

발 간 등 록 번 호

33-9750040-000442-01

헌법이론과 실무

2023-A-9

# 인공지능 법관 관련 헌법적 쟁점의 검토



Constitutional Research Institute



헌법재판소  
헌법재판연구원





# 인공지능 법관 관련 헌법적 쟁점의 검토

연구책임자 : 김 종 현 책임연구관(제도연구팀)



# 목 차



## I. 서론 / 1

1. 연구의 배경과 목적 .....	1
가. 문제의 소재 .....	1
나. 헌법적 검토의 필요성 .....	4
2. 연구의 대상과 범위 및 구성 .....	6
가. 연구의 범위 한정 .....	6
나. 연구의 구성 .....	7
3. 연구의 방법 .....	8

## II. 인공지능, 인공지능 법관 일반 / 10

1. 인공지능의 개념과 유형 및 알고리즘의 특성 .....	10
가. 인공지능의 개념과 유형 .....	10
1) 인공지능의 개념 .....	10
가) 선행연구, 제도의 정의 .....	10
(1) 지능 및 인공지능 개념정의의 어려움 .....	10
(2) 인공지능의 주류적인 개념정의 .....	10
(3) 인공지능의 개념에 관한 새로운 접근 .....	11
나) 검토 .....	13
2) 인공지능의 유형 .....	14
나. 알고리즘의 개념과 인공지능 알고리즘의 특성 .....	16
1) 알고리즘의 개념 .....	16
2) 인공지능 알고리즘의 특성 .....	16
다. 인공지능의 위험성 .....	17

라. 인공지능의 활용 현황 .....	19
2. 인공지능 법관 .....	21
가. 법관의 개념, 사법작용의 의미 .....	21
1) 사법의 개념, 특성 및 기능 .....	21
2) 헌법과 법률이 예정하는 법관 .....	21
3) 법관의 업무 .....	22
나. 사법 분야에서 인공지능의 활용 현주소 .....	22
1) 개관 .....	22
2) 국내 .....	23
3) 주요 국가들의 동향 .....	25
4) 법관의 사건처리 단계별 인공지능의 발전 현황 .....	27
가) 소송자료의 인식 .....	27
나) 소송자료 분석을 토대로 한 판단 .....	28
(1) 결론의 도출 .....	28
(2) 이유 기재 .....	28
다. 사법 분야 인공지능 관련 위험성과 윤리적·헌법적 문제의 증대성 .....	29
라. 현 단계에서 상정할 수 있는 인공지능 법관 .....	29
마. 인공지능 법관의 주요 특성 .....	33
1) 알고리즘의 편향성 .....	33
2) 알고리즘의 불투명성 .....	34
3) 귀납적 추론 방식의 사용 .....	34
4) 정보의 코드화를 전제한 데이터 기반성 .....	36
5) 인간의 의사 작용 배제 .....	37

### III. 인공지능 법관 관련 헌법적 쟁점들 / 39

1. 논의의 전제: 인공지능 법관 도입에 관한 기대 .....	39
가. 업무 효율성의 극대화 .....	39
나. 재판의 일관성, 공정성 및 정확성에 대한 기대 .....	40

1) 재판의 공정성 및 정확성 .....	40
2) 재판부별 판단의 편차 해소 .....	41
3) 과거 판단과의 원리적 일관성 .....	42
다. 국가예산절감 .....	43
라. 소결 .....	43
2. 민주주의 .....	44
가. 민주주의 일반 .....	44
나. 국가기관과 민주적 정당성 .....	45
다. 민주주의와 관련하여 인공지능 법관이 초래하는 문제 .....	45
3. 법치주의 .....	46
4. 재판청구권 .....	47
가. 법관에 의한 재판을 받을 권리 .....	47
나. 공개재판을 받을 권리 .....	49
다. 법률에 의한 재판을 받을 권리 .....	49
라. 공정한 재판을 받을 권리 .....	50
1) 사례: State v. Loomis 판결 .....	50
2) 선행연구 .....	51
3) 국제사회의 동향 .....	52
4) 토론: 투명성을 구비한 인공지능 알고리즘 .....	53
5. 법관의 독립성 .....	54
6. 재판제도 및 실무의 본질적 변화 가능성 .....	56
가. 합의체 .....	56
나. 심급제 .....	57
다. 판례변경 .....	58
라. 인공지능 법관(들)의 설계 및 배치 문제 .....	60
7. 인공지능 법관의 판단에 의한 기본권 침해 가능성 및 책임의 귀속 문제 .....	60
가. 편향성에 따른 부당한 차별 .....	60

1) 인공지능 편향성의 원인 .....	60
가) 정제되지 못하고 편향된 데이터 .....	60
나) 오류 및 편향성을 지닌 알고리즘에 입각한 판단 .....	61
2) 인공지능 편향성에 따른 기본권 침해 가능성 .....	62
3) 인공지능 편향성의 특성과 해약 .....	63
4) 인공지능 편향성 극복에 있어서의 어려움 .....	64
나. 판단과정의 불투명성과 기본권 침해 .....	65
1) 알고리즘 불투명성의 원인 .....	65
2) 판단과정의 불투명성과 관련한 헌법적 문제 .....	65
3) 알고리즘의 불투명성 관련 사례 .....	67
4) 토론: 불투명성 지적에 관한 비판 .....	67
다. 잘못된 판단과 책임 귀속 문제 .....	68
라. 기본권 침해 주장 및 증명이 어려워질 가능성 및 적법절차 관련 문제 ..	69
1) 기본권 침해 주장 및 증명이 어려워질 가능성 .....	69
2) 적법절차와 관련한 문제 .....	70
마. 개인정보침해 .....	71
바. 인간의 존엄성 .....	72
8. 사법작용의 본질과 인공지능 법관 .....	73
가. 가치판단 및 정성적 판단 .....	73
나. 인과관계와 상관관계 .....	74
다. 개별적·구체적 판단과 통계적·집단적 판단 .....	75
라. 수량화 문제 .....	77
마. 승복가능성 .....	77
1) 법관의 결정 정당화의 원천 .....	77
2) 당해 사안에 대한 판단이 아닌 과거 데이터에 입각한 판단 .....	78

#### IV. 인공지능 법관의 도입과 관련한 고려요소 및 제도적 방안 / 80

1. 사법 분야에서의 인공지능: 고위험 인공지능 시스템으로서 강력한 규제 .....	80
--	----



2. 사전규제의 중요성과 규제방안 .....	81
3. 사법데이터의 공개와 개인정보보호 .....	82
4. 설명 가능한 인공지능 알고리즘 .....	83
가. 설명 가능한 인공지능 논의의 배경 .....	83
나. 설명 가능한 인공지능의 문제점 .....	84
다. 절충적인 대안의 모색 .....	86
1) 투명성 요건을 충족하는 인공지능 알고리즘 .....	86
2) 제한된 투명성 .....	86
5. 보조적 역할에 한정하는 방안 .....	87
가. 보조적 역할만을 부여하는 방안에 관한 논의 .....	87
나. 보조적 역할 한정과 관련한 문제점 .....	88
1) 알고리즘의 인간 의존성 .....	88
2) 인간의 알고리즘 의존성 .....	89
3) 인간의 알고리즘 불신 내지 일관성 없는 신뢰도의 부여 .....	90
6. 당사자의 의사를 존중하는 방안 .....	90
7. 재판의 영역마다 인공지능 알고리즘 활용 여부 및 정도를 달리하는 방안 .....	91
가. 민사사건, 특히 소액사건의 경우 .....	91
나. 헌법재판과 형사재판의 경우 .....	92

## V. 결론 / 94

■ 참고문헌 .....	96
--------------	----



## 초 록

인공지능은 인간의 지적 능력을 그대로 구현하기보다는 인간이 지능으로써 수행해온 것들을 수행하는 데 방점이 있다. 현 단계에서 자연인 법관과 동일한 인공지능 법관을 상정하기는 어려우나, 인공지능이 제시하는 출력 값이 자연인 법관의 판단을 대체하는 단계에 이르러야 인공지능 법관이라고 할 수 있다. 인공지능 법관의 특성으로는 알고리즘의 편향성, 불투명성, 귀납적 추론방식, 데이터 기반성, 인간의 의사 작용 배제가 있다.

인공지능 법관은 업무효율 제고, 재판의 일관성·공정성에 대한 기대에서 논의되나 여러 헌법적 쟁점을 안고 있다. 인공지능 법관이 민주적 정당성을 담보할 수 있는지, 코드화된 법률의 처리 결과를 법률에 의한 재판이라고 할 수 있을지, 인공지능 알고리즘의 판단 과정과 근거가 불투명할 경우 공정한 재판을 받을 권리가 침해되지 않는지의 문제가 그러하다.

인공지능 법관은 헌법이 상정해온 사법제도의 틀에 관한 근본적 성찰을 요하는 문제다. 합의제, 심급제가 형해화될 우려가 없는지, 사회현실의 변화를 감지하지 못하고 기존 판례에 갇힐 가능성이 없는지 검토해야 한다.

현재의 약한 인공지능은 사법작용의 본질에 비추어 법관이 될 수는 없다고 판단된다. 인과관계가 아닌 상관관계를 판단하며, 개별적인 당해 사안에서 구체적 타당성 있는 결론을 도출하는 것이 아니기 때문이다. 인공지능의 판단 결과에 대한 구성원들의 승복가능성도 기대하기 힘들다.

그럼에도 인공지능을 재판에 활용한다고 할 때 양질의 데이터 확보, 훈련데이터 및 기본적인 작동 원리의 설명가능성 등이 구비되어야 한다. 재판에서 인공지능을 보조적으로 활용하는 것을 두고 인공지능 법관이라고 할 수는 없다. 사건의 성격 등을 고려해 인공지능 법관의 도입 여부를 입체적으로 논의할 여지도 있다.

**주제어:** 인공지능, 약한 인공지능, 인공지능 법관, 알고리즘, 불투명성, 공정한 재판을 받을 권리, 사법작용의 본질, 보조적 활용, 제한적 도입



# I. 서론

## 1. 연구의 배경과 목적

### 가. 문제의 소재

과학기술의 발전은 경제, 사회를 비롯한 인류의 모든 영역을 변화시키며 깊은 영향을 미친다. 증기기관과 기계화를 통한 제1차 산업혁명, 전기·통신기술의 진보를 바탕으로 한 제2차 산업혁명, 컴퓨터와 인터넷 등 정보통신기술을 기반으로 한 제3차 산업혁명에 이어 인공지능(AI) 등 최첨단 기술의 융합을 통한 제4차 산업혁명으로 이어지고 있다.<sup>1)</sup>

인공지능은 이미 우리 생활의 일부가 된 현실이다.<sup>2)</sup> 일각에서는 인공지능을 공적 영역에 광범위하게 도입하려는 경향이 지극히 자연스럽다고 평가한다.<sup>3)</sup> 그러나 인공지능과 관련한 매우 다양한 측면의 위험성을 지적하는 논자도 있다.<sup>4)</sup> 인공지능의 의미 및 활용 가능성에 관한 이해가 부족한 상황에서는 주로 윤리규범과 지침 형식으로 대응해 왔으나, 최근 미국이나 유럽연합 차원에서의 규제법 제정 움직임에서 보듯 관련 규제의 변화가 감지되기도 한다.<sup>5)</sup> 한국의 경우 아직 ‘AI법’이라는 단일법률을 두고 있지 않지만, 개인정보 보호법 등 기존의 법률을 개정하거나 새로운 법률을 추가 제정하는 등의 대응이 이루어지고 있다.<sup>6)</sup> 21세기 지능정보사회를 구현하는 것 등을 목적으로 지능정보화 기본법이 시행 중이기도 하다.

인공지능 시대 중요한 화두의 하나는 인공지능이 인간의 노동을 어느 정도 대체할 것인지의 문제다.<sup>7)</sup> 인공지능은 인간만이 할 수 있다고 여겨진 여러 작업을 대신하며, 무수

1) 김배원, 지능정보사회와 헌법, 인공지능(AI)의 발전과 헌법적 접근, 공법학연구, 제21권 제3호, 2020, 67쪽.

2) 이명민, 인공지능, 법, 그리고 민사소송, 안암법학, 제57호, 2018, 55쪽.

3) 김중권, 인공지능시대 알고크라시(Algocracy)에서의 민주적 정당화의 문제, 법조, 제69권 제5호, 2020, 203쪽.

4) 인공지능과 관련하여 보안, 불투명화, 제어상실, 사고, 범죄, 프라이버시·개인정보, 인간의 존엄성, 민주주의 등 여러 측면의 위험(risk)이 존재함을 환기한 연구로 강승규, 인공지능(AI)이 초래한 사회적 변화와 위험의 공존가능성에 대한 검토, 법이론실무연구, 제9권 제4호, 2021, 501쪽.

5) 홍석한, 유럽연합 ‘인공지능법안’의 주요 내용과 시사점, 유럽헌법연구, 제38호, 2022, 247-248쪽.

6) 김송옥, 인공지능 법제의 최신 동향과 과제, 김진우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 164-166쪽.

7) 관련한 대표적인 연구로 송성수, 인공지능은 인간의 일자리를 얼마나 대체할 것인가: 인공지능 시대의

히 많은 (헌)법적 쟁점을 안고 있다.<sup>8)</sup> 이러한 변화에 있어 법조제도 예외가 아니다.<sup>9)</sup> 정부와 법률시장, 개인이 사용하는 법률정보를 수집·분석·가공·제공·관리하는 서비스 내지 기술인 리걸테크(법정보기술) 시장은 급속히 확대되고 있다. 빅데이터를 학습하여 정확도를 향상시키는 인공지능은 법률서비스 고객과 법률가의 연결, 실사(due diligence), 법률 문서의 작성 자동화, 법률정보 검색, 소송결과의 예측 등 여러 분야에서 활용되고 있다.<sup>10)</sup> 인공지능이 법률시장에 들어오는 것은 피할 수 없는 대세라고도 평가된다.<sup>11)</sup>

인공지능 법관에 관한 논의는 다소 시기상조라고 볼 여지도 있다. 혹자는 인공지능 판사가 가능하려면 특정 사안에서 가장 상관관계가 높은 답을 제시하는 수준을 넘어 법적 논리를 제공할 수 있어야 하며, 이를 위해 법령과 판례가 완전히 기호논리학적으로 재구성되어야 하나, 가까운 시일에는 기술적으로 어려울 것이고 한다. 따라서 당분간 인공지능은 자연인 법관의 보조수단 내지 협업자로 남으리라고 전망한다.<sup>12)</sup> 인공지능이 아무리 발전해도 최종단계에서 의사결정을 하는 것은 사람이며, 인공지능은 법률가의 훌륭한 도구가 되어줄 것이라는 평가도 있다.<sup>13)</sup>

그런데 인공지능 판사라는 용어 자체는 이미 우리의 일상에 들어왔으며,<sup>14)</sup> 소액사건 등 제한된 영역에서 인공지능 법관, 로봇판사를 도입하려는 움직임이 발견된다.<sup>15)</sup> 학계

---

기술과 노동에 관한 시론, 코기토, 통권 제96호, 2022, 7-37쪽.

- 8) 인공지능 관련 헌법적 쟁점으로서 전자인격의 부여 문제, 기술적 실업, 프라이버시 침해, 국민의 안전보장, 재판청구권, 교육권, 자동행정과 법치주의, 민주주의에 대한 위협, 위험 거버넌스(governance) 등을 고찰한 연구로 김병록, 인공지능의 헌법적 쟁점과 과제, 법학논총, 제27집 제2호, 2020, 94-121쪽.
- 9) 최근 인공지능 챗봇인 챗 GPT(Generative Pre-trained Transformer)가 미국 미네소타대학 로스쿨의 4개 과목 시험(95개의 선택형 문항과 12개의 에세이로 구성)에서 C+학점을 받았다. 이는 낮지만 합격에 해당하는 성적이었다. 경영학에서 챗 GPT는 더 나은 성적을 거두었다. 와튼(Wharton)스쿨의 운영관리 교과목 시험에서, 챗 GPT는 B 내지 B-에 해당하는 성적을 거두었다. Samantha Murphy Kelly, ChatGPT Passes Exams From Law and Business Schools, CNN Business, 26 January 2023, <https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html>(최종방문일 2024. 2. 11.).
- 10) 이성진 외, 4차 산업혁명시대 법정보기술의 현황과 발전방안, 정보화정책, 제28권 제1호, 2021, 6쪽.
- 11) 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 75쪽.
- 12) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 45-46쪽.
- 13) 천경훈, [서초포럼] 챗GPT와 법률가의 미래, 법률신문, 2023. 2. 23, <https://www.lawtimes.co.kr/Legal-Opinion/Legal-Opinion-View?serial=185492&kind=BA10>(최종방문일 2024. 2. 11.).
- 14) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 39쪽.
- 15) 영국 민사사법위원회(Civil Justice Council)의 온라인 분쟁해결 고문단(Online Dispute Resolution Advisory Group)은 25,000파운드 미만의 사건에 대한 온라인 분쟁해결을 담당하는 HM 온라인 법원(HM Online Court)의 설립과 이를 통한 인터넷 기반 법원 서비스의 필요성을 주장하였다. 또 라트비아에서 인공지능

에서도 법관에 대한 강한 불신에 입각하여, 법관이 인공지능에 의하여 대체되어야 한다는 주장을 어렵잖게 찾아볼 수 있다.<sup>16)</sup> 또한 인공지능 알고리즘은 어느 순간 급속히 발전하며 그때 비로소 대응하려면 이미 늦을 수 있는바,<sup>17)</sup> 인공지능 법관에 관한 시론적 검토가 필요하다.<sup>18)</sup> 인공지능이 단시일에 출현하지는 않을 것이나, 점차 보조하는 지위로부터 나아가 자연인 법관의 판단을 선취하는 식으로 잠식해올 위험성을 경계한 연구도 있다.<sup>19)</sup>

인공지능 법관과 관련해서는, 재판 전 절차나 의사결정과정에서 인공지능기술을 활용하는 보조적 단계로부터, 자연인 법관을 완전히 대체하는 단계까지 논의가 가능하다.<sup>20)</sup> 비밀유지계약서(Non-Disclosure Agreement)처럼 정형화된 법률문서의 검토에 있어서는 인공지능이 이미 법조인보다 더 빠르고 정확하다는 실험결과도 있다.<sup>21)</sup> 한국 사회에서는 사법 불신과 함께 인공지능 법관에 관한 논의가 활발하다.<sup>22)</sup> 후자는 인공지능이 내리는 판단이 형사사법에서 한참동안 참고자료로 활용될 가능성이 높으나, 기술력 발달로 판단의 정확성이 제고될 경우 인간(형사)법관을 대체하는 시대가 찾아올지도 모른다고 전망한다.<sup>23)</sup>

---

능의 활용은 본래 활용도가 저조했던 소액사건을 활성화시키기 위한 사법정책의 일환으로 추진되고 있다고 한다. 정채연, 사법절차 및 사법서비스에서 인공지능 기술의 도입 및 수용을 위한 정책 연구, 사법정책연구원, 2021, 43, 75쪽.

- 16) 알고리즘의 의사결정이 자연인 법관의 판단을 대체할지 여부 및 시기에 관한 논쟁은 전혀 새로운 것이 아니라는 지적으로 Rebecca Crootof, “Cyborg Justice” and the Risk of Technological-Legal Lock-in, Columbia Law Review Online, Vol. 119, 2019, p. 236.
- 17) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 73쪽.
- 18) 인공지능 기술의 발전으로, 인공지능 법관이 머지않아 출현할 수 있으리라고 전망한 연구로 김일우, 고위험 인공지능시스템의 기본권 침해에 관한 헌법적 연구, 박사학위논문, 서강대학교, 2023, 94쪽.
- 19) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 14쪽. 인공지능 법관이 출현할 수 없다고 미리 단정하는 것이 속단일 수 있다고 한 전후재, 우리 민사재판의 미래에 관한 소고 - 디스커버리, 온라인 소송, 인공지능(AI)을 중심으로 -, 민사소송, 제27권 제1호, 2023, 132쪽도 참조.
- 20) Tania Sourdin, Judges, Technology and Artificial Intelligence: The Artificial Judge, Edward Elgar, 2021, p. 16.
- 21) 미국의 법률분야 스타트업 기업인 로직스(LawGeex)가 2018년 2월 수천 건의 비밀유지계약서 유형을 학습한 인공지능과 변호사 20명으로 하여금 비밀유지계약서 검토를 하도록 실험하였다. 이 실험에서 인공지능은 26초 만에 문서 5건의 검토를 마쳤으나, 변호사들은 평균 92분을 소요하였다. 인공지능의 검토 결과는 94%의 정확도를 보여, 변호사의 그것(85%)을 상회했다. 류인하, 사법부 신뢰 추락으로 떠오른 ‘존경하는 AI 판사님’, 경향신문, 2019. 4. 21, <https://www.khan.co.kr/national/national-general/article/201904210849001>(최종방문일 2024. 2. 11.).
- 22) 법률사무는 가치판단의 과정이며, 사법 불신은 인공지능으로 대체함으로써 해결할 문제가 아니라는 지적으로 성증탁, 제4차 산업혁명과 사회국가 실현 등에 대한 전망과 과제, 헌법학연구, 제27권 제1호, 2021, 370-371쪽.
- 23) 양천수, 인공지능과 법체계의 변화 - 형사사법을 예로 하여, 법철학연구, 제20권 제2호, 2017, 60쪽. 아직까지 인공지능은 자연인 법관의 기능 일부라도 온전히 대체할 수 있는 수준에 이르지 못했으나, 설명

여러 기술적, 사회윤리적인 이유를 들어, 법조 영역에서 인공지능의 활용에 부정적인 견해도 있다. 그러나 국내 법조계에서는 이미 법무법인이 법률인공지능 시스템을 도입하는가 하면,<sup>24)</sup> 인공지능에 기반한 형량예측 서비스가 출시되었다가 종료된 바 있다.<sup>25)</sup> 인공지능이 자연인 법률가의 영역을 기술적으로 담당할 수 있음을 보여준 사례다. 2017년 8월 19일부터 2022년 5월 9일까지 운영되었던 ‘청와대 국민청원’에서, 인공지능 판사를 도입해달라는 청원은 순하게 반복되었다.<sup>26)</sup> 인공지능이 당분간 자연인 법관을 완전히 대체하기는 어렵다는 전망이 우세하지만,<sup>27)</sup> 일부 법률가들은 인공지능 판사가 현실이라고 진단하기도 한다.<sup>28)</sup>

## 나. 헌법적 검토의 필요성

법관이 인공지능의 판단에 기초하여 판결·결정을 내릴 수 있는지, 나아가 인공지능이 법관으로서 재판을 담당할 수 있을지는 실로 중요한 헌법적 쟁점이며 이미 각국 법조계의 현안으로 다루어지고 있다. 미국 위스콘신 주 대법원은 형사재판에서 인공지능 알고리즘 자료를 근거로 한 지방법원 판결의 타당성을 인정하였으며,<sup>29)</sup> 동 판결은 인공지능 알고리즘의 판단에 기초한 사법판단의 타당성을 승인한 최초의 사례로도 평가된다.<sup>30)</sup>

---

가능성이 보완되고 편향성이 보정된다면 점차 자연인 법관을 대체하는 수준으로 변화가 진행되리라고 한 고유강, 법률인공지능이 변화시킬 재판의 모습: ‘인공지능 판사’의 기술적 대체가능성과 관련된 현황을 중심으로, 재판자료집, 제142집, 2021, 21쪽도 참조.

24) 서영상, 대륙이주, 국내로법 최초 ‘법률인공지능 시스템’ 도입, 법률신문, 2018. 2. 27, <https://m.lawtimes.co.kr/Content/Article?serial=140682>(최종방문일 2024. 2. 11.).

25) 안혜성, ‘로톡 형량예측서비스’ 종료 두고 법률플랫폼 갈등 증폭, 법률저널, 2021. 9. 15, <http://www.lec.co.kr/news/articleView.html?idxno=730329>(최종방문일 2024. 2. 11.).

26) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 4쪽.

27) 가령 김병록, 인공지능의 헌법적 쟁점과 과제, 법학논총, 제27집 제2호, 2020, 113쪽; 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 14쪽.

28) 양지열, 인공지능(AI)판사는 정의로운가?, Ai타임스, 2021. 8. 13, <http://www.aitimes.com/news/articleView.html?idxno=140115>(최종방문일 2024. 2. 11.). 컴퓨터 재판이나 인공지능을 이용한 사법판단 시도가 전혀 불가능하지는 않다고 한 김현석, 재판의 진행·판단에서 법관의 정서에 관하여, 법조, 통권 제733호, 2019, 291쪽도 참조. 다만 법관의 사실판단 및 결론도출 과정에서 감성적 요소를 배제할 수 없으며, 좋은 재판에는 감성의 적절한 개입이 필요하다고 한다. 중국적 효력을 인정하는 재판 알고리즘을 만들지 않고, 개별 사건별로 공방을 통해 감성과 이성, 직관과 분석 모두를 사용하여 재판하는 이유는 법가치와 법감정을 찾아내려는 노력에 대한 존중 때문이라라고 한다.

29) *Loomis v. Wisconsin*, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S.Ct. 2290 (2017).

30) 피고인(Eric Loomis)은 총격사건에 사용된 차량을 운전한 혐의로 기소되었다. 주 검찰은 미국 노스포인



2021년 10월 유럽의회 형사사법 인공지능 결의안<sup>31)</sup>은, 투명성과 책임성을 탑재한 인공지능이 도입되어 시민의 기본권에 아무런 해도 입히지 않는다는 것을 신뢰할 수 있을 때까지 형사재판에서 인공지능이 내린 결정을 그대로 사용해서는 안 된다고 선언하였다.<sup>32)</sup>

국내 학계에서도 형사정책에서 인공지능이 차지하는 비중이 점차 높아지리라는 예상이 유력하며,<sup>33)</sup> 양형위원회 또한 향후 양형 논의에서 인공지능을 전 과정에 걸쳐 활용할 예정이다.<sup>34)</sup> 재판업무에 인공지능을 활용하는 것과 병행하여, 인공지능 법관을 시범사업으로 진행해볼 것을 제안하는 연구도 있다.<sup>35)</sup>

인공지능과의 동거가 거스를 수 없는 현실이라면, 인공지능의 개발 및 활용에 관련한 규범을 확립하고 그 헌법적 한계를 설정함으로써 인공지능의 신뢰성을 확보할 필요가 있다.<sup>36)</sup> 유럽집행위원회의 독립 자문기관인 인공지능 고위 전문가 자문단(High-Level Expert Group on Artificial Intelligence)은 2019년 4월 8일 ‘신뢰할 수 있는 인공지능을 위한 윤리지침(Ethics Guidelines for Trustworthy AI)’을 발표하였는데, 신뢰할 수 있는 인공지능의 3요소로 합법적이고(lawful) 윤리적이며(ethics) 강건할(robust) 것을 제시하였다.<sup>37)</sup> 법조계에서의 인공지능 활용, 구체적으로는 인공지능 법관의 출현 가능성 또한 배제할 수 없는 현실에서,<sup>38)</sup> 이 연구는 인공지능 법관과 관련한 헌법적 쟁점들을 선행적으

---

트(Northpoint)사의 재범가능성 예측 프로그램 컴퍼스(COMPAS)의 위험 평가를 바탕으로 피고인의 재범 가능성이 높다고 주장하였고, 법관이 이를 양형 판단에서 인정하여 징역 6년을 선고하였다. 피고인은 적법절차원칙에 따른 권리가 침해되었음을 주장하며 항소하였으나 기각되었다. 상세히는 정채연, 사법절차 및 사법서비스에서 인공지능 기술의 도입 및 수용을 위한 정책 연구, 사법정책연구원, 2021, 24-25쪽.

31) Resolution of 6 October 2021 on Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial Authorities in Criminal Matters (2020/2016(INI)). (2020/2016(INI)).

32) 정소영, 형사사법에서 인공지능 사용에 대한 유럽의회 결의안 - ‘인공지능에 의한 결정’ 금지에 관하여 -, 형사정책, 제34권 제2호, 2022, 137쪽. 한편 유럽의회(European Parliament)는 2022년 4월 발표한 인공지능 법안 수정안 보고서에서, 금지되는 인공지능 관행(prohibited artificial intelligence practices)에 예측 치안을 추가해야 한다고 결론 내렸고, 이는 인공지능법안 제5조에 새로이 포함되었다. 정소영, 형사사법에서 인공지능 사용에 대한 유럽의회 결의안 - ‘인공지능에 의한 결정’ 금지에 관하여 -, 형사정책, 제34권 제2호, 2022, 141쪽.

33) 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 - 보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 42쪽.

34) 박수연, [인터뷰] 이상원 제9기 양형위원회 위원장 “형법은 사랑... 양형 논의도 사랑이 토대가 되어야”, 법률신문, 2023. 8. 14, [https://www.lawtimes.co.kr/news/190215\(최종방문일 2024. 2. 11.\)](https://www.lawtimes.co.kr/news/190215(최종방문일 2024. 2. 11.)).

35) 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 81쪽.

36) 이러한 취지의 서술로 조소영, 인공지능과 민주주의, 공법연구, 제49집 제2호, 2020, 159쪽.

37) High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019, p. 5.

38) “사법작용의 핵심으로서 재판도 인공지능기술로 대체할 가능성은 있다.” 정문식, 인공지능과 헌법, 한양

로 검토하는 시론(試論)의 의의를 가진다.

뒤에서 볼 것처럼 유럽의회는 2023년 6월 14일 인공지능법안을 채택하였으며, 유럽연합 집행위원회는 2023년 12월 8일 동 법안을 통과시켰다. 동 법안은 유럽의회와 회원국의 공식 승인을 거쳐 2026년부터 시행될 예정이다. 인공지능법안은 혜성처럼 갑자기 등장한 것은 아니었으며, 이전부터 진행되었던 인공지능 윤리규범과 규제정책에 관한 탄탄한 연구와 협력, 모니터링의 결과라고 평가된다.<sup>39)</sup> 인공지능 법관, 특히 가치판단의 성격이 강한 헌법재판에서의 인공지능 법관은 매우 중요한 헌법적 쟁점들을 내재하고 있는 바, 설령 인간과 동일한 수준의 인공지능이 조만간 출현해 법관의 역할을 하지는 않더라도 헌법재판소와 헌법재판연구원에서 보다 선제적으로 연구하며 대응할 필요가 있다.

## 2. 연구의 대상과 범위 및 구성

### 가. 연구의 범위 한정

오늘날 인공지능이 적용되지 않는 분야를 찾기가 더 어려울 정도로 인공지능은 사회 전반(바둑, 드론, 자율주행, 맞춤형 광고, 번역, 창작, 예측치안 등)에서 광범위하게 활용된다.<sup>40)</sup> 이 가운데 좁게는 인공지능 법관, 논의의 범주를 다소 확장하더라도 사법작용에서의 인공지능 활용과 관련한 헌법적 문제들을 검토한다.

인공지능 법관과 관련하여 논의할 수 있는 헌법적 쟁점은 무수히 다양하여<sup>41)</sup> 적절한 취사선택이 필요하다. 인공지능은 현재 비약적인 발달을 거듭하고 있는바, 그 효율성이나 기술적 정확성보다는 규범적, 특히 헌법적 차원의 논의에 주안점을 둔다. 하지만 인공지

---

대학교 법학연구소 편, 인공지능과 법, 도서출판 오래, 2023, 65쪽.

39) 김송옥, 인공지능 법제의 최신 동향과 과제, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 151쪽.

40) 김광수, 인공지능 알고리즘 규율을 위한 법제 동향 - 미국과 EU 인공지능법의 비교를 중심으로 -, 행정법연구, 제70호, 2023, 177쪽.

41) 인공지능의 헌법적 쟁점을 제시한 최근의 연구로 김종철, 인공지능의 헌법적 의의, 오병철 외, 인공지능과 법, 연세대학교 출판문화원, 2023, 18-40쪽. (1) 민주주의와 관련하여 정보접근성 강화로 실질적 공론이 가능해지고 정보편재에 의한 불평등한 정치적 지배구조를 극복할 수 있으나, 정보편향으로 민의가 왜곡되고 정보독재 및 감시가 극대화될 우려도 있다고 한다. (2) 다양한 기본권을 향유할 수 있는 여건이 향상될 수 있으나, 개인정보와 사생활이 침해되고 정보격차 등으로 평등과 관련한 문제가 제기될 수 있으며, 인공지능 살상무기 등 생명권과 인간의 존엄에 관한 문제가 불거질 수 있다고 한다.

능 알고리즘이 기술적으로 취약하고 미흡하다면,<sup>42)</sup> 규범적인 문제들에 더하여 인공지능 법관의 등장을 반대하는 유력한 논거가 될 수 있다. 가령 인공지능 알고리즘이 재범의 위험성 예측, 예측 치안(predictive policing) 등에서 활용되나 그 정확성이 객관적·중립적으로 엄밀히 검증되었다고 보기 어려운바,<sup>43)</sup> 이러한 현주소는 인공지능 법관 도입논의에서도 반드시 고려되어야 한다.

인공지능 법관이 현실화되기 위해서는 법관의 임용절차에 관한 헌법(제104조 제4항)과 법률(법원조직법 제41조 제3항), 개인정보보호법 등 여러 법제가 정비될 필요가 있으나 이 모두를 상세히 다루지는 아니한다. 인공지능 법관의 도입과 관련하여 특히 문제되는 헌법적 쟁점들에 주목하고자 한다.

법관의 역할은 재판에서의 판단에 국한되지 않는다. 법관은 사건을 관리하고 분쟁을 해결한다. 또한 법관은 중요한 공공 및 교육적 기능을 수행하고, 법원의 시스템 및 행정과 관련한 업무도 수행한다.<sup>44)</sup> 이 연구보고서는 법관의 역할 가운데 ‘재판’으로 그 논의의 범주를 한정하고자 한다.

## 나. 연구의 구성

먼저 인공지능과 인공지능 법관을 개관한다(II). 여기에는 인공지능의 개념 및 유형, 알고리즘의 특성, 인공지능의 활용 및 규제동향에 관한 논의가 포함된다. 또한 주요 국가들의 사법 분야에서 인공지능이 어떻게 활용되고 있는지 다루고, 법관의 사건처리 단계별로 인공지능의 발전 현황을 짚어본다. 특히 현재의 기술 수준에서 상정할 수 있으며 헌법적 검토를 요하는 쟁점으로서의 인공지능 법관이란 무엇인지를 검토한다.

---

42) 머신러닝과 같은 첨단기술이 반드시 뛰어난 기능, 높은 정확성을 보장하지 않는다고 한 고유강, 법률인공지능이 변화시킬 재판의 모습: ‘인공지능 판사’의 기술적 대체가능성과 관련된 현황을 중심으로, 재판자료집, 제142집, 2021, 14쪽 참조.

43) 김용재, 형사절차에서 인공지능 알고리즘의 활용가능성과 그 한계, LAW & TECHNOLOGY, 제16권 제4호, 2020, 9-10쪽. 혹자는 재범위험성을 예측하는 인공지능 알고리즘은 그 정당성을 차치하고 정확성 자체를 담보할 수 없다는 비판한다. 오류의 과학인 통계학이 제시하는 부정확한 가설은 다차원적이고 규범적 개념인 재범위험성을 설명하고 형벌을 정당화하는 데 본질적인 한계가 있다고 한다. 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 677쪽.

44) Tania Sourdin, Judge v Robot? Artificial Intelligence and Judicial Decision-Making, University of New South Wales Law Journal, Vol. 41, No. 4, 2018, p. 1133.

다음으로는 인공지능 법관과 관련한 주요 헌법적 쟁점들을 검토한다(III). 먼저 인공지능 법관에 대한 세간의 기대가 무엇을 배경으로 하는지, 인공지능의 활용이 재판에서 어떠한 순기능을 가질 수 있는지 살펴본다. 그리고 법치주의, 재판청구권, 법관의 독립성, 재판제도 및 실무의 본질적인 변화 가능성, 인공지능 법관의 판단에 의한 기본권 침해 가능성 및 책임의 귀속 문제, 사법제도의 본질과 관련한 문제 등을 순차적으로 논의한다.

이어 인공지능 법관의 도입과 관련하여 어떠한 요소들이 고려되어야 하는지, 제도적으로는 어떠한 방안을 강구할 수 있을지 모색해본다(IV). 특히 사법 분야에서 활용되는 인공지능 시스템을 고위험군으로 분류하는 학계 및 실무의 동향을 짚어보고, 사법데이터의 공개 및 개인정보 보호, 설명 가능한 인공지능 알고리즘에 관하여 서술한다. 인공지능을 재판에서 보조적 역할에 한정하거나 당사자의 의사를 존중하는 방안, 재판의 영역마다 인공지능 알고리즘의 활용 여부 및 정도를 달리하는 방안 등도 소개한다.

### 3. 연구의 방법

이 연구는 헌법학 및 인공지능과 관련한 국내외 단행본, 논문, 주석서 등을 참조하는 문헌 연구를 수행한다. 인공지능의 사법 분야 활용은 이미 현실화되었으며, 관련 선행연구도 무수히 집적되었다.

이 연구는 재판업무에서의 인공지능 활용에 관한 주요 국가들 및 국제사회의 동향을 활용하는 비교법적 연구를 수행한다. 인공지능의 발전은 전 세계적인 현상이며, 재판 등 사법절차에서의 인공지능 활용과 그에 대한 우려는 우리나라에서만 문제가 아니다. 물론 각국의 재판제도 및 운용이 상이하므로, 인공지능의 사법 분야 활용에 관한 접근 방식에도 차이가 있음은 감안해야 한다. 한 예로 2021년 10월 유럽의회는 형사사법에서 인공지능에 의한 결정을 금지하는 취지의 결의안을 발표하였는데, 영업 비밀을 존중하며 자율규제를 선호하는 미국과는 상반되는 태도라고 평가된다.<sup>45)</sup> 한국 사법체계는 유무죄의 판단과 양형이 별도로 분리되어 있지 않아, 인공지능의 재범위험성 평가를 법관이 참고할 경우, 편견을 가지고 사실인정을 할 우려가 있다는 지적도 제기된다.<sup>46)</sup>

45) 정소영, 형사사법에서 인공지능 사용에 대한 유럽의회 결의안 - ‘인공지능에 의한 결정’ 금지에 관하여 -, 형사정책, 제34권 제2호, 2022, 161쪽.

46) 김일우, 고위험 인공지능시스템의 기본권 침해에 관한 헌법적 연구, 박사학위논문, 서강대학교, 2023,

그런데 최근에는 미국과 유럽연합 사이에서도 인공지능에 관하여 지침 내지 가이드라인을 제시하는 것을 넘어 법적 규제를 가하려는 움직임이 관찰된다. 인공지능 기술의 파급력 및 세계적인 현상으로서의 성격을 고려할 때, 미국과 유럽이 공통적인 내용을 담은 유사한 규율로 수렴하리라는 전망도 제기된다.<sup>47)</sup> 사법 분야에서 활용되는 인공지능의 규제 환경이 국가별로 각기 다른바, 국제 표준(international standards)의 필요성이 커지고 있다는 지적도 있다.<sup>48)</sup> 이러한 측면에서, 비교법적 검토는 우리나라에서 인공지능 법관과 관련한 규제 여부 및 규제방안을 논의함에 있어 적잖은 시사점을 줄 것이다.

---

114쪽.

47) 김광수, 인공지능 알고리즘 규율을 위한 법제 동향 -미국과 EU 인공지능법의 비교를 중심으로-, 행정법연구, 제70호, 2023, 191쪽.

48) Benes Z. Aldana, 사법정의의 대변혁: 법률시스템에서의 AI 부상과 윤리적 고찰, 2023 사법연수원 국제 콘퍼런스 자료집, 사법연수원, 2023, 76-77쪽.

