도, 인공지능으 규제할 때에도 인간이 중심이 되어야 하며, 구분의 기준은 인간이 되어야 하는 것을 부인하여서는 안된다고 본다.[17]

인류는 인간을 중심으로하는 법 체계를 발전시켜왔고, 많은 혼란 속에서도 현재의 모습을 갖추고 있다. 인간은 사회적 유대관계를 맺고 사회 안에서 성장하며, 사회 속에서 고통과 행복을 느낀다. 인간은 노동 용구(用具)를 만들고 자연에 작용하여, 자연에 맞서는 주체성을 갖고 행동한다는 특징이 있다. 즉 인간은 매우 사회적인 동물이면서도 자연에 적응

15) 김연식, 위의 논문, 293면 이하 참고.

16) 탈인간중심주의적 관점에서 보자면, 인공지능에게 인격을 부여한다는 것은 완전히 새로운 방식으로 전통적 관점에서 인간과 사물의 관계를 벗어나기 때문에, 인공지능을 사람으로 대우하는 형태로 해석하지 않는다. Gunther Teubner, "Rights of Non-humans? Electronic Agents and Animals as New Actors in Politics and Law", Volume 33, Journal of Law and Society(2006), 503면

17) 필자의 의견과 반대로, 인격을 사람과 분리되는 것으로 이해하는 글로 양천수, "현대 지능정보사회와 인격성의 확장", 동북아법연구 제12권 제1호, 전북대학교(2018), 6면 이하 참고.

하는 것을 넘어 자연을 변화시키려고 하는 존재이기 때문에 사회불안 요소를 싫어한다고 하여도 그러한 불안을 극복하는 도전을 마냥 거부하지 않는다.[18] 따라서 그러한 인간이 가진 고유의 특성을 기반으로, 현재의 법체계에서도 새로운 기술을 바탕으로 한 인공지능 시대를 마냥 배척하기보다는 인간 중심적 태도 하에 새로운 기술에 대한 규제에 대응할 것이라고 판단된다.

인공지능과 관련하여 전자인간이나 전자인격(Electronic Personhood)과 같은 새로운 시도가 지속적으로 사회적·법제도적 이슈로 떠오르는 과정에서 이러한 정책의 타당성과 문제점에 대한 논의과정이 필요하다.[19] 가령, 법은 법인에게 인간과 동일한 행위를 할 것을 기대하지는 않는다. 규제의 전제가 되는 고통과 불편을 느끼게 되는 기준이 다르기 때문이다. 법인이 인격체로 일정한 영역에서 그 역할을 할 수 있다는 점에서 권리와 의무의 대상으로 기능한다고 하더라도 인간에게 고통스러운 일과 권장되는 일 또는 기대되는 행동과 그렇지 않은 행동은 인간만이 가진 고유한 특징 속에서 그 가치를 갖게 될 수밖에 없는 것과 마찬가지로 인공지능에 있어서의 특수성을 고려한 논의가 전개될 필요가 있다.[20] 따라서 투명한 의사결정 창구를 통해 결정되어가는 과정속에서 헌법이 구현하는 법치주의와 민주적 의사결정의 틀을 갖추게 되고 이는 정책에 대한 정당성을 부여하게 되는 일이기도 하게 된다.[21]

결국, 인공지능 이슈를 법·규제적으로 구체화시키는 과정에서 인공지능에게 공정, 책임, 정의와 관련된 윤리적 문제를 해결하고자 하는 국가 차

---

18) 물론 이 점이 동물에 대한 인간의 우월성을 의미하는 것은 아니다. 인간만이 품성을 지닌 유일한 존재가 아니며, 심리적으로 고통을 느끼는 유일한 존재도 아니라는 것을 여러 동물학자들은 밝히고 있다. 대표적으로 제인구달, 제인구달 침팬지와 함께 한 나의 인생, 사이언스북스, 2014, 148면 이하 참고.

19) 전자인간 또는 전자인격성에 관한 논의는 백수원, “전자인간 및 전자인격 인정에 따른 법적 논의와 시사점 고찰”, 외법논집 제43권 제4호, 한국외국어대학교 법학연구소(2019) 참고.

20) 법인격의 특성에 관하여서는 김민배, “AI 로봇의 법적 지위에 관한 쟁점과 과제-Bryson 등의 법인격 이론을 중심으로-”, 토지공법연구 제87집, 한국토지공법학회(2019), 797면 이하 참고.

21) 법적 인격은 매우 탄력적인 개념이므로 충분히 사회 현상과 기대에 따라 변화가능하다는 입장을 제시한 글로 Peter M. Asaro, “Robots and responsibility from a legal perspective”, The IEEE conference on robotics and automation, workshop on robo-Ethics, Rome. http://www.peterasaro.org/writing/ASARO%20Legal%20 Perspective.pdf(2007), 3면.

원의 움직임이 활성화될 수밖에 없고,[22] 이러한 지속적인 검증 과정을 거치면서 사회적 공감대를 확보할 수 있는 관련 규제가 제정될 수 있을 것으로 기대된다.

(2) 인공지능의 특징과 규제에 있어 고려되어야할 사항

앞의 내용을 종합적으로 고려할 때, 인공지능과 관련된 정책 또한 큰 틀에서는 현재의 법·규범내에서 규율되는 것이 바람직하다고 본다. 특히, 인공지능은 기계일 뿐인데 이들을 사람의 경쟁 상대로 인식하여 규제하려는 시도는 바람직하지 않다고 생각한다.[23] 과거 컴퓨터가 처음 개발되었을 때 컴퓨터에 대한 두려움이 있었던 것과 같이 인공지능도 현실화된 위험보다는 무지나 두려움에서 비롯된 과도한 우려가 제기되고 있는 것은 사실이다. 물론 컴퓨터나 인터넷, 스마트폰 등 각종 문명의 이기들은 그 새로움만큼 새로운 유형의 범죄나 권리 의무를 동반하고 있었고, 인공지능 또한 어떤 새로운 현상을 초래할 수 있다는 점에서는 이론의 여지가 없다. 특히 Uncanny valley[24]를 지나 인간유사로봇(Humanoid)이 발명되어 인공지능 뇌를 보유한 형태의 개발을 인정해야할 것인지가 그 논란의 핵심일 것이다. 그러나 단순한 인간을 보조하는 인공지능은 개발 독려의 대상일 뿐이고 몇 가지의 보완을 통해 현재의 법·제도 속에서 충분히 규제 가능할 것으로 판된된다.[25]

---

22) Ugo Pagallo, “Apples, oranges, robots: Four misunderstandings in today's debate on the legal status of AI systems”, The Royal Society Publishing(2018), https://www.researchgate.net/publication/328292234_Apples_oranges_robots_Four_misunderstandings_in_today%27s_debate_on_the_legal_status_of_AI_systems, 13면.

23) 같은 의견으로 김광수, “인공지능 규제의 법체계”, 토지공법연구 제93집, 한국토지공법학회(2021), 217면.

24) 불쾌한 골짜기: ‘거의 인간에 가까운’ 로봇에 대해 호감도가 떨어지는 영역을 나타낸다. 인간이 로봇이나 인간 아닌 것들에 대해 느끼는 감정에 관련된 로봇공학이론으로 로봇이 점점 더 사람의 모습과 흡사해질수록 인간이 로봇에 대해 느끼는 호감도가 증가하다가 어느 정도에 도달하게 되면 갑자기 강한 거부감으로 바뀌게 된다. 그러나 로봇의 외모와 행동이 인간과 거의 구별이 불가능할 정도가 되면 호감도는 다시 증가하여 인간이 인간에 대해 느끼는 감정의 수준까지 접근하게 된다는 이론이다. 위키백과: https://ko.wikipedia.org/wiki/%EB%B6%88%EC%BE%8C%ED%95%9C_%EA%B3%A8%EC%A7%9C%EA%B8%B0 참고.