## II. 인공지능, 인공지능 법관 일반

### 1. 인공지능의 개념과 유형 및 알고리즘의 특성

#### 가. 인공지능의 개념과 유형

##### 1) 인공지능의 개념

##### 가) 선행연구, 제도의 정의

##### (1) 지능 및 인공지능 개념정의의 어려움

인간의 각종 능력을 인간보다 더 뛰어나게 구사하는 인공물은 종전부터 있어왔지만(가령 사람보다 빠른 자동차, 사람보다 무거운 물건을 들어 올리는 중장비), 지능은 인간이 동물과 차별화되며 자연을 지배하는 능력이라는 점에서 인공지능은 주목할 필요가 있다. 인공지능은 자율성을 가지고 인간의 통제로부터 벗어날 여지가 있다는 점에서 우려를 자아내기도 한다.<sup>49)</sup>

그런데 문제는 인공지능을 정의하기가 그리 용이하지 않다는 점이다. 잠깐 사이에도 신기술이 출현하는 오늘날, 발전적 개념인 인공지능의 개념정의가 얼마 지나지 않아 낡은 개념이 될 우려가 있다. 또한 사회 각 영역마다 관련 기술이 달리 활용되며, 인공지능을 통하여 추구하는 목표에 대한 연구자들의 생각도 천차만별이다. 이러한 현실에서 인공지능을 일의적으로 개념정의하려는 시도는 성급한 일일 수 있다.<sup>50)</sup>

##### (2) 인공지능의 주류적인 개념정의

지능을 정의하기란 쉽지 않지만, 배우고 이해하는 능력 또는 새로운 상황에 대처하는 능력이라고 할 수 있다.<sup>51)</sup> 신경과학자들은 다양한 환경에서 복잡한 의사결정 문제를 해결

49) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 31쪽.

50) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 55-57쪽.

51) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 39쪽.

하는 능력이라고 정의한다.<sup>52)</sup> 가드너(Howard Gardner)는 지능을 독립적이지 않은 복수의 특성이 상호작용하는 체계로 규정하는 다중지능 이론(theory of multiple intelligences)이론을 창안하였으며, 오늘날 상당수 논자들은 다양한 환경에서 문제를 해결하는 일반적 능력을 지능 개념의 핵심요소로 본다.<sup>53)</sup>

인공지능이 일의적으로 정의되었다고 할 수는 없으나,<sup>54)</sup> 주요 개념정의에서는 “기계를 통해 인간의 지적 활동을 모사·구현”한다는 점을 공통적으로 포착하고 있다.<sup>55)</sup> 유네스코(UNESCO)는 인공지능의 단일한 개념정의가 존재하지 않는다고 하면서도, “인지, 학습, 추론, 문제해결, 언어 상호작용, 창의적인 작업수행과 같은 기능을 포함하여, 인간 지능의 특정한 기능들을 모사할 수 있는 기계”라고 풀이한다.<sup>56)</sup> 지능정보화 기본법 제2조 제4호가목에서는 지능정보기술의 하나로서, 전자적 방법으로 학습·추론·판단 등을 구현하는 기술을 규정한다. 소관위원회 심사 중인 법률안 중에서는 인공지능을 “학습, 지각, 판단, 자연언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현하기 위한 것”이라고 정의한 것이 있다.<sup>57)</sup>

### (3) 인공지능의 개념에 관한 새로운 접근

그런데 인공지능의 문제 해결과정이 사람이 사고하는 과정과 전혀 닮지 않았다는 지적이 있다. 인간이라면 당연히 아는 답을 인공지능은 모르며, 다만 답을 예측할 뿐이라는

52) 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 72쪽.

53) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 13-14쪽.

54) 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 72쪽.

55) 김웅재, 형사절차에서 인공지능 알고리즘의 활용가능성과 그 한계, LAW & TECHNOLOGY, 제16권 제4호, 2020, 4쪽; 이명민, 인공지능, 법, 그리고 민사소송, 안암법학, 제57호, 2018, 38쪽. 인공지능을 컴퓨터 등 기계를 통해 인간의 자연지능을 구현하거나 모방하는 기술이라고 정의한 연구로 이준일, 인공지능과 헌법 - 인공지능의 위험에 대한 헌법적 권리 차원의 대응 -, 헌법학연구, 제28권 제2호, 2022, 248쪽. 인공지능은 추론, 학습 등 인간 지능의 특정한 특성들(traits)을 모방할 수 있는 복잡한 컴퓨터 프로그램을 개발하는 일련의 이론과 기술이라고 한 Christian LExcellent, Artificial Intelligence versus Human Intelligence: Are Humans Going to Be Hacked?, Springer, 2019, p. 5도 참조.

56) Xianhong Hu et al., Steering AI and Advanced ICTs for Knowledge Societies: A Rights, Openness, Access, and Multi-stakeholder Perspective, UNESCO Publishing, 2019, p. 24.

57) 알고리즘 및 인공지능에 관한 법률안(의안번호: 2113509, 제안일자: 2021. 11. 24, 대표발의: 윤영찬, 2022. 3. 30. 과학기술정보방송통신위원회 상정) 제2조 제2호. “학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것”이라고 거의 유사하게 정의한 인공지능산업 육성 및 신뢰 확보에 관한 법률안(의안번호: 2118726, 제안일자: 2022. 12. 7, 대표발의: 윤두현, 2022. 12. 8. 과학기술정보방송통신위원회 상정) 제2조 제1호도 참조.

것이다. 그에 따르면 인공지능의 본질은 예측이라 해도 과언이 아니며, 인공지능 알고리즘은 예측기계라고 한다.<sup>58)</sup>

관련하여, 인공지능의 법적 개념에 관하여 사뭇 달리 접근한 연구는 주목할 만하다. 강현구(2021)에 따르면 인공지능에 접근하는 방식에는 계산심리학적 관점과 기계지능 관점이 있다. 전자는 사람과 동일한 방식으로 행동하는 프로그램을 만드는 데 주안점을 둔다. 후자의 경우 합리적으로 행동하는 합리적 에이전트(a rational agent)를 개발하는 데 중점을 두며, 주어진 과제를 성취하는 데 도움이 된다면 어떤 기술이든 도입한다. 그런데 합리성 기준은 수학적으로 잘 정의되어 있으며, 후자의 관점이 주어진 문제에 대하여 최상의 결과를 산출하는 실체를 개발할 수 있어 보편화되었다. 그렇다면 인공지능을 법적으로 정의할 때에는 인간의 능력을 얼마나 잘 구현했는지 보다 합리적 행동을 하는지에 중점을 두어야 한다.<sup>59)</sup> 이러한 논의에 입각하여, 동 연구는 인공지능을 “복잡한 환경 아래에서 데이터 분석을 통해서 환경을 인식하고, 이러한 데이터로부터 얻어진 정보와 학습된 지식에 기반을 둔 탐색으로 주어진 목적 달성에 가장 적합한 행동을 자율적으로 선택하여 실행하는, 인간에 의해 고안된 소프트웨어 또는 하드웨어 시스템”이라고 정의한다.<sup>60)</sup>

인간과 인공지능이 불확실성이라는 근원적 제약조건을 공유하면서 주어진 문제를 해결하기 위해 ‘제각기 다른 형태의 제한적 합리성을 구현’했다는 평가도 있다. 그에 따르면, 인간 지능이 현실에서 의사결정을 함에 있어서는, 종종 감정과 결부되고 무의식적 특성이 강하며 직관(intuition)이라고도 불리는 시스템1이 주도권을 가진다.<sup>61)</sup> 개별 정보를

---

58) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 161쪽.

59) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 40-41쪽. ‘합리적’이라는 표현은 수학적으로 정의될 수도 있으며, 무엇이 ‘인간적’이라는 말이 무엇인지에 관한 사변적인 논쟁을 피할 수 있다는 평가로 고훈수·정해빈·박도현, 인공지능과 차별, 저스티스, 통권 제171호, 2019, 204쪽.

60) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 43쪽. 다만 인공지능이 계속 발달하는 개념이므로 그 명확성의 정도가 완화될 수 있다고 한다. 머신러닝은 기본적으로 다량의 데이터와 컴퓨터 연산능력을 동원하여 주어진 목표를 달성하는 최적의 방법을 반복적으로 탐색하는 과정이지, 인간의 논리적 사고과정을 모사하는 것이 아니라고 한 고유강, 법률인공지능이 변화시킬 재판의 모습: ‘인공지능 판사’의 기술적 대체가능성과 관련된 현황을 중심으로, 재판자료집, 제142집, 2021, 22쪽도 참조.

61) 트롤리 딜레마, 구명정 딜레마 등에 관한 사고실험은 도덕 판단에 직관적 요소가 강하게 수반됨을 알려 준다고 한다. 행동주의 법경제학(Behavioral Law and Economics)은 의사결정과정에서 제한된 이성만족해야 한다고 주장하며, 뇌에 관한 과학적 연구를 통해 정서가 이성적인 판단에 필수적이라는 사실이 드러나기 시작했다고 한다. 김현석, 재판의 진행·판단에서 법관의 정서에 관하여, 법조, 통권 제733호, 2019, 264-266쪽.



순차적으로 처리하는 방식으로서 높은 수준의 주의와 결부되고 의식적 특성이 강하며 알고리즘이라고 불리는 시스템2가 시스템1의 오류를 통제·교정한다. 그런데 인공지능의 경우 시스템2가 중심이며, 시스템1이 이를 보완한다는 점에서 인간지능과 정반대의 면모를 지닌다고 한다.<sup>62)</sup>

이러한 맥락에서, 최근에는 인간 지능의 모사 내지 구현이 아닌, 합리적 결과 도출의 측면에서 인공지능의 개념에 접근하는 연구가 활발하다.<sup>63)</sup>

보다 실용적인 관점에서 인공지능을 수학적 함수의 일종으로 이해한 연구도 주목할 만하다. 김병필(2021)에 따르면 인공지능이란 인간이 미리 정해놓은 문제를 풀기 위해 입력 변수(예: 영어문장, 사진)를 출력변수(예: 한국어문장, 얼굴인지 여부의 판정)에 대응시키는 함수다. 머신러닝 기법은 데이터를 통해 함수를 학습한다는 특징이 있다고 한다.<sup>64)</sup>

#### 나) 검토

전술하였듯 인공지능은 매우 발전적 개념이며 따라서 일의적으로 그 개념을 정의하는 것은 성급할 수 있다.

생각건대, 인공지능은 주어진 문제의 인식·데이터의 분석·합리적인 판단을 통한 목표의 달성이란 ‘인간지능을 통하여 수행해온’ 일들(예: 기상예측, 재범위험성 평가, 바둑, 번역, 환자의 진단)을 처리하기 위하여 고안되었다.<sup>65)</sup> 다만 비행기를 설계·제작할 때 새가 하늘을 나는 원리·구조를 그대로 구현하는 것을 목적으로 하지 않듯, 인간의 두뇌·지적능력을 그대로 구현하는 데 방점이 있지는 않다.<sup>66)</sup> 그런데 아직은 요원하다고 하지만,

62) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 35-39, 49쪽.

63) 인공지능을 “현실의 환경을 인지하며, 자율적으로 작동하고, 변화에 적응하며, 목표를 만들고 추구하는 합리적 에이전트”라고 정의한 조한상·이주희, 인공지능과 법, 그리고 논증, 법과 정책연구, 제16집 제2호, 2016, 299쪽 참조. 인간의 행동을 그대로 흉내 내는 것이 인공지능인지에 대하여 이견이 제시되고 있음을 환기한다. 인공지능 연구는 인간을 닮은 장난감을 만들어내려는 것이 아니라, 종래 인간이 해온 일을 더 빠르고 정확하게 수행하는 무언가를 추구하는 것이라고 한다.

64) 김병필, 재판 예측 인공지능 기술의 현황과 한계, 저스티스, 통권 제182-2호, 2021, 92쪽. 마찬가지로 접근방식으로 서울대학교 산학협력단, 사법부에서의 인공지능(AI) 활용방안, 법원행정처, 2020, 4쪽.

65) 마빈 민스키(Marvin Minsky)는 인공지능을 “사람이 수행한다면 지능을 요하는 일들을 하는 기계를 만드는 과학”이라고 하였다. Marvin Minsky, Preface, in Marvin Minsky ed., Semantic Information Processing, MIT Press, 1968, p. v. OECD의 인공지능 권고는 인공지능 시스템을 “인간이 정의한 목적을 위하여, 실제 또는 가상 환경에 대해 영향을 미치는 예측, 추천 또는 의사결정을 도출할 수 있는 기계 기반 시스템”이라고 정의한다. OECD, Recommendation of the Council on Artificial Intelligence, 2020, OECD/LEGAL/0449, adopted on 22.06.2019. 번역은 홍석한, 유럽연합 ‘인공지능법안’의 주요 내용과 시사점, 유럽헌법연구, 제38호, 2022, 251쪽의 각주 6을 따랐다.

66) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8

인간과 같은 사고체계를 갖춘 ‘강한 인공지능’의 경우라면 인간지능을 구현한(구현하려고 하는) 것이라고 할 수 있겠다. 다만 인공지능 기술이 인간 지능의 불완전성과 오류까지 그대로 모사·구현하여야 하는지에 대하여는, 규범적 측면뿐만 아니라 경제적인 이유에서도 의문이 제기된다.<sup>67)</sup>

인공지능은 컴퓨터의 빠른 연산과 빅데이터의 등장, 딥러닝을 비롯한 신기술의 발달과 업계의 막대한 투자에 힘입어, 최근 분야를 막론하고 각광받으며 우리 일상에서 널리 활용되고 있다.<sup>68)</sup> 오늘날은 지능정보사회라고도 일컬어지는데, 데이터와 인공지능이 결합하고, 만물이 사물인터넷에 연결·통합되며 각 영역의 경계가 파괴된다고 평가된다.<sup>69)</sup>

## 2) 인공지능의 유형

일반적으로 인공지능은 약한 인공지능, 강한 인공지능, 초인공지능으로 분류된다.

약한 인공지능은 특정한 영역의 문제를 해결하는 데 특화된 인공지능으로, 바둑의 알파고 파고가 대표적이다. 패턴과 규칙을 입력해줄 필요가 있으며, 이를 벗어난 창조가 불가능하고 특정 부문에서 활용된다.

강한 인공지능은 일반 인공지능(general AI)이라고도 하는데, 인간처럼 지능을 범용적으로 확장할 수 있거나, 인간 수준의 인지능력을 보유한 인공지능을 일컫는다.

초인공지능은 특이점(singularity)이라고 하는, 모든 인류가 보유하는 인지능력의 총합과 같아지는 변곡점을 넘어선 인공지능을 말한다. 자유의지와 지각능력을 가지며 인간을 초월하는 수준의 인공지능으로, 인간의 명령을 거부할 수도 있다.<sup>70)</sup>

---

권 제2호, 2021, 42쪽. 현재 발전하고 있는 인공지능은 실제 인간 뇌 신경작용의 사고회로를 그대로 모방하기보다, 주어진 변수에서 어떤 결론이 도출되었는지 기존 자료를 관찰·분석하는 방법으로 학습하고 새로운 사건에 응용하는 구조라는 서술로 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 13쪽.

67) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 236쪽. 오늘날 인공지능 기술의 목적은 인간보다 훨씬 정확하고 빠르게 무언가를 예측하며 패턴을 인식함으로써 경제적 이윤을 창출하는 것이지, 인간과 똑같은 정신작용을 구현하는 데 있지는 않다고 판단된다.

68) 김건우, 인공지능윤리의 의의와 쟁점, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 40-41쪽.

69) 김배원, 지능정보사회와 헌법, 인공지능(AI)의 발전과 헌법적 접근, 공법학연구, 제21권 제3호, 2020, 70-71쪽.

70) 이상 인공지능의 유형에 관하여 상세히는 김배원, 지능정보사회와 헌법, 인공지능(AI)의 발전과 헌법적 접근, 공법학연구, 제21권 제3호, 2020, 74-75쪽; 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 5-6쪽; Daniel Ventre, Artificial Intelligence, Cybersecurity and Cyber

현재로서는 강한 인공지능의 도래가 가능하다는 증거도 불가능하다는 증거도 없으나,<sup>71)</sup> 오늘날의 인공지능은 모두 약한 인공지능이다.<sup>72)</sup> 무제한적인 데이터를 기반으로 스스로 학습해 나가는 인공지능이 강한 인공지능으로 진화할 가능성을 전면 배제할 수는 없다는 견해도 있으나,<sup>73)</sup> 강한 인공지능의 도래는 막연하다는 지적이 많다.<sup>74)</sup> 이에 인공지능에 관한 다수의 선행연구와 논의는 약한 인공지능을 중심으로 이루어지고 있다. 다만 강한 인공지능에 관하여도 법학에서 깊이 있는 검토가 이루어지고 있는바, 그러한 논의를 무용하거나 공상적이라고 치부할 필요는 없다고 생각한다.<sup>75)</sup>

이른바 제4차 산업혁명 시대 인공지능은 우리 사회의 여러 분야에서 각광받고 있으나, 다양한 기본권을 침해할 수 있다는 우려가 제기되기도 한다.<sup>76)</sup>

---

Defence, Wiley, 2020, pp. 60-61. 진화과학적 관점에서는 연산 작용을 넘어 스스로 생존하고 자기복제를 할 수 있어야만 지능이라고 부를 수 있기에, 이러한 분류가 적절한지 의문이라는 서술로 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 73쪽.

71) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 65쪽.

72) 김병필, 재판 예측 인공지능 기술의 현황과 한계, 저스티스, 통권 제182-2호, 2021, 91쪽.

73) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 65쪽.

74) 진정한 전문가들은 강한 인공지능의 출현이 터무니없다고 생각한다는 서술로 티로 하겐도르프, 인공지능윤리의 윤리학: 윤리가이드라인에 대한 평가, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 89쪽. 인간과 같은 사고체계를 갖춘 강한 인공지능에 관한 논의는 자칫 비현실적 구상이 될 위험이 있음을 지적한 연구로 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등への 적용에 관한 고찰, 형사법의 신통향, 통권 제51호, 2016, 216쪽. 강한 인공지능의 등장은 아직 막연하다는 서술로 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 5쪽.

75) 강한 인공지능의 출현을 확신하지 못하는 상태임을 인정하면서도, 강한 인공지능의 처벌가능성을 논의한 연구로 양천수, 인공지능과 법체계의 변화 - 형사사법을 예로 하여, 법철학연구, 제20권 제2호, 2017, 63쪽 이하. 강한 인공지능은 인간과 동등한 정신능력을 갖추고 있기에 형사사법 패러다임에 중대한 도전이 되며, 독일 형법학계에서 로봇 형법(Strafrecht für Roboter)이라는 이름으로 인공지능에 관한 형사처벌의 가능성이 논의되고 있음을 환기하고 있다.

76) 인공지능이 개인정보자기결정권, 표현의 자유, 평등권, 공정한 재판을 받을 권리 등 거의 모든 기본권을 침해할 가능성이 있다는 견해로 유승익, 인공지능이 인권과 민주주의에 미치는 영향과 인공지능법안의 쟁점, 민주법학, 통권 제82호, 2023, 83-85쪽. 특히 콘텐츠 관리(플랫폼 규칙을 위반한 게시물의 삭제 등), 사용자 프로파일링(이용자의 행동정보를 분석한 맞춤형 광고 제공 등)에 의한 표현의 자유, 개인정보자기결정권 제한 등을 언급하고 있다.

또한 인공지능에 행동·의사결정을 위임한 경우 ‘위임’이기 때문에 자유의 침해라고 말하기 힘들어지며, 행위·결정에 필요한 정보를 제공받은 경우에도 해당 정보를 ‘스스로’ 신뢰한 것이므로 자유의 외재적 제약이 있다고 주장하기 어려워진다고 한다. 유승익, 인공지능이 인권과 민주주의에 미치는 영향과 인공지능법안의 쟁점, 민주법학, 통권 제82호, 2023, 89-90쪽.

## 나. 알고리즘의 개념과 인공지능 알고리즘의 특성

### 1) 알고리즘의 개념

알고리즘은 십진법 전파에 큰 영향을 미친 교과서를 집필한 페르시아 수학자 알 콰리즈미(Muhammad ibn Mūsā al-Khwārizmī)를 기려 만든 용어이다.<sup>77)</sup> 알고리즘이란 일정한 목적(가령 문제해결)을 위해 수행해야 하는 일을 단계적으로 알려주는(지시하는) 명령어이다.<sup>78)</sup>

문제해결 방법이 미리 정해져 있지 않으면 그때그때 방법을 찾아야 하며, 종전 방법과 구별이 어려울 수 있다. 결과의 동일성이 확보되지 못할 불확실성에 대비하려면, 문제해결 방법을 여러 단계로 분절하고 순서를 부여해야 한다. 이는 ‘지시’에 의해 가능하며, 투입된 데이터를 처리해 결과를 산출한다. 컴퓨터를 통해 문제를 해결하는 알고리즘은 프로그래밍 언어로 표현될 필요가 있으며, 반복적인 자기지시로 구성되어 작동상의 폐쇄성을 갖는 시스템으로서의 성격을 가진다.<sup>79)</sup>

### 2) 인공지능 알고리즘의 특성

가장 널리 알려진 인공지능 프로그램의 하나인 알파고를 비롯하여, 오늘날 인공지능에는 몬테카를로 트리 탐색(Monte Carlo tree search: MCTS), 심층신경망 원리가 적용된다. 몬테카를로 트리 탐색이란, 초기상태에서 연산자를 적용하여 새로운 상태를 얻고, 이들에 다시 연산자를 가하여 또 다른 상태를 얻는 과정에서 최종 목표상태가 얻어지는 것이다.<sup>80)</sup> 심층신경망은 인공지능이 행동을 할 때 그 행동이 최종 목표상태 달성에 얼마나

---

77) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 155쪽; John C. Lennox, 2084: Artificial Intelligence and the Future of Humanity, Zondervan, 2020, p. 19.

78) 알고리즘 및 인공지능에 관한 법률안(의안번호: 2113509, 제안일자: 2021. 11. 24, 대표발의: 윤영찬, 2022. 3. 30. 과학기술정보방송통신위원회 상정)은 “문제를 해결하기 위한 일련의 순서적인 계산 또는 풀이 절차·방법을 구현하는 방법”이라고 정의한다(제2조 제1호). “주어진 문제를 풀기 위한 절차나 방법을 기술한 것”이라고 정의한 심우민, 알고리즘 규제와 입법 전략, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 239쪽도 참조.

79) 이상의 설명은 남중권, ‘알고리즘의 지배’와 ‘법의 지배’에 대한 법이론적 이해, 법과사회, 제71호, 2022, 288-292쪽을 요약·정리한 것이다.

80) 가령 바둑의 경우 비어 있는 바둑판이 초기상태이며, 바둑규칙에 의하여 승리조건을 충족시키는 것이 최종 목표 상태이다. 자율주행자동차의 경우 출발위치가 초기상태이고, 목적지에 도달한 것이 최종 목표 상태이며, 연산자는 자동차의 이동이다. 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른

가치 있는 행동인가를 판단하는 기준(평가함수)으로 작용하는데, 실제 신경의 구조와 유사점은 거의 없다고 한다.<sup>81)</sup>

이와 같은 원리에 따른 인공지능의 알고리즘은 종전의 그것과 다른 특성을 지닌다. 인공지능 알고리즘의 경우 달성해야 하는 목표 상태는 정해져 있지만, 문제를 푸는 규칙이 사전에 결정되어 있지 않으며 따라서 투명성이 결여되어 있다. 오늘날 인공지능은 잠정적인 규칙과 빅데이터를 입력한 계산 값과 정답을 비교하고, 그 오차만큼 종전 규칙을 개선하는 원리(학습)이다.<sup>82)</sup> 사람이 미리 제시한 문제풀이 규칙에 의하여 결과를 도출하는 것이 아니라, 방대한 데이터를 기반으로 자율적으로 형성한 심층신경망에 의해 판단한다는 점에서 그 의사결정과정의 마치 블랙박스라고 같다.<sup>83)</sup> 일상에서 블랙박스는 자동차나 항공기의 주행·비행 중 정보를 기록하는 작은 기계를 뜻하기도 하나, (자연과학, 공학, 컴퓨터 영역에서는) 그 내부적인 작동원리를 알지 못한 채 혹은 알 필요가 없는 가운데 입력 값과 출력 값을 볼 수 있는 시스템을 의미한다.<sup>84)</sup> 인공지능이 학습하고 활용한 빅데이터를 사람이 모두 분석하기는 불가능에 가까우며, 인공지능이 어떤 과정과 작동방식을 거쳐 규칙을 갱신하고 결과 값을 내놓았는지 추적하기도 어려워 블랙박스라고 한다.<sup>85)</sup>

## 다. 인공지능의 위험성

인공지능은 목표가 정해져 있으나 그 목표에 도달하는 방법이 미정이며 예측할 수 없다는 점에서 자율성을 가진다. 또한 전통적인 알고리즘과 달리 문제풀이 규칙이 사전에 결정되어 있지 않아, 마치 블랙박스처럼 투명성이 결여되어 있다. 빅데이터를 학습하여 패턴을 인식하고 무언가를 예측한다는 점에서, 개인정보자기결정권, 인격권, 사생활의 비

---

국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 34쪽.

81) 자세한 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 33-34쪽.

82) 출력층의 오차를 앞층으로 재분배하는 방법론을 오차역전파법(error back-propagation)이라고 한다. 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 83쪽.

83) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 35-36쪽.

84) 블랙박스의 다양한 사전적 의미에 관하여는 Cambridge Dictionary를 참조. [https://dictionary.cambridge.org/dictionary/english/black-box#google\\_vignette](https://dictionary.cambridge.org/dictionary/english/black-box#google_vignette)(최종방문일 2024. 2. 11.).

85) 이러한 현상은 머신러닝, 특히 딥러닝의 경우 더욱 심하며, 알고리즘이 복잡하고 강력할수록 더 불투명해질 가능성이 크다. 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신통향, 통권 제70호, 2021, 218쪽.

밀 등 여러 기본권이 침해될 위험성이 지적되기도 한다.<sup>86)</sup>

유럽연합 집행위원회가 2021년 4월 21일 발표한 ‘인공지능 법안’<sup>87)</sup>은 위험 수준에 따라 ‘금지된 인공지능’, ‘고위험 인공지능 시스템’, ‘낮은 위험 혹은 최소위험 인공지능’으로 나누고, 규율을 차등화하고 있다.<sup>88)</sup> 유럽의회는 2023년 6월 14일 인공지능법안을 찬성 499표, 반대 28표, 기권 93표로 채택하였다.<sup>89)</sup> 동 법안은 2023년 12월 8일 유럽연합 집행위원회까지 통과하였으며, 유럽의회와 회원국의 공식 승인을 거쳐 2026년부터 시행될 예정이다.<sup>90)</sup>

미국은 종래 가이드라인, 이니셔티브 등 비강제적인 방식으로 대응하며 자국의 패권과 경제적 이익, 국가안보를 강조해왔으나, 더 이상 혁신과 성장에만 초점을 두지 않고 부작용 및 역효과를 방지하려는 제도적 노력을 보이고 있다. 2022년 알고리즘 책임법안 (Algorithmic Accountability Act of 2022)은 관련 기업으로 하여금 자동화된 의사결정 시스템 및 증강된 의사결정 프로세스에 관한 영향평가를 하고 요약보고서를 제출하도록 하는 등 규제의 강제성과 구체성을 강화하고 있다.<sup>91)</sup>

유럽연합과 미국의 예에서 보듯 인공지능 알고리즘을 규율하는 일반법을 제정하려는 움직임이 활발하며, 한국에서도 인공지능산업육성 및 신뢰기반 조성 등에 관한 법률안이 과학기술정보통신위원회에서 심의 중이다.<sup>92)</sup>

---

86) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 35-36쪽.

87) 정식 명칭은 ‘인공지능에 관한 조화로운 규칙을 정립하고 일부 유럽연합 입법을 수정하기 위한 유럽의회 및 이사회의 규정을 위한 제안(Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts)’이다.

88) 상세히는 홍석한, 유럽연합 ‘인공지능법안’의 주요 내용과 시사점, 유럽헌법연구, 제38호, 2022, 252-257쪽.

89) Moreno Colombo, Phenotropic Interaction: Improving Interfaces with Computing with Words and Perceptions, Springer, 2023, p. x; 최창현, 유럽의회, 세계 최초 '인공지능 규제 법안' 채택...“더 엄격한 개인정보 보호, 더 엄격한 투명성, 협조하지 않을 경우 징벌적 벌금 부과”, 인공지능신문, 2023. 6. 15, <https://www.aitimes.kr/news/articleView.html?idxno=28270>(최종방문일 2024. 2. 11.).

90) 홍윤지, EU, 세계 첫 ‘AI 규제법’ 합의...한국식 AI 규제모델 마련해야, 법률신문, 2023. 12. 14, <https://www.lawtimes.co.kr/news/194071>(최종방문일 2024. 2. 11.).

91) 동 법안의 보다 상세한 내용은 홍석한, 미국 “2022 알고리즘 책임법안”에 대한 고찰, 미국헌법연구, 제34권 제1호, 2023, 85-95쪽 참조.

92) 인공지능산업 육성 및 신뢰 확보에 관한 법률안(의안번호: 2118726, 제안일자: 2022. 12. 7, 대표발의: 윤두현, 2022. 12. 8. 과학기술정보통신위원회 상정). 미국과 유럽연합의 인공지능 일반법 제정 움직임에 관하여 상세히는 김광수, 인공지능 알고리즘 규율을 위한 법제 동향 - 미국과 EU 인공지능법의 비교를 중심으로 -, 행정법연구, 제70호, 2023, 186-189쪽 참조.

## 라. 인공지능의 활용 현황

인공지능에 관한 연구는 1943년경 미국을 중심으로 시작되었다. 영국의 수학자이자 컴퓨터 공학자 앨런 튜링(Alan Turing)은 1950년 ‘계산 기계와 지능(Computing Machinery and Intelligence)’라는 논문을 발표했는데, 기술이 발전하면 컴퓨터가 언젠가 인간처럼 생각을 할 수 있다고 주장하였다.<sup>93)</sup> 인공지능(artificial intelligence)이라는 용어는 1956년 다트머스 회의(Dartmouth Conference)에서 존 맥카시(John McCarthy)에 의하여 주도되었다.<sup>94)</sup> 그런데 1970년대에 이르기까지, 인공지능에 관한 연구는 학술적인 것에 머물러 있었다. 1980년경 산업계에 이른바 전문가시스템(expert systems)<sup>95)</sup>이 도입되고 이에 상응하는 전문 하드웨어를 통해 인공지능 연구가 본격적으로 확산되었다. 컴퓨터에 지식과 정보를 학습시키는 연구가 이루어지면서 급물살을 탄 것이나, 그 데이터를 체계적으로 관리하고 인공지능에 적용함에 있어 한계가 드러났다.<sup>96)</sup>

그런데 1990년대 후반 인터넷이 등장하고 검색엔진을 통해 종전과 비교할 수 없을 만큼 방대한 데이터를 수집할 수 있게 되면서, 인공지능 연구가 다시 활발해졌다. 빅데이터를 분석하여 인공지능 스스로 학습하는 기계학습(머신러닝)의 형태로 진화한 것이다.<sup>97)</sup> 오늘날 널리 활용되는 인공지능은 머신러닝 방식에 의한 것인데, 컴퓨터 연산처리능력의

93) 또한 튜링은 기계가 지능을 가지는지 여부를 판별하는 간단한 테스트(튜링테스트)를 제안하였다. 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 17-18쪽.

94) Jaemin Lee, *Artificial Intelligence and International Law*, Springer, 2022, p. 6.

95) 전문가시스템은 (1) 지식과 경험의 데이터베이스화 (2) 의사결정 추론엔진 (3) 사용자 인터페이스로 구성되어 있었다. 주민식, 인공지능은 어떻게 발달해왔는가, 인공지능의 역사, 삼성SDS 인사이드리포트, 2017. 9. 15, [https://www.samsungsds.com/kr/insights/091517\\_cx\\_cvp3.html?referrer=https://www.google.com/](https://www.samsungsds.com/kr/insights/091517_cx_cvp3.html?referrer=https://www.google.com/)(최종방문일 2024. 2. 11.). 전문가시스템 연구는 이미 존재하는 지식과 상식을 모두 주입하면 인간 변호사나 의사처럼 추론하는 기계를 만들어낼 수 있다는 믿음으로부터 출발했다. 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 25쪽.

96) 주민식, 인공지능은 어떻게 발달해왔는가, 인공지능의 역사, 삼성SDS 인사이드리포트, 2017. 9. 15, [https://www.samsungsds.com/kr/insights/091517\\_cx\\_cvp3.html?referrer=https://www.google.com/](https://www.samsungsds.com/kr/insights/091517_cx_cvp3.html?referrer=https://www.google.com/)(최종방문일 2024. 2. 11.). 가령 권한을 넘은 표현대리의 경우에도 판례 사안 모두에 대하여 각각의 규칙을 만들고 이를 프로그램에 삽입하기란 사실상 불가능하다는 지적으로 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 61쪽. 1980년대 법률 분야에서 전문가시스템을 활용하려는 시도가 있었으나, 널리 활용되지 못하였다는 지적으로 서울대학교 산학협력단, 사법부에서의 인공지능(AI) 활용방안, 법원행정처, 2020, 7쪽.

97) 인공지능이 발달한 경과에 관하여 상세히는 최태우, 인공지능의 역사 그리고 GPU 기반 딥 러닝, CCTV 뉴스, 2016. 3. 15, <https://www.cctvnews.co.kr/news/articleView.html?idxno=45347>(최종방문일 2024. 2. 11.).

비약적인 향상과 빅데이터에 힘입은 바가 크다.<sup>98)</sup> 방대한 빅데이터를 토대로 상관관계를 분석하여 유의미한 정보를 추출하는 것은 인간의 능력을 벗어난다.<sup>99)</sup> 딥러닝은 머신러닝 방법론 가운데 하나로서, 뉴럴 네트워크가 진화한 것이다. 고전적인 머신러닝에서는 입력으로 들어가는 특징 값(feature)이 이미 만들어져 있거나 인간이 선택하는 과정이 필요하나, 딥러닝에서는 다단계의 계층적 학습과정을 통해 적절한 특징 값을 스스로 생성해 낸다.<sup>100)</sup> 딥러닝은 빅데이터를 학습하여 스스로 특징 값을 만들고, 사람이 인식하지 못한 숨은 특징도 찾아낸다.<sup>101)</sup>

오늘날 인공지능은 이용자의 취향을 고려한 맞춤형 광고, 안면인식기술, 바둑·체스와 같은 게임, 각종 문서의 작성, 그림 그리기, 일기예보, 자율주행 등 분야를 막론하고 널리 활용된다. 공적 영역에서도 인공지능은 널리 활용되는 추세이다. 행정기본법 제20조에 따라 행정청은 법률로 정하는 바에 따라 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 처분을 할 수 있다.<sup>102)</sup> 다만, 처분에 재량이 있는 경우는 그러하지 아니하다. 재범위험성 예측 등 사법 분야에서의 인공지능 활용에 관하여 상세히는 후술한다.

---

98) 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에의 적용에 관한 고찰, 형사법의 신동향, 통권 제51호, 2016, 215쪽.

99) 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에의 적용에 관한 고찰, 형사법의 신동향, 통권 제51호, 2016, 213-214쪽.

100) 딥러닝은 데이터의 복잡한 표현을 다른 간단한 표현들을 바탕으로 정의함으로써, 머신러닝이 아주 효율적으로 작동하도록 만든 것이다. 가령 사진에 자동차가 있는지 여부를 판단하는 딥러닝 인공지능은 동글다 또는 뾰족하다는 등의 간단한 특성을 먼저 짚어내고, 이를 바탕으로 더 복잡한 특성을 반복적으로 쌓아올려 최종적으로 자동차라는 특성(표현)이 데이터로부터 표현될 수 있도록 한다. 이러한 과정은 ‘동글다’, ‘뾰족하다’와 같은 간단한 특성을 판단하는 퍼셉트론(perceptron)을 층층이 쌓은 다층 퍼셉트론(Multi-Layer Perceptron)에 의하여 달성된다. 고훈수·정해빈·박도현, 인공지능과 차별, 저스티스, 통권 제171호, 2019, 208쪽.

101) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 41쪽.

102) 동 규정은 행정행위를 자동화함에 있어 인공지능 기술을 적용할 수 있음을 법률에 명시하였다는 점에서 의의가 크다. 오토 마이어(Otto Mayer)에 의하여 행정행위 개념이 정립될 무렵 예상하지 못했던 사정으로 인해, 행정이 혁신적인 변화를 맞이하게 된 것이다. 김찬희, 인공지능 기반 자동화 행정행위에 관한 헌법적 고찰, 헌법이론과 실무, 2023-A-3, 헌법재판소 헌법재판연구원, 2023, 3쪽.



## 2. 인공지능 법관

### 가. 법관의 개념, 사법작용의 의미

#### 1) 사법의 개념, 특성 및 기능

사법(司法)이란 구체적인 법적 분쟁이 발생한 경우에, 독립적 지위를 가진 법관이 법을 판단·선언함으로써 개인의 권리를 보호하고 법질서를 유지하는 국가작용을 일컫는다. 우리 헌법상 실질적 의미의 사법은 법원과 헌법재판소가 분장하고 있다.<sup>103)</sup>

입법과 행정이 적극적 국가권력임에 비하여, 사법은 당사자가 소를 제기해야 비로소 발동되는 소극적인 성격의 국가권력이다. 또한 독립성과 중립성은 사법의 본질적이고도 핵심적인 징표이며, 과제이기도 하다.<sup>104)</sup> 사법은 분쟁의 해결, 권리보호, 권력통제를 그 기능으로 하며, 이러한 기능을 수행함으로써 법치주의와 권력분립 수호자라는 지위 및 과제를 가지게 된다.<sup>105)</sup>

#### 2) 헌법과 법률이 예정하는 법관

헌법은 사법권이 법관으로 구성된 법원에 속하며(제101조 제1항), 또한 모든 국민은 헌법과 법률이 정한 법관에 의하여 법률에 의한 재판을 받을 권리를 가진다고 규정한다(제27조 제1항). 법관은 헌법과 법률에 의하여 그 양심에 따라 독립하여 심판한다(제103조). 법관은 탄핵 또는 금고 이상의 형의 선고에 의하지 아니하고는 파면되지 아니하며, 징계처분에 의하지 아니하고는 정직·감봉 기타 불리한 처분을 받지 아니한다(제106조 제1항). 한편 법관의 종류, 자격, 정년, 임기 등의 사항은 법원조직법에서 상세히 규정한다.<sup>106)</sup>

‘양심’에 따라 심판한다는 규정이나 신분보장에 관한 규정에 비추어, 헌법이 자연인을 법관으로 예정하고 있음을 알 수 있다. 법원조직법은 대법원장과 대법관이 20년 이상 판사·검사·변호사 등의 직에 있던 45세 이상의 ‘사람’ 중에서 임용한다고 규정하며(제42조 제1항), 판사는 10년 이상 일정한 직에 있던 ‘사람’ 중에서 임용한다고 규정한다(동조 제

103) 김하열, 헌법강의, 박영사, 2023, 915쪽.

104) 김하열, 헌법강의, 박영사, 2023, 916쪽.

105) 김하열, 헌법강의, 박영사, 2023, 916-917쪽.

106) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 82쪽.

2항). 법관의 결격사유(제43조), 정년(제45조)에 관한 법원조직법 규정 역시, 자연인 법관을 당연한 전제로 하고 있다.<sup>107)</sup>

### 3) 법관의 업무

헌법에 따라 사법권을 행사하는 법원은 원칙적으로 모든 법률상의 쟁송을 심판하는 권한을 가진다(법원조직법 제2조 제1항). 다만 행정기관에 의한 전심(前審)으로서의 심판 등은 예외이다. 법원은 그 밖에 등기, 가족관계등록, 공탁, 집행관, 법무사에 관한 사무 관장·감독 업무를 한다(동조 제3항).

따라서 법원을 구성하는 법관은 법률상의 쟁송을 심판하는 역할을 담당하는데, 법률상의 쟁송으로는 소송업무(민사, 형사, 가사, 행정 등)와 비소송업무(이혼·상속 시의 재산분할, 경매절차에 따른 배당 등 집행, 회생·파산절차에서의 관리·감독 등)가 포함된다.<sup>108)</sup>

## 나. 사법 분야에서 인공지능의 활용 현주소

### 1) 개관

인공지능을 활용하여 법적 판단을 내리려는 시도는 1970년대 후반부터 이루어졌다. 당시에는 주로 법률지식과 판단을 컴퓨터 데이터로 표현하기 위한 연구가 수행되었다. 1980년대에는 법률적 판단(legal reasoning)을 논리적 알고리즘으로 전환하는 것이 주된 목표였으며, 전문가시스템을 기본으로 하였다. 다만 이 당시 컴퓨터 연산능력에는 한계가 있었고, 무수히 많은 판단기준을 수기로 입력하는 것이 불가능하였다. 이후 한동안 사법 분야에서 인공지능은 유의미한 발전을 보이지 못하였다.<sup>109)</sup>

컴퓨터 연산처리능력의 비약적인 향상과 빅데이터의 집적, 머신러닝의 등장에 힘입어, 인공지능은 법률문서 작성, 법률자문, 소송의 진행, 승소가능성 예측, 사건에 적합한 법률

107) 인공지능 법관을 도입하려면 헌법과 법률 규정의 개정이 불가피하다는 지적으로 전후재, 우리 민사재판의 미래에 관한 소고 - 디스커버리, 온라인 소송, 인공지능(AI)을 중심으로 -, 민사소송, 제27권 제1호, 2023, 139쪽. 헌법 제105조, 제106조가 인공지능 법관에게 적용되지 않는다는 특례조항을 헌법에 추가하고, 법원조직법 제42조의 법관 자격을 변경하는 형태로 개정되리라고 전망한다.

108) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 85쪽.

109) 서치원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 '법의 제국'을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 3쪽.

사무소 추천 및 소개, 기계학습 등 다방면으로 활용된다. 자연인 변호사의 법률사무를 보조하는 인공지능은 이미 활용 단계에 있으며,<sup>110)</sup> 이른바 리걸테크가 법원실무에 이미 도래하여 법률서비스와 첨단 기술의 융합은 가속화되고 있다.<sup>111)</sup> 이른바 전자증거개시제도(e-Discovery)가 도입될 경우, 인공지능 활용의 필요성은 더욱 증가하리라고 전망된다.<sup>112)</sup> 다만 인공지능은 과거의 자료를 바탕으로 예상·자문을 하고 창조적이고 유연한 분석을 하는 단계에 이르지 못하였으며, 공감능력이 없고 우발적인 상황에 대처하는 전략이 부재하다는 한계가 지적된다.<sup>113)</sup>

사법 분야에서의 인공지능 활용은 분석 내지 예측 대상자의 여러 기본권과 관련한 문제를 내재하고 있다. 특히나 헌법적 쟁점이 많이 불거지는 것은 재범위험성 예측 프로그램이다. 미국에서는 재범위험성 예측 프로그램이 보호관찰 및 가석방 부서에서 사용되었으며, 이제는 구금 여부에 관한 판단 및 양형분야로 확장되고 있다. 미국은 지나치게 많은 인원의 수감 문제를 겪어왔는데, 이를 개선하려는 노력으로 각 주와 시들이 2010년대부터 알고리즘을 활용하는 위험성 평가도구(risk assessment tool) 도입을 검토해온 것이다. 인공지능 알고리즘을 활용하면 재범위험성이 낮고 공판에 성실하게 출석할 사람들을 가려내, 구속되거나 실형을 선고받아 구금되는 수감자의 수를 획기적으로 줄일 수 있다는 주장이 있어왔다.<sup>114)</sup>

## 2) 국내

한국의 경우 아직까지 인공지능 알고리즘을 통해 범죄·재범의 위험성을 예측하여 가석방·양형판단 등에 활용하고 있지는 않다. 다만 사법부에서의 인공지능 활용은 가시화되고 있다.

대법원은 차세대 전자소송시스템 구축(일명 ‘스마트법원 4.0’) 사업을 추진 중이며,

110) 영국의 대형로펌 알렌 앤 오버리(Allen & Overy)는 챗GPT 기술을 기반으로 한 인공지능 법률 챗봇 ‘하비’를 베타(시험) 버전으로 출시했는데, 계약서 초안 작성·판례분석·특정 판결의 예측 등을 수행한다. 다만 해당 로펌은 인공지능 챗봇의 답변을 곧이곧대로 받아들이지는 않으며, 자연인 변호사들이 사실 검증을 한다. 홍윤지, [취재수첩] AI와 법률가의 공생(共生), 법률신문, 2023. 2. 23, <https://www.lawtimes.co.kr/Legal-Opinion/Legal-Opinion-View?serial=185534&kind=BA14>(최종방문일 2024. 2. 11.).

111) 이명민, 인공지능, 법, 그리고 민사소송, 안암법학, 제57호, 2018, 37쪽.

112) 강영주, 인공지능을 활용한 민사 분쟁 해결에 관한 고찰, 안암법학, 제64권, 2022, 422쪽.

113) 강승식, 인공지능 판사, 과연 가능한가?, 헌법학연구, 제26권 제3호, 2020, 233-234쪽.

114) 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 910쪽.

2024년 서비스 개시를 목표로 한다. 현재의 재판업무시스템이 클라우드·빅데이터·인공지능 등 신기술 확산에 대응하기 어렵다는 문제의식 하에, 사법부 재판업무 전 과정의 표준화·자동화·간소화·지능화를 도모하고 있다. 구체적으로는 사법정보공개체계의 혁신, 국민중심사법서비스의 혁신, 지능형 사건관리 기반 재판사무 혁신, 디지털법원 실현을 위한 IT구조 개편을 추진 중이다.<sup>115)</sup>

차세대 전자소송시스템은 인공지능 알고리즘이 자연인 법관의 판단을 대체하는 등의 내용을 담고 있지는 않다. 다만 인공지능은 소송절차의 안내, 소장 작성 지원, 질의의 의도를 파악하는 통합검색 등에 활용될 예정이다.<sup>116)</sup> 또한 소송기록의 전면 디지털화가 이루어진다면, 향후 인공지능 법관의 훈련용 빅데이터로 활용될 여지가 있다.<sup>117)</sup>

[그림1: 대국민 사법서비스 개선 개요]<sup>118)</sup>



115) 차세대전자소송 추진단, 차세대전자소송 시스템 구축 사업 착수보고회 자료집, 2020, 1쪽. 당해 사업은 기본적으로 예산투입 과제이나, 세부과제에 따라서는 법률개정을 통한 입법적 뒷받침이 필요하다. 박준모, 스마트법원 4.0의 입법과제 - 예금 압류명령 전자정보 송신 제도 도입을 중심으로, NARS 현안 분석, 제99호, 국회입법조사처, 2019, 3쪽; 이성진, 대법원 “차세대 전자소송 시스템 구축한다”, 법률신문, 2018. 4. 12, <http://www.lec.co.kr/news/articleView.html?idxno=47291>(최종방문일 2024. 2. 11.).

116) 천대엽 법원행정처장은 2024년 1월 15일 취임사에서, 재판과 민원업무의 인공지능 활용을 위한 방안이 마련되어야 한다고 하였다. 또한 차세대 사법전산시스템의 시작과 고도화를 통해 재판업무의 시·공간적 한계를 극복하며, 미래세대의 가치와 시각에서 재판지연을 해소할 수 있는 창의적 방안도 연구·도입되어야 한다고 하였다. 이용경, 천대엽 법원행정처장, 법관 증원 및 사법부 예산 확보 약속, 법률신문, 2024. 1. 15, <https://www.lawtimes.co.kr/news/194998>(최종방문일 2024. 2. 11.).

117) 추진목표인 사법정보 공개체계 혁신에는 사법정보 공개 채널의 단일화, 판결문 등 사법정보 제공 확대 등의 내용이 담겨 있다. 차세대전자소송 추진단, 차세대전자소송 시스템 구축 사업 착수보고회 자료집, 2020, 7쪽.

118) 이성진, 대법원 “차세대 전자소송 시스템 구축한다”, 법률신문, 2018. 4. 12, <http://www.lec.co.kr/news/article>

[그림 2: 지능형 법원 도입 개요]<sup>119)</sup>



최근에는 판사들을 주축으로 법원 내에 ‘인공지능 커뮤니티’라는 연구회가 출범하였다. 33명의 회원이 참여하는 이 연구회의 목적은 재판업무 지원을 위한 인공지능 기술의 도입방안 및 인공지능 관련 법률문제 연구, 자료 소개, 세미나 개최 등을 통하여 인공지능 기술을 사법부에 도입해 재판업무의 충실화·효율화·신속화 기반을 마련하는 데 있다.<sup>120)</sup>

### 3) 주요 국가들의 동향

미국에서 인공지능은 민간 법률서비스를 중심으로 발전하고 있는데, 사법 시스템에서의 인공지능 활용은 주로 형사소송의 맥락에서 논의된다. 특히 통계처리를 통해 피의자·피고인·수형자의 재범위험성을 예측하는 인공지능 알고리즘이 보호관찰, 가석방, 양형 단계에서 널리 활용되고 있다.<sup>121)</sup> 후술하듯, 법관이 양형 판단을 할 때 인공지능 알고리즘의 재범위험성 평가 결과를 참고하는 것이 피고인의 적법절차에 따른 권리를 침해하는지 여부가 다투어진 사례도 있다.

View.html?idxno=47291(최종방문일 2024. 2. 11.)에서 발췌.

119) 이성진, 대법원 “차세대 전자소송 시스템 구축한다”, 법률신문, 2018. 4. 12, <http://www.lec.co.kr/news/articleView.html?idxno=47291>(최종방문일 2024. 2. 11.)에서 발췌.

120) 한수현, 법원 AI재판업무 지원 연구 위한 ‘인공지능 커뮤니티’ 출범, 법률신문, 2023. 11. 16, <https://www.lawtimes.co.kr/news/193149>(최종방문일 2024. 2. 11.).

121) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 89쪽.

오스트리아는 전자사법(e-justice) 분야에서 유럽 국가들 중 가장 선도적인 나라의 하나이다. 머신러닝 및 딥러닝 알고리즘을 신속하고 효율적인 문서분석 및 처리에 주로 활용하고 있으며, 사건을 신속하게 처리하고 해당 절차와 관련된 연결 관계를 분석하기 위해 사법부 내에 정보기술(IT) 전문가들의 지원이 이루어지고 있다. 판결을 공개할 목적으로 인공지능을 통한 익명화가 시행되기도 하였다.<sup>122)</sup>

프랑스 사법부는 비교적 일찍부터 예측적 사법에 관심을 표명해왔다. 2017년 렌(Rennes)과 두에(Douai) 고등법원은 ‘프레딕티스’라는 예측적 사법 소프트웨어를 시범적으로 적용하기도 하였다. 해당 소프트웨어는 해고 시 퇴직수당 액수를 산정하는 분야에 적용되었으나, 실제 상용화가 가능한 수준에는 이르지 못한다고 자체 평가되었다. 프랑스 사법부는 예측적 사법 시스템의 통계적 분성이 가능할 수 있도록, 사법 빅데이터를 구축하고 수집·분석된 데이터를 효율적으로 활용하는 방안을 고안하고 있다고 한다.<sup>123)</sup>

중국은 스마트 사법으로의 전환을 적극 추진하면서<sup>124)</sup> 2018년 기준으로 전국 법원에 100대 이상의 로봇을 보유하고 있었는데, 일부 로봇은 상법이나 노동관련 분쟁에서의 전문성을 갖추었다.<sup>125)</sup> 최고인민법원 정보센터의 쉬젠핑(Xu Jianfeng) 국장은 중국공정원(Chinese Academy of Engineering: CAE)이 발간하는 ‘전략연구’에 수록된 보고서에서, “스마트 법원 시스템이 중국 전역 모든 판사의 책상과 연결된다.”고 하였다. 머신러닝 기술로 구동되는 시스템이 참고용 판례를 자동 검색하고, 관련 규정을 추천하며, 법문서 초안을 작성하고, 평결에서 오류가 인지되면 이를 변경한다는 것이다. 그에 따르면 2019년부터 2021년까지 인공지능이 중국 법관들의 업무를 3분의 1 경감하였다고 한다. 최고인민법원의 요구에 따라 법관들은 모든 사건과 관련하여 인공지능의 상담을 받아야 하는데, 인공지능의 추천을 거부하는 판사에게는 기록과 감사를 위한 설명을 서면으로 제출하도록 요구한다고 한다.<sup>126)</sup>

---

122) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 110쪽.

123) 보다 상세히는 정채연, 사법절차 및 사법서비스에서 인공지능 기술의 도입 및 수용을 위한 정책 연구, 사법정책연구원, 2021, 44-47쪽.

124) 중국은 가장 적극적으로 사법, 재판 영역에 인공지능을 도입하고자 하는 국가의 하나로 꼽힌다. 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 98쪽.

125) Briony Harris, Could an AI Ever Replace a Judge in Court?, World Government Summit, 11 July 2018, <https://www.worldgovernmentsummit.org/observer/articles/2017/detail/could-an-ai-ever-replace-a-judge-in-court>(최종방문일 2024. 2. 11.).

126) 윤고은, 중국 판사들, 모든 사건에 대해 AI 상담 받는다, 연합뉴스, 2022. 7. 14, <https://www.yna.co.kr/view/AKR20220714150300074?input=1195m>(최종방문일 2024. 2. 11.); Tiffany Winfrey, China Uses Artificial

한편 에스토니아는 2019년 소액사건에서의 인공지능 법관 개발 프로젝트에 착수하였다고 알려지고 다수의 국내외 선행연구에서 인공지능 법관 도입과 관련한 대표적인 사례로 소개되었으나,<sup>127)</sup> 2022년 2월 16일 에스토니아 법무부는 그러한 프로젝트를 실시하거나 계획한 바 없으며 다만 법원의 업무부담을 경감하기 위한 정보통신기술 수단을 모색하고 있다고 하였다.<sup>128)</sup>

#### 4) 법관의 사건처리 단계별 인공지능의 발전 현황

판사는 사건기록(소장, 준비서면을 비롯한 주장서면들과 서증, 녹음파일, 영상파일 등 증거까지의 소송자료들을 포괄)을 검토한다. 이후 변론기일에는 당사자들, 소송대리인이 출석한 상태에서 핵심 쟁점에 관한 구술 공방을 주재한다. 마지막 변론기일이 끝나면, 그 때까지 제출된 주장과 증거를 토대로 판결문을 작성하며 선고함으로써 한 심급이 마무리된다. 더 크게 나누어본다면 (1) 읽고 알아듣는 단계와 (2) 생각하고 판단하는 단계로 대별된다.<sup>129)</sup>

##### 가) 소송자료의 인식

우리나라는 전자소송이 활성화되어 있으나, 소장이나 서증 등이 스캔되었더라도 데이터를 추출하기 위해 한 차례 더 문자정보로 변환해야 하는 경우가 많다.<sup>130)</sup>

일상어의 경우 자연어 처리기술이 비약적으로 발전하고 있으나,<sup>131)</sup> 아직까지 법률용어

---

Intelligence (AI) to Run Courts, Supreme Justices; Cutting Judges' Typical Workload By More Than a Third and Saving Billion Work Hours, The Science Times, 14 July 2022, <https://www.sciencetimes.com/articles/38760/20220714/china-now-runs-its-courts-supreme-justices-through-artificial-intelligence.htm>(최종방문일 2024. 2. 11.) 호주, 영국, 프랑스 사법부에서의 지능정보화 정책 동향에 대하여는 정채연, 사법절차 및 사법서비스에서 인공지능 기술의 도입 및 수용을 위한 정책 연구, 사법정책연구원, 2021, 32-48쪽.

127) Vigitė Vėbraitė and Goda Strikaitė-Latušinskaja, Digitalization of Justice in Lithuania, in Katarzyna Gajda- Roszczynialska ed., Impact of the COVID-19 Pandemic on Justice Systems, V&R unipress, 2023, pp. 230-231. 223-234; 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 108쪽 외 다수.

128) Maria-Elisa Tuulik, Estonia does not develop AI Judge, Republic of Estonia Ministry of Justice, 16 February 2022, <https://www.just.ee/en/news/estonia-does-not-develop-ai-judge>(최종방문일 2024. 2. 11.).

129) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 7쪽.

130) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 7-8쪽.

131) 자연어 처리(natural language processing)는 컴퓨터가 일상어로 사람들과 소통할 수 있도록 하는 시스템 개발에 주안점을 둔다. 이를 통해 컴퓨터는 텍스트를 읽고 음성을 들으며, 해석하고, 어떤 부분이 중요한지 판단할 수 있다. 이는 대규모의 자연어 데이터를 처리하고 분석할 수 있는 머신러닝을 통해

의 자연어 처리기술은 그 발전이 더디다. 그 원인으로는 체계화되지 않은 법률문서들을 데이터 객체로 변환하는 도구가 부족하다는 점 등이 지적된다. 현재의 자연어 처리 기술 수준에서 문장의 전체 맥락을 이해하고 법적 판단을 내리는 인공지능은 당분간 어렵다고 이야기된다.<sup>132)</sup>

## 나) 소송자료 분석을 토대로 한 판단

### (1) 결론의 도출

인공지능 판사는 법령과 판결문을 일차적으로 학습하는바, 한국의 경우 적어도 법관이 수십만 개의 판결·결정문을 전자적으로 해독할 수 있는 상태로 데이터베이스화되어 있다. 다만 빅데이터의 관점에서는 많은 양이 아닐 수도 있다.<sup>133)</sup>

문제는 판결문이 당해 사건의 정수(精髓)를 담고 있는가이다. 판결문 가운데에는 결론과 직접 관련된 사실만 간략히 기재한 것도 많으며, 민사 소액사건의 경우 이유 기재를 생략할 수도 있다(소액사건심판법 제11조의2 제3항).<sup>134)</sup> 또한 법관 개개인마다 판결문을 작성하는 스타일이 상이하며, 이 때문에 인공지능이 편향된 학습을 할 가능성도 있다. 소송자료의 경우 당사자나 제3자가 작성한 것이 상당 부분이며, 이들이 그 공개에 동의했다고 보기 어려우므로 인공지능 훈련 데이터로 제공하는 데 신중하여야 한다.<sup>135)</sup>

### (2) 이유 기재

사법절차는 설명과 설득이 가능해야 한다. 그런데 인공지능은 결론을 도출하는 과정을 설명하는 능력 측면에서 고전을 면치 못하고 있다. 사적 구제를 금지하고 분쟁을 해결하도록 설계된 사법제도에서 개별 결정은 정당화될 필요가 있는바, 수학적·통계적으로 가장 상관관계가 높은 답을 제시하는 것을 넘어 법적 논리를 제공할 수 있어야 한다.<sup>136)</sup>

---

가능하다. 최근에는 고급 공식을 사용하여 언어를 더 짧고 요소적인 조각들로 분해하고, 조각들의 관계를 파악하며, 언어 조각들이 어떻게 함께 의미를 만들어내는지 탐구한다. David M. Masuhara, *Artificial Intelligence and Adjudication: Some Perspectives*, Amicus Curiae, Vol. 111, 2017, p. 3.

132) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 8-9쪽.

133) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 9-10쪽.

134) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 10-11쪽.

135) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 11쪽.

136) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권



하지만 현재의 인공지능기술은 법적 논증 과정을 설득력 있게 설명하는 능력이 부족하다.<sup>137)</sup> 인공지능 머신러닝은 결론에 이르는 논증과정을 제대로 파악하기가 힘들며, 이를 블랙박스 문제라고도 한다. 후술하는 것처럼, 이 문제는 규범적 차원에서 공정한 재판을 받을 권리 등과 관련한 문제를 자아낸다.

## 다. 사법 분야 인공지능 관련 위험성과 윤리적·헌법적 문제의 중대성

각국과 국제사회는 인공지능과 관련하여 가이드라인, 권고 등으로 대처해 왔으나, 최근에는 인공지능 알고리즘을 규율하는 일반법을 제정하려는 입법적 움직임이 감지된다. 유럽연합의 인공지능법안은 위험(risk)을 크게 용인할 수 없는 위험, 고위험, 낮은·최소한의 위험으로 분류하는 위험기반 규제다. 법률의 해석·적용 등은 고위험 인공지능 시스템으로 분류된다.<sup>138)</sup> 고위험 인공지능은 적합성 평가를 받고 다양한 의무를 준수하도록 하고 있다.<sup>139)</sup>

인공지능의 사법 분야 활용에 있어 가이드라인을 제시하려는 국제사회의 노력도 이루어지고 있다. 가령 사법시스템과 사법 환경에서의 인공지능 이용에 관한 유럽 윤리헌장(European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment)은 5대 윤리원칙으로 기본권 존중의 원칙, 차별금지원칙, 품질과 보안의 원칙, 투명성·불편부당성·공정성 원칙, 이용자에 의한 통제 원칙을 제시하였다.<sup>140)</sup>

## 라. 현 단계에서 상정할 수 있는 인공지능 법관

오늘날 인공지능 알고리즘은 각국에서 법률문서 초안의 작성, 판결·결정의 검색, 관련 규정의 추천, 재범위험성 예측 등에 활용되고 있다. 이러한 단계를 넘어, 말 그대로 ‘인공

제5호, 2019, 11쪽; 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 144쪽.

137) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 144쪽. 인공지능 알고리즘은 주어진 사실관계에 어떠한 법적 판단이 나올지 꽤나 높은 확률로 예측할 수 있으나, ‘왜’ 그런지를 제대로 설명하지 못한다고 한다.

138) 법집행(law enforcement)은 높은 리스크를 가진 인공지능시스템으로 분류된다(동 법안 제6조 제2항).

139) 김희정, 인공지능 윤리와 형사정책적 시사점 - 형사사법에서의 인공지능 윤리 이슈를 중심으로 -, 형사정책, 통권 제69호, 2022, 86-87쪽.

140) European Commission for the Efficiency of Justice(CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Council of Europe, 2018, pp. 7-12.

지능 법관'이 현실화될 수 있을까?

일부 논자들은 장차 인공지능이 법관을 대체하리라고 보는데, 여기서 '인공지능 법관'은 대체로 자연인 법관을 대신하여 재판을 진행하고 판결문을 작성하는 인공지능을 가리키는 것으로 보인다. 하지만 어떤 경우에는 '사법적 분쟁해결에서 판단자의 역할을 수행하는 판결기계'나 '재판예측 시스템'을 의미하기도 한다.<sup>141)</sup>

학계에서는 자연인 법관을 완전히 대체하여 심문·신문 등 재판의 진행까지 담당하는 로봇 인공지능 법관은 현재의 기술로 불가능하며,<sup>142)</sup> 그보다 낮은 단계에서의 활용이 장려되어야 한다는 견해가 유력하다.<sup>143)</sup> 자연인 법관과 동일한 인공지능 법관을 구현하려면 인공지능이 특정 영역이 아닌 모든 측면에서 인간과 동일하거나 인간보다 우월한 수준의 강한 인공지능이 되어야 하는데, 가까운 시일 내에 강한 인공지능이 개발·활용될 가능성은 매우 희박하기 때문이다.<sup>144)</sup> 인공지능이 다양한 심리원칙, 특히 자유심증주의를 준수하며 사실관계를 확정하기 어려우리라는 평가도 유력하다.<sup>145)</sup> 사견으로도 인공지

141) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 137-138쪽.

142) 인간과 같은 사고체계를 갖춘 강한 인공지능에 관한 논의는 자칫 비현실적 구상이 될 위험이 있음을 지적한 연구로 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에의 적용에 관한 고찰, 형사법의 신동향, 통권 제51호, 2016, 216쪽. 지금까지 인공지능 발전은 주로 지능의 감각 및 인지 기능에 집중되었으며, 동작이나 행동 지능의 측면에서는 많이 미비하다고 한 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 60쪽도 참조.

143) 사건마다 소통과 공감능력을 바탕으로 당사자들의 목소를 경청하고 양심에 따른 고뇌를 거쳐 시대 정의에 부합하는 판단을 하는 인공지능 판사는 현재로서 공상과학에 가깝다는 평가로 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 909쪽. 현재 기술에 비추어 인공지능 변호사, 인공지능 판사는 공상과학에 더 가까우며, 낮은 단계에서의 활용이 오히려 더 권장될 필요가 있다는 주장으로 김병필, 재판 예측 인공지능 기술의 현황과 한계, 저스티스, 통권 제182-2호, 2021, 112쪽. 오세용 부장판사는 인공지능 법관의 출현이 향후에도 불가능하다고 단언할 수는 없겠으나, 현재의 기술수준 및 재판 환경으로서는 당분간 요원하다고 본다. 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 145쪽.

기일의 지정, 변론주재, 사실관계 파악 및 관련 법 적용, 판결문 작성 등에 이르기까지 인간과 동일한 수준의 판단과 움직임을 보이는 인공지능은 현재의 기술로 불가능하다고 한 김성용·정관영, 법조(法曹) 영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 76쪽도 참조.

144) 한애라, AI 판사?, 법률신문, 2020. 1. 20, <https://m.lawtimes.co.kr/Content/Opinion?serial=158857>(최종 방문일 2024. 2. 11.).

145) 자유심증주의란, 법원이 변론 전체의 취지와 증거조사의 결과를 참작하여 자유로운 심증으로 사회정의와 형평에 입각하여 논리와 경험의 법칙에 따라 사실주장이 진실한지 아닌지를 판단한다는 의미이다(민사소송법 제202조). 형사소송에서는 엄격한 증거재판주의(형사소송법 제307조)가 적용되므로 증거자료만이 판단자료가 되지만, 민사소송에서는 증거자료 외에 변론전체의 취지도 판단의 자료가 된다. 전원열, 민사소송법, 박영사, 2022, 364쪽. 그런데 변론 전체의 취지에는 당사자 또는 대리인의 주장내용 및 태도, 주장·입증의 시기, 공동당사자의 자백 및 태도 등이 모두 포함된다. 이시윤, 신민사소

능을 탑재하고 인간과 유사한 외형을 갖춘 채 재판과 관련한 모든 기능을 수행하는 인공지능 법관이 가까운 미래에 현실화되기는 쉽지 않으리라고 전망한다.<sup>146)</sup>

한편 인공지능의 작동원리나 일각에서 인공지능 법관을 기대하는 배경에 비추어볼 때, 자연인 법관과 ‘동일한’ 인공지능에 관한 논의는 다소 적절치 못한 것일 수 있다. 뒤에서 자세히 볼 것처럼 오늘날 주류적인 머신러닝 내지 딥러닝 방식의 약한 인공지능은 인간과 같은 추론을 통해 ‘답’을 내는 것이 아니며, 매우 높은 정확도로 예측하는 시스템이다. 인간은 무언가를 알지만, 인공지능은 이해하지 못한다. 그렇다면 법적 추론을 통해 사실 관계를 파악하고 관련규정의 의미를 이해하여 당해 사안에 포섭하는 작업은 적어도 현재의 인공지능 작동원리와는 상이하다. 또한 법관을 비롯한 인간은 기억용량이나 제한된 정보처리능력, 체력적 한계, 시공간적 제약으로부터 자유롭지 못하며, 유사한 사안에서도 상이한 결론을 도출하기도 한다. 논리 이외의 무언가<sup>147)</sup>에 의식적·무의식적으로 근거하여 판단하기도 한다.<sup>148)</sup> 법관은 평범한 인간으로서 정서의 영향을 받으며 휴리스틱 판단을 배제하는 것은 불가능하고, 오히려 감성의 역할이 기대되는 측면이 있다는 지적도 있다.<sup>149)</sup>

---

송법, 박영사, 2023, 364쪽. 관련하여, 인공지능 법관이 변론 전체의 취지를 고려하여 판단하려면 종합적 사고가 가능하여야 하는데 현 시점에서는 불가능하다는 평가가 유력하다. 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 72-73쪽. 인공지능에게서 자유로운 심증 형성을 기대할 수 없으며, 사회정의와 형평의 이념에 따른 판단을 구현하기도 어렵다고 한다.

146) 선행연구에 따르면, 외형적 활동은 로봇이, 내적인 활동은 인공지능이 각각 담당하면서 로봇과 인공지능이 결합하면 물리적·정신적 활동이 가능하게 된다. 안상원·박규동, 형사사법 절차상 인공지능(AI) 시스템 도입의 주요 논점, 범죄정보연구, 제16호, 2022, 138쪽. 인공지능을 탑재하고 인간과 유사한 외형을 가진 로봇 인공지능 법관을 생각해볼 수 있다.

그런데 아무렇게나 쓴 문서, 웅얼거리는 증인, 증언부언하거나 횡설수설하는 당사자들의 진술을 이해하며 재판을 진행하고 사실관계를 판단할 수 있으려면, 말 그대로 자연인이 할 수 있는 모든 것이 구현되는 수준에 이르러야 하지 않을까? 인공지능이 유의미한 학습을 하려면 판결문·결정문만을 데이터로 삼아서만 안 되고(판결문은 판사가 사실관계, 적용법조 등을 자신의 관점에서 정제한 문서), 사건 기록, 증거 일체, 구두변론 내용 모두가 그 대상이 되어야 할 것이다. 그런데 개인정보보호, 사생활, 전산화되지 않은 여러 사건들을 고려하면 단기간 내에는 어려울 수 있다.

147) 감정, 선입관, 개인적 호불호, 삶의 행적, 친분관계가 그 예이다. 심지어 금품·향응을 제공받고 편파적인 판단을 하는 경우도 있다.

148) 가령 노여움, 역겨움이라는 기본정서는 직관적인 사고를 촉발하여 엄벌 지향적인 성향을 높이는 반면, 두려움과 슬픔이라는 정서는 오판에 대한 염려에서 분석적 사고를 유발하고 관대한 결론을 지향하는 성향을 높인다고 한다. 김현석, 재판의 진행·판단에서 법관의 정서에 관하여, 법조, 통권 제733호, 2019, 254, 281-282쪽.

149) 김현석, 재판의 진행·판단에서 법관의 정서에 관하여, 법조, 통권 제733호, 2019, 278쪽. 변호사의 미래 또한 인공지능으로 대체할 수 없는 인간적인 자질(human qualities)에 달려 있다는 지적이 있다. 자연인 변호사는 고객의 말을 공감하고 경청하며, 분명하지 않을 수 있는 이야기를 파악하기 위해 전략

그렇다면 천문학적인 비용과 최첨단의 기술을 동원하여, ‘여러 한계를 지닌 자연인 법관과 동일한’ 인공지능 법관을 구현한다는 발상이 현실화될 가능성은 그리 높지 않다.<sup>150)</sup> 인공지능은 쉬지 않고 인간보다 훨씬 더 빠르게 방대한 정보를 처리하며, 높은 정확도로 무언가를 예측·감지하고, 동일·유사한 입력 값에 대하여 일관된 출력 값을 내놓기 때문에 각광받으며 상용화되고 있다.

그렇다면 인공지능 법관이라는 화두는 - 자연인 법관과 동일한 존재로 자리매김하는 문제라기보다는 - 방대한 정보를 빠르게 처리하여 일관되며 정확하게 무언가를 예측·인식하는 인공지능이 재판에서 어떠한 역할을 할 것인지의 문제로 파악해볼 수 있다. 만약 인공지능이 자연인 법관의 판단을 보조하고 참고자료를 제시할 뿐이라면, (현)법적 논쟁의 여지가 적다. 재판진행 등 모든 재판을 자연인 법관처럼 담당하는 인공지능에 관한 논의는 시기상조인 측면이 있으나, 인공지능 알고리즘이 제시하는 출력 값이 자연인 법관의 ‘판단을 대체’하는 단계에 이르러야 ‘인공지능 법관’이라고 칭할 수 있다. 다만 후술할 *State v. Loomis* 판결에서 보듯, 인공지능 알고리즘의 판단이 법관의 판단을 갈음하는 단계에 이르지 않더라도, 적법절차에 관한 권리의 침해 여부 등 여러 헌법적 쟁점이 불거질 수 있다.

가령 인공지능 알고리즘이 예측한 재범위험성을 자연인 법관이 그대로 수용해 취업제한, 신상정보등록 등 보안처분을 부과한다면, 해당 재판에서 인공지능이 사실상 법관으로서 기능하였다고 볼 수 있다. 인공지능 알고리즘이 산정한 과실비율에 따라 손해배상액이 정해지는 경우도 그러하다. 다만 여기서 ‘인공지능 법관’이란 헌법에 의하여 그 신분이 보장되며 양심에 따라 판단하는 자연인 법관 그 자체를 지칭한다기보다는, 과실비율이나 재범위험성 등에 관한 사법적 판단을 내린다는(구속력 있는 결과 값을 제시한다는) 의미이다.

이와 달리, 사람과 같은 방식으로 추론하며 감정을 가지는 강한 인공지능이 출현하고 여기에 헌법상 ‘법관’의 지위를 인정하는 것은 적어도 현행헌법 하에서는 받아들여지기 힘들며, 향후 개헌이 이루어져야 가능하리라 판단된다. 헌법은 법관이 ‘양심에 따라’ 심

---

을 세우고, 상상력과 판단력을 발휘하여 변론이 호소력을 가질 수 있도록 한다는 것이다. Melissa Love Koenig, Julie A. Oseid and Amy Vorenberg, *Ok Google, Will Artificial Intelligence Replace Human Lawyering*, *Marquette Law Review*, Vol. 102, No. 4, 2019, p. 1271.

150) 인공지능 법관 개발보다 훨씬 용이한 법률 데이터베이스 검색기능을 고도화하는 작업에도 엄청난 자금이 소요되었음을 지적한 오세용, *인공지능시대: 법관의 미래는?*, 박영사, 2022, 147쪽 참조.

판하며(제103조), ‘중대한 심신상의 장애’로 직무를 수 없을 때 법률이 정하는 바에 의하여 퇴직하게 할 수 있다고 규정한다(제106조 제2항). 법관의 신분 보장에 관한 헌법 제106조 제1항도 법관이 ‘자연인’임을 전제한 규정으로 읽힌다. 다만 전술하였듯 현재의 과학기술수준이나 경제논리를 고려할 때, 이러한 의미에서의 인공지능 법관이 가까운 미래에 현실화될 가능성은 낮다고 생각한다.

## 마. 인공지능 법관의 주요 특성

그렇다면 자연인 법관의 판단을 대체하여 재판에서 어떤 결과 값(예: 재범의 위험성 예측)을 내놓는 인공지능 법관은 어떠한 특성을 가지는가? 아직 인공지능 법관이 현실화되지 않아 그 특징을 면밀하게 포착하여 서술하기는 어려우나, 오늘날 주류적인 인공지능의 속성을 고려할 때, 다음과 같이 인공지능 법관의 특성을 생각해볼 수 있다.

### 1) 알고리즘의 편향성

적어도 오늘날의 인공지능(약한 인공지능)은 자유의지를 가지지 않으므로 편견이 성립할 여지가 거의 없으며, 입·출력 관계가 일관되어 무의식적인 편견이 개입할 여지가 적다. 설령 차별이 발생하더라도, 인공지능은 인간처럼 은폐에 능숙하지 않기 때문에 입·출력 경로에 용이하게 접근하여 규율할 수 있기도 하다.<sup>151)</sup>

그런데 알고리즘이 과학적이고 가치중립적이라는 시각도 있으나, 현실이 반드시 그렇지 않다.<sup>152)</sup> 인공지능의 의사결정은 인간이 볼 때 차별이라고 여겨지는 결과를 산출하고는 한다.<sup>153)</sup> 인공지능 알고리즘 편향성의 원인으로는 설계·개발자의 편향성, 사회에 내재하는 편향성이 정제되지 못한 학습 데이터, 오류 및 편향성을 지닌 알고리즘 등이 꼽힌다. 편향성 관련 헌법적 쟁점들에 관하여 상세히는 후술한다.

151) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 159쪽.

152) 알고리즘이 과학적이고 가치중립적이라는 것은 환상이라는 지적으로 김종철, 인공지능의 헌법적 의의, 오병철 외, 인공지능과 법, 연세대학교 출판문화원, 2023, 40쪽. 알고리즘 자체가 불투명한데다, 편향의 개연성이 높은 기성 정보에 의존한다는 점을 지적하고 있다.

153) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 160쪽.

## 2) 알고리즘의 불투명성

인공지능 알고리즘은 대부분 사기업의 영업비밀이기에 그 작동구조를 파악하기가 어렵다. Loomis v. Wisconsin 사건에서 피고인은 재범위험성을 예측하는 인공지능 알고리즘인 COMPAS의 소스코드와 가중치에 관한 정보를 요구하였지만, 법원은 COMPAS 알고리즘이 지적재산권법 및 기업비밀 관련법에 의하여 보호된다는 이유에서 이를 거부하였다.<sup>154)</sup>

설령 알고리즘을 모두 공개하더라도, 일반인 심지어 법관도 관련 전문지식이 부족하므로 그 원리를 이해하기 힘들다.<sup>155)</sup> 오늘날 주류적인 딥러닝 인공지능에서는, 데이터로부터 특징(feature)이라 불리는 독립변수를 추출하는 과정이 기계로 이전된다. 그에 따라 인공지능이 인간의 지식체계와 맺는 연관성이 줄어들고, 인공지능의 정보처리과정에 관한 인간의 이해가 급감한다.<sup>156)</sup> 또한 특정 인공지능 알고리즘의 핵심은 해당 개발자가 만들지만 복잡한 연산을 담당하는 라이브러리 등은 선발기업(예: 구글)에서 만든 라이브러리에 의존하는데, 그것이 너무도 복잡해 개발자들도 분석할 엄두를 내지 못한다.<sup>157)</sup> 일반인, 법관은 물론 관련 전문가, 심지어 당해 인공지능 법관 개발자조차 그 구체적인 알고리즘을 알기 어려울 수 있다는 것이다.<sup>158)</sup>

인공지능 알고리즘의 불투명성으로 말미암아, 인공지능 법관과 관련하여 제기되는 헌법적 쟁점들에 관하여 상세히는 뒤에서 살펴본다.

## 3) 귀납적 추론 방식의 사용

전문가시스템 등 지난날의 인공지능 기법은 ‘위에서 아래로’ 연역적인 추론(top-down

---

154) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 692쪽.

155) 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 383-384쪽.

156) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 84쪽.

157) 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 218-219쪽.

158) 이러한 지적으로 성중탁, 제4차 산업혁명과 사회국가 실현 등에 대한 전망과 과제, 헌법학연구, 제27권 제1호, 2021, 351쪽의 각주 21; 이병규, AI의 예측능력과 재범예측알고리즘의 헌법 문제 - State v. Loomis 판결을 중심으로, 공법학연구, 제21권 제2호, 2020, 185쪽.

deductive reasoning)에 기반을 두었다. 그것은 이미 존재하는 지식과 사례들로부터, 세계에 관한 추상적이고도 방대한 가설들을 정식화하는 것에서 시작한다. 그러나 오늘날 인공지능 기법의 원리는 이와 사뭇 다르며, ‘아래로부터 위로의’ 귀납적 추론방식(bottom-up inductive reasoning)에 입각하였다고 평가된다.<sup>159)</sup>

현대 인공지능 모델은 처음부터 완벽히 정의된 함수를 활용하지 않으며, 함수의 구조를 대강 정의한 뒤 데이터를 바탕으로 이를 갱신(update)한다. 함수를 갱신할 때에는 역전파(backpropagation)라는 편미분 기반의 방법을 활용하는데, 쉽게 말해 정답이 틀린 만큼 모델 전체에 피드백을 주는 것이다. 오답이 치명적이거나 오답에 관여한 정도가 클수록, 각 구성부분들이 더욱 큰 폭으로 갱신된다. 이러한 원리로부터 ‘학습(learning)’이라는 표현이 나온다.

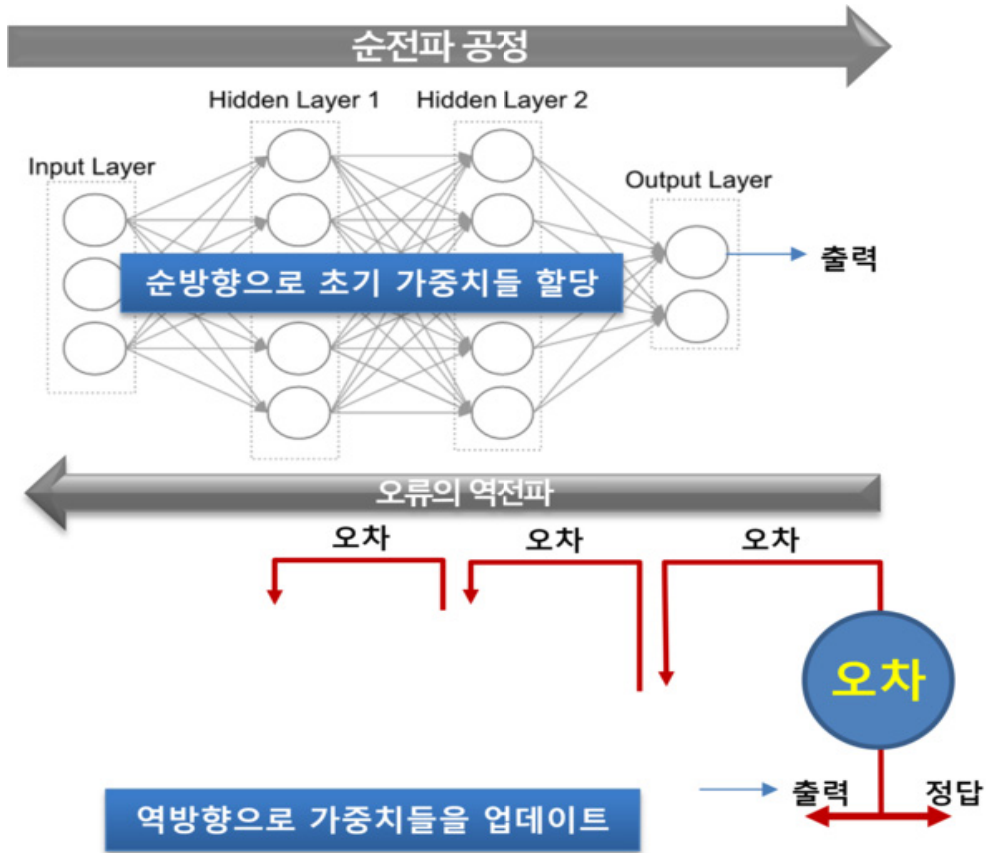
오늘날에는 딥러닝(deep learning)이라고 하는 인공지능이 주류적인데, 딥러닝은 극도로 귀납적인 방법론이다. 모델의 추론 결과는 전적으로 학습에 활용된 데이터에 좌우된다. 이상적인 구성의 데이터가 무한히 존재하고 그에 상응하여 무한히 깊은 층(layer)을 쌓을 수만 있다면 모든 상황에서 정확한 결과 값을 내놓겠지만, 현실은 그렇지 못하여 과소적합(underfitting) 또는 과대적합(overfitting)의 문제가 발생한다.<sup>160)</sup>

---

159) Alaa Khamis, Dipkumar Patel and Khalid Elgazzar, Deep Learning for Unmanned Autonomous Vehicles: A Comprehensive Review in Anis Koubaa and Ahmand Taher Aazr eds., Deep Learning for Unmanned Systems, Springer, 2021, p. 4.