25) 다만, 어떤 ‘결정을 내리는 능력’을 갖춘 인공지능에 대한 법적 평가를 어떻게 해야 할 것인지 고민될 수밖에 없으며 실제로도 유럽연합에서는 이러한 ‘결정을 내라는 능력’을 기

문제는 인공지능에 대한 공포는 인공지능을 법·제도적으로 포섭하려는 행위만큼이나 강렬하다는 점이다. 특히 인공지능이 인간의 학습을 참고하지만 이를 넘어선 결과물을 보여준다든지 스스로 학습을 통해 인간보다 더 나은 판단을 내린다거나 인간이 인공지능이 나타내는 결과물을 100% 이해하지 못하는 경우가 종종 발생하면서 더욱 공포심이 증대된 상황이다. 인공지능 기술에 대한 불충분한 이해로 인한 것이 첫 번째 이유이고 인공지능 기술 자체가 갖는 특성에 기인한 이유도 있다.[26]

인공지능기술에 대한 공포는 규제의 강력성을 수반할 수밖에 없는데, 사견으로 인공지능을 규제함에 있어서는 초기보다 매우 강력한 규제 중심적 대응보다는 인공지능의 형태와 특성상 각 개체 특성을 고려한 다각도적이면서도 신속히 대응 가능한 조직의 설립과 정책 마련이 매우 중요하다고 본다.[27] 이를 위해 다양한 집단의 의견을 청취할 수 있는 위원회 또는 의견제안소 등이 마련되는 것이 바람직할 수 있다는 의견을 제시하고자 한다.[28]

그 이유는 인간으로부터 학습하였고 인간과 유사한 사고체계를 갖추었

준으로 로봇의 자율성을 인정하고 있는데, 이 때 '결정을 내리는 능력'은 인공지능을 의미한다.

26) 인공지능 기술은 인간 정보처리 활동의 원리를 분석하는 기초 기술과 ICT를 통해 이를 구현하는 인공지능 SW 및 HW 기술을 말한다. 이러한 인공지능기술의 특징으로는 1) 무인 의사결정 : 인간의 고차원적 판단기능을 수행함으로써 기계가 독립된 주체로 활동하여 자동화 및 무인화 확산, 2) 실시간 반응 : 정보수집, 데이터 분석, 판단 추론 등 일련의 과정들이 ICT 기술(IoT, Cloud, Big Data, Mobile)을 통해 즉각 처리되어 실시간 응답 반응, 3) 자율 진화 : 딥러닝 등 기계 학습을 통해 스스로 진화하여 기계의 성능이 기하급수적으로 향상, 4) 만물의 데이터화 : 과거에는 보관 활용이 곤란했던 데이터(생체 행태정보, 비정형 정보 등)도 기계 학습 과정을 거쳐 의미 추출 가능 등을 들 수 있다. 이러한 4단계 중에서도 2단계 및 3단계를 거치면서 인간이 다소 이해할 수 없는 결정이 발생되곤 한다. 인공지능 기술에 관한 자세한 내용은 김일환, 앞의 논문, 39~40면 참고.

27) 가령, 인공지능에게 법적 책임을 지게 한다는 이유로 전통적인 자유형과 같은 형벌을 부과하는 것은 아무런 의미없는 제재일 것이다. 적어도 그들이 '고통'이라는 것을 느끼는 것을 전제로 할 때 이러한 제재는 의미가 있다. 인공지능에 대한 법적 지위 부여의 기준과 관련하여 '고통'을 중요 요소로 이해하는 글로 백수원, "헌법상 인공지능 규제의 범위와 방향에 대한 시론적 고찰", 미국헌법연구 제29권 제2호, 미국헌법학회(2018), 87면 이하 참고,

28) 기술의 변화는 항상 인간의 세계관에 영향을 미쳐왔으며, 새로운 진보로 인한 변화의 의미를 인지하고 이에 대응하는 것이 곧 인간의 역사였다. 같은 의견으로 이화순, "우리 인간의 의무는 '존재'하는 것이 아니고 '사는 것'이다."' 소프트웨어 세계 제6집, 한국정보기술전문가협회, 192면 이하 참고.

다고 하여도 인간과는 다른 인공지능만의 특성이 있다. 인공지능은 기호(CODE)를 통해 작동하며 알고리즘을 통해 판단을 하게 된다. 이러한 알고리즘에 의한 결정은 낡지도 퇴색되지도 않고, 흐려지지도 않는다면 점에서 매우 논리적이고 분명한 의사결정 방식이기는 하나, 인간의 기억저장과는 다르다는 것을 확인할 수 있다. 또한 사회적 학습을 통해 쉽게 바뀌기도 하고 개인의 편향적 생각으로 인해 잘못된 판단과 결정이 내려질 수 있는 인간의 인지체계와 달리 비교적 명료한 형태로 의사결정이 진행될 수밖에 없다. 따라서 그 의사결정이 옳든 그르든 그 결정은 매우 빠르고 즉각적일 가능성 또한 높은데, 이에 대한 충분한 검토가 이루어질 수 있는 규범적 장치가 필요하다고 보기 때문이다.

더하여, 이러한 문제에 대한 인식 때문에 오늘날에는 이러한 인공지능의 특성을 고려하여 사람에게 기대하는 수준의 윤리성을 인공지능도 갖추도록 요구하고 있기도 하다.[29] 인공지능의 알고리즘을 통한 의사결정에 더하여 윤리적 판단을 적용하고자 하는 이유는 결국 그 의사결정에 대한 신뢰성 강화를 위함이다. 인공지능은 결국 사람이 더 나은 결정을 하도록 돕는 기계이기 때문에 그 의사결정 전반에 대한 신뢰를 할 수 있어야 하기 때문이다. 즉, 규범적 틀을 만들어 위해서는 일련의 기준들이 필요한데, 이 기준은 인공지능의 특성을 고려하되, 인간 중심으로 신뢰를 확보하기 위한 법적 기준이어야 한다.

결국 인공지능을 전문으로 하는 규범적 장치가 필요하며 이러한 규범적 장치에는 인간 중심의 신뢰 확보를 위한 법적 기준이 마련되어야 하는데, 필자는 ①공정성(Fairness), ②설명가능성(Explain ability), ③견고성(Robustness), ④투명성(Transparency)을 제안하고자 한다.

‘공정성’은 서로 다른 집단을 차별하지 않는 ‘편견없는 알고리즘’을 의미한다. ‘설명가능성’은 인공지능이 왜 그런 결정을 내렸는지 설명할 수 있어야 한다는 것을 뜻하며, ‘견고성’은 오류에 대처하는 능력을 말한다. 인공지능은 학습한 것은 잘 알지만, 아주 작은 차이에도 큰 오차를 내기

29) IEEEInstitute of Electrical and Electronics Engineers)는 이미 2017년부터 인공지능 설계에 있어 윤리적 관심이 필요하다는 의견을 제시한 바 있다. Raja Chatila ; Kay Firth-Butterflied(ET AL), “The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems”, IEEE Robotics & Automation Magazine Volume 24 issue 1(2017), 110면 이하 참고.

도 한다. 마지막으로 요구되는 조건이 '투명성'인데, 이는 어떻게 데이터를 모으고 교육했는지, 개발자와 디자이너의 결정이 어땠는지 공개해야 한다는 것을 뜻한다. 위 4가지의 기준을 통해 인공지능의 판단에 대한 신뢰 확보가 가능하다고 보는데, 나머지 요건들의 전제가 되는 것이 바로 '투명성'이다.

기술 규제의 중심에 사람을 두고 위 네 기준을 I. 서론에서의 '아마존'과 같은 구체적인 사례에 적용해 볼 때, 인공지능의 인사채용시스템에서 왜 해당 사원이 배제되었는지나 왜 선정되었는지 사람에게 이해가능한 설명을 할 수 있어야 하며(설명가능성), 오류가 발생했던 알고리즘상의 편견을 배제하여야 할 것이다(견고성). 다만 편견을 지우는 과정도 주의가 필요하다. 일례로 성차별을 지우다 연령차별을 강화하는 경우가 생기는 경우도 있기 때문이다(공정성). 어떤 데이터를 사용하여 그러한 의사결정을 내렸는지, 해당 정보가 무엇인지 공개하고 이를 검증 및 소통하는 과정을 통해 편견의 이유나 원인을 찾을 수 있고 이를 개선하는 것이 가능해 진다(투명성)고 본다. 이 중에서도 투명성은 모든 요건의 전제가 된다는 점을 고려할 때, 그 법적 또는 헌법적 근거를 명확히 하는 것이 인공지능 규제의 법적 기초가 될 수 있다. 다음으로는 투명성의 법적 근거와 투명성이 내포하는 법적 의미를 확인하고자 한다.