160) 이상의 서술은 윤상원, 현대 인공지능의 원리와 한계, 법률신문, 2023. 12. 9, <https://www.lawtimes.co.kr/opinion/193984?serial=193984>(최종방문일 2024. 2. 11.)을 요약·정리하였다. 인공지능이 샘플 데이터를 지나치게 오랫동안 학습하는 경우 또는 모델이 너무 복잡한 경우 해당 데이터 세트에 들어 있는 무관 정보(노이즈)를 학습하기 시작한다. 모델이 노이즈를 기억하고 해당 학습 세트에 너무 가깝게 들어맞으면 과대적합 상태가 된다. 이와 달리 인공지능의 학습 시간이 불충분하거나 입력 변수의 양이 충분치 않아 입력 변수와 출력 변수 사이의 유의미한 관계를 확인하지 못할 때 과소적합이 일어난다. 자세히는 <https://www.ibm.com/kr-ko/topics/overfitting>(최종방문일 2024. 2. 11.) 참조.

[그림3: 순전파 공정과 오류의 역전파 공정]<sup>161)</sup>



#### 4) 정보의 코드화를 전제한 데이터 기반성

당초 인공지능의 개념은 확률과 통계학을 근간으로 세상의 여러 문제를 해결하려는 시도의 하나로 1950년대 태동하였다. 인공지능 알고리즘은 코드로 구성되는데, 코드를 보면 전형적으로 수식 형태가 그 뿌리가 된다.<sup>162)</sup>

오늘날 각광받는 인공지능의 위력은 그것이 예측이라는 데 있다. 머신러닝 내지 딥러닝 기법은 기존 빅데이터를 학습하여<sup>163)</sup> 예측 결과를 내놓는데, 모델이라고도 하는 예측

161) 이재호, 딥러닝의 해석을 위한 시스템다이내믹스의 활용, 스마트스 경제신문, 2022. 4. 6, <https://www.sbr.ai/news/articleView.html?idxno=1717>(최종방문일 2024. 2. 11.).

162) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 159쪽.

163) 빅데이터는 매일 약 250경 바이트씩 증가할 정도로 막대하며, 이는 인간의 인지능력 범위를 벗어난다.



규칙을 살펴보면 입력과 출력 사이의 패턴을 기술해주는 방정식이 존재한다. 데이터를 활용하여 적절한 예측 규칙을 찾아나가는 과정을 ‘모형 훈련시키기’라고도 부른다.

가령 지도학습 머신러닝의 경우, 훈련 데이터는 여러 특성 데이터와 그 특성 데이터의 결과의 짝으로 구성된다. 훈련이란 결국 특성 요소들을 입력변수로 넣고, 은닉층에서의 연산으로 결과가 나오면 실제 결과(정답)와 비교하여 오차를 줄여나가는 식이다. 따라서 새로운 (양질의) 데이터가 축적될수록 예측이 점차 향상된다.<sup>164)</sup>

데이터로부터 예측 규칙을 추출하는 데 신경망(neural network)을 활용하면서, 인공지능의 혁신적인 발전이 이루어졌다. 신경망이라는 표현을 사용하지만, 데이터의 복잡한 패턴을 기술할 수 있는 복잡한 방정식(입력을 출력에 대응시키는 시스템)일 뿐 인간의 뇌와 동일한 일을 하는 것은 아니다.<sup>165)</sup>

아직 현실화가 요원하다고 평가받는 강한 인공지능을 차치한다면, 현재 주류적인 머신러닝 내지 딥러닝 기법을 활용한 인공지능 판사 역시 정보의 코드화를 전제한 데이터에 기반을 두고 일정한 결과 값을 내놓을 것이다. 그리고 인공지능 판사의 학습에 활용되는 데이터가 양질이며 풍부할수록 더 정확한 판단을 내리리라고 예상할 수 있다. 반면 데이터가 부족하거나 질적으로 대표성이 결여되어 있다면, 이를 기반으로 하는 인공지능은 편향성을 띄게 될 위험이 크다.<sup>166)</sup>

## 5) 인간의 의사 작용 배제

한국을 위시한 주요 국가들과 국제기구들은 인공지능 윤리 가이드라인과 원칙을 제시하여 왔으며, 최근 유럽연합과 미국에서는 인공지능 알고리즘을 규율하는 일반법을 제정하려는 움직임도 활발하다.

인공지능 관련 데이터의 사용과 알고리즘 설계에 있어서는 선제적인 윤리적·제도적 대응이 필요하다고 지적되는바, 이는 인공지능이 결정과 행위를 자동화하는 기술이며 결

---

오로지 컴퓨터만이 빅데이터로부터 무언가를 배우고 의미를 파악할 수 있다. 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 170쪽.

164) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 163-165쪽.

165) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 165-166쪽.

166) 김찬희, 인공지능 기반 자동화 행정행위에 관한 헌법적 고찰, 헌법이론과 실무, 2023-A-3, 헌법재판소 헌법재판연구원, 2023, 24쪽.

정과정에서 인간을 배제하기 때문이다.<sup>167)</sup> 인공지능을 종전 컴퓨터 기술과 구별하는 것대의 하나가 ‘자율성(autonomy)’인데, 인간의 직접적 지시 없이 기계가 획득하고 분석한 정보를 토대로 독립적으로 행동한다는 점이 바로 자율성의 속성이다.<sup>168)</sup> 인공지능에 관한 연구가 시작될 무렵에는 자율성이 인공지능의 주요한 개념표지가 아니었으나, 이후 관련 연구가 진척되면서 인공지능 자율성으로부터 여러 법적인 쟁점들이 도출되고 인공지능의 자율성에 관한 규율의 필요성을 인식하기에 이르렀다.<sup>169)</sup>

전술하였듯 오늘날 인공지능 관련 기술 수준 및 작동원리, 경제논리를 고려할 때 자연인 법관과 전적으로 동일한 인공지능에 관한 논의는 다소 비현실적일 수 있으나, 인공지능 알고리즘이 제시하는 출력 값이 자연인 법관의 판단을 대체하는 단계에 이르러야 사실상 법관으로서 기능하였다고 볼 수 있다. 이처럼 자율성을 가지고 자연인 법관의 판단을 대체하는 ‘인공지능 법관’의 판단과정에서는 인간의 의사 작용이 배제되며, 그와 관련한 여러 헌법적 쟁점들이 불거질 수 있다.

---

167) 김희정, 인공지능 윤리와 형사정책적 시사점 - 형사사법에서의 인공지능 윤리 이슈를 중심으로 -, 형사정책, 통권 제69호, 2022, 83쪽.

168) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 183쪽. 물론 인공지능 알고리즘이 처음부터 완전한 자율성을 가질 수는 없으며, 설계자가 몇 가지 초기 지식과 학습능력을 주입해야 한다.

169) 김찬희, 인공지능 기반 자동화 행정행위에 관한 헌법적 고찰, 헌법이론과 실무, 2023-A-3, 헌법재판소 헌법재판연구원, 2023, 17쪽.

### Ⅲ. 인공지능 법관 관련 헌법적 쟁점들

사법 불신과 함께 인공지능 법관의 도입을 촉구하는 견해도 존재한다. 특정 분야에서는 인공지능 법관이 가능할 수 있다는 전망도 있다.<sup>170)</sup> 그러나 인공지능 시스템의 도입은 사법제도 자체의 이념적·철학적 토대를 무너뜨릴 수 있다는 우려도 제기된다.<sup>171)</sup> 자연인 법관의 판단을 대체하거나 자연인 법관과 동일한 역할을 하는 인공지능 법관이 도입되려면, 여러 헌법적인 문제에 대한 해명이 필요하다.

#### 1. 논의의 전제: 인공지능 법관 도입에 관한 기대

인공지능이 우리 사회에서 널리 활용되며 법조계 역시 그 예외가 아님을 앞서 보았다. 그런데 우리 사회, 심지어 실무자들 사이에서도 ‘인공지능 법관’에 관한 기대가 적지 않다. 왜 그럴까? 이하에서는 인공지능 법관과 관련한 법적 쟁점들을 검토하기에 앞서, 인공지능 법관에 대한 기대 및 논의의 배경을 짚어보기로 한다.

##### 가. 업무 효율성의 극대화

인공지능 법관의 이점이라면, 판단이 신속하고 효율적으로 이루어지리라는 점이다. 법관의 숫자 부족과 사건의 증대로 말미암은 판결·결정의 지연은 사법제도의 장애와 공적 거래비용 증대를 초래하며, 특히 형사재판에서의 선고가 늦어진다면 이는 피해자에 대한

---

170) 인공지능과 데이터 기반 기술의 발달로 인해 알고리즘 모델이 더 적합한 재판 환경이 조성되었다고 한다. 형사재판에서 디지털증거(digital evidence)의 비중이 늘기 시작하였는데, 이러한 증거에는 차원 높은 시용성이 부여되며, 반대신문권이나 대면권 보장만으로 증거의 허용성이나 증명력을 탄핵하기가 어렵다고 한다. 인공지능이 생성한 증거가 주로 사용되는 재판에서는, 전통적인 사실인정 절차에서와 같은 복잡한 알고리즘이 구현될 필요가 없으며, 인공지능 기반 사법체계의 혁신은 이러한 사실인정, 증거영역에서부터 모색되리라고 전망한다. 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 610-611쪽.

171) 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 587쪽.

사법 정의의 실현이 지연됨을 의미한다. 컴퓨터가 정보를 즉시 처리하기 때문에, 알고리즘은 자연인보다 빠르게 사안을 해결할 수 있다. 적시에 판결이 선고되면, 선고가 지연됨으로 말미암아 발생하는 많은 부작용을 개선할 수 있으리라 기대된다.<sup>172)</sup>

인공지능 법관까지는 아니더라도, 인공지능을 재판의 보조도구로 활용할 경우 업무 효율성이 제고되리라 기대된다. 가령 다수의 기존 판결문을 분석하는 역할을 인공지능이 수행한다면, 판사가 유사한 사건 판결을 검색하고 비교하는 데 투입하는 시간과 노력이 절약될 수 있다.<sup>173)</sup>

## 나. 재판의 일관성, 공정성 및 정확성에 대한 기대

### 1) 재판의 공정성 및 정확성

오늘날 딥러닝 방식의 인공지능은 결론에 도달하는 과정을 설명하기 어려워, 이용자에게 일종의 블랙박스라고 평가된다. 인공신경망<sup>174)</sup>은 입력층과 출력층 사이에 존재하는 은닉층(hidden layer)이 생명인데, 은닉층에서의 오류보정과 가중치 변경과정을 비전문가가 이해하기는 힘들다는 것이다.<sup>175)</sup> 그럼에도 현실에서는 인공지능 범죄예측 시스템의 도입을 주저하지 않는 경향이 감지되는데, 이는 인공지능 예측결과의 정확성이 그 결과에 이르는 과정에 대한 신뢰도까지 증가시킨 것이라고 한다.<sup>176)</sup>

재판의 객관성과 중립성은 재판절차의 공정성을 보장하는 중요한 요소일 뿐만 아니라,

---

172) 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 223-224쪽. 우리나라의 경우 법관의 숫자가 부족하며 1인의 법관이 너무 많은 사건을 담당해야 하는 현실이 재판지연으로 이어진다는 비판으로 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학 논문집, 제45집 제1호, 2021, 61-62쪽. 인공지능 법관이 법관 수의 절대적 부족이라는 문제를 해결하는 하나의 방법이 될 수 있다고 한다. 알고리즘을 활용하면 신속하게 분쟁을 해결하여 법관 숫자 부족 문제를 해결할 수 있다는 유사한 견해로 강영주, 인공지능을 활용한 민사분쟁 해결에 관한 고찰, 안암 법학, 제64권, 2022, 409쪽.

173) 한애라, AI 판사?, 법률신문, 2020. 1. 20, <https://m.lawtimes.co.kr/Content/Opinion?serial=158857>(최종 방문일 2024. 2. 11.).

174) 인간 뇌의 신경망은 다수의 뉴런이 연결된 형태로 구성되어 있는데, 신경망의 네트워크와 정보전달체 계로부터 영감을 얻어 이를 수학적으로 모델링한 것이 인공신경망이다. 서울대학교 산학협력단, 사법 부에서의 인공지능(AI) 활용방안, 법원행정처, 2020, 11-12쪽.

175) 전문가, 심지어 설계자라도 그 구체적인 작동과정을 다 파악하기란 쉽지 않다.

176) 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에서의 적용에 관한 고찰, 형사법의 신통향, 통권 제51호, 2016, 233쪽.

법원의 판단이 공중의 인정을 받기 위한 전제조건이기도 하다. 그러나 일선 실무는 개인적인 선호, 편향, 과로, 부패 등의 문제로 인해 자연인 법관이 언제나 중립적이고 객관적이며 공정하다고 기대하기 어려움을 보여준다.<sup>177)</sup> 최근 학계에서는, 방대한 자료를 분석해 결론을 추론하는 슈퍼컴퓨터에 기초한 인공지능이 인간보다 더 공정하고 정확하게 재판을 할 수 있다는 주장이 제기된다. 충돌하는 이익을 조절하고 합리적인 솔루션을 찾는 문제에 있어, 욕망을 가진 인간이 개입하기보다 인공지능에게 맡기는 것이 더 나을 수 있다는 것이다. 보다 근본적으로는 지금까지 존재해온 근대국가의 모델이 수명을 다해가고 있으며, 인공지능에 기초한 국가모델은 인간을 더 행복하게 만들고 국가의 본래적 기능 수행에도 훨씬 더 적합한 모델이 되리라고 평가한다.<sup>178)</sup>

미국의 형사재판에서는 인종차별에 관한 논란이 있어왔다. 77,000명 이상 피고인의 형량을 분석한 결과, 동일 법원에서 동일 범죄를 저지른 흑인 피고인들은 백인 피고인들보다 12% 중형을 선고받은 것으로 나타났다고 한다. 알고리즘의 주요 장점 가운데 하나로, 형사판결 선고에 있어 인종차별의 영향을 없앨 수 있다는 점이 꼽힌다.<sup>179)</sup> 인간과 달리, 알고리즘은 본능적이고 무의식적인 편견이 없으며 의도적으로 누군가를 차별할 수 없다. 다만 형량 결정의 편견을 없애려면, 자연인 법관의 암묵적 편견과 그것이 반영된 결과가 변수에 포함되지 않도록 설계하여야 한다.<sup>180)</sup>

## 2) 재판부별 판단의 편차 해소

『인공지능시대 법관의 미래 모습』을 집필한 오세용 부장판사는 인공지능 판사가 도입되더라도 또 다른 불만과 불신이 발생할 가능성이 크다고 본다. 가령 누적된 데이터를

177) Zichun Xu, Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities, Applied Artificial Intelligence, Vol. 36, No. 1, 2022, p. 1030. 이른바 사법행정권 남용 의혹이 불거진 이후 재판 불신이 크게 증가하였으며, 청와대국민청원 게시판에 인공지능 판사를 요청하는 다수의 청원이 접수되었다고 한 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 63-64쪽 참조.

178) 정종섭, 헌법학원론, 박영사, 2022, iii쪽.

179) 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 224-225쪽. 미국에서 재범위험성 예측도구의 적극적으로 도입한 배경의 하나로, 통계에 기반한 인공지능 알고리즘은 내집단 선호를 보이지 않아 인종차별을 하지 않으리라는 기대가 있었다고 한 박혜진, 양형과 인공지능 - 국민의 법감정 수렴, 위험 판단, 양형 분석을 중심으로 -, 법조, 제72권 제4호, 2023, 265쪽도 참조.

180) 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 225-226쪽.

학습해 결론을 도출하는 작동원리상, 현재 사회의 분위기를 제대로 반영하지 못하고 선례에 묶여 있을 가능성이 크다는 것이다. 하지만 그는 판사의 역할을 완전히 대체하는 인공지능은 시기상조라고 하면서도, 양형 업무에서 인공지능이 활용될 수 있다고 본다. 인공지능을 활용하면 재판부별 양형 편차가 감소하여, 사법 시스템에 대한 국민적 신뢰를 높일 수 있으리라는 것이다.<sup>181)</sup>

미국에서는 인공지능 알고리즘에 입각한 재범위험성 예측도구를 양형 판단에 적극 도입하고 있는데, 그 배경으로는 양형편차가 논의된다. 어떤 판사를 만나느냐, 같은 판사라도 어떤 상황이나에 따라 다른 판단을 받을 수 있다는 문제의식이 있었던 것이다. 인공지능 알고리즘은 인간 판사와 달리 동일한 입력에 대하여 같은 결과를 출력하므로, 이를 활용하여 양형을 한다면 유사한 사건에서 일관성을 기할 수 있다고 본 것이다.<sup>182)</sup>

유사한 사안에서 재판 결과의 편차가 줄어들고 일관성이 확보된다 함은, 곧 판결·결정의 예측가능성이 높아진다는 의미이기도 하다. 이렇게 되면, 최종적인 판단을 받아보기 이전에 조정이나 화해가 성립할 가능성이 높아질 수도 있다.<sup>183)</sup>

### 3) 과거 판단과의 원리적 일관성

과거 결정과의 원리적 일관성 측면에서 인공지능이 자연인 판사보다 장점을 가진다는 연구도 있다. 그에 따르면 인공지능은 기본적으로 축적된 판례를 학습하고 이를 관통하는 일관된 논리, 즉 알고리즘을 형성하기 때문에 자연스레 원리적인 일관성을 획득할 수 있다. 자연인 법관의 판단은 법관 개인의 가치관이나 성향, 심리상태, 심지어 당사자의 외모에 의하여도 영향을 받는다, 인공지능은 판단의 일관성에 있어 자연인 법관보다 확실한 장점을 갖는다고 한다.<sup>184)</sup> 비합리적인 인간과 달리 컴퓨터는 일관된 판단에서 벗어

---

181) 이배운, “판사를 AI로 교체하자”현직 판사에게 물어보니, 이데일리, 2023. 7. 24, <https://www.edaily.co.kr/news/read?newsId=02551846635677800&mediaCodeNo=257&OutLnkChk=Y>(최종방문일 2024. 2. 11.). 참고로 미국의 현행 사법 시스템에서 주어진 사건의 사실에 관한 판단에서 불평등이 발생하며, 이는 개별 법관의 암묵적 편견과 뿌리 깊은 신념의 영향 때문이라는 지적이 있다. 상세히는 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 222-223쪽.

182) 박혜진, 양형과 인공지능 - 국민의 법감정 수렴, 위험 판단, 양형 분석을 중심으로 -, 법조, 제72권 제4호, 2023, 264-265쪽.

183) 한애라, AI 판사?, 법률신문, 2020. 1. 20, <https://m.lawtimes.co.kr/Content/Opinion?serial=158857>(최종방문일 2024. 2. 11.)도 참조.

184) 서치원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 ‘법의 제국’을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 13쪽.

나지 못하며, 폭넓은 형평성을 달성할 수 있다는 지적도 있다.<sup>185)</sup>

## 다. 국가예산절감

인공지능 알고리즘을 재판에 활용할 경우 법관들이 적절한 판결을 선고하는 데 드는 시간이 줄어들고, 법원을 운영하며 법관을 보좌하는 데 필요한 인적·물적 비용을 감소시킬 수 있다고 논의된다.<sup>186)</sup>

## 라. 소결

인공지능 법관은 재판 업무의 효율성을 증대하고 재판의 일관성, 공정성 및 정확성을 향상시키리라는 기대 속에서 그 도입과 관련한 논의가 이루어지고 있다. 2020년 11월경부터 2021년 1월경까지 인공지능 기술 전문가 24명, 현직 법관 26명, 미래전략 전문가 35명을 설문조사한 결과(복수응답 가능)도 이와 유사한 인식을 보여준다. 인공지능 법관 도입 시 얻을 수 있는 장점으로 재판결과의 일관성과 예측가능성 확보(51표), 공정하고 엄정한 재판(58표), 법관 개인의 성향이나 가치관에 따른 편향성 제거(55표), 신속한 재판(62표), 자연인 법관의 감소로 인한 국가예산 절감(28표)이 꼽혔다.<sup>187)</sup>

하지만 인공지능 법관에 관한 우려의 목소리도 나온다. 위의 설문조사에서는 인공지능 법관을 도입할 경우의 문제점으로 개개인의 구체적인 사정을 살피지 못할 것(46표), 알고리즘의 불투명성·편향성(58표), 시대변화에 따른 판례변경의 어려움(47표), 선례 없는 사건 처리의 어려움(50표), 오작동·해킹 등으로 부당한 결론에 대한 공포(50표)가 지적되었다.<sup>188)</sup> “인공지능이 보석 및 양형 결정 시 편향성을 제거해 준다고 보는가?”라는 질문에 관한 법관 381명의 답변을 분석한 2023년 미국의 연구에서도, 법관들이 알고리즘에 대한

185) 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 223쪽. 인공지능은 중립적·객관적이므로 판사 개인의 경험·편견·지식의 한계에 따른 편차와 자의를 줄일 수 있다고 한 한애라, AI 판사?, 법률신문, 2020. 1. 20, <https://m.lawtimes.co.kr/Content/Opinion?serial=158857>(최종방문일 2024. 2. 11.)도 참조.

186) 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 224쪽; 강영주, 인공지능을 활용한 민사분쟁 해결에 관한 고찰, 안암법학, 제64권, 2022, 409쪽.

187) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 258-260쪽; 오세용, 인공지능(AI)이 앞당긴 미래 사법의 변화 그리고 법관, 2023 사법연수원 국제콘퍼런스 자료집(지정토론편), 사법연수원, 2023, 117쪽.

188) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 258-260쪽.

혐오(algorithm aversion)를 경험하며, 인공지능으로 말미암아 편향성이 심화되는 것을 우려하고, 법관이 컴퓨터에 의하여 대체되는 것을 염려하는 것으로 나타났다. 다만 법관들은 인공지능이 보석, 양형에 관한 판단에서 정보 제공의 측면에서 잠재적 도움이 될 수 있다고 믿고 있으며, 인공지능이 실증 테스트를 거치고 가이드라인을 준수해야 한다고 하였다.<sup>189)</sup>

## 2. 민주주의

### 가. 민주주의 일반

헌법은 제1조 제1항에서 “대한민국은 민주공화국이다.”라고 명시하여, 우리나라가 민주국가이고 공화국가임을 분명히 한다. 민주주의에 관하여 다양한 이론과 견해가 주장되며, 민주주의만큼 보편적이고 다양한 의미로 사용되는 용어도 찾아보기 어렵다.<sup>190)</sup> 하지만 민주국가란 국가와 지배의 정당성이 국민에 있음을 의미한다.<sup>191)</sup>

어원상으로도 민주주의는 국민에 의한 지배를 뜻한다. 민주주의를 뜻하는 용어인 democracy는 프랑스어 démocratie가 16세기에 영어로 유입된 것이다. 그 출처는 고대 그리스의 demokratia인데, 그 어근의 의미는 인민(demos)과 지배(kratos)이다.<sup>192)</sup>

이때 국가와 지배의 정당성을 담보하는 것은 우리 사회의 구성원인 ‘국민’이다. 현행헌법은 국가권력의 기원을 밝히는 제1조 제2항에서 민주주의 원리를 구체화하고 있는데, 국민주권주의가 바로 그것이다. 민주주의에서 정치적 지배는 선형적으로 또는 전통적으로 정당화되는 것이 아니라, 정치적 지배를 받는 자의 동의에 의해 정당화된다. 이로써 국민주권주의는 국민에 의한 지배를 정당화한다.<sup>193)</sup>

---

189) Anna Fine, Stephanie Le and Monica K. Miller, Content Analysis of Judges’ Sentiments Toward Artificial Intelligence Risk, *Criminology, Criminal Justice, Law & Society*, Vol. 24, Iss. 2, 2023, p. 31. 이 연구는 Benes Z. Aldana, 사법정의의 대변형: 법률시스템에서의 AI 부상과 윤리적 고찰, 2023 사법연수원 국제콘퍼런스 자료집, 사법연수원, 2023, 65-67쪽에도 소개되고 있다.

190) 한수용, 헌법학, 법문사, 2022, 113쪽.

191) 정종섭, 헌법학원론, 박영사, 2022, 233쪽.

192) 데이비드 헬드, 박찬표 옮김, 민주주의의 모델들, 후마니타스, 2014, 17쪽.

193) 한수용, 헌법학, 법문사, 2022, 113쪽. 그런데 직접민주주의에는 여러 현실적인 어려움이 따르며, 주권자인 국민은 선거로 대표자를 선출하여 대의기관을 구성하고 신임을 부여한다. 또 주기적으로 정치적 심판을 통해 권력을 통제한다. 김종현, 범죄자의 선거권 제한의 헌법적 검토 - 선거권 제한 입법의 위



## 나. 국가기관과 민주적 정당성

대한민국의 주권은 국민에게 있고 모든 권력은 국민으로부터 나오므로(헌법 제1조 제2항), 모든 국가기관은 국민으로부터 직·간접적으로 민주적 정당성을 부여받아 구성되어야 하고, 이러한 민주적 정당성은 국가기관의 권한 행사의 원천이 된다. 국민이 직접 선출하는 국회의원으로 구성된 국회와 대통령뿐만 아니라, 국회와 대통령의 관여로 선출되는 비선출 권력인 사법부에도 간접적으로 민주적 정당성이 부여된다.<sup>194)</sup> 대법원장은 ‘국회의 동의’와 ‘대통령의 임명’(헌법 제104조 제1항), 대법관은 ‘대법원장의 제청’, ‘국회의 동의’와 ‘대통령의 임명’(헌법 제104조 제2항)을 통해 간접적으로 민주적 정당성을 부여받는다. 대법원장과 대법관이 아닌 법관의 경우 ‘대법관회의의 동의’를 얻어 ‘대법원장이 임명’함으로써(헌법 제104조 제3항) 간접적으로 민주적 정당성을 부여받는다.<sup>195)</sup>

## 다. 민주주의와 관련하여 인공지능 법관이 초래하는 문제

민주적 정당성은 일반적으로 인적 정당성과 사물적·내용적 정당성을 통해 매개된다. 국가권력을 행사하는 공직자가 선거에 의하여 직접 선출되거나 국민에 의하여 선출된 국가기관에 의해 임명됨으로써 인적 민주적 정당성의 요청을 충족한다. 민주주의는 그 또한 우리사회의 구성원이기도 한 국민(자연인)에 의한 지배를 바탕으로 한다는 점에서, 인공지능 법관 등 국가권력행사에 인공지능을 도입하는 것은 인적 민주적 정당성의 차원에서 당연히 문제를 초래한다.<sup>196)</sup>

우리 헌법질서는 ‘인공지능’이라는 새로운 존재를 국가권력의 원천이자 공동체의 구성원인 국민으로 상정한 바 없다. 또한 향후 헌법과 법률이 개정되는 경우는 별론으로 하고, 사법절차에 활용될 수 있는 인공지능 알고리즘에 직접적·간접적으로 민주적 정당성

헌심사 및 법관의 개별적 선고 필요성에 관한 논의를 소재로 -, 공법연구, 제51집 제1호, 2022, 28쪽.

194) 다만 사법부와 국회·대통령은 그 민주적 정당성의 크기가 다르다고 할 수 있다. 비례입헌주의(proportional constitutionalism)는 시민을 대표하는 공직자가 가치의 조정문제에 관한 한 “민주적 정통성의 크기만큼 결정하라.”고 주문한다. 객관적 정답이 없는 도덕적 조정문제의 결정은 민주적 정통성이 큰 정치부문에서 이루어져야 하며, 법원은 정치과정에서 결정되지 않은 나머지를 결정해야 한다고 한다. 조홍식, 물경시정치 - 비례입헌주의를 주장하며 -, 서울대학교법학, 제49권 제3호, 2008, 115쪽.

195) 현재 2021. 10. 28. 2021헌나1, 판례집 33-2, 321, 347-348.

196) 김중권, 인공지능시대 알고크라시(Algocracy)에서의 민주적 정당화의 문제, 법조, 제69권 제5호, 2020, 186-187쪽.

을 부여한 바도 없다. 자연인 법관은 그 또한 우리 사회의 한 구성원이며 간접적으로나마 민주적 정당성을 확보하고 있음을 생각할 때, 인공지능 알고리즘이 자연인 법관을 전적으로 대체하게 된다면 민주주의 관점에서 허용될 수 있는지가 문제될 것이다.

### 3. 법치주의

법치주의는 우리 헌법의 기본원리다. 법치주의란 국가권력을 법에 기속시킴으로써, 국가권력의 자의와 남용을 막고 국민의 자유와 권리를 보장하려는 원리다.<sup>197)</sup>

혹자는 인공지능이 기술적 편익에 기초하여 법치를 강화할 수 있다고 본다. 인공지능의 판단은 법관의 그것보다 투명하며, 오늘날 재판에 존재하는 차별 요인<sup>198)</sup>을 제거함으로써 동종 사건에서의 일관성 및 형평성을 달성할 수 있다는 것이다. 또한 재판의 지연을 방지하고 신속히 판결을 선고함으로써 법치주의를 촉진할 수 있으며, 미국 등에서 문제되는 재판에서의 인종차별도 감소하리라고 전망한다.<sup>199)</sup>

그러나 인공지능 법관이 도입된다면, 법치주의와 관련하여 적잖은 우려가 현실화될 가능성이 있다. 법이란 구성원 대다수가 준수하고자 할 때에만 효과를 가지는데, 인공지능의 법적 의사결정이 불공정하거나 예측불가능하다고 인식될 경우 법치주의가 훼손되리라는 전망이 있다.<sup>200)</sup> 심지어 개발자조차 인공지능의 구체적인 알고리즘을 알기 어려운 경우도 발생할 수 있다.<sup>201)</sup>

---

197) 정연주, 헌법학원론, 법영사, 2023, 65쪽.

198) 인공지능은 인간과 달리 본능적이고 무의식적인 편견이 없으며, 의도적으로 차별할 수 없다고 한다. 다만 암묵적인 편견에 의하여 차별을 초래하는 변수가 포함되지 않도록 알고리즘을 설계하여야 한다고 한다. 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 225-226쪽.

199) 정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 221-229쪽.

200) Lisa Webley, 권현영 옮김, 인공지능(AI)과 법조 분야: 윤리적·규제적 고려사항, 경제규제와 법, 제12권 제2호, 2019, 77-78쪽. 인공지능의 구체적인 산출과정을 아무도 알지 못하거나 소수의 전문가만 알 수 있다면, '명확하게 규정되고 대외적으로 공포된 법에 의한 지배'의 의미가 상당히 퇴색되리라는 지적으로 성중탁, 제4차 산업혁명과 사회국가 실현 등에 대한 전망과 과제, 헌법학연구, 제27권 제1호, 2021, 371쪽. 그 경우 법에 의한 지배가 아닌 '코드에 의한 지배'로 대체되어 버릴지도 모른다고 한다. 그런데 인공지능이 판사의 역할을 한다는 사실 자체가 문제될 것은 없다는 주장도 있다. 자연인 법관이 판결을 내린다고 하여 우리가 법의 지배가 아닌 법관의 지배를 받는 것은 아니라는 취지다. 서치원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 '법의 제국'을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 7쪽.

201) 성중탁, 제4차 산업혁명과 사회국가 실현 등에 대한 전망과 과제, 헌법학연구, 제27권 제1호, 2021,

그렇다면 법치주의의 관점에서, 인공지능 법관의 알고리즘을 인간이 설계하거나 적어도 인간이 감수해야 한다고 요구할 수 있다. 그런데 이처럼 인간의 개입을 요구하는 접근방식은 ‘단지 누구라도 인간이 개입되어 있다면’ 법의 지배 요건이 충족되는 것처럼 보이게 함으로써, 법치 관념이 다시 인치 관념으로 회귀하는 반작용을 초래할 여지가 있다.<sup>202)</sup>

## 4. 재판청구권

헌법이 보장하는 자유와 권리는 빈틈없는 사법적 권리구제수단과 절차까지 보장할 때 비로소 완성된다. 재판청구권은 국가에 대하여 재판을 청구할 권리를 말한다.<sup>203)</sup> 그런데 인공지능 법관이 현실화될 경우, 재판청구권의 여러 측면에서 위헌성 문제가 불거질 수 있다.

### 가. 법관에 의한 재판을 받을 권리

헌법 제27조 제1항은 헌법과 법률이 정한 법관에 의하여 재판을 받을 권리를 보장한다. 법관에 의한 재판을 받을 권리를 보장한다고 함은 법관이 사실을 확정하고 법률을 해석·적용하는 재판을 받을 권리를 보장한다는 뜻이고, 만일 그러한 보장이 제대로 이루어지지 아니한다면, 헌법상 보장된 재판을 받을 권리의 본질적 내용을 침해하는 것으로서 우리 헌법상 허용되지 아니한다.<sup>204)</sup>

법률로 인공지능에게 법관의 자격을 부여한다면, 인공지능이 법관으로서 재판을 할 수 있는지가 문제된다. 그런데 재판청구권에 관한 헌법 제27조에서는 ‘헌법과’ 법률이 정한 법관이라고 규정하므로, 법률로 인공지능에게 법관 자격을 인정하더라도 인공지능 법관을

---

351쪽의 각주 21. 그런데 이와 달리, 개발자 혹은 누군가가 인공지능 법관의 알고리즘 메커니즘을 완전하게 파악하고 있다면 또 다른 문제가 발생할 수도 있다. 누군가 인공지능 법관이 어떤 판결을 할 것인지를 대단히 구체적으로 예측할 수 있다면(유·무죄, 양형, 손해배상 책임의 인정여부나 손해액 산정 등을 거의 정확히 예견), 그(개발자 또는 인공지능에 아주 정통한 자)가 법률시장을 사실상 좌우할 가능성이 있다.

202) 남중권, ‘알고리즘의 지배’와 ‘법의 지배’에 대한 법이론적 이해, 법과사회, 제71호, 2022, 312쪽.

203) 김학성·최희수, 헌법학원론, 피앤씨미디어, 2023, 775쪽.

204) 헌재 2020. 12. 23. 2019헌바353, 공보 291, 134, 137.

‘헌법이 정한 법관’이라고 할 수 있는지가 여전히 문제된다.<sup>205)</sup> 인공지능에 의한 재판을 인정하기 위해서는 개헌과 법률개정이 필요하다는 지적이 제기되는 것은 이 때문이다.<sup>206)</sup>

인공지능 판사를 헌법상 ‘법관’으로 볼 수 없다는 입장에서는, 인공지능에 의한 재판이 법관에 의한 재판을 받을 권리를 침해하며 법치주의원리에 따른 재판을 무의미하게 만든다고 다룰 것이다.<sup>207)</sup> 전술하였듯 강한 인공지능이 출현하고 여기에 헌법상 법관의 지위를 인정하려면 개헌이 이루어져야 한다고 판단되나, 가까운 미래에 현실화될 가능성은 낮다고 생각된다. 그런데 오늘날 고도로 발달한 알고리즘은 (강한 인공지능의 단계는 아니지만) 인간의 결정을 지원하는 것을 넘어 의사결정을 대체하는 수준에 이를 수도 있다고 평가되는바, 알고리즘에 의하여 사실관계의 확정, 법률 적용 및 기타 판단(가령 작량 감경)이 이루어진다면, 그것이 법관에 ‘의한’ 재판인가?<sup>208)</sup>

최근 학계에서는 법관을 비롯한 의사결정주체들이 인공지능 알고리즘이 도출한 결과에 지나치게 영향을 받을 위험성(자동화 편향: automation bias)을 경계하는 논의가 있다.<sup>209)</sup> 보다 근본적으로 알고리즘이 일정한 정보를 처리하여 결과를 산출하는 것을 ‘재판’이라고 할 수 있는지도 문제다. 적어도 현 단계의 인공지능 알고리즘은 사람이 하는 방식으로 자연어를 이해하지 못하며,<sup>210)</sup> 누군가가 인공지능 외부의 법령, 판례, 학설 등을 코드화해 입력하고 이를 현실세계에 맞추어 업데이트하여야 하는 까닭이다.<sup>211)</sup>

---

205) 김배원, 지능정보사회와 헌법, 인공지능(AI)의 발전과 헌법적 접근, 공법학연구, 제21권 제3호, 2020, 94쪽.

206) 정문식, 인공지능과 헌법, 한양대학교 법학연구소 편, 인공지능과 법, 도서출판 오래, 2023, 66쪽.

207) 이러한 문제를 경감하는 방안의 하나로써, 가령 인공지능 법관에 의하여 1심 재판을 하되 이에 불복하는 경우 자연인 법관(들)이 담당하는 항소심에서 다룰 수 있도록 하는 제도도 구상해 볼 수 있다.

208) 다수의 논자들은 인공지능의 편견, 부정확성에 관한 문제를 제기한다. 그런데 인공지능이 판단한 내용·결과의 타당성과는 별개의 문제로서, 우리는 같은 사람에 의한 판단이라는 정통성을 요구하는 듯하다는 분석이 있다. 이병규, AI의 예측능력과 재범예측알고리즘의 헌법 문제 - State v. Loomis 판결을 중심으로, 공법학연구, 제21권 제2호, 2020, 186쪽.

209) 상세히는 김웅재, 형사절차에서 인공지능 알고리즘의 활용가능성과 그 한계, LAW & TECHNOLOGY, 제16권 제4호, 2020, 10쪽; 이상용, 알고리즘 규제의 원리와 구조, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 196쪽.

210) 김병록, 인공지능의 헌법적 쟁점과 과제, 법학논총, 제27집 제2호, 2020, 123쪽의 각주 144. 적용법조, 사건기록, 당사자들의 진술 등을 이해하지도 못한 채 기계학습 등으로 구현된 알고리즘을 통해 결과를 산출한 것을 두고 ‘재판’이라고 할 수 있을지가 문제된다.

211) 인공지능이 구현된 소프트웨어 에이전트가 자율적인 행위자로서 법관의 역할을 할 수 있는지가 문제된다. 인공지능은 학습능력을 가지고 있어 부분적으로 인간프로그래머에 의존하지 않고, 사전에 확정된 프로그램으로부터 벗어날 수 있다. 그리하여 예측 불가능한 결정을 발하는 능력을 가진다. 그렇다면 인공지능은 기계시스템과 달리 순수한 도구는 아닌 셈이다. 김중권, 인공지능시대 알고크라시(Algocracy)에서의 민주적 정당화의 문제, 법조, 제69권 제5호, 2020, 188-189쪽.

## 나. 공개재판을 받을 권리

공개재판이라 함은 재판의 공정성을 확보하기 위하여 재판의 심리와 판결을 일반 국민에게 공개하는 것을 말한다. 헌법 제27조 제3항과 제109조가 규정하는 재판의 공개는 제도의 보장이면서 동시에 공개재판을 받을 권리를 기본권으로 보장하는 것이기도 하다.<sup>212)</sup>

그런데 판결기계(Judgement Machine)에 의한 재판을 위해서는 형사재판의 대원칙인 공개재판과 변론주의가 수정될 수밖에 없다는 지적이 있다.<sup>213)</sup> 형사피고인은 상당한 이유가 없는 한 지체 없이 공개재판을 받을 권리를 가지는데(헌법 제27조 제3항), 그러한 재판이 공개재판을 받을 권리를 침해하지 않는지 문제될 수 있다.

## 다. 법률에 의한 재판을 받을 권리

헌법 제27조 제1항은 모든 국민이 ‘법률에 의한’ 재판을 받을 권리를 보장한다. 그런데 인공지능 법관이 재판을 하려면, 사안에 직접 적용되거나 관련이 있는 법령(법률 포함)을 코드화해 입력해 주어야 한다. 문제는 법률과 코드가 공통의 메타언어(상위언어)를 가지고 있지 않으며, 코드화하는 과정에서 그 지향점이 변화를 겪을 수 있다는 점이다. 특히 재량 여지나 판단 여지를 부여하는 규정처럼, 법률이 곧바로 코드로 번역될 수 없는 경우는 더욱 그러하다.

또한 현실에서 법률을 코드로 번역하는 작업은 의회가 아니라 (민간전문가를 활용하는) 집행부에 의하여 수행된다.<sup>214)</sup> 인공지능 법관이 ‘코드화된 법률’을 사안에 적용하여

---

하지만 인공지능 법관이 가능하려면 결국 (인공지능 외부의) 법령, 판례, 학설 등을 코드화해 입력하고 이를 현실세계에 맞추어 업데이트해 주어야 한다. 오류의 발생가능성, 과학기술의 발전에 따라 인간(자연인)이 알고리즘을 수정·업데이트하는 경우도 있을 터이다. 스스로 외부 세계를 인식하고 비판적으로 소화하는 자연인과 동일하게 보기 어려운 측면이다.

212) 정종섭, 헌법학원론, 박영사, 2022, 911쪽.

213) 판결기계는 재판의 진행, 각종 심문·신문까지 하는 것은 현 단계에서는 도저히 달성할 수 없는 비현실적인 구상이며, (형사재판에서) 인공지능 판사는 사실관계 인정-법률 적용-양형결정-판결문 작성의 기능을 수행할 것이라고 한다. 양종모, 인공지능에 의한 판사의 대체 가능성 고찰, 홍익법학, 제19권 제1호, 2018, 10-11쪽.

214) 상세히는 김중권, 인공지능시대 알고크라시(Algocracy)에서의 민주적 정당화의 문제, 법조, 제69권 제5호, 2020, 192쪽.

결과를 산출하는 것을 두고 ‘법률에 의한’ 재판이라고 할 수 있을지가 문제된다.<sup>215)</sup>

## 라. 공정한 재판을 받을 권리

헌법에 ‘공정한 재판’에 관한 명문의 규정이 없지만, 재판청구권이 국민에게 효율적인 권리보호를 제공하기 위해서는 법원에 의한 재판이 공정하여야 할 것은 당연하므로, ‘공정한 재판을 받을 권리’는 헌법 제27조의 재판청구권에 의하여 함께 보장된다고 보아야 한다.<sup>216)</sup>

공정한 재판을 위해서는 당사자주의와 구두변론주의가 보장되어 공소사실에 대한 답변과 입증 및 반증 등 공격·방어권이 충분히 보장되어야 한다.<sup>217)</sup> 특히 형사재판에서 피고인의 방어권이 보장되려면 법관의 판단과 그 근거를 이해하고, 반박할 수 있어야 한다. 그런데 개발사(민간기업)의 영업비밀이라는 이유로 판단근거가 공개되지 않는다면, 피고인의 방어권 행사가 가능한지 의문이 제기된다.

### 1) 사례: State v. Loomis 판결

미국의 State v. Loomis 사건에서, 법원은 COMPAS 프로그램의 재범위험성 평가결과를 양형 판단에 참작하였다. COMPAS는 루미스의 재범위험성이 높다고 판단하였고, 그러한 평가를 참고하여 루미스에게 징역 6년이 선고되었다. 루미스는 중형 선고가 부당하다고 항소하면서, 인공지능 알고리즘에 대한 확인 혹은 이의제기 기회가 주어지지 않았고, 이에 따라 적법절차에 따른 권리가 침해되었으며, 해당 알고리즘의 판단과정 및 작동방식이 공개되지 않는다면 양형결정에 따를 수 없다고 주장하였으나 받아들여지지 않았다.<sup>218)</sup>

루미스는 (1) 정확한 정보에 따라 양형을 선고받을 적법절차상의 권리 침해, (2) 개인별 양형을 받을 권리 침해, (3) 성별에 따라 선고받지 않을 권리 침해를 주장하였다. 그러나 법원은 (1) 위험성 점수를 확인하고 다룰 수 있었다는 점, 질문과 답변을 검증할 기

---

215) 법을 알고리즘화한 전문가집단이 실질적으로 입법, 행정, 사법의 기능을 수행함으로써, 법적으로 해당 권한을 보유한 기관의 형식적 지위와 실제 사이에 괴리가 발생할 수 있다는 지적으로 남중권, ‘알고리즘의 지배’와 ‘법의 지배’에 대한 법이론적 이해, 법과사회, 제71호, 2022, 310-311쪽.

216) 헌재 2013. 9. 26. 2012헌바23, 판례집 25-2상, 708, 712.

217) 전광석, 한국헌법론, 집현재, 2023, 531쪽.

218) Loomis v. Wisconsin, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S.Ct. 2290 (2017).

회가 주어졌으며, COMPAS가 정확하다는 여러 주(state)의 연구가 있다는 이유로 정확한 정보에 따른 양형을 받을 권리 침해를 인정하지 않았다. (2) COMPAS가 집단적 측면에서 위험성을 예측하는 도구이기는 하나, 양형의 결정적 요소가 아니므로 개인별 양형 선고를 받을 적법절차 권리 침해도 인정하지 않았다.<sup>219)</sup> (3) COMPAS가 성별을 고려한 것은 위험성 평가의 정확도를 높이고, 그 결과 피고인을 비롯한 사법체계 전반에 이익이 되므로 적법절차에 따른 재판을 받을 권리를 침해한다고 볼 수 없다고 판단했다.<sup>220)</sup>

다만 위스콘신 주 대법원은 COMPAS의 재범위험성 예측결과를 법관에게 제시할 때, 그 한계를 인식시키기 위한 경고문구<sup>221)</sup>가 첨부되어야 한다고 판시했다. 미국의 탐사보도 매체인 프로퍼블리카(ProPublica)는 COMPAS가 재범위험성이 높다고 판단한 피고인 7,000명을 분석한 결과 그 가운데 20%만이 재범을 저질렀음을 지적하였다. 또한 흑인 피고인이 백인 피고인보다 약 2배가량 재범위험성이 높다고 판단되었음을 지적하였다. 이후 COMPAS 알고리즘의 속성 및 정확성에 관한 논쟁은 더욱 격화되었다.<sup>222)</sup>

## 2) 선행연구

신체의 자유라는 기본권의 중대한 제한 여부가 결정되는 형사재판에서 활용되에도, 인공지능 알고리즘의 작동원리를 공개하도록 강제하지 못한다는 점은 법치국가원칙을 고려

219) 일반적으로 미국 법원들은 본래 예방 또는 법 집행 목적으로 개발된 인공지능 시스템이라 하더라도, 판결이 자동화된 위험성 평가에만 배타적으로 의존하지 않는 한 인공지능 시스템을 선고에까지 폭넓게 사용하는 데 동의하는 것으로 보인다는 지적으로 Giuseppe Contissa and Giulia Lasagni, When It Is (Also) Algorithms and AI That Decide on Criminal Matters: In Search of an Effective Remedy, *European Journal of Crime, Criminal Law and Criminal Justice*, Vol. 28, No. 3, 2020, p. 288.

220) 이 사건의 주요 사실관계 및 판시사항은 Han-Wei Liu, Ching-Fu Lin and Yu-Jie Chen, *Beyond State v Loomis: Artificial Intelligence, Government Algorithmization and Accountability*, *International Journal of Law and Information Technology*, Vol. 27, No. 2, 2019, pp. 122-141; 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 917-919쪽에 정리되어 있다.

221) (1) COMPAS는 상용 소프트웨어로 재범위험성 점수의 산정 방식이 공개되지 않는다는 점, (2) COMPAS의 점수는 집단 데이터에 의존하므로 고위험 개인들을 식별할 수 없다는 점, (3) 미국 전역의 데이터 샘플에 기초하며, 위스콘신 주민들에 대한 교차검증이 없었다는 점, (4) 일부 연구자들은 COMPAS가 소수자 출신 범죄자들의 재범위험성을 높게 예측한다고 문제를 제기한 점, (5) COMPAS는 본래 교정국의 선고 후 처분에 조력하기 위해 개발된 것이라는 점(양형목적으로 개발되지 않았다는 점)

222) Jana Soukupova, *AI-Based Legal Technology: A Critical Assessment of the Current Use of Artificial Intelligence in Legal Practice*, *Masaryk University Journal of Law and Technology*, Vol. 15, No. 2, 2021, pp. 293-294.

할 때 진지한 법이론적 성찰을 요청한다고 평가된다.<sup>223)</sup>

인공지능 알고리즘에 의하여 재범위험성이 예측되고 그에 따라 선고형이 결정될 경우 재판의 공정성 및 공정성에 대한 신뢰가 깨진다는 비판이 있다. 피고인은 법관이 당해 재판의 진정한 결정권자가 아니라고 생각할 것이며, 판단의 기초자료나 가정(assumption), 편향성을 탄핵할 기회도 가지지 못할 것이라는 이유에서다.<sup>224)</sup>

김용재(2020)에 따르면 양형판단에 제출되는 증거는 엄격한 증명의 대상이 아니므로 COMPAS의 재범위험성 예측결과 등을 제출할 수 있으나, 여기에는 공정한 재판을 받을 권리에 따른 한계가 있다. 공정한 재판이라고 할 수 있으려면 모든 증거자료가 공개·조사되고 실질적 방어권이 보장되어야 한다. 따라서 기계에 의한 재판의 위험(자동화 편향)을 차단하고, 예측 알고리즘의 작동원리를 공개하고 설명하여 실질적 반박기회를 피고인 측에 부여해야 한다. 그리고 입법론적으로는 형사소송법이나 형사소송규칙에 명문의 규정을 마련하는 것이 분쟁의 소지를 줄이고 명확성을 담보할 수 있어 바람직하다고 한다.<sup>225)</sup>

혹자는 인공지능의 재범위험성 예측의 정확도에 따라, 공정한 재판을 받을 권리의 침해 여부를 달리 판단할 여지가 있다고 본다. 장진환(2022)은 인공지능 머신러닝(딥러닝)의 복잡성과 불투명성을 고려할 때, 설명 가능한 인공지능 모델이 개발되지 않은 현 단계에서 인공지능 재범위험성 평가도구를 공판과정에서 활용할 수는 없다고 한다. 그러나 공정한 재판을 받을 권리는 인공지능에 의한 재범위험성 평가를 통해 얻어질 공익과 비교형량의 대상이 되는바, 향후 인공지능 재범예측의 정확성이 지금보다 훨씬 높아질 경우 이를 도입하더라도 공정한 재판을 받을 권리를 본질적으로 침해한다고 평가하지 않을 여지가 있다고 한다.<sup>226)</sup>

### 3) 국제사회의 동향

2018년 사법시스템과 사법 환경에서의 인공지능 이용에 관한 유럽 윤리헌장은 형사

---

223) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 128-129쪽.

224) 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 395쪽.

225) 김용재, 형사절차에서 인공지능 알고리즘의 활용가능성과 그 한계, LAW & TECHNOLOGY, 제16권 제4호, 2020, 20-21쪽.

226) 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 -보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 64-67, 71쪽.



사건에서 공정한 재판을 보장하기 위해 민감정보에 따른 차별을 방지해야 하므로, 인공지능의 사용을 최대한 유보하여야 한다고 하였다.<sup>227)</sup> 유럽의회가 형사사법 인공지능 결의안을 발표하기 전인 2021년 7월 LIBE 위원회<sup>228)</sup>가 자체 보고서를 발간하였는데, 공정한 재판을 받을 권리가 유럽연합 기본권 헌장 및 유럽인권협약상의 권리임을 확인하였다. 그리고 형사사법에서 인공지능 기술을 사용할 때, 형사사법절차의 모든 단계에서 방어권의 실질을 박탈하는 조치로 이어져서는 안 됨을 특히 유념해야 한다고 지적하였다.<sup>229)</sup>

#### 4) 보론: 투명성을 구비한 인공지능 알고리즘

State v. Loomis 판결에서 위스콘신 주 대법원은 개인별 양형 선고를 받을 적법절차에 따른 권리 침해를 인정하지 않았는데, 이는 COMPAS가 양형의 결정적 요소가 아니라는 이유에서였다. 이를 반대해석하면, 인공지능 재범위험성 예측이 양형을 결정하는 요소로 기능할 경우 적법절차원칙 위반이라는 평가도 가능하다. 관련하여 양형에 결정적 영향을 미치지 않을 때에만 인공지능 재범위험성 평가가 허용된다면, 도대체 그러한 위험평가를 왜 사용하는지 의문이 제기되기도 한다. 선고절차의 주요 목표는 피고인을 수감해야 하는지 여부와 수감기간(형량), 피고인이 지역 공동체에서 적절히 감독받을 수 있을지 여부를 판단하는 것이기 때문이다.<sup>230)</sup>

사법절차에서 활용되는 인공지능 알고리즘이 모두 비공개되어 불투명한 것은 아니다. 미국의 공판 전 절차(pretrial justice system)에서 피고인의 구속 재판 여부를 결정하는 데 활용되는 공공안전평가(Public Safety Assessment: PSA)는 9가지 위험요소에 관한 통계를 바탕으로 향후 공판에 불출석할 가능성, 새로운 범죄를 저지를 가능성 등을 예측한다. 이들 위험요소들은 피고인의 연령, 범죄의 내용, 전과 등 통계 데이터를 통해 파악될 수

---

227) European Commission for the Efficiency of Justice(CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Council of Europe, 2018, p. 5.

228) 유럽의회의 상임위원회 중 하나로 정식 명칭은 시민자유, 사법, 내무 위원회(Committee on Civil Liberties, Justice and Home Affairs)이다.

229) Committee on Civil Liberties, Justice and Home Affairs, Report on Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial Authorities in Criminal Matters, 2020/2016(INI), 13 July 2021, p. 25. 상세히는 정소영, 형사사법에서 인공지능 사용에 대한 유럽의회 결의안 - ‘인공지능에 의한 결정’ 금지에 관하여 -, 형사정책, 제34권 제2호, 2022, 151-152쪽.

230) John Villasenor and Virginia Foggo, Artificial Intelligence, Due Process and Criminal Sentencing, Michigan State Law Review, Vol. 2020, NO. 2 2020, p. 337.

있는 것들로 구성되어 있으며, 위험요소와 가중치 및 점수 산정방식 모두가 완전히 투명하게 대중에게(피고인 본인에게도) 공개된다.<sup>231)</sup>

## 5. 법관의 독립성

법관은 헌법과 법률에 의하여 양심에 따라서 독립하여 재판한다(헌법 제103조).<sup>232)</sup> 공정한 재판을 위하여 법관은 재판을 함에 있어 누구의 간섭, 지시로부터도 벗어나야 한다.<sup>233)</sup>

헌법 제103조가 규정하는 양심의 의미와 관련하여 주관적 양심설, 객관적 양심설, 절충설이 대립하나,<sup>234)</sup> 설령 법관에게 요구되는 직무상의 양심이라 해도 그 본질인 내면성과 주관성을 배제할 수는 없다. 양심에 따른 판단과 행위는 행위주체의 차원에서 이루어지며, 개인의 직접적인 자기 확신으로부터 비롯되기 때문이다.<sup>235)</sup> 법관이라는 주체가 자기 확신에 의하여 판단하고 행위(판결)를 하였다면, 이는 양심에 따른 독립적 재판이라고 할 수 있을 것이다.

그런데 인공지능 법관이 법적 판단을 내리는 데 필요한 자료의 취사선택이 자연인에 의하여 이루어진다면 이는 결국 인공지능 법관에 대한 ‘간섭’이며, 순전히 인공지능 판사에 의한 재판으로 볼 수 없을지도 모른다. 자연인 인간이 선택적으로 입력하는 자료에 근거하여 판단을 내리는, 정확히는 어떤 패턴을 추출하거나 확률적 위험성 등을 예측하는 인공지능 법관이 ‘그 양심에 따라 독립하여’ 심판하는 헌법상 법관의 상(像)에 부합하는지도 문제될 수 있다.

---

231) 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 911-915쪽.

232) 동조의 ‘양심’에 대하여 헌법 제19조의 그것과 다른 객관적 양심이라는 견해가 다수이다. 그러나 최근에는 법관이 두 개의 이질적 양심을 소유할 필요가 없고, 스스로 두 개의 양심을 가진다고 가정 혹은 착각하더라도 재판 수행에 그다지 유용하지 않다는 견해가 개진된다. 강석정, 법관은 두 개의 양심을 가져야 하는가? - 헌법 제103조 법관의 ‘양심’에 관하여, 사법, 제41호, 2017, 193쪽.

233) 김하열, 헌법강의, 박영사, 2023, 928쪽. 특정한 판결에 의문을 표하거나 담당 판사의 파면을 요구하는 국민청원이 있었음을 환기하면서, 여론이 인공지능 법관의 판단에 영향을 주기는 어렵다고 한 연구로 서지원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 ‘법의 제국’을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 8-10쪽.

234) 상세히는 한수용, 헌법학, 법문사, 2022, 1375쪽.

235) 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 72-73쪽.

한편 헌법 제103조는 헌법과 법률을 법관의 양심에 대한 규범적·객관적·외부적 통제 장치로 강조한다.<sup>236)</sup> 이처럼 판사의 양심이 헌법과 법률의 제약을 받으면서 판사는 심리적으로 고민하게 되며, 이는 서구 근대법 체계가 예정한 갈등이다.<sup>237)</sup> 그런데 인공지능 알고리즘이 이러한 갈등, 고뇌 속에서 양심에 따른 판단을 할 수 있을까?

전술한 것처럼 오늘날 인공지능은 인간의 지능을 통해 수행해온 일들을 처리하고자 고안되었으나, 인간의 두뇌와 지적 작용을 그대로 구현하고자 하지는 않는다. 다수의 논자들이 인공지능을 데이터 분석을 통해 목적 달성에 적합한 행동을 선택·실행하는 합리적 에이전트, 예측기계, 수학적 함수의 일종으로 정의하는 것은 이 때문이다. 현 단계의 인공지능이 입력 값과 출력 값의 함수, 어떤 패턴의 발견이나 결과의 통계적 예측 프로그램에 지나지 않는다면,<sup>238)</sup> 적어도 현행헌법이 규정될 무렵 상정한 양심을 가졌다고 보기는 쉽지 않다. 자연인과 달리 인공지능에는 양심이 존재하지 않는다면, 헌법 제103조에서 ‘그 양심에 따라’라는 문구를 삭제하지 않을 경우 인공지능 법관의 출현이 위험적 요소를 지닌다는 지적이 제기된다.<sup>239)</sup> 인공지능 법관은 당사자들과 감성적으로 교감하고 직업적 양심에 따라 사건을 치열하게 고민하는 판사라고 할 수 없어 한계를 지닌다는 견해도 있다.<sup>240)</sup>

인공지능 법관이 아니라 인공지능 알고리즘의 판단, 예측결과를 참고하는 경우에조차 법관의 독립성이 문제될 수 있다. 가령 법관이 COMPAS를 활용하여 양형판단을 한 경우에, 법관이 실제로 그 점수를 얼마나 고려했는지 확인할 길은 없다. 상당히 앞서가는 과학기술이라는 인식이 만연한 인공지능 알고리즘이 제시한 예측결과는 상당한 영향력을 가지기 십상이다. 그런데 영업비밀이라는 등의 이유로 인공지능 알고리즘의 작동방식이 불투명하다면, 재범위험성에 관하여 강력한 의견을 표출하는 검증되지 않은 자격 없는 판사를 법정에 끌어들이는 결과가 될 수도 있다.<sup>241)</sup>

236) 강석정, 법관은 두 개의 양심을 가져야 하는가? - 헌법 제103조 법관의 ‘양심’에 관하여, 사법, 제41호, 2017, 187쪽.

237) 강석정, 법관은 두 개의 양심을 가져야 하는가? - 헌법 제103조 법관의 ‘양심’에 관하여, 사법, 제41호, 2017, 190쪽.

238) 인간과 같은 사고체계를 갖춘 강한 인공지능의 개발은 기술적으로 요원하며, 경제적 차원에서도 의문이 제기됨을 전술하였다.

239) 전후재, 우리 민사재판의 미래에 관한 소고 - 디스커버리, 온라인 소송, 인공지능(AI)을 중심으로 -, 민사소송, 제27권 제1호, 2023, 140-141쪽.

240) 강승식, 인공지능 판사, 과연 가능한가?, 헌법학연구, 제26권 제3호, 2020, 237쪽.

241) 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집,

## 6. 재판제도 및 실무의 본질적 변화 가능성

### 가. 합의체

재판에는 단독사건도 있지만 여러 법관이 합의부를 구성하여 하는 사건도 있다. 현행헌법은 9인의 재판관으로 구성된 헌법재판소를 상정하며(제111조 제2항), 대법원도 대법원장과 대법관으로 구성되므로(제5장), 복수의 판단주체에 의한 재판을 상정한다고 풀이된다.

그런데 인공지능 법관이 현실화된다면, 상이한 판단을 하는 여러 법관(재판관)이 합의체를 구성하고 다양한 의견(다수의견, 반대의견, 보충의견 등)을 개진하며 논박하는 재판 자체가 근본적으로 달라질 수도 있다고 전망된다.<sup>242)</sup> 인공지능 알고리즘의 판단이 법관의 그것을 대체하는 경우, 더 나아가 인공지능을 탑재한 로봇 법관(들)이 모든 재판을 주재하는 경우 모두 이러한 문제가 불거질 수 있다.

판결에는 개인의 경험을 통해 형성된 호·선입견·본능·습관·신념이 반영되는바 이는 무의식적이고 불가피하며, 재판부 구성의 다양성은 가치의 독단을 견제하고 판결의 설득력을 제고할 수 있다고 평가된다.<sup>243)</sup> 그런데 상이한 경험과 배경을 가진 다양한 판단주체가 의견을 개진하고(다양성), 그것이 당사자들의 변론, 사회적 여론, 가치관의 변화에 따라 변화해가는(가변성) 재판을 인공지능에 기대할 수 있는가?<sup>244)</sup>

만약 인공지능 법관 모두가 동일한 알고리즘을 가지고 학습(머신러닝)한 것이라면, 동일한 사안에서 같은 결론을 도출할 터이므로 굳이 여럿을 둘 이유가 없다. 서로 다른 알고리즘으로 작동하는 인공지능 법관을 여럿 두는 방법도 있겠으나,<sup>245)</sup> 이 또한 많은 헌법적·정책적 문제를 초래할 수 있다. 그만큼의 다양한 인공지능 법관을 구비할 수 있을

---

2021, 922-924쪽.

242) 헌법은 9인의 재판관으로 구성된 헌법재판소를 상정하며(제111조 제2항), 대법원도 대법원장과 대법관으로 구성되므로(제5장) 복수의 판단주체에 의한 재판을 상정한다고 풀이된다.

243) 강승식, 인공지능 판사, 과연 가능한가?, 헌법학연구, 제26권 제3호, 2020, 248-249쪽.

244) 사건은 어떤 판사를 만나는 같은 결론에 이르러야 한다는 주장이 있으나, 이는 현실에 부합하지 못하다. 또한 법관들 사이의 견해차를 폄하하거나 성급한 의견일치를 종용할 우려가 있다. 박은정, 법해석 방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 70쪽.

245) 번역의 경우에도 구글 번역기, 파파고(Papago), 딥엘(DeepL) 등 여러 인공지능 번역기가 동일한 문장을 조금씩 다르게 번역한다. 이에 관하여 상세히는 이예린, ‘신조어 강자’ 파파고·‘다국어 지원’ 구글 번역기·‘번역의 정석’ 딥엘[ICT], 문화일보, 2023. 12. 11, <https://www.munhwa.com/news/view.html?no=2023121101032007275001>(최종방문일 2024. 2. 11.).

지, 어떤 알고리즘을 무슨 기준에 따라 선정할 것인지 답하기가 쉽지 않다.<sup>246)</sup> 특정한 알고리즘을 가진 법관들을 일정 수 배치하는 것 자체가 이미 사건의 전반적인 향배를 좌우할 수 있다는 점도 문제된다.

## 나. 심급제

법원에서 사실판단과 법률적용을 잘못함으로써 인하여 발생하는 기본권의 침해를 사법절차를 통해 구제하기 위해서, 재판의 심급제도가 마련되어 있다. 헌법은 법원의 조직을 대법원과 각급법원으로 계층화하고 있으며 비상계엄하의 단심군사재판을 예외현상으로 규정하고 있는바, 심급제도는 우리 사법절차에서 불가결한 요소다.<sup>247)</sup>

그런데 동일한 알고리즘에 따라 작동하는 인공지능 법관들로 각급 법원이 구성될 경우, 사실상 심급제<sup>248)</sup>가 형해화될 우려가 있다. 심급제는 다양한 경험과 사법철학을 가진 각급 법원 법관들의 판단이 동일하다면 어느 정도 신뢰할 수 있다는 점에서도 그 의의가 있다.

그런데 상급심 인공지능 법관의 알고리즘이 하급심 인공지능 법관과 동일할 경우, 같은 사실관계를 입력하면 같은 결과가 나오므로 (하급심 판결 이후 사실관계, 적용법조에 근본적인 변화가 있는 경우가 아니라면) 항소, 상고 등은 무의미하다.<sup>249)</sup>

관련하여 하급심은 인공지능 법관, 상급심은 자연인 법관이 담당하는 구조로 재판제도를 구성하거나, 인공지능은 1심 이전의 전심 절차에서만 분쟁의 조기 해결에 기여하도록 만들 수 있다는 견해가 있다.<sup>250)</sup> 그런데 설령 상급심 법원을 자연인 법관이 담당하더라도 문제는 남는다. 하급심 법원이 어떠한 과정을 거쳐 판결하였는지가 투명하고 설명 가

---

246) 인공지능이 서로 다른 다양한 가치관을 개별적으로 가지리라 기대하기는 어려우며, 모두 동일한 알고리즘으로 작동하는 형태가 될 가능성이 높다는 평가로 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 76쪽.

247) 허영, 한국헌법론, 박영사, 2023, 1159쪽.

248) 대한민국헌법 제101조 ② 법원은 최고법원인 대법원과 각급법원으로 조직된다.

249) 이러한 취지의 주장으로 고유강, 법률인공지능이 변화시킬 재판의 모습: ‘인공지능 판사’의 기술적 대체가능성과 관련된 현황을 중심으로, 재판자료집, 제142집, 2021, 29쪽.

250) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 289쪽. 김도훈(2021)에 따르면, 인공지능 또한 오류가능성을 지니며 가변성은 재판의 본질적 속성이므로, 인공지능 법관을 도입하더라도 상소제도는 여전히 필요하다. 이 연구는 인공지능이 정책적 판단능력을 요하는 법률심을 담당하는 것은 적절치 못하다며, 사실심 그 중에서도 1심만 담당하도록 할 것을 제안한다. 그리고 항소심은 자연인 법관이 담당하도록 하면, 통제적 기능을 실현할 수 있다고 한다. 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 77쪽.

능해야 이를 파기할지 유지할지 여부를 판단할 수 있는데, 하급심 인공지능 판사의 판단 과정이 블랙박스과 같다면, 상급심 자연인 법관이 이를 어떻게 파악하여 파기 또는 유지할 것인지가 문제된다. 상급심 자연인 법관이 하급심 인공지능 판사의 판단에 지나치게 경도되어 동일한 판단을 할 가능성도 배제할 수 없다.

#### 다. 판례변경

“선례 없는 사건이나 시대상에 맞춰 판례를 바꿔야 하는 사건에서 현 AI가 제대로 된 판단을 할 수 있을지 의문” (오세용 인천지방법원 부장판사)<sup>251)</sup>

종전 사건들의 법리와 사실관계 등을 인공지능 법관의 학습 데이터로 삼는다면, 사회적 여건과 가치관의 변화, 선행 판결·결정에 관한 반성적 분석을 통한 판례변경이 가능한가? 한 번 형성된 법리가 고착될 우려는 없는가?

생각건대, 사실관계가 동일·유사한 사안이라면 인공지능 판사의 알고리즘이 전면적으로 변화하지 않는 한 판례변경 등은 어려울 것이다.<sup>252)</sup> 이와 관련하여 인공지능의 판단은 기존 판결들과 일관성을 가지겠지만, 새로운 관점의 변화를 반영하지 못하며 판결의 퇴보로 이어질 수 있다는 지적이 있다.<sup>253)</sup> 법원의 판결이 종전에 다른 법원이 내린 판결들의 평균적 계산으로 수렴하게 될 가능성도 있다.<sup>254)</sup>

혹자는 약한 인공지능이 자율성을 가지지 못한다 하더라도 그 동안 정립된 해석방법과 범죄체계론 및 판례를 학습함으로써 죄책판단을 할 수 있다고 하면서도, 기존의 판례나 해석론이 없는 난제(hard case)에 있어서는 인공지능이 제대로 죄책을 판단하기 어려

251) 이배운, “판사를 AI로 교체하자” 현직 판사에게 물어보니, 이데일리, 2023. 7. 24, [https://www.edaily.co.kr/news/read?newsId=02551846635677800&mediaCodeNo=257&OutLnkChk=Y\(최종방문일 2024. 2. 11.\)](https://www.edaily.co.kr/news/read?newsId=02551846635677800&mediaCodeNo=257&OutLnkChk=Y(최종방문일 2024. 2. 11.)). 인공지능이 과거 판례에 반영되지 않았던 법률 개정상황, 기술변화 등 사회발전과 변화에 적절히 대처하여 당사자들이 모두 수용할 수 있는 결론을 내릴지 의문이라는 지적으로 정문식, 인공지능과 헌법, 한양대학교 법학연구소 편, 인공지능과 법, 도서출판 오래, 2023, 66쪽.

252) 같은 입장으로 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 13쪽.

253) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범설계, 법학논총, 제27집 제1호, 2020, 136쪽. 인공지능만으로는 기존 법리를 변경하는 판결이 어려움을 환기한 조원용, 대의제 국회 입법과정에서 인공지능(AI)의 역할: 공익의 ‘발견’과 ‘결정’을 중심으로, 가천법학, 제11권 제3호, 2018, 62-63쪽의 각주 36도 참조.

254) 서울대학교 산학협력단, 사법부에서의 인공지능(AI) 활용방안, 법원행정처, 2020, 63쪽.

을 것이라고 한다.<sup>255)</sup> 아직 규율이 확립되지 않고 변화해 가는 분야에서, 기존 데이터를 학습한 인공지능 법관이 변화에 적절히 대응하지 못하고 새로운 결론을 내는 데 지나치게 소극적일 가능성을 환기한 연구도 있다.<sup>256)</sup>

국제적인 차원에서도, 인공지능에 의한 사법적 판단이 자연인 법관을 선례로부터 벗어나지 못하게 할 위험성이 지적된다. 유럽평의회(Council of Europe) 산하의 사법효율을 위한 유럽위원회(European Commission for the Efficiency of Justice)에서는 2018년 사법시스템과 사법 환경에서의 인공지능 이용에 관한 유럽 윤리헌장을 채택하였는데, 판사들의 미래 선택을 선례들에 가둘 수 있음을 환기하였다. 그리고 판례법의 결정화(crystallization of case law) 위험성 및 법관의 공정성과 독립성에 부정적인 영향을 미칠 가능성을 지적하였다.<sup>257)</sup>

혹자는 인공지능 판사가 인간판사보다 많은 장점을 가진다고 인정하면서도, 바로 이 문제 때문에 법관의 역할을 ‘대체’할 수는 없고 ‘보완’하는 데 만족해야 한다고 한다. 그에 따르면 인공지능 판사는 여론에 기속되는 문제가 없으며, 축적된 판례를 학습한 까닭에 과거 결정과의 일관성 면에서 인간 판사보다 나올 수 있다. 또한 학습 속도가 월등히 빠르고 언어적 제한에 구애되지 않기 때문에, 국내뿐 아니라 세계 각국의 제도와 판례를 학습하며 비교법적으로 문제를 해결할 수도 있다. 그러나 인공지능은 그 구성상 기존의 판례에 기속되는 문제가 있으며, 이는 잘못된 판단의 지속을 방치하는 결과로 이어질 수 있다고 한다. 인간의 가장 큰 덕목은 잘못을 고치고 새로 시작할 수 있는 용기인데, 인공지능 법관에게서는 이러한 역할을 기대할 수 없다는 취지이다.<sup>258)</sup>

기존의 판례로부터 벗어나기 어렵다는 문제에 더하여, 유사한 사안에 관한 상이한 판결례가 존재할 경우 인공지능 법관이 적절히 선택하기가 어려우리라는 점도 지적될 수 있다.<sup>259)</sup>

---

255) 양천수, 인공지능과 법체계의 변화 - 형사사법을 예로 하여, 법철학연구, 제20권 제2호, 2017, 58-59쪽.

256) 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 13쪽. 인공지능이 알고리즘에서는 소외된 소수자에 대한 독창적 배려와 결단을 기대하기 힘들다고 한 김일우, 고위험 인공지능시스템의 기본권 침해에 관한 헌법적 연구, 박사학위논문, 서강대학교, 2023, 118쪽도 참조.

257) European Commission for the Efficiency of Justice(CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Council of Europe, 2018, p. 67.

258) 서치원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 ‘법의 제국’을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 8-20쪽.

259) 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 68쪽.

## 라. 인공지능 법관(들)의 설계 및 배치 문제

모든 인공지능이 같은 사안에서 동일한 결론에 이르지 않을지도 모른다. 알고리즘이 상이하고 학습 데이터가 다르다면, 서로 다른 결과를 도출할 것이다.<sup>260)</sup> 하지만 재판부의 합의, 심급제, 판례변경 가능성을 위하여 서로 다른 알고리즘으로 구현된 인공지능 법관을 설계해 배치하는 것 또한 여러 규범적, 현실적 문제로부터 자유롭지 못하다. 각각의 인공지능 법관들이 어떤 알고리즘, 성향을 가지도록 할 것인가? 누가 그것을 결정하는가? 특정 알고리즘과 성향을 가진 인공지능 판사를 배치함으로써, 사실상 사건의 결론이 좌우되는 셈은 아닌가? 모두 답하기 어려운 문제이다.

## 7. 인공지능 법관의 판단에 의한 기본권 침해 가능성 및 책임의 귀속 문제

### 가. 편향성에 따른 부당한 차별

#### 1) 인공지능 편향성의 원인

전술하였듯 알고리즘이 과학적이고 가치중립적이라는 관점도 있지만, 인공지능 알고리즘은 인간이 볼 때 차별이라고 여겨지는 결과를 산출하고는 한다. 인공지능 편향성의 원인으로서는 다음과 같은 것들이 있다.

#### 가) 정제되지 못하고 편향된 데이터

미국 등에서 인공지능은 예측 치안(predictive policing)에서 활용되고 있다. 그런데 경찰이 수집한 범죄 관련 데이터는 이를 표준화할 절차와 방법이 마련되어 있지 못하며, 작성과정에 관한 책임 있는 통제도 없다. 예측치안 알고리즘 공급업체에서 관련 데이터를 정제하기 위한 노력을 하는 것도 아니다.<sup>261)</sup>

인공지능의 학습 및 판단을 위하여 투입되는 데이터 자체가 사회에 존재하는 편향을

260) 전술하였듯 번역의 경우에도 구글 번역기, 파파고(Papago), 딥엘(DeepL) 등 여러 인공지능 번역기가 동일한 문장을 조금씩 다르게 번역한다.

261) 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 388쪽.



반영하기 십상이며,<sup>262)</sup> 개발자의 의도적인 조작이 없더라도 편향성을 지녀 부당한 차별 취급으로 이어질 가능성이 있다. 관련 데이터가 적을 경우 부정적 예측의 가능성이 높아져, 소수집단에 대한 예측결과가 상대적으로 나쁠 가능성도 있다.<sup>263)</sup>

데이터가 ‘디지털화되는 과정’에서 편향성이 개입될 수도 있다. 또한 데이터는 그 수집 주체·시기·방법·목적 등에 따라 그 특성이 달라지기도 하며, 관련 양상을 전체적으로 정확하게 반영하지 못할 수 있다. 가령 허리케인 샌디(Sandy)에 관한 대부분의 트윗이 뉴욕 맨해튼 지역에서 나왔는데, 마치 맨해튼이 태풍 피해의 중심지처럼 착각하게 만들 수 있다.<sup>264)</sup>

한 선행연구에서는 데이터 공정성과 관련하여 측정오류와 표본편향 문제를 지적한다. ① 측정오류에는 예측변수에 대응하는 특성값이 잘못 측정된 경우(feature bias)와, 결과 변수(y)에 대응하는 결과값이 잘못 측정된 경우(label bias)가 있다. 실제 피의자가 마약판매를 한 횟수를 알기 어려워, 마약판매로 ‘체포된 횟수’를 기준으로 하는 경우가 전자이다. 실제 재범을 하는 것을 직접 관찰하기 어렵기에, 재범을 하여 유죄판결을 받았는지를 결과값으로 하는 경우가 후자이다. ② 표본편향(sample bias) 오류는 가령 학습데이터의 인구분포와 알고리즘이 적용되는 지역의 인구분포가 상이한 경우에 발생할 수 있다.<sup>265)</sup>

#### 나) 오류 및 편향성을 지닌 알고리즘에 입각한 판단

여러 변수 사이의 관계를 파악하기 위해 가령 4차원 이상의 공간을 구성하는 알고리즘은 인간이 지각할 수 있는 범위를 넘어선다. 인간이 이해할 수 있는 차원으로 표현하자면 차원을 축소해야 하는데, 몇 개의 변수가 대표되는 과정에서 오류, 편향이 불가피할 수 있다. 가령 성별, 연령, 국적, 피부색 등의 몇몇 특성으로 인간을 정형화해 판단할 가능성이 크다.<sup>266)</sup>

---

262) 수집된 데이터가 반드시 통계적 편향성 없이 완벽한 대표성을 갖추었다고 확신하기 어렵다는 지적으로 박도현, 인공지능 공정성의 이론과 실제, 법철학연구, 제26권 제2호, 2023, 26쪽.

263) 특정 집단에 관한 데이터가 매우 부족하다면 이는 주류적 집단을 중심으로 수집된 것으로서 구조적으로 신뢰하기 어렵다는 지적으로 이준일, 인공지능과 헌법 - 인공지능의 위험에 대한 헌법적 권리 차원의 대응 -, 헌법학연구, 제28권 제2호, 2022, 365쪽.

264) 양종모, 인공지능 알고리즘의 편향성, 불투명성이 법적 의사결정에 미치는 영향 및 규율 방안, 법조, 제66권 제3호, 2017, 76-78쪽.

265) 박혜진, 양형과 인공지능 - 국민의 법감정 수렴, 위험 판단, 양형 분석을 중심으로 -, 법조, 제72권 제4호, 2023, 269-270쪽.

266) 남중권, ‘알고리즘의 지배’와 ‘법의 지배’에 대한 법이론적 이해, 법과사회, 제71호, 2022, 301쪽.

인공지능 알고리즘은 의도적으로 편향되어 있을 수 있다. 관련하여 기술의 작동, 관행, 인식, 해석은 인공지능과 데이터, 기술을 다루는 사람들과 조직이 중립적이지 않음을 감안하여 평가하여야 한다는 주장이 제기된다.<sup>267)</sup> 이와 달리, 설계자 자신도 모르게 알고리즘이 편향되어 있을 수도 있다.<sup>268)</sup> 편향된 알고리즘에 기인한 오류·오판에 따라, 평등권을 비롯한 각종 기본권이 침해될 가능성도 배제할 수 없다.<sup>269)</sup>

## 2) 인공지능 편향성에 따른 기본권 침해 가능성

미국의 탐사보도 매체 프로퍼블리카(ProPublica)는 2016년 ‘기계의 편견(Machine Bias)’이라는 제목의 기사에서, 재범위험성을 예측하는 COMPAS 프로그램이 인종적으로 편향되어 있으며 정확성이 떨어진다는 보고서를 발표하여 주목받았다.<sup>270)</sup> 또한 COMPAS에서는 성별과 연령 변수가 중요한데, 젊은 남성의 위험지수가 높게 평가될 우려가 있다고 지적된다.<sup>271)</sup> 미국에서 지금까지 흑인의 강력범죄율이 백인의 그것보다 높았는데, 이는 흑인 피고인이 동일·유사한 범죄를 저지른 백인 피고인보다 더 높은 형을 선고받고 가석방될 확률이 줄어드는 결과를 초래할 수 있다.<sup>272)</sup>

267) 마크 코켈버그, 배현석 옮김, 인공지능은 왜 정치적인 수밖에 없는가 - AI의 정치학과 자유, 평등, 정의, 민주주의, 권력, 동물과 환경, 생각이음, 2023, 117쪽.