3. 투명성 원칙에 입각한 인공지능 규제의 핵심적 요건

투명성에 관한 내용이 법적 근거를 가질 수 있다면, 보다 법·규제적인 면에서 정당화될 수 있을 것이다. 인공지능 시대의 특징에 알고리즘에 의해 많은 의사결정이 이루어질 수 있다는 점에서 토론에 의한 설득이나 개인의 취향이나 가치를 존중하는 기존의 민주주의 사회에서의 특성이 사라질 수 있다는 우려가 포함될 수 있다. 그러나 알고리즘의 활용이 증대될수록 이를 적절하게 통제하는 법의 역할은 더욱 강조될 수밖에 없다.[30] 이에 따라 인공지능과 관련된 규제가 법규범으로서의 어떤 틀을 갖추고 이를 보편적 기준으로 삼는데 매우 유용할 수 있는 헌법적 근거

30) 인공지능 시대일수록 '법의 지배(rule of law)'는 강화되어야 한다는 글로 김광수, 앞의 논문, 231면 이하 참고.

도 필요하다. 우리 헌법상 명문의 규정은 없으나 기존의 헌법원리나 헌법원칙을 통해 투명성 원칙을 확인할 수 있다는 관점에서 투명성을 헌법원칙으로 이해하는 입장이 있는데,[31] 그 내용이 아직 헌법재판소 등을 통해 명시적으로 확인된 것은 아니라고 할지라도 새로운 헌법원칙으로 인정할만한 합리적 논증이 가능하다면 이러한 내용을 인공지능 규제에 차용하여 인공지능에 대한 규제가 진일보 할 수 있을 것으로 기대된다.

원래 헌법상 투명성 원칙은 주권의 소재가 국민에게 있음을 확인하고 국가권력행사의 정당성을 확보하는 과정에서 강조되는 헌법원칙이다. 헌법상 투명성 원칙이 강조하는 바는 무엇보다 대표자에 의한 의사결정에 대해 일체감을 가질 수 있도록 하는 것과 신뢰관계를 유지하는 것에서 출발한다. 이는 대표자의 활동과 국민의 의사가 국민에게 충분히 알려지고 대표자에게 자유롭게 전달되고, 국민이 직접 의사결정에 참여하는 상호 교류를 통해 대표자의 결정과 국민의 의사가 일치하는 효과를 발생시킬 수 있다. 결국 국가권력을 민주적으로 통제하고 대표자의 권력을 견제하기 위해 투명성 원칙이 필요하다고 평가할 수 있을 것이다.

이러한 투명성 확보를 위한 요소들을 보다 구체적으로 설시하자면, 모든 사람들의 정보획득과 이용 가능성을 편리하게 하여야 한다는 '접근성', 명료하고 평이한 용어로 표현되어 그 내용을 쉽게 파악할 수 있어야 한다는 '이해성'. 필요한 즉시 제공될 것이 요구되는 '적시성' 및 관심을 두고 있는 사안에 대해 이해할 수 있는 내용으로 이루어져야한다는 의미로 '관련성'과 내용과 질이 신뢰할 수 있는 완전한 형태이어야 함을 뜻하는 '질과 신뢰성' 등이 포함될 것을 의미한다. 이는 곧 국가기관의 의사결정과 그에 따른 접근 권한의 행사 결과가 공개되고 국가기관이 보유한 정보에 대해 국민이 최대한 접근할 수 있도록 하여야 한다는 점을 기초로 하며, 그러한 권한 행사의 절차가 투명해야 한다는 점과 공개되는 정보가 국민의 입장에서 이해가 가능하도록 질적인 완전성을 갖추어야 한다는 점, 권한행사의 과정이 열려서 폭넓은 국민의 참여가 보장되어야 한다는 점, 책임관계를 명확하게 하고 책임추궁을 위한 제도적 장치가 마련되어 있어야 한다는 점을 뜻하게 된다.

---

31) 백수원, "헌법상 투명성 원칙에 관한 연구", 박사학위 논문, 성균관대학교(2010) 참고.

결국, 국민에게 국가기관의 의사결정 결과와 과정이 공개되는 것, 그 공개가 국민의 입장에서 충분하고 정확하게 이해할 수 있도록 완전하게 이루어질 것, 국민의 의사가 국가기관에 자유롭고 유효하게 전달되고 정치과정의 개방성을 통하여 국민에게 직접적, 간접적으로 참여할 수 있는 기회가 부여되는 것, 대표자가 국민의 신뢰에 위배되는 활동을 하는 경우 이를 바로잡고 통제함으로써 책임을 갖게 하는 것으로 정리할 수 있다.

앞서 언급한 바와 같이 이러한 투명성 원칙의 내용은 비록 헌법전에 명문의 단어는 없지만 완전히 새로운 것이 아니라. 민주주의원칙, 법치주의원리와 같은 헌법의 기본원리 및 알권리와 표현의 자유, 개인정보자기결정권 등과 같은 기본권 이념에 내포된 이념으로 보고 있다.

인공지능과 관련된 법·제도를 구현함에 있어 이러한 헌법상 투명성의 원칙이 중요한 법적 근거가 될 수 있을 것으로 기대되며, 인공지능을 넘어 어떤 새로운 기술에 대응하는 법·제도를 사회 속에 수렴시킬 때 필요한 근거로 기능할 수 있을 것으로 판단된다.[32]

III. 인공지능 관련 정책 기준의 한계와 법·규범화 가능성 검토

1. 4가지 정책 기준(투명성·설명가능성·공정성·견고성)의 한계와 의의

(1) '투명성'과 '설명가능성'의 한계 및 의의

디지털 혁신으로 과거에는 쉽게 접근할 수 없었던 정보에 대한 접근이 용이해 진 것은 사실이다. 그러나 인공지능과 관련하여서는 이러한 단편적인 정보에 대한 접근성은 큰 의미가 없다. 인공지능 생산자 측에서 인공지능 기술 설계 및 인공지능의 딥러닝 결과에 대한 예상 인지 수준 등을 공개하여야만 어느 정보 접근이 이루어졌다고 평가할 수 있을 것이다. 인공지능의 구조상 다양한 집단이 협력하여 개발될 수 있는데, 이러

32) 인공지능에 의한 결정으로 헌법의 전통적 가치들이라고 할 수 있는 민주주의나 법치주의가 훼손될 수 있다는 지적이 있다. Karen Yeung, "Why Worry about Decision-Making by Machine?", ALGORITHMIC REGULATION, Oxford University Press(2019), 4면 이하 참고. 그러나 이러한 가치들을 인공시대에도 포섭할 수 있도록 하여야 한다는 것이 필자의 입장이다.