268) 이준일, 인공지능과 헌법 - 인공지능의 위험에 대한 헌법적 권리 차원의 대응 -, 헌법학연구, 제28권 제2호, 2022, 365쪽.

269) 인공지능의 분석에 따라 중범죄를 저지르리라 예측되는 자들을 목록화해 집중적으로 감시할 경우, 실제 선량한 시민이 예비범죄자로 낙인찍히는 등의 문제가 있을 수 있다고 한 이병규, AI의 예측능력과 재범예측알고리즘의 헌법 문제 - State v. Loomis 판결을 중심으로, 공법학연구, 제21권 제2호, 2020, 173쪽 참조.

270) Julia Angwin et al., “Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And It’s Biased Against Blacks.”, ProPublica, 23 May 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>(최종방문일 2024. 2. 11.). 프로퍼블리카가 수행한 통계적 분석의 핵심은 인종별로 통계 프로그램의 오류율이 상이하다는 점이다. 높은 위험(8점 이상)으로 분류되었지만 실제로 재범을 저지르지 않은 비율(False Positive Rate)이 백인은 23.5%였으나, 흑인은 44.9%였다고 한다. 낮은 위험(7점 이하)으로 분류되었으나 실제로 재범을 저지른 경우(False Negative Rate)가 백인은 47.7%였는데, 흑인은 28.0%에 불과하였다고 한다. 서울대학교 산학협력단, 사법부에서의 인공지능(AI) 활용방안, 법원행정처, 2020, 58-59쪽.

271) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 680쪽.

272) 이지나, 형사사법에서의 인공지능 예측프로그램 활용 제언: 인종적 편향 가능성 회피를 중심으로, 원광법학, 제39권 제2호, 2023, 63쪽.

### 3) 인공지능 편향성의 특성과 해악

편향된 데이터에 따라 인공지능 알고리즘이 차별적인 판단(예: 유색인종 피고인의 높은 유죄 선고율, 형량)을 하고, 그것이 또 인공지능 알고리즘이 학습하는 입력변수(데이터)가 되는 ‘악순환’이 발생할 수도 있다.<sup>273)</sup> 수학자 캐시 오닐(Cathy O’Neil)은 인간의 편향성이 일단 훈련 데이터나 알고리즘에 개입할 경우 재차 입력에 반영될 수 있다고 하면서, 인공지능이 창발적 해악을 유발하는 경로로 피드백 루프(feedback loop)를 제안하였다.<sup>274)</sup>

만약 인공지능의 의사결정이 인간의 편향성을 거울처럼 반영할 뿐이라면, 독자적인 규율은 불필요하며 문제의 근원인 인간의 행동을 규제해야 한다는 반론이 제기될 수 있다.<sup>275)</sup> 그런데 선행연구가 지적하였듯, 인공지능은 인간의 의사결정에서와 다른 창발적 해악을 초래할 수 있다고 평가된다. ① 인간도 상호작용을 통해 편향성을 강화하지만, 인공지능은 같은 시간에 훨씬 더 많은 편향성을 무수하게 쏟아낸다.<sup>276)</sup> ② 인간은 편향성을 회피하는 선호를 가지고 있다. 자신의 마음 속 깊이 자리한 편향성을 그대로 반영하기보다 보정을 가하려는 습성이 있다는 것이다. 그런데 인공지능은 이 같은 인간의 편향성 회피 선호를 누락하여, 인간 편향성을 실제보다 과다하게 평가할 수 있다. ③ 인공지능은 판단의 편차를 현격히 감소시킴으로써, 자신과 다른 견해를 가진 구성원과 교류할 가능성을 낮출 수 있다.<sup>277)</sup>

인공지능의 편향성을 인식하기가 쉽지 않을 수도 있다. 가령 인공지능의 학습 데이터에 젠더, 인종과 같은 기준에 관한 명시적인 언급이 없다면, 편향을 인식하기 쉽지 않을

---

273) 이러한 문제는 결국 데이터 자체의 결함과, 그 데이터를 입력변수로 사용하여 스스로 학습하며 가중치를 부여하는 인공지능 알고리즘의 문제가 맞물린 것이다. 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 388-389쪽. 인공지능에 입각한 예측치안에서도 이 같은 문제가 발생한다. 인공지능 알고리즘이 주로 유색인종, 빈민이 거주하는 지역을 우범지대로 지정하고 그 결과 해당 지역에 더 많은 경찰이 출동해 경범죄를 적발한다. 그곳에 거주하는 빈민, 흑인, 히스패닉이 체포되고, 이는 다시 해당 지역을 순찰해야 할 필요성을 강화한다. 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 701쪽.

274) 박도현, 인간 편향성과 인공지능의 교차, 서울대학교 법학, 제63권 제1호, 2022, 161쪽. (원문 확인)

275) 인공지능의 편견과 차별 위험이 과장되거나 오도되기 쉽다고 한 이상용, 알고리즘 규제의 원리와 구조, 김진우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 216쪽 참조.

276) 인공지능은 잠도 자지 않고 망각하지도 않으며 감정적이지도 않다. 그리고 오늘날 인공지능의 작동 속도는 인간 행위의 그것보다 훨씬 빠르다.

277) 박도현, 인간 편향성과 인공지능의 교차, 서울대학교 법학, 제63권 제1호, 2022, 157-162쪽.

것이다. 그리고 이 같은 문제점을 인식하지 못한다면, 해결책이 없을 수도 있다.<sup>278)</sup>

#### 4) 인공지능 편향성 극복에 있어서의 어려움

편향성을 극복하기 위한 방안으로 인공지능 알고리즘 설계 단계에서의 윤리적 책무 강화, 표준화 작업, 인간의 언어 및 논증구조를 통한 설명가능성 확보, 알고리즘에 대한 감사 등이 제시된다.<sup>279)</sup> 하지만 인공지능 알고리즘의 편향성 문제를 해결하기란 상당히 어렵다. ‘공정성’에 관한 관점의 차이가 있으며, 상이한 입장을 모두 충족시키기 힘들다는 문제가 있다.<sup>280)</sup> 가령 현실을 그대로 반영하지 못한다면 불공정하다고 보는 입장이 있는가하면, 현실을 고스란히 반영하는 것이 불공정하다고 여기는 입장도 있다.<sup>281)</sup>

일각에서는 재범위험성 평가 알고리즘인 COMPAS가 인종차별적인 효과를 초래하는 편향성을 가진다고 비판한다. 그러나 제조사인 Equivant(당시 Northpointe)는 평가의 정확도가 흑인과 백인 피고인 모두에게서 동일하며 따라서 인종 중립적이라고 반박하였다. 흑인이 백인보다 더 많이 재구속되는 현실에서, ‘결과에 있어서의 공정성’과 ‘예측의 동등성’을 모두 충족하기란 불가능하다.<sup>282)</sup>

알고리즘의 편향성을 인위적으로 교정하려는 시도가 또 다른 불공정을 초래할 여지도 있다. 결과를 산출함에 있어 각종 변수의 우선순위 및 가중치를 변경하는 개입 자체가 편향적 관점을 주입하는 것이 될 수도 있다.<sup>283)</sup>

인공지능의 불공정성을 교정하기 위한 노력이 여타 가치와 상충할 여지도 있다. 가령

---

278) 마크 코켈버그, 배현석 옮김, 인공지능은 왜 정치적인 수밖에 없는가 - AI의 정치학과 자유, 평등, 정의, 민주주의, 권력, 동물과 환경, 생각이음, 2023, 95쪽.

279) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 134-135쪽.

280) (1) 현실에서 여성의 재범위험이 남성보다 낮음에도 인종이나 성별 내지 그 대리지표를 아예 고려하지 못하도록 하는 경우, 여성의 재범위험성이 상대적으로 과다 평가될 수 있다. (2) 위양성률과 위음성률이 동일해야 한다는 기준(분류동등성)도 있지만, 위양성률의 차이는 모집단의 특성을 반영하는 경우가 더 많다는 점에서 비판이 가해진다. 즉 흑인 피고인이 백인 피고인보다 재범 확률이 높기 때문에 위양성률이 더 높게 나타날 수 있다는 얘기다. (3) 예측된 위험점수가 같다면 실제 재범확률이 같아야 한다는(calibration) 기준도 있는데, 분류동등성을 심각하게 침해하는 경우가 있다. 박혜진, 양형과 인공지능 - 국민의 법감정 수렴, 위험 판단, 양형 분석을 중심으로 -, 법조, 제72권 제4호, 2023, 268-269쪽.

281) 박도현, 인공지능 공정성의 이론과 실제, 법철학연구, 제26권 제2호, 2023, 23쪽.

282) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 128-129쪽.

283) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 131쪽.

인공지능 의사결정의 공정성을 확보하고자 민감정보를 확인하려는 것이 개인정보보호법상 최소수집의 원칙(제3조 제1항)을 해칠 우려가 있으며, 공정성 지표의 향상이 정확성의 축소를 일정 정도 동반하곤 한다.<sup>284)</sup>

## 나. 판단과정의 불투명성과 기본권 침해

### 1) 알고리즘 불투명성의 원인

인공지능 알고리즘은 사기업의 영업비밀로서 그 작동구조가 공개되지 않는 경우가 많다. *Loomis v. Wisconsin* 사건에서 피고인은 COMPAS의 소스코드와 가중치에 관한 정보를 요구하였지만 법원이 거절하였음을 앞서 보았다.

전술한 것처럼 알고리즘을 모두 공개하더라도, 관련 지식이 부족한 사람이라면 그 원리를 파악하기가 어렵다. 또한 특정 인공지능 알고리즘의 핵심은 해당 개발자가 만들지만 복잡한 연산을 담당하는 라이브러리 등은 구글을 위시한 선발기업에서 만든 라이브러리에 의존하는데, 그것이 너무도 복잡해 개발자들도 분석할 엄두를 내지 못한다. 요컨대 일반인, 법관은 물론 인공지능 법관의 개발자조차도 그 구체적인 알고리즘을 알기 어려울 수 있다는 것이다.

### 2) 판단과정의 불투명성과 관련한 헌법적 문제

인공지능 알고리즘의 불투명성이 언제나 법적인 문제를 발생시킨다고 보기는 힘들다. 가령 로보어드바이저( robo-advisor)처럼 투자수익만 내면 되는 경우에는 투자결정 과정 자체가 크게 문제되지 않을 수 있다. 하지만 당사자의 권리의무에 관한 판단을 하는 인공지능 법관의 경우, 그 결정과정이 투명하게 설명되어야 한다.<sup>285)</sup> 그런데 법조 영역에서 활용되는 인공지능 알고리즘은 기업의 영업비밀이며 자산적인 성격을 가지기에 공개되지 않는 경우가 많다.

인공지능의 이러한 ‘블랙박스’와 같은 속성은 공정한 재판을 받을 권리와 관련한 문제를 자아낼 수 있다. 알고리즘의 결론은 마치 델포이의 신탁처럼 나타날 수 있고, 판단의

284) 박도현, 인공지능 공정성의 이론과 실제, 법철학연구, 제26권 제2호, 2023, 43쪽.

285) 양종모, 인공지능에 의한 판사의 대체 가능성 고찰, 홍익법학, 제19권 제1호, 2018, 19쪽.

기초가 된 자료(빅데이터), 가정(presumption), 편향성 등에 관하여 탄핵할 기회가 사실상 주어지지 않기 때문이다.<sup>286)</sup>

인공지능 알고리즘의 불투명성이 불평등을 초래하리라는 견해도 개진된다. 가령 대부분의 피고인은 전문지식이 부족해 재범위험성 예측 알고리즘을 이해하기 어렵는데, 전문가들의 도움을 받을 수 있는 부유한 피고인과 그렇지 못한 피고인 간의 불평등이 심화될 것이라는 것이다.<sup>287)</sup>

투명성 문제를 위해 ‘설명 가능한 인공지능’이 시도되고 있지만, 대체로 설명가능성과 예측의 정확도가 반비례한다는 문제가 있다.<sup>288)</sup> 다시 말해, 설명요구권은 당해 인공지능의 정확도를 떨어뜨릴 수도 있다.<sup>289)</sup> 투명성을 위해 알고리즘을 공개할 경우 ‘조작’으로 인하여 당초 목표가 좌절될 우려도 있다.<sup>290)</sup>

다만 최근에는 해석이 가능하면서도 블랙박스 모델만큼 성능이 우수한 모델들이 존재한다는 연구결과를 제시하면서, 형사절차에서는 해석 가능한 유리박스(glass box) 알고리즘을 사용하는 것이 헌법적으로 요구된다는 주장도 제기된다.<sup>291)</sup> 더 적은 변수를 사용하여 비전문가들도 손쉽게 해석할 수 있으면서도 같은 결과를 달성하는 ‘착한 알고리즘’이 있음을 논증한 연구도 있으며,<sup>292)</sup> 최근에는 자신이 내린 결론의 이유를 자연어로 설명할 수 있는 머신러닝 알고리즘이 고안되기도 하였다.<sup>293)</sup>

---

286) 이러한 취지의 비판으로 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 395쪽. 그런데 블랙박스 속성이 인공지능 의사결정에서만 문제 아닐 수 있다. 진정으로 블랙박스라도 같은 것은 인간의 직관적 의사결정이라는 서술로 박도현, 인간 편향성과 인공지능의 교차, 서울대학교 법학, 제63권 제1호, 2022, 165쪽.

287) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 695쪽.

288) 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 381쪽.

289) 김송옥, 인공지능 법제의 최신 동향과 과제, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 177-178쪽.

290) 가령 조세법 적발과 관련한 예측모델의 기법이 투명하게 공개되면, 이를 악용하여 조세를 회피할 우려가 있다. 범죄예측모델의 알고리즘이 밝혀질 경우, 예측 시 고려되는 각종 행동을 하지 않음으로써 이를 무력화시킬 수도 있다. 양종모, 인공지능 알고리즘의 편향성, 불투명성이 법적 의사결정에 미치는 영향 및 규율 방안, 법조, 제66권 제3호, 2017, 94쪽.

291) 상세히는 박혜진, 양형과 인공지능 - 국민의 법감정 수렴, 위험 판단, 양형 분석을 중심으로 -, 법조, 제72권 제4호, 2023, 271-272쪽. 설명이 불가능한 알고리즘이 성능이 훨씬 뛰어나다는 이유로 형사사법 절차에 도입하려면, 도입을 주장하는 측에서 높은 증명책임을 충족시켜야 한다고 한다.

292) Elaine Angelino. 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 695쪽에서 재인용.

293) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의

### 3) 알고리즘의 불투명성 관련 사례

복역 중 가석방을 신청하였으나 거부당한 로드리게스(Glenn Rodríguez) 사건은 인공지능 알고리즘의 불투명성을 잘 보여준다. 로드리게스는 뉴욕 북부에 소재한 교정시설에 수감되어 준수한 교정 기록을 보였다. 그러나 재범위험성을 판단하는 COMPAS가 로드리게스의 위험성을 높게 판단하였기에, 가석방심사위원회에서는 가석방 기각결정을 정당화하였다. 하지만 COMPAS의 알고리즘은 영업 기밀이었으며, 가석방심사위원회는 COMPAS가 어떻게 위험성의 정도를 측정하는지 알지 못하였다. 로드리게스는 COMPAS의 평가에 입력된 정보 가운데 하나에서 오류의 존재를 확인하였으나,<sup>294)</sup> 입력된 정보의 가중치를 몰랐기에 그러한 오류가 어떤 결과를 초래했는지 설명할 수 없었다.<sup>295)</sup> 다른 재소자들의 위험성과 교차 검증한 끝에, 로드리게스는 그의 높은 위험성 점수가 인종적 편견에 영향을 받았을 수 있는 교도관의 주관적인 관점에 따른 것임을 밝혔고 조기 석방되었다. 형사사법 및 예측치안(predictive policing)에서 인공지능의 차별적인 이용 사례의 하나로 꼽힌다.<sup>296)</sup>

### 4) 토론: 불투명성 지적에 관한 비판

혹자는 머신러닝 시스템의 개념을 고려할 때, COMPAS 시스템은 입력과 출력만 알 뿐 가중치 설정 등에 관하여 알 수 없다는 비판이 말도 안 되는 공격이라는 느낌이 든다고 한다. 머신러닝 알고리즘의 특징은 은닉층에서의 가중치 변경을 통해 적절한 결론을

---

분석, 사법, 제64호, 2023, 696쪽. 다만 의료적 진단에 관한 연구였다고 한다.

294) 교도소 공직자가 자신에 대한 COMPAS의 질문 가운데 하나에 실수로 답변한 것을 발견하였다. Jenny Jiao, The Pandora's Box of the Criminal Justice System, *Juris*(Duke's Undergraduate Law Magazine), 25 September 2017, <https://dukeundergraduatelawmagazine.org/2017/09/25/the-pandoras-box-of-the-criminal-justice-system/>(최종방문일 2023. 12. 31.).

295) Rebecca Wexler, When a Computer Program Keeps You in Jail, *The New York Times*, 13 June 2017, <https://www.nytimes.com/2017/06/13/opinion/how-computers-are-harming-criminal-justice.html>(최종방문일 2024. 2. 11.). 로드리게스는 동일한 오류로 다시 평가를 받은 다른 재소자의 재범위험성 점수가 유의미하게 하향되었음을 알게 되었으나, 입력 값의 가중치가 영업비밀이어서 해당 오류가 자신의 재범 위험성에 증대한 영향을 미쳤다는 점을 증명할 수 없었다. Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, *Stanford Law Review*, Vol. 70, No. 5, 2018, pp. 1370-1371.

296) Bernd Carsten Stahl, Doris Schroeder and Rowena Rodrigues, *Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenge*, Springer, 2022, p. 12.

도출하도록 하는 학습과정이라는 것이다. 오늘날 인공지능은 법률가의 사고과정을 모방한 알고리즘이라는 아이디어를 포기했고, 데이터의 무차별 대입 처리를 활용하는 양적 접근 방식으로 전환되었다고 한다.<sup>297)</sup>

그에 따르면, 영업비밀로 보호되는 소스코드의 공개에 집착하는 것은 기본적인 이해 부족이 초래한 잘못이다. 1995년에서 2000년 사이에 개발된 COMPAS의 인공지능 알고리즘은 대단할 것이 없고, 중요한 것은 오히려 훈련 데이터라고 한다.<sup>298)</sup> 또한 사기업 제품으로서 그 소스코드가 영업비밀로 보호된다는 것을 문제 삼는 것도 적절치 않다고 주장한다.<sup>299)</sup>

최근에는 투명성의 무제한적인 강요는 부적절하다고 하면서, 차별 등 인공지능이 초래하는 법적 문제에 대응하려는 시도가 나타난다. 우리 사회는 이미 증명책임(burden of proof) 법리를 통해 인간 행동의 불투명성 문제를 해결해 왔다는 것이다. 이 연구는 가령 민사상 손해배상 소송의 경우에도, 관련법에 명시된 증명책임 전환 규정을 활용한다면 부분적으로라도 인공지능의 차별 문제를 해결할 수 있다고 주장한다. 결론적으로 인공지능이 불투명하다고 하여, 인공지능의 차별 문제가 재판으로 구제될 수 없는 근본적인 한계라고 단정하기 어렵다고 한다.<sup>300)</sup>

#### 다. 잘못된 판단과 책임 귀속 문제

인공지능의 행위로 인한 책임을 누구에게 귀속시킬 것인가가 간단치 않다는 점도 문제다. 인공지능 개발자, 인공지능 알고리즘을 구동한 자에게 책임이 있다고 볼 여지도 있지만, 인공지능 그 자체에 책임을 귀속시켜야 한다는 입론도 가능하며, 현실적으로 누구에게도 책임을 묻기 어렵다는 관점도 있을 수 있다. 인공지능 법관의 잘못된 의사결정으로 기본권 침해(가령 실제로 범죄를 저지르지 않았는데 유죄판결을 받는 경우)가 발생하

297) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 344쪽.

298) 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 345-346쪽.

299) 정부가 쓰는 모든 물건을 직접 제조하거나 정부가 사용하는 모든 소프트웨어의 저작권을 가져야 하는 것이 아니라고 한다. 마이크로소프트의 오피스 소스코드는 영업비밀로 보호되는데, COMPAS의 소스코드를 공개해야 한다는 것은 차별이며 시비라고 한다. 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률연구재단, 2022, 329쪽.

300) 고학수·정해빈·박도현, 인공지능과 차별, 저스티스, 통권 제171호, 2019, 236-240쪽.



더라도, 인공지능 법관이 학습한 방대한 빅데이터 가운데 어떤 자료에 편향성이 있었는지, 은닉층을 포함하는 딥러닝 과정의 어떤 부분에서 잘못된 출력 값이 도출되었는지 추적하기란 대단히 어렵기 때문이다. 오늘날 인공지능 알고리즘 대부분은 오픈소스 라이브러리를 활용하여 만들어진다는 점을 생각한다면, 책임을 귀속시킬 설계·제작자의 개념 자체도 그리 명확하지 않을 수 있다.

또한 학계에서는 기본권 보호의무와 관련하여, 종래 문제되는 것은 인간의 행위에 귀속되는 위험이었으나 인공지능의 경우 인간 아닌 실체에 의한 기본권 침해가 문제된다는 지적도 제기된다.<sup>301)</sup>

다만 인공지능 법관이 공적 기관인 법원에 소속되어 있으므로,<sup>302)</sup> 중국적으로는 국가가 인공지능 법관으로 인한 책임을 부담할 가능성이 크다는 견해가 유력하다. 그에 따르면 국가는 사고의 빈도수, 피해규모 등에 따라 국가예산으로만 처리하거나, 보험가입 내지 기금 조성을 통해 위험에 대비할 것이라고 한다.<sup>303)</sup>

## 라. 기본권 침해 주장 및 증명이 어려워질 가능성 및 적법절차 관련 문제

### 1) 기본권 침해 주장 및 증명이 어려워질 가능성

미국 수정헌법 제4조는 영장이 상당한 이유(probable cause)에 근거하여야 한다고 규정한다. 상당한 이유를 요구함으로써, 동조는 영장발부를 적절히 제한하며 시민들의 재산과 프라이버시가 (체포·압수·수색으로부터) 안전하리라는 점을 확신시킨다. 한국 형사소송법은 구속의 사유로서 피고인이 죄를 범하였다고 의심할 만한 상당한 이유가 있어야 한다고 규정한다(제70조 제1항). 그리고 구속사유를 심사함에 있어 재범의 위험성을 고려하도록 규정한다(동조 제2항). 영장에 의한 체포의 경우에도 피의자가 죄를 범하였다고 의심할 만한 상당한 이유가 있어야 하며(제200조의2 제1항), 구속의 경우에도 상당한 이유를 요한다(제201조 제1항).

301) 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제 8권 제2호, 2021, 48쪽. 인공지능이 자율성을 가지며 의사결정과정의 투명성이 결여되어 있어, 사업자의 행위로 당연히 귀속시킬 수 없게 됨을 환기하고 있다.

302) 자연인 법관과 유사하게 재판을 주재하는 로봇 법관이 아니라 인공지능 알고리즘이라고 하더라도 그것이 법원에 소속되어 작동한다는 점에서는 다르지 않을 것이다.

303) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 291쪽.

그런데 인공지능 법관을 비롯하여 형사사법 시스템에서 인공지능의 판단에 의존하게 될 경우, 가령 인공지능의 위험성 판단결과가 수사기관의 합리적 의심을 뒷받침함으로써 종전에는 정당화되기 힘들었던 검문·수색이 정당화될 여지가 있다. 이러한 검문·수색이 수정헌법 제4조에 반한다고 단정하기 어려워질 수도 있다.<sup>304)</sup> 인공지능의 판단에 입각한 체포, 구속에 대하여도 ‘상당한 이유’가 없음을 다투기가 더욱 힘들어질 수 있다. 인공지능에 관한 지식의 부족, 알고리즘의 복잡성, 알고리즘 비공개 등으로 말미암아, 피고인 측에서 인공지능이 죄를 범하였다고 의심할 만한 상당한 이유를 어떻게 판단하였는지 알기 어렵기 때문이다. 법관이 인공지능의 분석을 토대로 재범위험성이 높다고 판단해 취업제한, 신상공개 등의 보안처분을 부과한 경우에도, 그 부과 of 당부를 다투기가 어려워질 수 있다.

## 2) 적법절차와 관련한 문제

인공지능의 블랙박스वाद도 같은 속성 및 설명가능성 문제로 인하여, 최근 학계와 실무에서는 적법절차 관련 이슈가 활발하게 논의되고 있다.<sup>305)</sup>

헌법 제12조 제1항은 법률과 적법한 절차에 의하지 아니하고는 처벌·보안처분 또는 강제노역을 받지 아니한다고 하여 적법절차원칙을 규정하고 있는바, 적법절차원칙은 형사소송절차에 국한되지 않고 모든 국가작용 전반에 대하여 적용된다. 적법절차원칙에서도 출할 수 있는 중요한 절차적 요청 중의 하나로 당사자에게 적절한 고지를 행할 것, 당사자에게 의견 및 자료 제출의 기회를 부여할 것을 들 수 있다. 그러나 적법절차원칙이 구체적으로 어떠한 절차를 어느 정도로 요구하는지는 일률적으로 말하기 어렵고, 규율되는 사항의 성질, 관련 당사자의 사익, 절차의 이행으로 제고될 가치, 국가작용의 효율성, 절차에 소요되는 비용, 불복의 기회 등 다양한 요소들을 형량하여 개별적으로 판단할 수 밖에 없다.<sup>306)</sup>

적법절차원칙은 설령 동일한 결과에 이르더라도 그 과정에서 당사자에게 공정한 절차와 기회를 부여할 것을 요구한다. 가령 형사재판에서 피고인은 자신이 원하는 결과를 얻

304) 이러한 지적으로 이지만, 형사사법에서의 인공지능 예측프로그램 활용 제언: 인종적 편향 가능성 회피를 중심으로, 원광법학, 제39권 제2호, 2023, 67쪽. 적법절차조항이 더 약화되는 결과를 초래할 수 있다고 한다.

305) 가령 Forrest, Katherine B., AI시대의 사법정의, 패러다임의 전환, 사법의 대응, 2023 사법연수원 국제 콘퍼런스 자료집, 사법연수원, 2023, 40-41쪽.

306) 헌재 2020. 2. 27. 2015헌가4, 판례집 32-1상, 1, 10.

지 못하더라도, 법관 앞에서 의견을 진술하고 이의를 제기할 기회를 가짐으로써 결과에 영향을 미칠 수 있었다고 느낀다. 그런데 그 작동원리를 알 수 없는 인공지능 법관이 판결을 내리거나, 인공지능의 판단에 기초하여 법관이 판결할 경우 피고인은 진정한 의사 결정과정에 참여한다고 느끼기 힘들다.<sup>307)</sup> 일련의 과정이 적법한 절차(due process)의 요구를 충족하는지 여부를 측정할 수 없기 때문에, 다수의 법학자들은 블랙박스와도 같은 알고리즘에 의지하는 국가권력의 판단이 그 자체로(per se) 절차적 적법절차원칙에 반한다고 본다.<sup>308)</sup>

전술하였듯 미국 위스콘신 주 대법원은 State v. Loomis 판결에서, 인공지능 알고리즘에 대한 확인 혹은 이의제기 기회가 주어지지 않아 적법절차에 따른 권리가 침해되었다는 주장을 받아들이지 않았다. COMPAS가 집단적 측면에서 위험성을 예측하는 도구이지만, 양형의 결정적 요소가 아니므로 개인별 양형 선고를 받을 적법절차 권리가 침해되지 않았다고 본 것이다. 하지만 당사자의 입장에서는, 법관이 인공지능의 판단을 단지 참고하였는지 전적으로 수용하였는지 여부를 확인할 길은 없다. 위스콘신 주 대법원의 판결을 반대해석하면, 인공지능의 판단(재범위험성 예측)이 양형을 결정짓는 요소였다면 적법절차원칙 위반이라는 평가가 가능하다.

## 마. 개인정보침해

오늘날 인공지능은 빅데이터를 학습하여 추론의 정확도를 향상시킨다. 인공지능 법관이 정확한 판결을 할 수 있으려면, 종전의 방대한 사건들에 관한 데이터(판결·결정문, 관련 증거 등)가 확보되어야 하는데,<sup>309)</sup> 개인정보 침해의 소지가 있다.<sup>310)</sup>

307) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 697쪽. Ric Simmons, p. 1086.

308) Alfred R. Cowger Jr., The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms, Lexington Books, 2020, p. 41.

309) 다양한 리걸테크 서비스의 수준을 향상시키기 위해서는 제한적으로 공개되어온 판결문 등 법률 관련 빅데이터의 공개 문제 해결이 선행되어야 한다는 지적으로 이성진 외, 4차 산업혁명시대 법정정보기술의 현황과 발전방안, 정보화정책, 제28권 제1호, 2021, 14쪽. 강민구 부장판사는 “AI 산업 증진을 위해 판결문의 전면적·실시간 공개, 익명화 최소 원칙이 입법을 통해 도입돼야 한다.”고 하였다. 홍윤지, [인터뷰] ‘디지털전도사’ 강민구 고법부장판사, “한국 법조, 생성형 AI 기술 신속히 도입해야”, 법률신문, 2023. 5. 31, <https://www.lawtimes.co.kr/news/187980>(최종방문일 2024. 2. 11.).

310) 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 - 보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 59쪽. 유사한 지적으로 강현구, 인공지능의 법적

최근 학계에서는 입법론적으로 범죄예방·수사·기소·형 집행 등 목적의 정보를 개인정보보호법의 대상에서 제외함과 동시에, 그 정보의 보호, 프로파일링 허용여부 및 기준, 영장주의 적용 여부 등을 명시할 필요가 있다는 주장이 제기된다. 그에 따르면 범죄예측 및 형사사법절차 중에서도 구체적으로 어떤 영역에서 인공지능 알고리즘을 활용하는지에 따라 요구되는 데이터의 범위가 달라져야 한다고 한다. 또한 범죄와 관련자의 유형별로, 치안 및 형사사법절차의 단계별로 인공지능 알고리즘의 도입 범위와 수준을 조정하여야 한다고 한다.<sup>311)</sup>

## 바. 인간의 존엄성

근본적으로 인공지능이 자연인 법관의 판단을 대체하고 국민이 인공지능의 판단에 구속되어야 한다면 자율적 존재인 인간의 존엄성에 반하는 것은 아닌지도 문제된다.<sup>312)</sup> 우리의 운명을 기계에 일임하는 것에 대하여, 인간의 존엄과 가치를 우선시하는 생각 때문에 본능적인 거부감을 가질 수 있다는 견해도 개진된다.<sup>313)</sup>

관련하여, 최근 대두되는 ‘인간중심적 인공지능’ 개념을 주목할 만하다. 인공지능은 인간의 자율성을 대체하는 것이 아니라 보완·강화하는 수단이 되어야 하며, 인간은 인공지능을 통해 획득된 정보를 토대로 스스로의 판단과 행동을 할 수 있어야 한다는 것이다. 그리고 이는 자기결정에 바탕을 둔 자기지배로서의 민주주의를 실현하기 위한 본질적 요청이라고 한다.<sup>314)</sup> 인공지능이 공익판단을 위한 자료제공과 중간 판단을 하는 것은 무방

---

정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 51쪽. 참고로 개인정보처리자는 당초 수집 목적과 합리적으로 관련된 범위에서 정보주체에게 불이익이 발생 하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부 등을 고려하여 대통령령으로 정하는 바에 따라 정보주체의 동의 없이 개인정보를 이용할 수 있다(개인정보보호법 제15조 제3항).

311) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 146-147쪽.

312) 2020년 과학기술정보통신부가 발표한 「인공지능(AI) 윤리기준」은 사람 중심의 인공지능을 위한 최고 가치인 인간성(Humanity)을 위한 3대 기본원칙으로 인간의 존엄성 원칙, 사회의 공공선 원칙, 기술의 합목적성 원칙을 제시하였다. 과학기술정보통신부, 과기정통부, 사람이 중심이 되는 「인공지능(AI) 윤리기준」 마련, 과학기술정보통신부 보도자료, 2020. 12. 23, <https://www.msit.go.kr/bbs/view.do?sCode=user&mPid=112&mId=113&bbsSeqNo=94&nttSeqNo=3179742>(최종방문일 2024. 2. 11.).

313) 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 609쪽.

314) 윤수정, 인공지능과 민주주의에 관한 소고, 헌법재판연구, 제8권 제2호, 2021, 20-22쪽. 인공지능에 대한 절대적 의존과 자동화된 인공지능 시스템이 결합하면, 인간의 자율성과 자기결정에 근거한 자기지

하나, 인공지능이 공익을 결정하는 것은 주권의 이양(포기 내지 강탈)이어서 현행 헌법질서에서 허용될 수 없다는 견해도 있다.<sup>315)</sup>

## 8. 사법작용의 본질과 인공지능 법관

### 가. 가치판단 및 정성적 판단

재판에서 정량적인 요소만을 고려할 수는 없으며,<sup>316)</sup> 가치판단 및 정성적인 판단이 필요하다. 근본적으로 인공지능이 과잉금지원칙에서의 목적의 정당성·법익의 균형성 심사 와 같은 가치판단을 할 수 있을지, 계량화가 곤란한 판단작용에 적합한지가 문제된다.<sup>317)</sup> 법에서는 ‘야간’, ‘소음’, ‘위험한 물건’처럼 그 내포와 외연이 확정적이지 않은 개념들이 존재한다. ‘음란한’, ‘건전한’, ‘불명예스러운’과 같이, 가치판단에 근거해야만 그 의미를 파악할 수 있는 개념들도 적지 않다.<sup>318)</sup>

그런데 법 해석의 올바른 기준을 정립하고자 하는 시도들은 지금까지 성공하지 못했으며, 이는 여러 해석방법들이 상호 이질적이기 때문이라고 풀이된다. 결국 해석방법론을 통해 불확실성을 떨쳐내기는 힘들다. 그렇다면 법관의 판단은 자신의 결론이 불확실성을 면할 수 없음을 알면서도 스스로의 전인격적인 책임 하에 끝을 내는 것이며, 순수한 앞

---

배의 실현이 요원해진다고 주장한다.

315) 조원용, 대의제 국회 입법과정에서 인공지능(AI)의 역할: 공익의 ‘발견’과 ‘결정’을 중심으로, 가천법학, 제11권 제3호, 2018, 65쪽. 인공지능 재범위험성 평가도구가 지금보다 훨씬 정확성을 향상시킬 경우 법관이 그 평가결과를 의무적으로 참고하도록 하는 것이 타당하나, 그 판단결과에 구속되도록 하는 것은 적절치 않다는 지적으로 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 -보안처분 선고절차에서의 적용가능성을 중심으로-, 형사정책, 제34권 제1호, 2022, 69-70쪽.

316) 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 609쪽. 인간에 대한 의사결정에서 정량요소만 고려해서는 안 된다고 한 양종모, 법 분야 위협의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 390쪽도 참조.

317) 법률가가 다루는 개념 상당수는 계량화가 곤란하며, ‘상당한 이유’, ‘합리적인 의심’, ‘법익의 균형’ 등을 수치화해 컴퓨터에 집어넣을 방법이 잘 떠오르지 않는다는 지적으로 유형용, [칼럼] AI법관의 도래, 법률저널, 2016. 1. 15. 이론적으로는 특정한 결론과 특정한 사실 간의 상관관계를 설명하는 함수를 만들 수 있다면 거기에 변수를 집어넣고 확률을 계산할 수 있겠으나, 지금까지 집적된 무수한 판결에 나타난 사실관계를 하나하나 분석해 적절한 변수를 할당하고 회귀분석을 하는 것이 가능할지의 문제를 지적하고 있다.

318) 상세히는 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 77쪽.

의 문제라기보다 실천이 함께 요구되는 활동이다.<sup>319)</sup> 지성과 경험을 신뢰하는 법학자들은 법에서 중요한 것이 논리가 아니라 법관의 가치판단과 경험이라고 강조해 왔으며, 현대 법철학 또한 법언어의 불확정성·모호성으로 인하여 법해석이 법관의 재량과 창조적 활동의 영역이라는 논점에서 출발한다.<sup>320)</sup>

입법과정에서의 인공지능 활용에 관한 선행연구 가운데에는, 사회과학적 분석과 전문가의 견해를 통해 공익을 ‘발견’하는 법률안의 경우 인공지능을 효율적으로 활용할 수 있으나, 상이한 이해관계 내지 가치관의 대립 속에 공익을 ‘결정’해야 하는 법률안의 경우 인공지능의 역할이 제한적이라고 파악한 것이 있다.<sup>321)</sup>

## 나. 인과관계와 상관관계

상관관계와 인과관계는 다르다. A와 B가 항상 동시에 존재한다 하여 A가 B의 원인이라고 단정할 수 없다. 가령 적당량의 음주를 하는 사람이 건강이 좋다는 연구결과가 나오더라도, 술이 건강의 원인이라고 단언할 수 없다는 것이다.<sup>322)</sup> 그런데 민사, 형사를 비롯한 재판에서는 주로 ‘인과관계’가 문제되지만, 적어도 현 단계에서의 인공지능은 막대한 데이터를 토대로 패턴을 분석하며 상관관계 머신의 성질을 가지므로<sup>323)</sup> 재판의 본질에 부합하지 않는다고 볼 여지가 있다.<sup>324)</sup>

---

319) 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 80쪽.

320) 박상철, 인공지능 등 정보기술을 통한 법의 기술적 구현 과정의 혁신, 정보법학, 제22권 제2호, 2018, 147쪽.

321) 조원용, 대의제 국회 입법과정에서 인공지능(AI)의 역할: 공익의 ‘발견’과 ‘결정’을 중심으로, 가천법학, 제11권 제3호, 2018, 71-75쪽.

322) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 62쪽. 재범위험성 예측 알고리즘은 데이터 사이의 패턴이나 상관관계를 알려주지만 인과관계를 확정하기에 근본적으로 불충분하다고 한 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 675쪽도 참조.

323) 인공지능은 문제해결과정에서 철저히 상관관계에 주목할 뿐, 인간이라면 당연히 중시할 인과관계에 대한 고려가 없다는 지적으로 양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위협 예측, 흥진기법률연구재단, 2022, 182쪽.

324) 책임소재를 증명하기 위해서는 counterfactual(but-for-test)의 문제, “가령 A가 B를 밀지 않았으면, B는 죽지 않았을까?”와 같이 실제 사실과 다른 세계를 가정해 문제를 해결해야 한다. 그런데 법률 분야에서는 인과관계를 규명하기 위해 통제된 실험(의약품 효능 평가를 위해 실험군과 대조군을 두고 하는 실험 등)을 하기가 매우 힘들다. 김병필, 재판 예측 인공지능 기술의 현황과 한계, 저스티스, 통권

오늘날 기계학습 알고리즘은 이미지 파일을 읽어 동물의 이름을 맞추고, 필적을 읽어 알파벳을 맞추며 뛰어난 예측력으로 세상을 놀라게 한다. 그러나 오늘날 인공지능이 실제 사람처럼 지성으로 고민하여 판단하는 것은 아니며, 통계모델을 만들어 주어진 데이터에 적용함으로써 예측을 하는 것이다.<sup>325)</sup> 머신러닝 인공지능은 논리적 방식으로 결과를 도출하지 않으며, 알고리즘 설계자는 왜 그러한 결과가 나왔는지 설명할 수 없다는 지적이 제기된다.<sup>326)</sup>

법적 판단과 결론은 사실을 확인한 다음, 거기에 필요한 법을 적용함으로써 이루어진다. 이런 과정을 거친 판결을 받아들일 수 있는 것은 재판의 결과가 정당하다는 근거 때문이며, ‘논증’이 그 결과를 충족시킨다. 이처럼 재판이 논증, 즉 논리적 증명의 결과이기 때문에 정당성을 가진다면,<sup>327)</sup> 빅데이터의 학습을 통해 높은 확률로 무언가를 예측한다고 하여 그것이 재판으로서의 정당성을 담보하기는 어렵다.