한 구조는 정보 접근에 있어 취약성을 드러낼 수밖에 없다. 투명성 부족은 프로그램이 다양해지고 하드웨어 기술이 복잡해질수록 공동의 작업으로 인해 더욱 가중될 수밖에 없다. 어떤 구성 요소가 어떻게 작동하는지에 대한 지식이 부족한 경우가 빈번히 발생할 것이고, 알고리즘의 경우에도 인공지능 학습 알고리즘의 생성에 관여한 프로그래머조차도 자동학습으로 수정되는 프로그램을 파악하지 못한다는 문제점을 안게 된다. 인공지능 기술에 있어 기본적인 사항이라고 볼 수 있는 복구기술(Reverse Enginerring)과 같은 정보 기술 시스템 또한 매우 높은 수준의 전문성과 복잡한 절차로 구현되기 때문에 정보 습득에 한계를 보일 수밖에 없다. 게다가 인공지능의 알고리즘의 경우에는 일종의 영업비밀로 치부되기 때문에 공개가 거절되는 것이 당연시 된다는 점도 인공지능 기술에 있어 투명성을 저해하는 요소가 된다.[33]

실제로, 인공지능 기술은 비교적 비슷한 프로세스로 진행되는 단순한 과정을 거치지만 일부 기술의 경우에는 설명이 어려운 것이 사실이고, 특히 딥러닝 기술과 같은 인풋과 아웃풋 사이의 관계 속에서 데이터를 분류하는 기술에 대해 투명성을 가지기는 매우 어렵다. 그러나 이러한 인공지능 또한 적어도 어떻게 그러한 결정을 내리게 되었는지 설명하는 것은 가능하여야 한다. 기술적 어려움이 있다는 점을 고려하더라도 가능한 한 투명하게 공개함으로써 신뢰를 확보하는 것이 장기적으로 보았을 때 해당 산업 발전에 더 이바지할 것이다. “Transparent AI is explainable AI”라고 주장한 예와 같이 결국 투명성은 설명가능성과 상호 호응하며, 기술 전반에 관하여 인간이 종합적으로 검증할 수 있어야 비로소 확보될 수 있다. 인공지능의 의사결정이나 그러한 결정의 생성 경위가 반드시 설명 가능하여서 인간이 이해할 수 있어야 한다는 것이 인공지능 정책에 있어 출발점이자 도착점이라고 할 수 있다.[34]

---

33) Thomas Wischmeyer&Timo Rademacher, “Regulating Artificial Intelligence”, Springer(2020), 17면.

34) 현재 기술의 발전에 따른 새로운 제도의 도입에 있어 ‘제도’와 ‘기술’ 등 여러 요인에 집중하면서 ‘인간’을 중심 주제에서 제외시키고 있다는 문제점을 지적한 글로 김광억, “인간의 존재에 대한 새로운 지평:전환적 성찰과 융합적 전망”, 지식의 지평 29호, 한국학술협의회(2020), 14면 이하 참고.

(2) '공정성'과 '견고성'의 한계와 의의

인공지능에 대해 우리는 흔히 이들 인공지능이 객관적일 것으로 생각한다. 그러나 현실적으로 인공지능은 알고리즘이 가진 편향성이나 학습 진행 방향에 따라 매우 편향될 가능성을 안고 있다.[35] 또한 인공지능이라고 해서 언제나 정답을 내지는 못한다. 때때로 실수도 발생할 수 있는데, 투명성이 없이는 이러한 오류 발생에 대한 확인이 사실상 불가능하다. 인공지능을 매우 다양한 영역에 투입시킬 수 있다는 점을 고려할 때, 생명을 해치거나 다른 큰 위해를 가할 수 있는 가능성도 충분히 존재한다. 이러한 위험에 대해 숨기려고 하거나 다소 소극적인 방식으로 해결하려고 할수록 사람들이 느끼는 두려움은 커질 수 있고, 사회적 불안이 가중될 수 있다. 따라서 바람직한 해결을 위해서 이러한 잠재적 위험에 대한 공개는 필요하다. 공개를 통해 문제점을 노출시키고 이에 대한 공론화된 대응책을 마련할 수 있는 법·제도적 장치를 마련하는 것이 인공지능 활용의 전제가 되어야 함은 당연한 사실이다. 게다가 이러한 개방적 정책이 오히려 인공지능을 반대하는 사람들을 설득할 수 있는 매우 적절한 대응일 수 있다. 바람직한 규제로 인공지능 산업을 더욱 확장시킬 수 있고 잠재적 위험에 대한 리스크를 줄이는 효과도 기대할 수 있다. 뿐만 아니라 완전히 새로운 산업의 쇠락을 불러올 수도 있고 정책이나 규제에 있어 완전히 새로운 패러다임을 적용시키는 방향 전환을 일으킬 수도 있다.[36] 이러한 과정에서 투명성 확보는 산업 및 정책 전반에 대한 공정성과 견고성을 공고히 하여 위험을 예측하고 대비할 기회를 제공하는 역할을 할 것이다.

---

35) 본질적으로 인간 자체가 잘못을 저지르기 쉬운 존재이며, 이에 대한 현실 인식을 통해 상호 이해와 소통이 생겨난다. 인공지능에 대해서도 이러한 시선이 필요하다. 인간 존재의 부족함과 존재론에 관하여 K.J. Vanhoozer(역:조형국), "인간존재, 가능성 그리고 시간", 철학과 문화 제12집, 한국외국어대학교 철학연구소(2006), 197면 이하 참고.

36) 예컨대, 인공지능을 채용한 자율주행 차량에 대한 사고를 모두 회사가 책임지겠다는 자동차회사의 방침이 현실화된다면 결국 자동차보험은 매우 위축될 수 있을 것이다. 자동차회사 볼보(VOLVO)는 2015년 "자율주행 모드에서 사고가 발생한다면 우리가 모두 책임을 질 것"이라고 밝히면서 "자동차를 해킹하는 행위는 형사 범죄로 간주될 것"을 강조하였다. 그러면서도 "바이러스 침투로 해킹이 발생하더라도 애플이나 MS로 책임을 돌리지 않겠다"고도 밝힌 바 있다. 산업 패러다임의 변화를 엿볼 수 있는 발언이라고 평가할 수 있다. BLOTER, "볼보 "무인차 사고 나면 책임지겠다"" BLOTER(2015.10.08.), https://www.bloter.net/newsView/blt201510080006

2. 4가지 정책 기준의 법·규범화 가능성

인공지능 규제에 있어 위 4가지 정책 기준의 의의를 확인하였으나, 결국 그 실현은 기술의 영향이 클 수밖에 없다. 인공지능 기술 소스 자체를 공개하는 것이 하나의 방안이 제시되고 있다. 최근 이슈가 되고 있는 AI 투명성 (AI Transparency)에 관한 이론은 '기술 공개를 통한 인공지능 규제'라는 관점에서 인공지능이 투명성을 높이고자 한다. 인공지능이 '어떤 방식'으로 '무엇을 결정'하는지 공개되어야 한다는 개념으로 기술 공개를 통해 인공지능에 대한 이해를 높이는 것을 목적으로 한다.[37] 그러나 이러한 기술 공개에 우려를 표명하는 학자도 있는데, 인공지능 기술을 공개하여 이용자 친화적인 인공지능(User-Friendly AI)이 탄생하게 되면 결국 무책임한 인공지능 이용자도 증가할 것이기 때문에 더 큰 고민거리가 될 수 있다는 지적이 그것이다.

그러나 이러한 오픈 소스 방식의 정책을 통해 결국은 보다 나은 인공지능 기술에 대한 이해와 그 이해에 기초한 규제를 찾아낼 수 있을 것으로 생각한다. 사람보다는 적겠지만 인공지능도 실수를 하기 때문에 왜 그러한 실수가 일어나는지 공개적으로 알아야만 그에 합당한 규제도 나올 수 있다고 본다.[38] 결국 인공지능이 제시하는 결과물에 대한 의미있는 추론이나 계획 등을 설명하기 위해서 알고리즘의 최적화가 필요하고 이러한 알고리즘 최적화는 오픈 소스(open source) 방식의 투명성 확보를 통해 가능하다고 본다.[39]

더하여, 인공지능을 효과적으로 관리하고 규제하기 위하여 알고리즘 제작단계에서부터 문제가 발생했을 때 대응 조치 및 폐기에 이르기까지

---

37) 인공지능관련 논의는 결국 알고리즘을 중심으로 진행되는데, 알고리즘은 인간 뇌의 구조 및 기능과 관계가 없고 매우 다른 것으로 이해하는 입장이 있고, 반대로 인간의 인식과 뇌의 활동과 인공지능의 알고리즘이 그 과정상 유사하다고 보는 입장도 있다. 류지웅, "인공지능(AI)로봇의 법적 문제에 관한 연구-EU의 RoboLaw의 입법동향을 중심으로-", 토지공법연구 제78집, 한국토지공법학회(2017), 305~305면 참고.

38) Richard Roovers, "A Call for Transparency and Responsibility in Artificial Intelligence", Deloitee(2019), 7~8면.

39) 오픈 소스 방식은 평등주의, 능력주의, 자기조직화라는 세 가지 원칙에 기반한 협업으로 정의될 수 있다. Riehle, D(ET AL), "Open Collaboration within Corporations Using Software Forges", Volume26 issue2, IEEE Software(2009), 52-58면. 자유로운 협업을 통해 더 나은 소프트웨어를 만들어 내고 품질을 향상시키기 위해 주로 활용된다.

전 과정을 통해 투명성이 확보되어야 한다.[40] 투명성을 기초로 하는 인공지능 규제를 법·규범화하고 이를 적용하기 위해서는 결국 어느 정도의 공개나 설명 또는 공정, 견고성이 있는 것을 투명성을 갖춘 것으로 판단할 수 있는지에 관한 문제이고, 이에 관한 구체적인 가이드라인이 필요한 이유이기도 하다. 결국 인공지능 제조사 및 사용자가 이해할 수 있고 납득할 수 있는 구체적 가이드 마련을 통해 투명성은 현실화되는 것이고, 이러한 기준은 향후 재판 또는 조정 과정에서 적법 여부에 관한 검토에서부터 입증책임의 문제까지 영향을 주는 중요한 정책 결정 방향이라고 할 수 있다. 따라서 아래에서는 인공지능 규제지침(가이드라인) 마련을 위한 전제로 투명성 확보를 위한 구체적 구성 요소를 확인하고자 한다.

① 설명가능성(Explainability)

인공지능 기술을 다루는 기업은 인공지능의 정보를 담고있는 이른바 '블랙박스'를 열어야 한다. 기업 내부에서 무슨 일이 일어나고 있는지, 어떠한 자료와 데이터를 가지고 의사소통이 이루어지는지를 소비자들가 보기 쉬운 곳에 공개할 필요가 있다. 이를 통해 인공지능의 결정이 공정하고 정확하며, 신뢰할 수 있음을 사용자에게 알려야 한다.[41]

② 의사소통(Communication)

인공지능이 더욱 보편화되기 위해서, 그리고 투명성을 확보하기 위해서 기업은 인공지능 제품이 주는 명백한 혜택과 부작용을 사용자들에게 명확한 메시지로 제공하여야 한다. 인공지능 제품이 어떻게 이러한 혜택을 제공하는지 설명하여야 하고 이를 통해 사용자가 겪을 불편이나 감수

40) 인공지능 반려로봇의 생애 전 단계에서의 관리를 강조한 글로 백수원, "전자인격체로서 반려로봇(섹스로봇)의 규제 기준 마련을 위한 시론적 고찰", 법학논총 제32권 제3호, 국민대학교 법학연구소(2020), 125면 참고.

41) 글로벌 온라인 지불 시스템을 운영하는 페이팔(PayPal)의 수석 부사장인 Sri Shivananda는 "고객이 인공지능 서비스를 제공하는 플랫폼이나 기업을 신뢰할 수 있을 때, 그들이 체험하는 경험의 이면에 있는 인공지능 또한 신뢰하기 시작합니다."라고 하였다. 결국 인공지능을 제공하는 기업과 소비자와의 신뢰가 쌓이면, 인공지능 기술을 수월하게 받아들일 수 있다고 본다.

할 위험을 알려야 한다. 결국 사용자가 접하는 시스템이나 실행 방식, 문제 해결 방법 등에 다양한 요소에 대해 사용자와 기업이 상호 소통할 수 있어야 신뢰를 형성할 수 있게 된다고 본다.

③ 인공지능의 목적(Purpose)

인공지능은 학습을 통해 발전하면서 사용자의 생활을 변화시킨다. 거주지역, 수입, 자주가는 공간 등 다양한 요소들을 수치화 해 대상을 분석하여 필요한 서비스를 제공하게 된다. 다만, 이 과정에서 사회적 인식이나 윤리와 같은 사항이 배제될 수 있는데, 이로 인해 사용자가 생각했던 '목적'과 다른 결과가 나올 수 있다.[42] 그렇게 된다면 사용자는 인공지능 기술에 대한 신뢰를 잃어버리게 될 수 있다. 따라서 인공지능의 행동에 대한 오해가 생기지 않도록, 프로그래밍된 인공지능의 '목적'을 공개해야 한다.[43]

④ 개인정보보호: 데이터 프라이버시(Data Privacy)

오늘날 디지털 비즈니스 영역에서 중요한 이슈 중 하나는 바로 개인정보보호와 보안이라고 할 수 있다. 해킹이나 개인정보 유출을 목적으로 하는 다양한 악성앱 등으로 인한 사생활 침해나 개인정보 유출이 걱정되는 수준이다. 지능정보사회일수록 데이터가 어떻게 수집되고 어디로 전달되는데 대한 공포가 수반될 수 있는데, 투명성, 보안수준, 사용자에 대한 관심 등을 고려한 정책이 요구된다.

---

42) 다만, 순간적인 판단이 요구되는 딜레마적 상황(A를 치면 B가 다치는 상황이고 B를 치면 A가 다치는 상황)에서 인공지능의 대응은 자칫 윤리적 논쟁으로 번질 수 있다. 이를 해결하는 인공지능의 상황 판단에 대하여 어떤 알고리즘이 더 정당하다고 할 수 있을지 검토하는 일은 제조사 입장에서도 매우 어려운 결정이다.

43) 미국 승차 공유 서비스 기업 리프트(Lyft)의 자전가 및 스쿠터 책임자 인 데이비드 포스터(David Foster)는 이어서 그는 "목적없는 AI는 가치가 없으며, 인공지능이 우리 삶에 가치를 더하지 않는다면 사용자들은 그것을 단순히 우리 삶에 침입하는 것으로 보고 거부할 것"이라고 말했다. 그는 기술의 '목적'을 투명하게 공개하여 소비자가 직접 서비스의 결과를 평가할 수 있도록 해야 한다고 주장하기도 하였다.

⑤ 인터페이스(Interface)

사용의 용의성을 의미하는 인터페이스의 편의성 또한 투명성에 있어 중요한 요소이다. 인공지능 제품 조작이 직관적이고 작동하기 쉽게 만들어졌다는 것 자체만으로도 사용자에게 신뢰감을 주기 때문이다. 인공지능 기술에 대한 거부감을 줄이는데 이러한 사용의 용의성이 유의미한 이유이다.

3. 인공지능 규제 현황과 한계 및 대응방안

(1) 해외 인공지능 관련 가이드라인 등 검토

투명성은 규제의 목적을 명확히 하는데 도움이 되고, 규제대상의 순응성을 확보하는 장점이 있는 반면에 불필요한 규제에 대한 논의를 촉발시킬 수도 있다.[44] 따라서 각 국의 상황을 검토하고, 관련 규제를 만들거나 적용시키는 식으로 대처해가는 접근이 안전하다. 현재까지 진행된 인공지능 관련 세계적 추이를 비교 검토하여 바람직한 규제 방안을 마련하는 것이 바람직할 것이다.

2019년 5월 개최된 OECD 각료이사회(MCM)에서 신뢰 가능한 인공지능을 위한 가이드라인인 'OECD AI 권고안'을 OECD 회원국 만장일치로 공식 채택하였는바, OECD AI 권고안에서는 인공지능과 관련된 용어를 우선 정의했으며, 신뢰 가능한 인공지능 구현을 위한 5가지 원칙과 5가지 정책 및 국제협력 방향으로 구성되어 있다. 신뢰 가능한 인공지능 구현을 위한 5가지 원칙 중 하나가 투명성 확보, 설명 가능성 확보(Transparency and explainability)이다. 이를 구체적으로 설명하면 인공지능 행위주체는 인공지능 시스템에 대한 이해를 증진시키고 감춰진 것이 없도록 투명성과 설명 가능성을 확보하여야 한다. 이를 위해 인공지능 행위주체는 인공지능 시스템의 전체 맥락을 잘 설명하는, 최신의 의미가 있는 정보를 제공해야 한다. 이를 통하여 이해관계자가 인공지능 시스템에 대한 이해를 촉진하고, 일터에서 이뤄지는 인공지능 시스템과의 상호작용을 인식하게 하고, 인공지능 시스템의 결과에 대하여 이의를

44) 김일환·백수원, "글로벌 IT기업의 투명성 활동에 관한 헌법적 검토", 법학논총 제23권 제1호, 조선대학교 법학연구소(2016), 147면.