## 다. 개별적·구체적 판단과 통계적·집단적 판단

오늘날 널리 활용되는 인공지능은 빅데이터에 기반을 둔 머신러닝 방식이다. 과거의 방대한 데이터로부터 상관관계·패턴을 추출하고, 이를 특정한 사안에 적용하는 형태다. 이러한 방식은 범죄발생지역의 확률적 예측에 입각한 경찰력의 배치와 같은 ‘일반적인 공권력 행사’, ‘미래에의 대응’에는 비교적 적합할지도 모른다.<sup>328)</sup>

하지만 (기본권 침해가 확실시되는 경우에 제기하는 헌법소원, 추상적 규범통제, 향후 발생할 위험을 방지하기 위한 가처분 등은 별론) 재판은 이미 발생한 ‘개별’ 사안에서, 구체적 타당성 있는 결론을 도출하는 것이다. 연령·성별·가정환경·고용상태·재산·종교·

---

제182-2호, 2021, 107-108쪽.

325) 박상철, 인공지능 등 정보기술을 통한 법의 기술적 구현 과정의 혁신, 정보법학, 제22권 제2호, 2018, 141쪽.

326) 이상용, 알고리즘 규제의 원리와 구조, 김진우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 222쪽.

327) 차병직, 검증의 마당, 법률신문, 2023. 1. 5, <https://m.lawtimes.co.kr/Content/Opinion?serial=184254>(최종방문일 2024. 2. 11.).

328) 불심검문에 인공지능 알고리즘의 범죄예측 결과를 적용할 수 있다는 견해로 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에서의 적용에 관한 고찰, 형사법의 신동향, 통권 제51호, 2016, 234-235쪽. 불심검문의 경우 비교적 낮은 정도의 혐의를 요하며, 경찰관의 직감에 의하는 것보다는 빅데이터에 의한 범죄예측이 정확할 수 있다는 점을 언급하고 있다.

범죄전력 등 무수히 많은 인자(因子)들이 각종의 법적 사건의 통계적 경향을 제법 정확하게 드러낼 수는 있겠으나, 그것이 ‘지금 문제되는 그 사건’의 사실관계(가령 특정인이 범죄를 저질렀는지 여부, 특정 계약상 의무의 이행 여부)를 말해주지는 않는다.<sup>329)</sup> 법 분야는 개인적 특성이 중요한데도 인공지능 알고리즘에서는 이것이 무시되며, 정형화된 판단이 이루어질 우려가 있다는 지적이 제기된다.<sup>330)</sup>

선행연구에 따르면 가령 교통사고 사건 판결문을 대량으로 학습시킨 인공지능 알고리즘이 수많은 매개변수에 대하여 비교적 정확한(가령 80~90%) 판단결과를 제시할 수도 있다. 그러나 그러한 알고리즘이 보험사의 보험료 책정이나 사고 시 합의 도출의 자료로 활용될 수는 있어도 판결의 근거로는 쓸 수 없다고 한다. 판결이란 통계적인 답을 도출하는 것이 아니라, 당해 특정사건에서 구체적인 정의를 찾아내는 일이기 때문이다.<sup>331)</sup> 이와 관련하여, 예측이 법을 대체해서는 안 되며 ‘패턴’이 법원(法源)이 되어서는 안 된다는 지적이 있다.<sup>332)</sup>

이러한 문제는 특히 형사사건의 경우에 더욱 간과할 수 없는 것이다. ‘나와 닮은 타인’이 나에게 형벌을 부과하는 것을 정당화해주지 못한다. 형벌이란 피고인이 저지른 행위를 바탕으로, 구체적인 개인의 역사와 맥락을 심리해 그에게 책임을 묻는 과정이기 때문이다. 그런데 빅데이터에 기반을 둔 오늘날의 인공지능 알고리즘은 구조적으로 개인적·구체적인 특성의 고려를 소홀히 할 우려가 있다.<sup>333)</sup>

329) 가령 연령, 성별, 전과, 취업상태, 경제적 여건, 거주지역 등 여러 측면에서 데이터 분석에 의할 때 범죄의 위험성이 대단히 낮은 사람이라도, 문제되는 사건의 범인이 아니라고 단정할 수는 없다.

330) 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 390쪽; 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 609쪽.

331) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 62쪽.

332) 김병필, 재판 예측 인공지능 기술의 현황과 한계, 저스티스, 통권 제182-2호, 2021, 110쪽.

333) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 679-681쪽. 인공지능 알고리즘이 개인의 개성과 특성을 배제한 채 집단의 특성만을 적용한다는 유사한 지적으로 손은지, 인공지능의 헌법적 위험과 대응적 규범체계 정립에 관한 연구, 박사학위논문, 숙명여자대학교, 2023, 62쪽.

다만 최근에는 머신러닝 방식이 아니라 비지도학습과 강화학습을 통해 스스로 학습하는 인공지능이 언젠가는 법관의 업무를 대체할 수 있으리라는 전망도 제기된다. 인공지능망을 트윈 모듈로 연결하여 쌍방을 경쟁시키는 ‘적대적 생성 네트워크(GAN: Generative Adversarial Network)’ 모델에 따라 대립 당사자인 원고-피고를 상징하여, 소송절차에 관한 데이터 학습을 시키고 서로 경쟁시킴으로써 인간 못지않은 인공지능 법관의 출현이 가능하리라고 한다. 전후재, 우리 민사재판의 미래에 관한 소고 - 디스커버리, 온라인 소송, 인공지능(AI)을 중심으로 -, 민사소송, 제27권 제1호, 2023, 137쪽.



## 라. 수량화 문제

인공지능 알고리즘이 일정한 결과를 산출하려면 계량화가 필요하며, 결과 또한 수치로 나타날 수 있다. 그런데 법 분야에서 의사결정에 기여하는 표지 가운데에는 정성적인 부분이 많은바,<sup>334)</sup> 인공지능 법관은 측정 불가능한 것을 측정하려 들 수 있다. 또한 예컨대 인공지능이 신의성실의 원칙, 권리남용의 법리, 공서양속, 조리 등의 추상적인 불확정개념을 해석·적용할 수 있을지 의구심이 존재한다.<sup>335)</sup> 어떤 기준으로 점수화·수치화할 것인지도 문제다. 가령 어떤 사람의 재범위험성이 ‘10’이면, 재범위험성이 ‘1’인 사람보다 10배 더 확률이 높은 것인가?<sup>336)</sup>

증명도를 절대적인 확률 문제로 여겨온 선행연구가 있으나, 최근에는 민사·형사재판에서의 증명도를 수량화(quantification)할 수 없다는 견해가 유력하게 개선된다.<sup>337)</sup> 0과 1로 구성되는 컴퓨터로 구동되는 인공지능 법관은 결국 주어진 자료(증거, 변론)에 입각하여 가령 원고의 주장이 사실일 가능성을 수치화할 것인데, 이것이 증명책임 및 증명도에 관한 현재의 재판실무에 부합하는지는 다소 의문이 있다.

## 마. 승복가능성

### 1) 법관의 결정 정당화의 원천

전술하였듯 법적인 판단은 순수한 앎의 영역이라기보다 실천의 영역이기도 하다. 법관이 내리는 결정을 정당화하는 원천은 법관의 지위(형식적 권한)라기보다 타당한 근거의 제시에 있다.<sup>338)</sup> 법관은 판결에서 최상의 근거가 제시되는 논증을 하여야 하며, 불확실성

---

334) 특정한 상황에서 임대인의 반환청구권 행사가 임차인에게 지나치게 가혹하고 급작스런 시기에 이루어져 신의성실의 원칙에 반하는지 여부처럼 불확정개념의 해석·적용에 있어, 인공지능의 능력이나 역할이 어디까지인지 아직 잘 가늠이 서지 않는다고 한 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 14쪽 참조.

335) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 144쪽.

336) 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 391-392쪽.

337) 김종현, 헌법재판에서의 증명책임, 저스티스, 통권 제183호, 2021, 55-56쪽.

338) 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 83-84쪽. 소수의 사법부 엘리트에 의한 지배를 권위모델만으로 정당화하기란 불가능하다고 한다.

속에서도 정당성을 확보하기 위한 노력은 포기될 수 없다.<sup>339)</sup> 그런데 머신러닝 방식으로 구현되는 인공지능이 제시하는 결론은 ‘지금 이 사건’에 관한 논증이라기보다 과거 빅데이터 분석을 통해 드러나는 확률·패턴이라는 점에서, 정당화의 원천을 가지기 쉽지 않다.

한편 판결과 결정의 영향을 받는 시민의 입장에서는, 성공적인 논증만으로 그 해석이 정당화되지 않는다. 법해석이 주관에 의존할 수밖에 없고 만일 다른 법관의 판단을 받았더라면 결과가 달라질 수 있었음에도, 판단에 수긍하는 데에는 법관에 대한 신뢰가 전제된다. 그 또한 한 사람의 시민인 법관이 동료 시민들에 대하여 가지는 배려와 존중의 원칙, 예의·성실·신중·절제·용기·공정·의사소통능력 등 사법적 덕목을 지향하는 태도가 없다면 종교적인 판단이 정당화될 수 없다.<sup>340)</sup> 인공지능 법관에 반대하는 입장에서는, 사회 구성원들이 알고리즘에 입각한 의사결정권자에게 존경심과 경외감을 가질 수 있을지 의문을 표한다.<sup>341)</sup>

## 2) 당해 사안에 대한 판단이 아닌 과거 데이터에 입각한 판단

당사자 개인의 행위 등 사실관계가 아닌 과거의 빅데이터 및 집단적 속성에 근거하여, 복잡하고 불투명한 알고리즘의 연산을 거쳐 확률론적으로 나타나는 판결·결정에 당사자들이 승복하리라고 기대하기가 어렵다.<sup>342)</sup>

비단 재판의 영역이 아니더라도, 사람들은 ‘지금’ ‘나의’ 문제에 대하여 ‘과거’, ‘집단에 관한’ 자료를 토대로 판단하는 것에 쉽사리 수긍하지 못한다. 가령 실제 수학능력시험을 치르지 않고 나의 과거 여러 시험성적, 지능지수, 연령, 성별 등을 입력해 인공지능

---

339) 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 84-85쪽. 다만 인공지능 법관의 판단이 이론에 근거한 추론인지 여부는 중요하지 않다고 본 연구가 있어 눈길을 끈다. 물리학 이론을 잘 모르는 당구선수가 물리학 법칙을 아는 것처럼 게임을 하듯, 인공지능은 이론에 근거하였는지 여부와 무관하게 똑같이 작동할 것이라는 이유에서다. 서치원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 ‘법의 제국’을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 12-13쪽. 다만 이 연구는 기존 판례에 지속되는 문제점 때문에 인공지능이 자연인 법관을 대체할 수 없다고 본다.

340) 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 97-98쪽.

341) Ray Worthy Campbell, Artificial Intelligence in the Courtroom: The Delivery of Justice in the Age of Machine Learning, Colorado Technology Law Journal, Vol. 18, No. 2, 2020, p. 348.

342) 수학을 기초로 한 확률론과 통계학에 기반을 둔 기계의 결정에 승복할 수 있을지에 의문을 표한 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 609쪽 참조.

알고리즘을 통해 점수를 산출하거나,<sup>343)</sup> 나와 유사한 각종 속성을 가진 집단의 성적에 관한 빅데이터에 입각해 (치르지도 않은 시험의) 성적을 산출한다면 받아들이기 쉽지 않을 것이다.

---

343) 영국에서는 코로나 팬데믹으로 인하여 고교졸업시험이 취소되고, 기존의 데이터(종전 성적)에 근거해 학점을 부여하는 알고리즘을 도입해 성적을 산출했고, 그 결과 원하는 대학 진학에 실패한 사례가 있다고 한다. 상세히는 김희정, 지능정보화 시대의 알고리즘 차별에 대한 법적 소고 - 미국의 알고리즘 차별 사례를 중심으로-, *홍익법학*, 제21권 제3호, 2020, 511-512쪽.

## IV. 인공지능 법관의 도입과 관련한 고려요소 및 제도적 방안

유럽평의회 산하의 사법효율을 위한 유럽위원회는 2018년 12월 3일 사법시스템과 사법 환경에서의 인공지능 이용에 관한 유럽 윤리헌장을 채택하였다. 동 헌장은 사법시스템에서의 인공지능 이용에 있어 5대 원칙을 제시하였다. (1) 기본권 존중, (2) 차별금지, (3) 품질과 보안, (4) 투명성·불편부당성·공정성, (5) 이용자에 의한 통제<sup>344)</sup>가 그것이다.<sup>345)</sup> 이처럼 국제사회에서는 사법시스템에서 인공지능을 활용하는 것을 전제로, 가이드라인 내지 규범을 제시하고자 노력하고 있다. 이하에서는 인공지능 판사(혹은 재판에서의 보조적 활용)가 합헌적이기 위한 요건 내지 원칙, 재판에서 인공지능의 적절한 활용 정도, 선결 과제 등을 살펴본다.

### 1. 사법 분야에서의 인공지능: 고위험 인공지능 시스템으로서 강력한 규제

학계의 논의 및 국내외 제도 동향을 살펴보면, 인공지능 법관, 혹은 그러한 수준에 미치지 못하더라도 사법 분야에서 활용되는 인공지능 시스템은 ‘고위험 인공지능 시스템’으로 분류하여야 한다는 입장이 유력하다.<sup>346)</sup> 재판이 기본권과 전체 법질서에 미치는 중대한 영향을 고려할 때, 이러한 견해는 원론적으로 타당하다고 생각한다.

윤영찬 의원이 대표발의한 알고리즘 및 인공지능에 관한 법률안에 따르면, ‘고위험인

344) 디지털 지배를 어떻게 민주적으로 정당화할 수 있을지가 논의된다. 인간행위의 우위에 의한 정당화 구상에 따르면, 과소금지의 의미에서 인공지능 시스템에서 인간이 원천 배제되어서는 안 된다. 베텔스만 재단(Bertelsmann Stiftung)이 발간한 『알고리즘의 규칙』은, 인간의 제어가능성이 확고하게 될 수 없다면 알고리즘 시스템 도입이 포기되어야 한다고 강변한다. 유럽연합 데이터보호기본명령(DSGVO)은 전적으로 자동처리에 의거한 결정에 따르지 않을 권리를 보장한다(제22조 제1항). 상세히는 김중권, 『인공지능시대 알고크라시(Algocracy)에서의 민주적 정당화의 문제, 법조, 제69권 제5호, 2020, 200쪽.』

345) European Commission for the Efficiency of Justice(CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Council of Europe, 2018, pp. 7-12. 각각의 원칙에 대하여 상세히는 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리헌장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 『저스티스, 통권 제172호, 2019, 47-50쪽.』

346) 가령 김일우, 『고위험 인공지능시스템의 기본권 침해에 관한 헌법적 연구, 박사학위논문, 서강대학교, 2023, 200쪽.』 피고인의 민감한 개인정보를 수집·분석하여 범죄의 위험성을 평가하는 프로파일링에 관한 인공지능 시스템의 활용에는 더욱 신중해야 한다고 덧붙인다.

공지능’이란 국민의 생명, 신체의 안전 및 기본권의 보호에 중대한 영향을 미치는 인공지능을 말하는데, 여기에는 ‘수사 및 기소 등 기본권을 침해할 수 있는 국가기관의 권한 행사에 이용되는 인공지능’이 포함된다.<sup>347)</sup> 유럽연합 집행위원회의 인공지능 법안은 인공지능시스템을 위험성의 정도에 따라 분류하는데, 형사사법영역에서의 재범예측시스템이나 법집행기관의 사실판단, 관련법조 적용 등에 활용되는 인공지능시스템은 고위험으로 분류된다.<sup>348)</sup>

## 2. 사전규제의 중요성과 규제방안

인공지능 알고리즘의 법적 의사결정이 광범위한 영향력을 가지며, 편향성 및 차별여부의 확인이 어렵거나 불가능하다는 이유에서, 사후규제가 부적절하며 사전규제에 주안점을 두어야 한다는 견해가 있다.<sup>349)</sup>

법치주의적인 관점에서, 인공지능 법관의 알고리즘을 인간이 설계하거나 적어도 인간이 감수해야 한다고 요구할 수 있다.<sup>350)</sup> 그런데 혹자는 이처럼 인간의 개입을 요구하는 접근방식은 ‘단지 누구라도 인간이 개입되어 있다면’ 법의 지배 요건이 충족되는 것처럼 보이게 함으로써, 법치관념이 다시 인치 관념으로 회귀하는 반작용을 초래할 여지가 있다고 본다.<sup>351)</sup> 생각건대 인공지능 법관의 알고리즘을 인간이 설계·감수하는 것은 물론, 인공지능 알고리즘에 의한 재판 자체가 법치국가원리 및 적법절차원칙에 부합하는 것이어야 한다.

347) 알고리즘 및 인공지능에 관한 법률안(의안번호: 2113509, 제안일자: 2021. 11. 24, 대표발의: 윤영찬, 2022. 3. 30. 과학기술정보방송통신위원회 상정) 제3조 제5호.

348) Annex III High-Risk AI Systems Referred to in Article 6(2)(제6조 제2항에 관한 고위험 인공지능시스템 부속서 III).

349) 양종모, 인공지능 알고리즘의 편향성, 불투명성이 법적 의사결정에 미치는 영향 및 규율 방안, 법조, 제66권 제3호, 2017, 81쪽 이하. 개발과정의 사전통제 및 실전 투입 전 사용승인, 알고리즘의 분석·평가를 통한 규제, 윤리적 설계 등의 방안을 제시한다.

350) 이병규(2020)는 미국 위스콘신 주의 State v. Loomis 판결과 유럽 일반개인정보보호법(General Data Protection Regulation: GDPR)에서 ‘인간 관여의 원칙’이라는 공통점이 발견된다고 한다. 그런데 최종적으로 인간의 책임 하에 판단된다는 이유만으로, 개인이 알고리즘의 불공정한 평가로부터 자유로울 수 있다는 생각은 지나친 낙관이라고 한다. 일반개인정보보호법은 인간의 관여가 명목적일 가능성을 고려해, 알고리즘 등의 투명성 확보를 중시한다고 한다. 이병규, AI의 예측능력과 재범예측알고리즘의 헌법 문제 - State v. Loomis 판결을 중심으로, 공법학연구, 제21권 제2호, 2020, 184쪽.

351) 남중권, ‘알고리즘의 지배’와 ‘법의 지배’에 대한 법이론적 이해, 법과사회, 제71호, 2022, 312쪽.

### 3. 사법데이터의 공개와 개인정보보호

인공지능 알고리즘의 불투명성 및 편향성과 관련하여, 학계에서는 인공지능 알고리즘의 학습에 필요한 사법데이터를 공개해야 한다는 견해가 유력하며 국제사회에서도 이러한 방향으로의 규제 움직임이 감지된다. 가령 유럽연합 인공지능법안에 따르면, 고위험 인공지능 시스템<sup>352)</sup>은 사용자가 시스템의 도출결과를 해석하고 적절히 사용할 수 있도록 설계·개발되어야 한다(제13조 제1항). 고위험 인공지능 시스템의 경우, 적절한 디지털 형식 또는 사용자가 접근하고 이해할 수 있는 완전하고 정확하며 명확한 관련 정보 (concise, complete, correct and clear information)가 제공되어야 한다(제13조 제2항).

사법 데이터는 인공지능 알고리즘이 처리할 수 있는 원 데이터(raw data)로 제공되어야 할 것이다. 재판의 심리와 판결을 공개하도록 하는 헌법 제109조의 문언, 데이터 일부만을 공개할 경우 인공지능 학습에 필요한 정보량이 적어지고 데이터가 왜곡될 우려를 고려할 때, 판결은 망라적으로 공개하여야 한다.<sup>353)</sup>

한편 인공지능 알고리즘 개발에 필요한 사법 데이터를 사법부가 독점해서는 곤란할 것이다. 그렇게 될 경우, 당해 인공지능 알고리즘의 판단결과가 타당한지를 검증하고 반박할 방법이 없게 되어 절차권이 중대하게 침해될 수 있기 때문이다.<sup>354)</sup>

다만 상당부분 당사자가 작성한 것인데다, 당사자가 그 공개에 동의한 바 없는 소송기록의 공개에는 신중을 기해야 한다.<sup>355)</sup> 프로파일링의 위험을 고려할 때, 일반인 전체에 무조건적으로 전산 처리가 가능한 형태의 데이터베이스를 공개하는 데에는 신중하여야 할 것이다.<sup>356)</sup> 또한 법관, 전문가 성명을 공개할 경우, 이들의 성향을 분석하거나 소위

---

352) 전술하였듯 여기에는 재범예측시스템이나 법집행기관의 사실판단·관련법조 적용 등에 활용되는 인공지능시스템이 포함된다.

353) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 55쪽; 고유강, 법률인공지능이 변화시킬 재판의 모습: ‘인공지능 판사’의 기술적 대체가능성과 관련된 현황을 중심으로, 재판자료집, 제142집, 2021, 26쪽.

354) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 56쪽.

355) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 55쪽. 유사한 입장으로 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 11쪽.

356) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차

‘포럼 쇼핑’이 초래되고 법관 개인에 대한 공격이 이루어질 수 있음을 고려해야 한다.<sup>357)</sup>

## 4. 설명 가능한 인공지능 알고리즘

### 가. 설명 가능한 인공지능 논의의 배경

인공지능의 작동 기제(입력된 데이터로부터 왜, 어떻게 특정 결과를 얻었는지)를 인간이 제대로 이해하지 못하는 블랙박스 속성이 문제된다. 인공지능이 인간의 의사결정을 보조하는 수준이라면 그로 인한 문제가 크지 않겠으나,<sup>358)</sup> 인간이 인공지능에 중요한 의사결정 자체를 위탁할 여지가 있다.<sup>359)</sup>

그런데 인공지능이 복잡하고 강력할수록 불투명성이 커지며, 전문가 심지어 개발자조차 정확하고 구체적인 작동 기제를 이해하기 어려울 수 있다. 이에 인공지능의 의사결정이 부당하다고 생각해도 이의를 제기하기 어려운 문제가 생긴다.<sup>360)</sup>

이와 관련하여, 설명 가능한 인공지능(explainable AI, XAI) 개념이 논의된다.<sup>361)</sup> 설명

---

에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 57쪽.

357) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 59쪽.

358) 인공지능에 의한 재범의 위험성 판단이 지금보다 훨씬 정확해질 경우 법관이 그 판단결과를 의무적으로 참고하도록 하되, 판단결과가 법관을 구속해서는 안 된다는 견해로 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 - 보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 69-70쪽. 법관의 독립성, 법관에게 재판을 받을 권리의 문제, 인공지능마다 판단결과가 다를 가능성, 설계자나 회사끼리의 담합 가능성 등을 이유로 제시한다.

359) 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 209-210쪽.

360) 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 218-219쪽. 형사사건에서 변호인이 증인의 시력을 문제 삼아 목격 자체의 신빙성을 탄핵하는 것은 어렵지 않으나, 복잡한 수학적 알고리즘의 산출결과를 탄핵하기란 어렵다는 지적으로 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에의 적용에 관한 고찰, 형사법의 신동향, 통권 제51호, 2016, 231-232쪽.

361) 인공지능 기술의 맥락에서 설명가능성에 관한 연구는 1970년대 중반부터 진행되었으며, 설명 가능한 인공지능이라는 표현은 2004년 도입되었다. 하지만 설명 가능한 인공지능에 관한 관심이 크게 제고된 것은 최근의 일이다. 2012년 당시 설명가능성을 다룬 문헌은 50개 미만이었으나, 2019년에는 설명 가능한 인공지능을 다룬 출판물이 300건을 상회하였다. 하지만 설명 가능한 인공지능의 개념이 아직 애매한 상태로 남아 있다는 지적이 제기된다. 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 251쪽; Charles Kerrigan and Oliver Vercoe, Explainable AI and Responsible AI, in Charles Kerrigan ed., Artificial Intelligence: Law and Regulation, Edward Elgar,

가능성과 투명성 개념이 완전히 동일한 것은 아니라는 견해가 있다. 투명성이 정보 획득 차원에서 나타나는 정보의 비대칭성 교정에 관한 개념이라면, 설명가능성은 정보처리과정에서 발생하는 인지적 비대칭성의 교정에 초점을 맞춘 개념이라는 것이다.<sup>362)</sup>

다만 오늘날 각국의 규제 논의에서, 인공지능 알고리즘의 불투명성 문제는 정보 획득 차원으로 국한되지는 않는다. 특정한 결과 값이 어떻게, 왜 산출되었는지 파악하기 힘들 때에도 인공지능의 불투명성 문제가 언급된다. 그리고 인공지능 작동기제의 불투명성으로 초래되는 문제의 대안으로, 설명 가능한 인공지능이 떠오르고 있다. 인공지능의 의사 결정과정을 설명할 수 있다면 그 판단결과에 대하여 다툼 길이 생기며, 근본적인 원인을 지적해 해당 코드를 수정하거나 문제의 데이터를 제거함으로써 편향성으로부터 보호될 수 있다고 주장된다. 또한 오류의 원인을 제공한 자에게 책임을 추궁하고, 헌법적 가치인 투명성과 절차적 정당성, 인간의 존엄성을 충족할 수 있다고 한다.<sup>363)</sup>

## 나. 설명 가능한 인공지능의 문제점

그런데 설명 가능한 인공지능에도 여러 문제점이 있다. (1) 소스코드 등을 공개하도록 한다면 기업의 영업비밀 공개를 강요하는 것으로 평가될 수 있다.<sup>364)</sup> (2) 그리고 일반인은 물론 전문가, 개발자조차 인공지능의 구체적인 작동 원리를 파악하기 힘든 경우가 있는바, 소스코드 등을 전면 공개하더라도 일반인의 입장에서 불투명성이 해소되었다고 볼 수는 없다. (3) 그렇다고 설계 단계에서부터 복잡성을 줄여 설명 가능한 단순모델 알고리즘을 구현할 경우, 정밀한 기제를 통해 방대한 데이터를 분석함으로써 정확한 판단을 하는 머신러닝의 장점을 포기하는 셈이 된다.<sup>365)</sup> 다시 말해, 현 단계에서 머신러닝으로 개

---

2022, p. 528.

362) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 233-234쪽.

363) 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 220-224쪽.

364) 2019년 3월 미국 아이다호 주 의회가 제정한 형사사법절차에서의 위험성평가 알고리즘 사용에 관한 법은 지적재산권에 기한 비공개 항변을 배척하고, 위험성평가도구에 쓰이는 모든 문서, 데이터기록 및 정보공개를 요구한다. 그러나 기업과 이해당사자의 재산권을 중대하게 제한하여 상당한 논란이 예상되며, 미국 전역에서 이러한 제도가 도입될 수 있을지에 회의적인 시각이 많다고 한다. 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 926-927쪽.

365) 재범의 위험성을 예측하는 COMPAS의 판단은 법률 비전문가 인간의 판단보다 결코 더 정확하지 않다는 연구결과가 있다. 다만 훨씬 더 많은 데이터를 통해 훈련이 가능하며 더 다양한 고려요소를 아우



발된 인공지능 시스템에 설명가능성을 부여하기 위해서는 효율성을 어느 정도 희생해야 한다는 것이다.<sup>366)</sup> 다만 뒤에서 보는 것처럼, 학계 및 실무에서는 사법절차, 특히 형사사법절차에 있어서는 설명가능성 및 투명성을 정확성보다 우선하는 견해가 유력한 듯하다. (4) 인공지능의 알고리즘이 설명 가능하다고 해도, 입력되는 데이터에 내재한 사회구조적인 편향성이 해결되는 것은 아니다. (5) 인공지능의 작동 기제가 설명 가능할 정도로 투명해진다면, 사용자들에 의하여 악용될 가능성(가령 인공지능이 상관관계를 파악하지 못하도록 조작적 행동을 함)이 있다. (6) 근본적으로 인간의 사고와 의사결정이 반드시 설명 가능한지도 생각해 볼 문제다. 인간은 때로는 언어·논리로 설명이 불가능한 의사결정도 하며, 말도 안 되는 설명을 만들어내기도 한다는 것이다.<sup>367)</sup>

혹자는 이와 관련하여, 법 분야는 생명이나 자유 등에 중대한 결과를 초래할 수 있으므로 설명 가능성에 주안점을 둘 수밖에 없고 따라서 딥러닝 등 첨단기법을 포기해야 한다고 한다. 다소 정확성이나 성능이 떨어지더라도 설명 가능성이 있는 의사결정 트리(tree) 등의 알고리즘이 법 분야와 궁합이 맞다는 것이다.<sup>368)</sup> 법원이 사법절차에서 인공지능을 활용하고자 한다면, 효율성을 다소 희생하더라도 설명가능성을 우선 충족해야 한다는 견해도 같은 맥락이다.<sup>369)</sup> 전반적으로 학계 및 실무에서는 사법절차, 특히 형사사법절차의 경우 설명가능성을 판단의 정확성이나 효율성보다 중시하는 견해가 유력한 듯하다.<sup>370)</sup>

그런데 의사결정 트리 등의 단순한 인공지능 알고리즘은 자연인 법률가들의 업무부담

---

를 수 있고, 최고의 임상전문가가 알고리즘 설계단계에서부터 참여할 수 있으므로, 향후에는 인공지능의 판단이 더 우수해질 것이라고도 전망된다. 그러나 재범의 위험성과 관련한 변수가 워낙 많고, 특정할 수 없는 고려요소가 존재하는데다, 빅데이터 자체가 집적되지 않은 유형의 범죄도 있기 때문에, 정확도를 아주 획기적으로 향상시키는 것은 쉽지 않으리라고 한다. 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 - 보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 56-61쪽.

366) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 67쪽.

367) 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 225-231쪽.

368) 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 231-234쪽.

369) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 67쪽.

370) 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 928-929쪽.

경감이나 부정확한 판단의 문제를 해소하는 데 별반 도움이 되지 못할지도 모른다. 법관에 의한 재판을 받을 권리, 심급제도 형해화 등 여러 문제의 소지를 안고 있는데다 정확성·성능마저 탁월하지 못한 ‘단순한 인공지능 알고리즘’이 어떤 실천적인 의의를 가질 수 있는지 의문이 제기될 수 있다.

## 다. 절충적인 대안의 모색

### 1) 투명성 요건을 충족하는 인공지능 알고리즘

미국의 여러 주에서 공판 전 절차에서의 구속 여부 판단에 활용되고 있는 공공안전평가(Public Safety Assessment: PSA)에서 보듯, 머신러닝 기법을 활용한 인공지능 알고리즘이 투명성 요건을 충족하는 경우도 있다. PSA의 경우 작동방식의 투명성이 상당히 보장되어, 판사가 그 정확성을 심사하고 구속여부 판단에 어느 정도 반영할 것인지 결정할 수 있다고 평가된다. 다만 상식적인 수준의 9가지 요소들만 고려하고 있어 정교한 예측 정확성을 갖추었는지 의문이 제기되기도 한다.<sup>371)</sup>

### 2) 제한된 투명성

최근에는 현실적·제도적 불투명성이 줄어들수록 반드시 이로운지 의문이며 기술적 불투명성의 여러 여건들을 감안해야 하므로, 투명성의 최대화가 아닌 최적화가 요구된다는 견해가 개진된다. 투명성이 기존 강자들에게 보다 많은 정보를 제공하여 권력관계를 공고화하는 결과를 초래할 수 있고, 이용자들이 투명하게 공개된 알고리즘 작동방식을 악용할 가능성이 존재하기 때문이다.<sup>372)</sup>

미국에서는 보다 절충적인 견해로서 ‘제한된 투명성’ 개념이 제시되기도 한다. 자격을 갖춘 소수의 전문가단체만이 위험성 예측 프로그램 작동방식을 검증할 수 있도록 하여, 투명성보장과 영업비밀 보호를 조화시킬 수 있다는 것이다.<sup>373)</sup>

371) 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 934쪽.

372) 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 238-239쪽.

373) 조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 927쪽.

## 5. 보조적 역할에 한정하는 방안

### 가. 보조적 역할만을 부여하는 방안에 관한 논의

최근 ‘인간중심적 인공지능’이 대두되고 있다. 인공지능은 인간의 자율성을 대체하는 것이 아니라 보완·강화하는 수단이 되어야 하며, 인간은 인공지능을 통해 획득된 정보를 토대로 스스로의 판단과 행동을 할 수 있어야 한다는 것이다. 그리고 이는 자기결정에 바탕을 둔 자기지배로서의 민주주의를 실현하기 위한 본질적 요청이라고 한다.<sup>374)</sup> 인공지능이 공익판단을 위한 자료제공과 중간 판단을 하는 것은 무방하나, 인공지능이 공익을 결정하는 것은 주권의 이양(포기 내지 강탈)이어서 현행 헌법질서에서 허용될 수 없다는 견해도 있다.<sup>375)</sup>

그렇다면 인공지능 자체에 법관의 지위를 부여하지는 않고, 자연인 법관이 재판에 있어 보조적인 역할만을 부여하는 방안을 생각해 볼 수 있다.<sup>376)</sup> 주현경·정채연(2020)은 법적 정당성과 규범적 당위성이 계산적 확실성·수학적 합리성에 의하여 대체될 위험을 경고한다. 그에 따르면 인공지능은 (형사)재판에서 보조적 역할만을 수행해야 하며, 인간의 도덕·규범적 판단을 완전히 대체할 수 없다. 인공지능 알고리즘은 규범적 가치체계를 이해하지 못하며, 판단과정을 인간의 언어로써 논증할 수 없는데, 법관이 자신의 판단근거를 소명하는 것이 법치국가의 핵심이기 때문이다. 또한 인공지능 알고리즘의

374) 윤수정, 인공지능과 민주주의에 관한 소고, 헌법재판연구, 제8권 제2호, 2021, 20-22쪽. 인공지능에 대한 절대적 의존과 자동화된 인공지능 시스템이 결합하면, 인간의 자율성과 자기결정에 근거한 자기지배의 실현이 요원해진다고 주장한다.

375) 조원용, 대의제 국회 입법과정에서 인공지능(AI)의 역할: 공익의 ‘발견’과 ‘결정’을 중심으로, 가천법학, 제11권 제3호, 2018, 65쪽. 인공지능 재범위험성 평가도구가 지금보다 훨씬 정확성을 향상시킬 경우 법관이 그 평가결과를 의무적으로 참고하도록 하는 것이 타당한, 그 판단결과에 구속되도록 하는 것은 적절치 않다는 지적으로 장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 - 보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 69-70쪽.

376) 사법 영역에서 인공지능이 수행하는 역할은 법관을 대체하는 것이 아닌 보조하는 것이어야 한다는 주장으로 정채연, 사법절차 및 사법서비스에서 인공지능 기술의 도입 및 수용을 위한 정책 연구, 사법정책연구원, 2021, 94쪽. 이수진(2023)은 가사소송에서 인공지능과 온라인분쟁해결이 이혼 당사자를 대면함으로써 받는 감정적 어려움을 방지하고 사건결과 예측을 통해 합리적인 합의를 도모하게 할 수 있다고 한다. 그러나 인공지능이 법관의 의사결정을 완전히 대체할 수는 없으며, 자연인 법관의 법적 판단을 보조하는 역할에 한정하여야 한다고 주장한다. 이수진, 가사재판에서 인공지능과 온라인분쟁해결의 의의, 사법, 제64호, 2023, 106-113쪽.

판단은 기존의 판결·결정들과 일관성을 가지겠지만 새로운 관점의 변화(판례변경 등)를 기대할 수 없으며, 선례가 내재한 편견과 편향성이 영속화될 우려가 있다. 법관에 의한 재판은 받을 권리, 법관의 독립성을 고려할 때에도, 인공지능 알고리즘이 자연인 법관과 대등한 범위·수준에서 활용되는 것은 무리라고 한다.<sup>377)</sup> 인공지능이 법관 집단의 경향성을 넘어서는 판단을 할 수 없고, 따라서 자연인 법관을 완전히 대체하지는 못하리라는 유사한 전망도 있다.<sup>378)</sup>

하지만 인공지능 알고리즘은 사회구조적인 문제들을 들추어내고, 인간의 부주의·실수를 감소시키며 자의적 판단 기타 재판에서의 부조리(전관예우, 유전무죄 등)를 극복하는 순기능도 가지는데, 보조적인 역할을 할 수 있다. 가령 법관은 자신의 판단이 인공지능 알고리즘의 그것과 불일치하면 숙고하게 될 것이며, 이를 통해 법적 논증의무가 강화될 수 있다.<sup>379)</sup> 인공지능 알고리즘의 검토의견을 당사자가 거부할 경우 참고하지 않거나, 당사자가 원할 경우 참고하되 국민참여재판 배심원의 평결(권고적 효력)과 유사하게 활용할 수 있다고 전망한다.<sup>380)</sup>

## 나. 보조적 역할 한정과 관련한 문제점

그런데 이처럼 인공지능 알고리즘이 자연인 법관을 보조하며 참고자료를 제시할 뿐이라면, 이를 ‘인공지능 법관’이라고 부르는 힘들 것이다. 인간의 정보처리와 인공지능 알고리즘의 정보처리가 서로 복잡하게 얽혀 있어, 인공지능이 현실에서 보조적인 지위에 머물지 않을 위험성도 간과할 수 없다.

### 1) 알고리즘의 인간 의존성

인공지능 시스템은 일정한 목적을 위하여 설계된 프로그램에 따라 효율적·자율적으로

377) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 130-131쪽.

378) 권경휘, 법의 결정성, 인공지능 그리고 법관의 미래, 법학연구, 제32권 제2호, 2022, 359쪽.

379) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 139-140쪽; 권경휘, 법의 결정성, 인공지능 그리고 법관의 미래, 법학연구, 제32권 제2호, 2022, 358-359쪽.

380) 주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 151-152쪽.

결과물을 생성할 뿐이므로, 알고리즘의 작업은 설계자가 의도한 목적에 구속될 수밖에 없다는 기술적 한계를 지닌다. 머신러닝 인공지능의 ‘학습’이란 입력한 데이터로부터 결과를 도출할 때의 가중치를 지속적으로 업데이트하여 실제와 예측의 편차를 줄여나가는 과정일 뿐이며, 학습과정에는 개발자의 설계가 필연적으로 전제된다.<sup>381)</sup>

머신러닝에 기반을 둔 오늘날의 인공지능 알고리즘에서는 데이터의존성이 강화되어, 질 낮은 데이터가 걸러지지 못한다. 또한 인공지능 알고리즘은 특수하고 이례적인 사안에서도 일반적인 판단을 반복하며, 인간이 해온 선례에 갇혀버릴 수 있다. 판례를 변경하거나 법령의 폐지를 검토해야 할 때도 있는데, 인공지능은 자연인 법관이 해온 결정(훈련 데이터)에 의존하며 그로부터 벗어나기 힘들다.

## 2) 인간의 알고리즘 의존성

인간과 인공지능이 팀을 이루어 효과적으로 재판을 수행하기 위해서는, 인간이 기계에 대하여 적절히 조정된 정도의 신뢰를 부여해야 한다는 지적이 제기된다. 만약 인간이 모순된 증거가 존재하거나 명백히 불공정한 결과가 산출됨에도 불구하고 인공지능 시스템을 과도하게 신뢰한다면, 그는 자신에게 요구되는 역할을 수행하지 못하는 셈이다.<sup>382)</sup>

그러나 보조적 역할에 한정하더라도 인공지능 알고리즘을 도입할 경우, 법관의 책임의식이 분산되고 법관이 다른 인지적 노력을 소홀히 할 수도 있다. 또한 기계의 판단을 과대평가하는 자동화편향, 인공지능 알고리즘이 산출한 결과를 강력한 기준으로 삼아 새로운 정보를 수용하지 않거나 미세하게만 조정하는 앵커링 효과(Anchoring effect)가 나타날 수도 있다.<sup>383)</sup> 설령 인간이 의사결정과정의 최종 단계에 자리해 있더라도, 인공지능의 판단이 결정적인 영향을 미친다는 주장도 제기된다.<sup>384)</sup> 인간은 인공지능의 제안을 고려하며 ‘자율적’으로 판단했다고 생각할지 모르나, 실제로는 유도된 선택일 가능성이 높다

---

381) 손은지, 인공지능의 헌법적 위험과 대응적 규범체계 정립에 관한 연구, 박사학위논문, 숙명여자대학교, 2023, 14-15쪽. 인간과 동등한 능력을 가지고 범용으로 활용되는 강한 인공지능, 인간을 뛰어넘는 초인공지능은 아직 현실화되지 않았으므로 차치한다.

382) Rebecca Crootof, “Cyborg Justice” and the Risk of Technological-Legal Lock-in, Columbia Law Review Online, Vol. 119, 2019, pp. 243-244.

383) 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 702-706쪽.

384) 손은지, 인공지능의 헌법적 위험과 대응적 규범체계 정립에 관한 연구, 박사학위논문, 숙명여자대학교, 2023, 273쪽.