제기할 수 있어야 하며 이해하기 쉬운 정보와 예측 또는 결정의 근거로 사용된 논리 등 설명을 요구할 수 있어야 한다.[45]

한편, EU 집행위원회(European Commission)에서도 이른바 '신뢰할 수 있는 인공지능 윤리 가이드라인'을 발표하였다.[46] 동 가이드라인에서는 투명성을 높이기 위한 전제로 추적가능성과 설명가능성 및 의사소통을 요구하였다.

OECD 권고안 및 EU 집행위원회(European Commison)에서 제시하는 투명성을 강화하기 인공지능 관련 정책을 정리해보면, i)인공지능(AI)의 사용을 고지, ii)설명을 들을 권리의 인정, iii)정보에 대한 접근권 인정, iv)종합적인 AI 투명성 정책 그리고 v)AI 기반 결정의 내역을 기록하고 이를 감독하는 전문 감독기관의 설치로 해석 가능하다.[47]

전문가들은 알고리즘이 내린 의사결정으로 개인의 권리가 침해받을 경우 그에 대한 설명을 요구할 수 있는 권리를 보장해야 한다고 보고 있다.[48] 현재는 알고리즘이 특정한 의사결정을 했을 때 그 결정 과정에 어떤 변수를 계산 요소로 포함시키고, 변수 간의 관계를 어떻게 설정했는지를 알 수 없기 때문에 이를 보여주고 설명해 달라고 요청하는 권리를 법적으로 인정할 수 있을 것인가가 논의의 출발점이 될 것이다.

유럽연합의 개인정보보호규정(GDPR)에는 이런 권리를 '설명받을 권리', '프로파일링(알고리즘으로 개인의 특성을 분석하는 기법)을 거부할 권리'로 인정하고 있으며, 나아가 독일에서는 자동결정장치(automated decision making system:ADMs)를 적용하는 경우에는 그 장치의 로직(logic)과 분류기(classifiers) 및 입력 데이터에 적용되는 가중치를 공개할 것을 요구하는 입법이 필요하다는 의견이 주장되고 있다. 프랑스에서도 알고리즘을 이용한 결정의 경우 상대방은 그 특성에 관한 설명을 제

45) 한국정보화진흥원, "신뢰 가능 AI 구현을 위한 정책 방향 - OECD AI 권고안을 중심으로 -", 한국정보화진흥원(2019), 29면 이하.

46) 한국정보화진흥원, "인공지능 윤리가이드라인 - 일본과 사례를 중심으로 EU-", 한국정보화진흥원(2019), 26면 이하 참고.

47) Thomas Wischmeyer, 앞의 논문, 76면 이하 참고.

48) 인공지능에 의한 의사결정에 대한 거부감은 주로 그 알고리즘의 비공개성에서 기인한다. 예를 들면, 인공지능 판사를 도입하는 일이 재판의 공공성과 공개성을 침해할 것이기 때문에 신뢰할 수 없다는 지적이 있다. Michael Kirby, "The Future of Courts-Do They Have One?", Journal of Law and Information Science Volume 9 Issue 2(1998), 143면 참고.

공받을 필요가 있고 그 작동원리에 대한 투명성이 필요하다는 주장이 제기되고 있다.[49]

기업의 자율적 AI 규제가 아니라 국가기관이 나서서 이용자를 보호하는 법적 규범을 명확히 해야 한다는 의견도 있다. 호주 국가인권위원회(Australian Human Rights Commission)는 2019년 12월 발표한 보고서에서 "기업의 자율적 조치들은 특히 인권보호 측면에서 그 효과성이 의심을 받고 있다"며 "새로운 기술의 사용이 초래할 수 있는 심각한 위험성을 고려할 때 자율적 규제 틀은 그 자체로 인권을 보호하고 증진하는 데 불충분하다"고 밝혔다. 따라서 인공지능이 특히 개인의 권리에 중대한 영향을 미치는 행정적·산업적인 의사결정에 사용될 때 마땅히 그 의사결정에 대해서 당사자들이 근거를 밝히도록 요구하고, 설명을 들을 수 있어야 한다"며 "호주 인권위의 입장은 그보다 한 단계 더 나아가 법적 책임을 요구하고, 문제 제기를 할 수 있도록 법 규범을 마련해야 한다는 것"을 강조한다. 결국 인공지능과 관련하여 자율규제가 필요한 것이 아니라 법이 필요하다는 점을 강조하고 있는 것이다.[50]

그러나 기업들은 알고리즘을 자신들의 지적재산권이나 영업비밀이라 주장하면서 공개를 거부할 수 있다.[51] 네이버의 경우 언론사들의 기사배열 알고리즘 공개 요구에 대해 영업비밀 등을 이유로 부정적인 입장을 밝힌 바 있다. 오늘날 기본권의 대사인적 효력이 대국가적 효력만큼이나 기본권 효력의 보장적 관점에서 중요한 이유는 결국 이러한 거대 민간세력의 개인에 대한 영향력이 커지고 있음에도 개인 대 개인간의 일로만 치부하여 개인의 기본권 보장에 소홀할 수 있기 때문이다.[52]

---

49) Thomas Wischmeyer, 위의 글, 82~97면 참고.

50) James Eyers, "We need laws about AI, not self-regulation", FINANCIAL REVIEW(2019.12.17.), https://www.afr.com/technology/we-need-laws-about-ai-not-self-regulation-20191216-p53k8r

51) 기업이 스스로의 영업을 더 좋은 환경에서 구현하기 위해 영업 지침 등을 보다 사회윤리에 부합하도록 개정하는 경우도 분명히 있다. 구글, 네이버 등 인터넷 기업은 자율적 가이드라인을 만들어 사용자에게 신뢰감을 갖도록 하는 등 이해충돌적 상황이나 비윤리 비난으로부터 탈피하고자 한다. Wolfgang Hoffmann-Riem, "Artificial Intelligence as a Challenge for Law and Regulation", REGULATING ARTIFICIAL INTELLIGENCE, Springer(2020), 8면 이하 참고.

52) 주영재, "인공지능에 방울 달기, 알고리즘 투명성·설명권 보장으로" 경향신문(2020.02.15.), http://news.khan.co.kr/kh_news/khan_art_view.html?art_id=202002151423001

(2) 인공지능 관련 규범적 한계점과 대응 방안 검토

우리 정부가 2020년 12월 인공지능 국가전략을 발표하고 '사람 중심의 AI'를 강조했지만, 구체적인 실행 계획은 후일로 미뤘다. 기본적으로 AI 규제 논의가 지지부진한 데는 'AI 군비경쟁'도 한몫하고 있다. 국가별로 치열한 AI 산업 역량 강화에 더하여 군사적 목적의 이용에 대한 경쟁까지 가속화되는 상황에서 AI에 대한 투명성 확보가 어려울 것은 당연한 일인지도 모른다. AI 경쟁 속에서 AI를 윤리적으로 사용해야 한다는 주장은 한가로운 소리로 여겨질 수 있다는 뜻이다. 전문가들은 알고리즘이 내린 의사결정으로 개인의 권리가 침해받을 경우 그에 대한 설명을 요구할 수 있는 권리를 보장해야 한다고 주장하지만, 알고리즘이 특정한 의사결정을 했을 때 그 결정 과정에 어떤 변수를 계산 요소로 포함시키고, 변수 간의 관계를 어떻게 설정했는지를 알 수 없다. 이를 설명해 달라고 요청하는 권리를 법적으로 인정할 수 있을 것인가 하는 문제는 기술적인 검증이 보다 강하게 요구되는 바이다.

또한, 기업들은 인공지능 알고리즘을 자신들의 지적재산권이나 영업비밀이라 주장하면서 공개를 거부할 수 있다. 네이버의 경우 언론사들의 기사 배열 알고리즘 공개 요구에 이런 이유를 들어 부정적인 입장을 밝힌 바 있다. 결국 기업의 목표가 '이윤 추구'라는 기본 도식이 바뀌지 않는다면 '착한 인공지능'은 얼마든지 '이윤을 극대화하는 인공지능'으로 대체될 수 있다는 우려가 설득력을 갖는다.

따라서 기업의 자율적 인공지능 규제만으로는 분명한 한계가 있다. 물론 기업 스스로가 윤리의식을 가지고 투명성 확보가 가능한 형태로 운영하기를 기대하고 싶지만 기업의 자율적 조치들은 특히 인권 보호 측면에서 그 효과성이 의심을 받고 있다. 인공지능이 새로운 기술이고, 이러한 새로운 기술의 사용이 초래할 수 있는 심각한 위험성을 고려할 때 자율적 규제 틀은 그 자체로 인권을 보호하고 증진하는 데 불충분하다.

결국 인공지능이 특히 개인의 권리에 중대한 영향을 미치는 행정적·산업적인 의사결정에 사용될 때 마땅히 그 의사결정에 대해서 당사자들이 근거를 밝히도록 요구하고, 설명을 들을 수 있어야 하며, 한 단계 더 나아가 법적 책임을 요구하고, 문제를 제기할 수 있도록 법·제도를 마련해야 한다.[53] 아울러 법치주의는 상호 견제와 감시에 의하여 이루어진다는

점을 고려한 정책이 마련될 필요가 있다. 국가의 권력을 입법, 행정 및 사법으로 나누는 목적도 서로 견제를 하여 국민의 권익 보장과 사회 질서 유지에 기울어짐이 없도록 하기 위해서이다. 인공지능에 관련하여 그 업무를 입법부와 행정부 그리고 사법부가 고유한 영역에서 법령에 이를 추후에 심사하는 순환 구조가 이루어져야하는 것은 당연하고, 이에 더하여 인공지능의 특성상 보다 전문적인 감독이 필요하다. 법의 해석과 적용은 생각보다 단순하지 않으며, 창조성과 유연성을 요구하는 경우가 빈번하다. 법률 정보의 단순 제공이나 분류·가공하는 정도가 현 단계에서 인공지능에게 기대되는 역할이다.[54] 인공지능이 법률 사무를 수행하는 경우도 있다고 하지만 법조인을 완전히 대체하지 못할 것이라는 주장이 설득력을 얻는 이유도 바로 이러한 점 때문이다. 결국은 공감 능력을 통해 법의 해석과 적용은 완성된다고 보는데,[55] 이 점에서 인공지능 규제에 투명성을 통한 소통이 필요한 이유와 연결된다. 인공지능 관련 규제 기관은 단순한 분쟁 조정을 넘어 결과적·절차적·내용적·책임의 투명성까지 두루 심사·판단하는 기능을 수행하게 될 것이다. 최소한 긴급한 침해의 예방과 저지를 위해서도 인공지능의 특수성을 고려한 규제 기관이 필요하다고 본다.[56]

---

53) 인공지능의 활용은 인간 생활에 이익을 주지만 위험과 위협이 더 커질 수 있음을 경고한 글로 Thomas Wischmeyer, Regulating Artificial Intelligence, Springer(2020) 참고.

54) 김기영, “인공지능 활용 법률정보서비스의 발전과 관련 쟁점”, 문화미디어엔터테인먼트법 제11권 제1호, 중앙대학교 문화미디어엔터테인먼트법연구소(2017), 98면.

55) 강승식, “인공지능 판사, 과연 가능한가?”, 헌법학연구 제26권 제3호, 한국헌법학회(2020), 234면.

56) 김광수, 앞의 논문, 239면.

IV. 결론

과학 기술과 인간 중 누가 멀리 내다볼 수 있는, 소위 넓고 긴 안목의 진화에서 그리 현저한 성취에 앞서고 있는 것일까. 과학 기술을 창조한 것은 인간이지만 인간의 사고체계 도전할 수 있을 정도의 과학적 진보는 인간에게 두려움까지 안겨주고 있다. 인간은 구체성을 갖추지 않은 미래에 대하여 예측할 수 있는 능력이 있다. 즉 생명이 걸린 위험이 다가오기 아주 훨씬 전에도 그리고 아주 먼 거리에서 아무런 상관성도 없는 듯이 보이는 작은 조짐으로부터도 인간은 예측 능력을 발휘하여 불안을 느끼고 그에 대한 대응책을 마련한다. 그러나 점차 과학과 기술에 예측 능력을 양도함으로써 어떤 사태가 임박했을 때 기술을 사용하여 임기응변적 반응기제를 기민하게 동원하는 것에 익숙해지기 시작하였다. 인간은 보다 확실한 예측을 원했고, 그로 인해 스스로 판단하기보다는 기계나 과학적 기술 또는 시스템적 사고에 익숙해져 갔다.

이로 인해 인간의 직관력은 다소 약화되고 과학 의존도는 심화되어 갔다. 또한 멀리 떨어진 혹은 먼 미래에 대응하고자 하는 생각보다는 눈에 보이는 현상에 대응하고자 하는 생각이 더 설득력을 얻었다. 현실에 충실해야 한다는 가치관이 '지금 여기'를 강조함으로써 지구와 인류 전체의 운명을 바라보는 먼 시간적 전망에 소홀하며 인공지능과 관련된 이슈에 있어서도 기술technology 영역의 문제로만 보고 선도적 규제보다는 발견된 문제를 해결하는 방향의 규제를 선호한다.

그러나 인공지능 기술의 탄생으로 우리는 매우 새로운 기술과 새로운 윤리기준을 세울 수 있는 혼돈의 시기를 겪고 있다. 이 혼돈의 시기의 끝이 언제인지 모를 만큼 빠르게 인공지능시대가 도래할 수 있다는 두려움도 있고, 기술이 법의 영역에서 어떻게 도입·적용·구현될 수 있을지에 대한 고민이 무르익어가고 있다.[57] 그러나 한편으로는 현재가 곧 인공지능 규제를 올바른 방향으로 가져갈 수 있는 순간이라고 볼 수도 있다.

인공지능을 활용함에 있어 효과적인 규제를 논의하는 것은 이제부터라

57) Robin Widdison, "Electronic Law Practice: An Exercise in Legal Futurology", Volume 60 Issue 2, Modern Law Review(1997), 143면

고 볼 수 있다. 투명성은 헌법원칙인 투명성 원칙에서 근거하며, 인공지능 규제의 바람직한 방향을 제시할 수 있는 하나의 방안이 될 수 있다고 본다. 본문에서 구체적으로 인공지능 투명성 강화를 위한 정책의 5단계로 AI의 사용을 고지, 설명을 들을 권리의 인정, 정보에 대한 접근권 인정, 종합적인 AI 투명성 정책 그리고 AI 기반 결정 내역을 기록하고 이를 감독하는 전문 감독기관의 설치 등을 제시하였다.

한계가 예상되는 점도 있으나 헌법의 투명성원칙에 기초한 규제가 인공지능 산업 발전적 측면에서나 규제의 효율적 측면에서 가장 주요한 가치임을 공감하고 이를 추진해야 한다. 로봇에 대한 규제는 규제 영역에 있어 새로운 패러다임을 요구한다는 점에서 큰 도전 과제이다. 그러나 기술과 규제 사이의 간극을 해소하는 일이 인간을 위하여 법이 할 수 있는 최선의 배려일 것이다.

(논문투고일: 2021. 6. 5., 심사개시일: 2021. 6. 9., 게재확정일: 2021. 6. 23.)

백수원

인공지능, 투명성, 공개성, 설명가능성. 공정성

【참 고 문 헌】

Ⅰ. 국내문헌

강승식, 인공지능 판사, 과연 가능한가?, 헌법학연구 제26권 제3호, 2020

김광수, 인공지능 규제의 법체계, 토지공법연구 제93집, 2021

김광억, 인간의 존재에 대한 새로운 지평:전환적 성찰과 융합적 전망, 지식의 지평 제29호, 2020

김기영, “인공지능 활용 법률정보서비스의 발전과 관련 쟁점”, 문화 미디어 엔터테인먼트법 제11권 제1호, 2017

김민배, AI 로봇의 법적 지위에 관한 쟁점과 과제-Bryson 등의 법인격 이론을 중심으로-, 토지공법연구 제87집, 2019

김민호·이규정·김현경, 지능정보사회의 규범설정 기본원칙에 대한 고찰, 성균관법학 제28권 제3호, 2016

김연식, 과학 기술의 발달에 따른 탈인간중심주의 헌법학 모색, 헌법학연구 제26군 제4호, 2020

김일환, 지능정보사회에서 헌법의 역할과 기능, 성균관대학교 제32권 제3호, 2020

_____· 백수원, 글로벌 IT기업의 투명성 활동에 관한 헌법적 검토, 法學論叢 제23권 제1호, 조선대학교 법학연구소, 2016

_____· 백수원, 헌법상 투명성원칙에 관한 시론적 고찰, 공법연구 제39집 제2호, 2010

류지웅, 인공지능(AI)로봇의 법적 문제에 관한 연구-EU의 RoboLaw의 입법동향을 중심으로-, 토지공법연구 제78집, 2017

박기주, 제4차 산업혁명 시대 인공지능(로봇)의 법적 지위-새로운 인(人) 개념의 법적 설계 가능성을 중심으로, 한국지식재산연구원, 2017

백수원, 전자인격체로서 반려로봇(섹스로봇)의 규제 기준 마련을 위한 시론적 고찰, 국민대학교 법학논총 제32권 제3호, 2020

_____, 전자인간 및 전자인격 인정에 따른 법적 논의와 시사점 고찰, 한국외국어대학교 외법논집 제43권 제4호, 2019

_____, 헌법상 인공지능 규제의 범위와 방향에 대한 시론적 고찰, 미국헌법연구 제29권 제2호, 2018

_____, 헌법상 투명성 원칙에 관한 연구, 성균관대학교 박사학위 논문, 2010

양천수, 현대 지능정보사회와 인격성의 확장, 동북아법연구 제12권 제1호, 2018

이화순, 우리 인간의 의무는 '존재'하는 것이 아니고 '사는 것'이다. 소프트웨어 세계 제6집 한국정보기술전문가협회, 2003

정 완, 인공지능 로봇에 관한 법적 고찰, 경희법학 제55권 제4호, 2020

제인구달(역:박순영), 제인구달 침팬지와 함께 한 나의 인생, 사이언스북스, 2014

한국정보화진흥원, 인공지능 윤리가이드라인 - 일본과 사례를 중심으로 EU -, 2019

_____, 신뢰 가능 AI 구현을 위한 정책 방향 - OECD AI 권고안을 중심으로-, 2019

Ⅱ. 국외문헌

Gunther Teubner, Rights of Non-humans? Electronic Agents and Animals as New Actors in Politics and Law, 33 Journal of Law and Society 497, 2006

John Chipman Gray(edited by Davide Campbell and Philip Thomas), The nature and sources of the law. Routledge. 1997

Karen Yeung and Martin Lodge, Algorithmic Regulation: An Introduction, in: Karen Yeung / Martin Lodge, Algorithmic Regulation, Oxford, 2019

Karen Yeung, Why Worry about Decision-Making by Machine?, in: Karen Yeung / Martin Lodge, Algorithmic Regulation, Oxford University Press, 2019

Kevin Vanhoozer(역:조형국), 인간존재, 가능성 그리고 시간, 철학과 문학 제12집, 2006

Peter Asaro, Robots and responsibility from a legal perspective, T

he IEEE conference on robotics and automation, workshop o n roboethics, Rome. http://www.peterasaro.org/writing/ASARO%20Legal%20 Perspective.pdf, 2007

Raja Chatila ; Kay Firth-Butterflied ; John C. Havens ; Konstantino s Karachalios, The IEEE Global Initiative for Ethical Consid erations in Artificial Intelligence and Autonomous Systems, I EEE Robotics & Automation Magazine Volume 24 issue 1, March 2017

Richard Roovers, A Call for Transparency and Responsibility in Ar tificial Intelligence, Deloitee, 2019

Riehle, D.; Ellenberger, J.; Menahem, T.; Mikhailovski, B.; Natchet oi, Y.; Naveh, B.; Odenwald, T."Open Collaboration within C orporations Using Software Forges, IEEE Software, Volume2 6 issue2, 52-58, 2009

Robin Widdison, Electronic Law Practice: An Exercise in Legal Fut urology, Volume 60 Issue 2, 1997

Thomas Wischmeyer, Artificial Intelligence and Transparency: Ope ning the Black Box, in: Thomas Wischmeyer

Thomas Wischmeyer·Timo Rademacher, Regulating Artificial Intellig ence, Springer, 2020

Ugo Pagallo, Apples, oranges, robots: Four misunderstandings in t oday's debate on the legal status of AI systems, The Royal Society, 2018, https://www.researchgate.net/publication/328292234_Apples_oranges_robots_Four_misunderstandings_in_today%27s_debate_on_the_legal_status_of_AI_systems

Wolfgang Hoffmann-Riem, Artificial Intelligence as a Challenge for Law and Regulation, in: Thomas Wischmeyer Timo Rademach er (editors), Regulating Artificial Intelligence, Springer, 2020

Ⅲ. 기타

경향신문, http://news.khan.co.kr/kh_news/khan_art_view.html?art_id=202002151423001

두산백과, https://terms.naver.com/entry.naver?docId=1135955&cid=40942&categoryId=31611

위키백과, https://ko.wikipedia.org/wiki/%EB%B6%88%EC%BE%8C%ED%95%9C_%EA%B3%A8%EC%A7%9C%EA%B8%B0

BLOTER, https://www.bloter.net/newsView/blt201510080006,

FINANCIAL REVIEW, https://www.afr.com/technology/we-need-laws-about-ai-not-self-regulation-20191216-p53k8r,

Abstract

The direction of AI regulation based on the Constitutional Transparency Principle

Baek, Soowon(Professor of Songwon University)

Who is ahead of such remarkable achievements in the far-sighted, so-called broad-sighted evolution of science and technology or humans? Humans are the creators of science and technology, but scientific progress that can challenge the human mind system is causing fear in humans. Humans have the ability to predict the future that is not concrete. That is, even long before a life-threatening danger approaches, and even from the seemingly irrelevant omen from a great distance, human beings use their predictive ability to sense anxiety and prepare countermeasures. Gradually, however, by delegating predictive powers to science and technology, they began to become accustomed to using technology to agilely mobilize ad-hoc mechanisms when an event was imminent. Humans wanted more certain predictions, and as a result, they became accustomed to machines, scientific technology, or systemic thinking rather than judging themselves.

This weakened human intuition somewhat and deepened scientific dependence. In addition, the idea of responding to visible phenomena was more persuasive than thinking about responding to distant or distant futures. By emphasizing the "here now" value of being true to reality, we are focusing on the distant time outlook of the earth and the fate of the entire human world, and prefer to see only as a problem in the area of technology

in the issue of artificial intelligence and to regulate the direction of solving discovered problems rather than leading regulations.

Although limitations are expected, regulations based on the transparency principle of the Constitution must be pursued with the understanding that the most important value in terms of the development of the AI industry and the effectiveness of regulations. Regulation of robots is a huge challenge in that it requires a new paradigm in the regulatory domain. However, bridging the gap between technology and regulation would be the best care the law can do for humans.

Baek, Soowon

Artificial intelligence, Transparency, Openness, Electronic personality, Explainability