고도 지적된다.<sup>385)</sup>

### 3) 인간의 알고리즘 불신 내지 일관성 없는 신뢰도의 부여

이와 반대로 자연인 법관이 인공지능 알고리즘을 불신하고 그 판단결과를 받아들이려 하지 않는다면, 우리는 무용하거나 심지어 유해한 인프라에 과도한 투자를 하는 것일지도 모른다. 자연인 법관이 때로는 알고리즘의 판단결과를 신뢰하여 받아들이고 때로는 불신하여 무시한다면, 그러한 불일치(discrepancy)는 자연인 법관이 홀로 판단하는 경우보다 더 많은 예측 불가능성을 초래할 수 있다.<sup>386)</sup>

## 6. 당사자의 의사를 존중하는 방안

당사자가 반대할 경우 전통적인 방식의 재판(자연인 법관에 의한 재판)을 하는 방안을 생각할 수 있다.<sup>387)</sup> 유럽연합 일반개인정보보호법(GDPR)은 프로파일링 등 본인에 관한 법적 효과를 초래하거나 이와 유사하게 본인에게 중대한 영향을 미치는 자동화된 처리에만 의존하는 결정의 적용을 받지 않을 권리를 규정하는데(제22조), 향후 인공지능이 법관을 대체할 수 있을 정도로 진전될 경우 인공지능 법관의 부당한 판결에 대응할 수 있는 보호방안으로서 중요한 의의를 가질 수 있다.<sup>388)</sup> 다만 이 규정은 자동화된 처리에 ‘만’ 의존하는 결정의 적용에서만 문제되는바, 전체 의사결정과정 가운데 매우 간단한 업무만 자연인을 거치도록 하는 등 법제도를 우회할 가능성이 문제된다.<sup>389)</sup>

2024년 3월 15일부터 시행되는 개정 개인정보 보호법도 자동화된 결정에 대한 정보주

---

385) 손은지, 인공지능의 헌법적 위험과 대응적 규범체계 정립에 관한 연구, 박사학위논문, 숙명여자대학교, 2023, 32쪽.

386) Rebecca Crootof, “Cyborg Justice” and the Risk of Technological-Legal Lock-in, Columbia Law Review Online, Vol. 119, 2019, p. 244.

387) 당사자가 반대할 경우 전통적인 방식에 따른 재판을 하는 것을 인공지능 법관과 관련한 대안의 하나로 제시한 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 64쪽 참조.

388) 이러한 취지의 지적으로 김일우, 고위험 인공지능시스템의 기본권 침해에 관한 헌법적 연구, 박사학위논문, 서강대학교, 2023, 180-181쪽.

389) 인공지능 알고리즘에만 기반을 둔 행위의 대상이 되지 않을 권리(이의권)는 상당히 제한적인 상황에서서만 사용될 수 있으며, 실효성을 확보하기 어려울 가능성이 높다고 한 고학수·정해빈·박도현, 인공지능과 차별, 저스티스, 통권 제171호, 2019, 252쪽 참조.

체의 권리 등을 규정한다(제37조의2). 다만 완전히 자동화된 의사결정 자체를 금지하지는 않으며, 정보주체에게 자신의 권리 또는 의무에 중대한 영향을 미치는 경우에 한하여 이를 거부하고 그에 대한 설명, 인적 개입에 의한 재처리를 요구할 수 있는 권리를 부여한다. 또한 완전히 자동화된 의사결정에만 적용되므로, 일부 인적 개입이 이루어진 경우에도 거부권 등을 주장할 수 없다.<sup>390)</sup>

당사자의 의사를 재판의 어느 단계에서 존중할 것인지와 관련해서는 여러 방안을 생각해볼 수 있다. 유럽연합의 일반개인정보보호법(GDPR) 제22조가 규정하는 취지와 같이, 처음부터 인공지능 알고리즘의 판단을 거부하는 것도 가능하다. 이와 달리 가령 민사소액사건에서 인공지능이 판단을 내리고, 이에 불복하는 당사자가 이의를 신청하는 등의 제도를 구상해볼 수도 있겠다.

## 7. 재판의 영역마다 인공지능 알고리즘 활용 여부 및 정도를 달리하는 방안

### 가. 민사사건, 특히 소액사건의 경우

포괄적·추상적 개념의 화용론적(pragmatic) 해석<sup>391)</sup>과 정치적 고려가 중요한 공법에서 보다, 구체적 법률요건에 관한 의미론적(semantic) 해석<sup>392)</sup>이 중요한 사법(私法)과 형사법에서 정보기술의 응용 가치가 크리라는 예측이 있다.<sup>393)</sup> 한국에서 인공지능의 사법절차 활용 논의는 주로 민사절차와 관련하여 이루어진다.<sup>394)</sup> 주요 국가들은 소액사건의 경우 신속성과 경제성을 추구할 수 있는 다양한 특례를 인정한다.<sup>395)</sup> 당사자가 처분 가능한 영역(예: 이혼, 매매 등 민사적 법률행위), 전체 법질서와는 거리가 있는 영역, 당사자의

---

390) 고환경, 인간 중심 AI와 완전히 자동화된 의사결정에 대한 대응권, 법률신문, 2023. 12. 21, <https://www.lawtimes.co.kr/opinion/194236>(최종방문일 2024. 2. 11.).

391) 전후 관계의 맥락에 따른 해석.

392) 문자 그대로의 의미에 따른 해석.

393) 박상철, 인공지능 등 정보기술을 통한 법의 기술적 구현 과정의 혁신, 정보법학, 제22권 제2호, 2018, 135쪽.

394) 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리현장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 42쪽.

395) 소액사건은 일반적으로 단독판사가 담당하므로 법관들 사이의 합의가 불필요하다는 점도 이러한 논의의 배경으로 판단된다. 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 79쪽.

진술이나 당사자가 제출한 서류에 주로 의지하는 재판에서라면, 인공지능 법관의 도입 혹은 인공지능 알고리즘 판단의 활용도가 더 클 수 있으리라고 생각된다. 관련하여, 민사 소송 가운데 소액사건의 이행권고결정이나 지급명령 사건은 인공지능에 의한 처리가 비교적 용이하다는 평가가 있다.<sup>396)</sup>

선행연구 가운데 매우 구체적으로 소액사건에 인공지능 법관을 도입하는 방안을 제시한 것이 있어 주목된다. 그에 따르면 인공지능은 수단이며 통제의 대상이므로 모든 사건에서 자연인 법관을 대체할 수는 없다. 다만 특정한 유형의 정형화된 사건, 당사자의 합의로 분쟁해결이 가능한 사건의 경우 인공지능 법관이 재판하는 대상사건으로 고려할 수 있다. 현재 소액사건의 기준인 3,000만원은 너무 고액이며, 민사본안사건 중 매우 높은 비중을 차지하므로 보다 사건의 범위를 제한하고 당사자에게 선택권을 부여할 수 있다고 한다.

구체적으로 소가 300만 원 이하 사건은 원칙적으로 인공지능 법관이 재판하며, 소가 1,000만 원 이하 사건은 원칙적으로 자연인 법관이 재판하되 당사자 일방의 신청과 상대방의 동의가 있으면 인공지능 법관이 재판할 수 있게 하자고 제안한다.<sup>397)</sup>

## 나. 헌법재판과 형사재판의 경우

반면 전체 범질서와 관련되며 고도의 가치판단이 개입하는 헌법재판이라든지, 당사자(특히 피의자, 피고인)의 기본권을 크게 제한하는 형사재판에서라면 인공지능 법관의 도입 내지 인공지능 알고리즘 판단결과의 활용에 더욱 신중을 기해야 할 수 있다. 헌법재판의 경우 인공지능 법관에 의한 대체가능성이 더욱 축소된다고 한 선행연구가 있다. 위헌 여부를 판단할 때 과거의 판례나 법률 문언뿐만 아니라 당대 시대상을 반영하여야 하는데, 인공지능이 이러한 역할을 할 수 있을지 회의적이라는 이유에서다.<sup>398)</sup> 인공지능 유

396) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 140쪽. 경매배당이나 개인회생절차도 인공지능에 의한 처리가 가능한 분야라고 한다. 강영주(2022)에 따르면, 민사절차는 사적 자치에 기반을 두고 당사자의 자율적 의사가 중요시되므로 인공지능을 활용하여 분쟁을 해결하는 정책이 더욱 부각된다. 특히 소액사건은 법관들의 합의가 불필요하며, 신속성과 경제성이 보다 중시되므로 인공지능을 도입하여 시범사업을 시도하기에 적합하다고 한다. 강영주, 인공지능을 활용한 민사 분쟁 해결에 관한 고찰, 안암법학, 제64권, 2022, 415, 419쪽.

397) 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 79-80쪽.

398) 손영선, 인공지능 법관의 도입 가능성 검토 - 법관의 독립과 양심을 중심으로 -, 연세법학, 제43호,



립사법윤리헌장은 민감 정보(sensitive data)에 근거한 차별을 방지하고, 공정한 재판을 받을 권리에 부합하기 위해 형사 사안(criminal matters)에서 인공지능의 활용에 극히 신중해야 한다는 입장이다.<sup>399)</sup> 혹자는 선례가 거의 없는 사건이나 새로운 쟁점이 나타난 사건의 경우에도, 귀납적 판단에 전적으로 의지하는 인공지능보다는 자연인 법관에게 맡길 가능성이 높다고 전망한다.<sup>400)</sup>

인공지능 법관의 출현 가능성을 사안에 따라 달리 접근하는 탄력적인 선행연구가 있어 주목된다. 인공지능이 법관의 재판영역을 전부 대체하는 것은 아직 불가능하며 보조 내지 제한적 실현만이 가능하지만, 요건이나 쟁점이 단순한 사건들에 있어서는 자연인 법관의 역할을 대체할 가능성이 충분하다는 것이다.<sup>401)</sup> 단순하고 정형화되어 있어 인공지능이 법관의 판단을 대신하도록 점차 보완해나갈 수 있는 분야로는 교통사고, 음주운전, 과태료 처분, 범칙금, 경범죄 사건 등을 들고 있다.<sup>402)</sup>

---

2023, 62-63쪽.

399) European Commission for the Efficiency of Justice(CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Council of Europe, 2018, p. 5. 한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리헌장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 47쪽. 특히 사형이나 장기 징역형 등 중형을 부과할 가능성이 높은 형사사건의 경우에도 자연인 법관에게 그 결단을 맡기려는 경향이 높으리라고 한 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 288-289쪽도 참조.

400) 오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022, 288쪽.

401) 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 87-88쪽.

402) 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 75-76쪽.

## V. 결론

오늘날 인공지능은 사법 분야에서 널리 활용되며 인공지능에 의한 법관의 대체 가능성도 논의되고 있다. 유럽연합의 인공지능법안이 그 동안의 꾸준한 연구에 기초하였음을 고려할 때, 가까운 시일에 자연인 법관과 동일한 수준의 인공지능 법관이 현실화되지는 않더라도 헌법재판소 및 헌법재판연구원 차원에서 선제적인 검토가 필요하다.

최근 논의되는 인공지능 법관은 그 개념에서부터 적잖은 논쟁거리를 안고 있다. 인공지능이란 무엇이며 그 특징이 어떠한지 부터가 답하기 쉽지 않은 문제이다. 이 연구보고서에서는 인공지능이 인간의 두뇌와 지적 능력을 그대로 구현하는 데 방점이 있지는 않으나 인간이 지능으로써 수행해온 것들을 수행하기 위하여 고안되었다는 점에 착안하였다. 현 단계에서 자연인 법관과 전적으로 동일한 인공지능 법관을 상정하기는 어렵겠으나, 인공지능 알고리즘이 제시하는 출력 값이 자연인 법관의 판단을 대체하는 단계에 이르러야 인공지능 법관이라고 할 수 있다고 보았다. 그리고 그 특징으로는 알고리즘의 편향성, 불투명성, 귀납적 추론 방식, 데이터 기반성, 인간의 의사 작용 배제를 제시하였다.

인공지능 법관은 업무 효율성 제고, 재판의 일관성과 공정성에 대한 기대에서 논의되나, 법치주의, 민주주의를 비롯한 헌법적인 문제들로부터 자유롭지 못하다. 재판의 결론을 도출하는 인공지능 알고리즘을 ‘헌법이 정한 법관’이라고 할 수 있는가? 인공지능 법관이 판단을 내리는 데 필요한 자료(훈련데이터뿐만 아니라 소송기록 등)가 자연인에 의하여 선택적으로 입력된다면, 그 양심에 따라 독립하여 심판하는 헌법상 법관이라 할 수 있는가? 인공지능 법관은 직접적·간접적으로 민주적 정당성을 부여받은 존재라고 할 수 있는가?

인공지능 법관은 기본권의 침해와 관련해서도 여러 어려운 쟁점을 지닌다. 판결기계에 의한 재판이 공개재판을 받을 권리를 침해하지는 않는가? 컴퓨터가 읽을 수 있도록 코드화한 법률을 사안에 적용하여 결과를 산출한 것이 법률에 의한 재판인가? 결론이 어떻게 도출되었는지 알 수 없다면, 가령 형사 피고인의 방어권이 보장되지 못해 공정한 재판을 받을 권리 침해가 아닌가? 데이터와 알고리즘의 편향성으로 말미암아 평등권이 침해될 우려는 없는가?

인공지능 법관은 우리 헌법이 상정해온 사법제도의 틀과 본질에 관한 근본적인 성찰을 요하는 문제이기도 하다. 인공지능은 여러 법관이 합의체를 구성하여 다양한 의견을

개진하며 논박하는 재판의 지형을 근본적으로 바꾸고, 심급제를 형해화할 우려가 없는가? 사회현실의 변화를 감지하고 새로운 결론을 도출하지 못한 채 기존 판례에 갇힐 가능성은 없는가?

현재의 약한 인공지능은 빅데이터를 학습하여, 통계적 상관관계 및 패턴을 파악해 특정 문제를 예측하는 것이다. 인간의 지능을 동일하게 혹은 그 이상으로 구현하는 강한 인공지능, 초인공지능이 등장하고 이를 헌법상의 법관으로 인정하는 것은 별론으로 하고, 현재의 머신러닝·딥러닝 인공지능이 ‘사법작용의 본질에 비추어’ 법관이 될 수는 없다고 생각한다. 오늘날의 약한 인공지능은 가치판단 및 정성적인 고려를 하지 않으며, 인과관계가 아닌 상관관계를 판단한다. 또한 개별적인 당해 사안에서 구체적인 타당성 있는 결론을 도출하는 것이 아니라, 과거의 데이터에 입각하여 추출한 확률 내지 패턴을 당해 사안에 적용하여 예측한다.

법관이 내리는 결정을 정당화하는 원천은 타당한 근거의 제시에 있는데, 인공지능의 결론 도출은 지금 이 사건에 관한 분석에 입각한 논증의 산물이 아니므로 승복가능성도 기대하기 어렵다. 법관이 신뢰받는 이유는 그 또한 구성원의 한 사람으로서 여러 사법적 덕목을 가지기 때문이기도 하다.

이러한 문제점들이 존재함에도 불구하고 인공지능을 법관의 판단에 참고하고, 나아가 언젠가 인공지능 법관이 도입될 가능성이 있다고 할 때, 여러 제도적 안배가 필요할 것이다. 인공지능이 학습할 수 있는 양질의 데이터 확보를 위해 사법데이터가 (개인정보 등을 침해하지 않는 선에서) 공개되어야 한다. 비전문가가 복잡한 알고리즘 모두를 이해할 수는 없으나, 인공지능 법관(혹은 법관이 참고하는 인공지능)이 활용하는 훈련데이터 및 기본적인 작동 기제는 설명할 수 있어야 한다. 다만 설명가능성을 위해 지나치게 성능을 희생할 경우, 부정확한 인공지능을 굳이 활용할 이유가 무엇인지가 문제된다.

재판에서 인공지능을 보조적으로 활용하는 방안이 있으나, 이를 ‘인공지능 법관’이라고 할 수는 없다. 인공지능 법관을 도입하되 당사자가 원치 않으면 자연인 법관에 의한 기존의 재판을 받도록 하거나, 자연인 법관으로만 구성된 상급심에 항소를 보장하는 방안도 있다. 모든 영역에서 인공지능 법관의 도입 가부를 일률적으로 판단하기보다, 사건의 성격 등을 고려해 입체적으로 접근할 여지도 있겠다. 가령 가치판단의 성격이 강하며 기본권이나 전체 법질서에 중대한 영향을 미치는 헌법재판, 형사재판의 경우 활용에 더욱 신중을 기하고, 단순하고 정형화된 사건부터 해 나가는 방안이 그것이다.

## 참 고 문 헌

### 1. 단행본

김하열, 헌법강의, 박영사, 2023.

김학성·최희수, 헌법학원론, 피앤씨미디어, 2023.

양종모, 4차 산업혁명 시대의 인공지능 알고리즘에 의한 법 분야 위험 예측, 홍진기법률 연구재단, 2022.

오세용, 인공지능시대: 법관의 미래는?, 박영사, 2022.

이시윤, 신민사소송법, 박영사, 2023.

전광석, 한국헌법론, 집현재, 2023.

전원열, 민사소송법, 박영사, 2022.

정연주, 헌법학원론, 법영사, 2023.

정종섭, 헌법학원론, 박영사, 2022.

한수용, 헌법학, 법문사, 2022.

허 영, 한국헌법론, 박영사, 2023.

데이비드 헬드, 박찬표 옮김, 민주주의의 모델들, 후마니타스, 2014.

마크 코켈버그, 배현석 옮김, 인공지능은 왜 정치적인 수밖에 없는가 - AI의 정치학과 자유, 평등, 정의, 민주주의, 권력, 동물과 환경, 생각이음, 2023.

(Coeckelbergh, Mark, The Political Philosophy of AI, Polity Press, 2022.)

Colombo, Moreno, Phenotropic Interaction: Improving Interfaces with Computing with Words and Perceptions, Springer, 2023.

Cowger, Alfred R. Jr., The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms, Lexington Books, 2020.

Hu, Xianhong et al., Steering AI and Advanced ICTs for Knowledge Societies: A Rights, Openness, Access, and Multi-stakeholder Perspective, UNESCO Publishing, 2019.

Lee, Jaemin, Artificial Intelligence and International Law, Springer, 2022.

- Lennox, John C., 2084: Artificial Intelligence and the Future of Humanity, Zondervan, 2020.
- Lexcellent, Christian, Artificial Intelligence versus Human Intelligence: Are Humans Going to Be Hacked?, Springer, 2019.
- Minsky, Marvin ed., Semantic Information Processing, MIT Press, 1968.
- Sourdin, Tania, Judges, Technology and Artificial Intelligence: The Artificial Judge, Edward Elgar, 2021.
- Stahl, Bernd Carsten, Schroeder, Doris and Rodrigues, Rowena, Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenge, Springer, 2022.
- Ventre, Daniel, Artificial Intelligence, Cybersecurity and Cyber Defence, Wiley, 2020.

## 2. 논문

- 강석정, 법관은 두 개의 양심을 가져야 하는가? - 헌법 제103조 법관의 ‘양심’에 관하여, 사법, 제41호, 2017, 153-198쪽.
- 강승규, 인공지능(AI)이 초래한 사회적 변화와 위협의 공존가능성에 대한 검토, 법이론 실무연구, 제9권 제4호, 2021, 495-524쪽.
- 강승식, 인공지능 판사, 과연 가능한가?, 헌법학연구, 제26권 제3호, 2020, 229-262쪽.
- 강영주, 인공지능을 활용한 민사분쟁 해결에 관한 고찰, 안암법학, 제64권, 2022, 399-431쪽.
- 강현구, 인공지능의 법적정의와 인공지능기술 발달에 따른 국가의 기본권 보호의무, 헌법재판연구, 제8권 제2호, 2021, 29-64쪽.
- 고유강, 법관업무의 지원을 위한 머신러닝의 발전상황에 대한 소고, LAW & TECHNOLOGY, 제15권 제5호, 2019, 3-17쪽.
- 고학수·정해빈·박도현, 인공지능과 차별, 저스티스, 통권 제171호, 2019, 199-277쪽.
- 고유강, 법률인공지능이 변화시킬 재판의 모습: ‘인공지능 판사’의 기술적 대체가능성과 관련된 현황을 중심으로, 재판자료집, 제142집, 2021, 1-32쪽.
- 권경휘, 법의 결정성, 인공지능 그리고 법관의 미래, 법학연구, 제32권 제2호, 2022, 321-366쪽.

- 김건우, 인공지능윤리의 의의와 쟁점, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 39-76쪽.
- 김광수, 인공지능 알고리즘 규율을 위한 법제 동향 - 미국과 EU 인공지능법의 비교를 중심으로 -, 행정법연구, 제70호, 2023, 173-201쪽.
- 김도훈, 민사소송상 인공지능 법관에 관한 소고, 법학논문집, 제45집 제1호, 2021, 57-88쪽.
- 김배원, 지능정보사회와 헌법, 인공지능(AI)의 발전과 헌법적 접근, 공법학연구, 제21권 제3호, 2020, 67-108쪽.
- 김병록, 인공지능의 헌법적 쟁점과 과제, 법학논총, 제27집 제2호, 2020, 87-135쪽.
- 김병필, 재판 예측 인공지능 기술의 현황과 한계, 저스티스, 통권 제182-2호, 2021, 89-117쪽.
- 김성용·정관영, 법조(法曹)영역으로의 인공지능 도입에 대한 제언 - 영업비밀 침해사건을 중심으로 -, LAW & TECHNOLOGY, 2019, 68-89쪽.
- 김송옥, 인공지능 법제의 최신 동향과 과제, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 148-184쪽.
- 김웅재, 형사절차에서 인공지능 알고리즘의 활용가능성과 그 한계, LAW & TECHNOLOGY, 제16권 제4호, 2020, 3-23쪽.
- 김일우, 고위험 인공지능시스템의 기본권 침해에 관한 헌법적 연구, 박사학위논문, 서강대학교, 2023, 1-224쪽.
- 김종철, 인공지능의 헌법적 의의, 오병철 외, 인공지능과 법, 연세대학교 출판문화원, 2023, 11-46쪽.
- 김종현, 범죄자의 선거권 제한의 헌법적 검토 - 선거권 제한 입법의 위헌심사 및 법관의 개별적 선고 필요성에 관한 논의를 소재로 -, 공법연구, 제51집 제1호, 2022, 25-60쪽.
- 김종현, 헌법재판에서의 증명책임, 저스티스, 통권 제183호, 2021, 39-74쪽.
- 김중권, 인공지능시대 알고크라시(Algocracy)에서의 민주적 정당화의 문제, 법조, 제69권 제5호, 2020, 181-207쪽.
- 김현석, 재판의 진행·판단에서 법관의 정서에 관하여, 법조, 통권 제733호, 2019, 254-306쪽.
- 김희정, 인공지능 윤리와 형사정책적 시사점 - 형사사법에서의 인공지능 윤리 이슈를 중

- 심으로 -, 형사정책, 통권 제69호, 2022, 79-110쪽.
- 김희정, 지능정보화 시대의 알고리즘 차별에 대한 법적 소고 - 미국의 알고리즘 차별 사례를 중심으로-, 홍익법학, 제21권 제3호, 2020, 501-524쪽.
- 남중권, ‘알고리즘의 지배’와 ‘법의 지배’에 대한 법이론적 이해, 법과사회, 제71호, 2022, 283-321쪽.
- 박기주, 인공지능 시대, 사법제도의 사회적 가치 실현체계의 법적 기초 연구, 사법, 통권 제59호, 2022, 581-625쪽.
- 박도현, 인간 편향성과 인공지능의 교차, 서울대학교 법학, 제63권 제1호, 2022, 139-175쪽.
- 박도현, 인공지능과 해악: 창발적 해악론을 중심으로, 박사학위논문, 서울대학교, 2021, 1-334쪽.
- 박도현, 인공지능 공정성의 이론과 실제, 법철학연구, 제26권 제2호, 2023, 5-52쪽.
- 박상철, 인공지능 등 정보기술을 통한 법의 기술적 구현 과정의 혁신, 정보법학, 제22권 제2호, 2018, 133-162쪽.
- 박은정, 법해석방법에서 본 헌법 제103조 법관의 독립과 양심, 서울대학교 법학, 제57권 제2호, 2016, 63-103쪽.
- 박혜진, 양형과 인공지능 - 국민의 법감정 수렴, 위험 판단, 양형 분석을 중심으로 -, 법조, 제72권 제4호, 2023, 252-293쪽.
- 서치원, 인공지능은 법의 제국의 왕자가 될 수 있는가 - 드워킨의 ‘법의 제국’을 중심으로 분석한 인공지능의 판사 대체 가능성 -, 연세 의료·과학기술과 법, 제9권 제2호, 2018, 1-26쪽.
- 성중탁, 제4차 산업혁명과 사회국가 실현 등에 대한 전망과 과제, 헌법학연구, 제27권 제1호, 2021, 343-381쪽.
- 손은지, 인공지능의 헌법적 위험과 대응적 규범체계 정립에 관한 연구, 박사학위논문, 숙명여자대학교, 2023, 1-355쪽.
- 송성수, 인공지능은 인간의 일자리를 얼마나 대체할 것인가: 인공지능 시대의 기술과 노동에 관한 시론, 코기토, 통권 제96호, 2022, 7-37쪽.
- 심우민, 알고리즘 규제와 입법 전략, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 225-258쪽.

- 안상원·박규동, 형사사법 절차상 인공지능(AI) 시스템 도입의 주요 논점, 범죄정보연구, 제16호, 2022, 135-155쪽.
- 양종모, 법 분야 위험의 예측을 위한 인공지능 알고리즘과 관련된 법적 쟁점, IT와 법연구, 제22집, 2021, 371-412쪽.
- 양종모, 인공지능 알고리즘의 편향성, 불투명성이 법적 의사결정에 미치는 영향 및 규율 방안, 법조, 제66권 제3호, 2017, 60-105쪽.
- 양종모, 인공지능에 의한 판사의 대체 가능성 고찰, 홍익법학, 제19권 제1호, 2018, 1-29쪽.
- 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등예의 적용에 관한 고찰, 형사법의 신동향, 통권 제51호, 2016, 210-242쪽.
- 양종모, 재범의 위험성 예측 알고리즘과 설명가능성 및 공정성의 문제, 형사법의 신동향, 통권 제70호, 2021, 207-240쪽.
- 양천수, 인공지능과 법체계의 변화 - 형사사법을 예로 하여, 법철학연구, 제20권 제2호, 2017, 45-76쪽.
- 유승익, 인공지능이 인권과 민주주의에 미치는 영향과 인공지능법안의 쟁점, 민주법학, 통권 제82호, 2023, 77-106쪽.
- 윤수정, 인공지능과 민주주의에 관한 소고, 헌법재판연구, 제8권 제2호, 2021, 3-27쪽.
- 이명민, 인공지능, 법, 그리고 민사소송, 안암법학, 제57호, 2018, 35-59쪽.
- 이병규, AI의 예측능력과 재범예측알고리즘의 헌법 문제 - State v. Loomis 판결을 중심으로, 공법학연구, 제21권 제2호, 2020, 169-191쪽.
- 이상용, 알고리즘 규제의 원리와 구조, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 187-224쪽.
- 이성진 외, 4차 산업혁명시대 법정정보기술의 현황과 발전방안, 정보화정책, 제28권 제1호, 2021, 3-21쪽.
- 이수진, 가사재판에서 인공지능과 온라인분쟁해결의 의의, 사법, 제64호, 2023, 95-144쪽.
- 이슬아, 인공지능 판사 앞의 7가지 숙제 - 재범위험성 예측 알고리즘을 둘러싼 과학기술적·법적 논의 분석, 사법, 제64호, 2023, 665-714쪽.
- 이준일, 인공지능과 헌법 - 인공지능의 위험에 대한 헌법적 권리 차원의 대응 -, 헌법학연구, 제28권 제2호, 2022, 347-383쪽.



이지나, 형사사법에서의 인공지능 예측프로그램 활용 제언: 인종적 편향 가능성 회피를 중심으로, 원광법학, 제39권 제2호, 2023, 57-73쪽.

장진환, 인공지능재범위험성평가도구의 도입가능성과 규범적 문제점 - 보안처분 선고절차에서의 적용가능성을 중심으로 -, 형사정책, 제34권 제1호, 2022, 41-77쪽.

전후재, 우리 민사재판의 미래에 관한 소고 - 디스커버리, 온라인 소송, 인공지능(AI)을 중심으로 -, 민사소송, 제27권 제1호, 2023, 79-148쪽.

정문식, 인공지능과 헌법, 한양대학교 법학연구소 편, 인공지능과 법, 도서출판 오래, 2023, 41-67쪽.

정소영, 형사사법에서 인공지능 사용에 대한 유럽의회 결의안 - ‘인공지능에 의한 결정’ 금지에 관하여 -, 형사정책, 제34권 제2호, 2022, 137-172쪽.

정영화, 인공지능과 법원의 분쟁해결 - 최근 영미법국가들의 인공지능 법제, 홍익법학, 제21권 제1호, 2020, 209-247쪽.

조소영, 인공지능과 민주주의, 공법연구, 제49집 제2호, 2020, 147-167쪽.

조원용, 대의제 국회 입법과정에서 인공지능(AI)의 역할: 공익의 ‘발견’과 ‘결정’을 중심으로, 가천법학, 제11권 제3호, 2018, 49-80쪽.

조재현, 인공지능 알고리즘 기반 위험성 예측 프로그램의 형사재판 활용 문제, 재판자료집, 제142집, 2021, 903-940쪽.

조한상·이주희, 인공지능과 법, 그리고 논증, 법과 정책연구, 제16집 제2호, 2016, 295-320쪽.

조홍식, 물경시정치 - 비례입헌주의를 주창하며 -, 서울대학교법학, 제49권 제3호, 2008, 97-126쪽.

주현경·정채연, 범죄예측 및 형사사법절차에서 알고리즘 편향성 문제와 인공지능의 활용을 위한 규범 설계, 법학논총, 제27집 제1호, 2020, 115-162쪽.

틸로 하겐도르프, 인공지능윤리의 윤리학: 윤리가이드라인에 대한 평가, 김건우 편저, 인공지능 규제 거버넌스의 현재와 미래, 파이돈, 2023, 77-110쪽.

한애라, “사법시스템과 사법환경에서의 인공지능 이용에 관한 유럽 윤리헌장”의 검토 - 민사사법절차에서의 인공지능 도입 논의와 관련하여, 저스티스, 통권 제172호, 2019, 38-79쪽.

홍석한, 미국 “2022 알고리즘 책임법안”에 대한 고찰, 미국헌법연구, 제34권 제1호,

- 2023, 73-107쪽.
- 홍석한, 유럽연합 ‘인공지능법안’의 주요 내용과 시사점, 유럽헌법연구, 제38호, 2022, 243-282쪽.
- Webley, Lisa, 권헌영 옮김, 인공지능(AI)과 법조 분야: 윤리적·규제적 고려사항, 경제 규제와 법, 제12권 제2호, 2019, 69-80쪽.
- Campbell, Ray Worthy, Artificial Intelligence in the Courtroom: The Delivery of Justice in the Age of Machine Learning, Colorado Technology Law Journal, Vol. 18, No. 2, 2020, pp. 323-350.
- Contissa, Giuseppe and Lasagni, Giulia, When It Is (Also) Algorithms and AI That Decide on Criminal Matters: In Search of an Effective Remedy, European Journal of Crime, Criminal Law and Criminal Justice, Vol. 28, No. 3, 2020, pp. 280-304.
- Crootof, Rebecca, “Cyborg Justice” and the Risk of Technological-Legal Lock-in, Columbia Law Review Online, Vol. 119, 2019, pp. 233-251.
- Fine, Anna, Le, Stephanie and Miller, Monica K., Content Analysis of Judges’ Sentiments Toward Artificial Intelligence Risk, Criminology, Criminal Justice, Law & Society, Vol. 24, Iss. 2, 2023, pp. 31-46.
- Kerrigan, Charles, and Vercoe, Oliver, Explainable AI and Responsible AI, in Charles Kerrigan ed., Artificial Intelligence: Law and Regulation, Edward Elgar, 2022, pp. 527-533.
- Khamis, Alaa, Patel, Dipkumar and Elgazzar, Khalid, Deep Learning for Unmanned Autonomous Vehicules: A Comprehensive Review in Anis Koubaa and Ahmand Taher Aazr eds., Deep Learning for Unmanned Systems, Springer, 2021, pp. 1-24.
- Koenig, Melissa Love, Oseid, Julie A. and Vorenberg, Amy, Ok Google, Will Artificial Intelligence Replace Human Lawyering, Marquette Law Review, Vol. 102, No. 4, 2019, pp. 1269-1304.
- Liu, Han-Wei, Lin, Ching-Fu and Chen, Yu-Jie, Beyond State v Loomis: Artificial Intelligence, Government Algorithmization and Accountability, International

- Journal of Law and Information Technology, Vol. 27, No. 2, 2019, pp. 122-141.
- Masuhara, David M., Artificial Intelligence and Adjudication: Some Perspectives, *Amicus Curiae*, Vol. 111, 2017, pp. 2-15.
- Soukupova, Jana, AI-Based Legal Technology: A Critical Assessment of the Current Use of Artificial Intelligence in Legal Practice, *Masaryk University Journal of Law and Technology*, Vol. 15, No. 2, 2021, pp. 279-300.
- Sourdin, Tania, Judge v Robot? Artificial Intelligence and Judicial Decision-Making, *University of New South Wales Law Journal*, Vol. 41, No. 4, 2018, pp. 1114-1133.
- Vėbraitė, Vigita and Strikaitė-Latušinskaja, Goda, Digitalization of Justice in Lithuania, in Katarzyna Gajda-Roszczyńska ed., *Impact of the COVID-19 Pandemic on Justice Systems*, V&R unipress, 2023, pp. 223-234.
- Villasenor, John and Foggo, Virginia, Artificial Intelligence, Due Process and Criminal Sentencing, *Michigan State Law Review*, Vol. 2020, NO. 2 2020, pp. 295-354.
- Wexler, Rebecca, Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System, *Stanford Law Review*, Vol. 70, No. 5, 2018, pp. 1343-1430.
- Xu, Zichun, Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities, *Applied Artificial Intelligence*, Vol. 36, No. 1, 2022, pp. 1025-1045.

### 3. 언론보도내역

- 고환경, 인간 중심 AI와 완전히 자동화된 의사결정에 대한 대응권, *법률신문*, 2023. 12. 21.
- 과학기술정보통신부, 과기정통부, 사람이 중심이 되는 「인공지능(AI) 윤리기준」 마련, *과학기술정보통신부 보도자료*, 2020. 12. 23.
- 류인하, 사법부 신뢰 추락으로 떠오른 ‘존경하는 AI 판사님’, *경향신문*, 2019. 4. 21.
- 박수연, [인터뷰] 이상원 제9기 양형위원회 위원장 “형법은 사랑... 양형 논의도 사랑이

- 토대가 되어야”, 법률신문, 2023. 8. 14.
- 서영상, 대륙이주, 국내로펌 최초 ‘법률인공지능 시스템’ 도입, 법률신문, 2018. 2. 27.
- 안혜성, ‘로톡 형량예측서비스’종료 두고 법률플랫폼 갈등 증폭, 법률저널, 2021. 9. 15.
- 양지열, 인공지능(AI)판사는 정의로울까?, Ai타임스, 2021. 8. 13.
- 유형용, [칼럼] AI법관의 도래, 법률저널, 2016. 1. 15.
- 윤고은, 중국 판사들, 모든 사건에 대해 AI 상담 받는다, 연합뉴스, 2022. 7. 14.
- 윤상원, 현대 인공지능의 원리와 한계, 법률신문, 2023. 12. 9.
- 이배운, “판사를 AI로 교체하자”현직 판사에게 물어보니, 이데일리, 2023. 7. 24.
- 이성진, 대법원 ‘차세대 전자소송 시스템 구축한다’, 법률신문, 2018. 4. 12.
- 이예린, ‘신조어 강자’파파고·‘다국어 지원’구글 번역기·‘번역의 정석’딥엘[ICT], 문화일보, 2023. 12. 11.
- 이용경, 천대엽 법원행정처장, 법관 증원 및 사법부 예산 확보 약속, 법률신문, 2024. 1. 15.
- 차병직, 검증의 마당, 법률신문, 2023. 1. 5.
- 천경훈, [서초포럼] 챗GPT와 법률가의 미래, 법률신문, 2023. 2. 23.
- 최창현, 유럽의회, 세계 최초 ‘인공지능 규제 법안’ 채택...“더 엄격한 개인정보 보호, 더 엄격한 투명성, 협조하지 않을 경우 징벌적 벌금 부과”, 인공지능신문, 2023. 6. 15.
- 최태우, 인공지능의 역사 그리고 GPU 기반 딥 러닝, CCTV뉴스, 2016. 3. 15.
- 한수현, 법원 AI재판업무 지원 연구 위한 ‘인공지능 커뮤니티’ 출범, 법률신문, 2023. 11. 16.
- 한애라, AI 판사?, 법률신문, 2020. 1. 20.
- 홍윤지, [인터뷰] ‘디지털전도사’ 강민구 고법부장판사, “한국 법조, 생성형 AI 기술 신속히 도입해야”, 법률신문, 2023. 5. 31.
- 홍윤지, [취재수첩] AI와 법률가의 공생(共生), 법률신문, 2023. 2. 23.
- 홍윤지, EU, 세계 첫 ‘AI 규제법’ 합의...한국식 AI 규제모델 마련해야, 법률신문, 2023. 12. 14.
- Angwin, Julia et al., “Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And It’s Biased Against Blacks.”, ProPublica, 23

May 2016,

Harris, Briony, Could an AI Ever Replace a Judge in Court?, World Government Summit, 11 July 2018.

Kelly, Samantha Murphy, ChatGPT Passes Exams From Law and Business Schools, CNN Business, 26 January 2023.

Tuulik, Maria-Elisa, Estonia does not develop AI Judge, Republic of Estonia Ministry of Justice, 16 February 2022.

Winfrey, Tiffany, China Uses Artificial Intelligence (AI) to Run Courts, Supreme Justices; Cutting Judges' Typical Workload By More Than a Third and Saving Billion Work Hours, The Science Times, 14 July 2022.

#### 4. 보고서, 논평, 서평, 토론회 자료집 기타

김찬희, 인공지능 기반 자동화 행정행위에 관한 헌법적 고찰, 헌법이론과 실무, 2023-A-3, 헌법재판소 헌법재판연구원, 2023, 1-80쪽.

박준모, 스마트법원 4.0의 입법과제 - 예금 압류명령 전자정보 송신 제도 도입을 중심으로, NARS 현안분석, 제99호, 국회입법조사처, 2019, 1-21쪽.

서울대학교 산학협력단, 사법부에서의 인공지능(AI) 활용방안, 법원행정처, 2020, 1-182쪽.

오세용, 인공지능(AI)이 앞당긴 미래 사법의 변화 그리고 법관, 2023 사법연수원 국제콘퍼런스 자료집(지정토론문), 사법연수원, 2023, 114-122쪽.

정채연, 사법절차 및 사법서비스에서 인공지능 기술의 도입 및 수용을 위한 정책 연구, 사법정책연구원, 2021, 1-160쪽.

주민식, 인공지능은 어떻게 발달해왔는가, 인공지능의 역사, 삼성SDS 인사이드리포트, 2017. 9. 15.

차세대전자소송 추진단, 차세대전자소송 시스템 구축 사업 착수보고회 자료집, 2020.

Aldana, Benes Z., 사법정의의 대변혁: 법률시스템에서의 AI 부상과 윤리적 고찰, 2023 사법연수원 국제콘퍼런스 자료집, 사법연수원, 2023, 49-88쪽.

Committee on Civil Liberties, Justice and Home Affairs, Report on Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial

Authorities in Criminal Matters, 2020/2016(INI), 13 July 2021, pp. 1-33.

European Commission for the Efficiency of Justice(CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Council of Europe, 2018, pp. 1-77.

Forrest, Katherine B., AI시대의 사법정의, 패러다임의 전환, 사법의 대응, 2023 사법연수원 국제콘퍼런스 자료집, 사법연수원, 2023, 19-48쪽.

High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission, 2019, pp. 1-35.

OECD, Recommendation of the Council on Artificial Intelligence, 2020, OECD/LEGAL/0449, adopted on 22.06.2019.

## 인공지능 법관 관련 헌법적 쟁점의 검토

2024年 3月 19日 印刷

2024年 3月 26日 發行

발행: 헌법재판소  
헌법재판연구원

인쇄: 성문인쇄사(02·2272·7553)

<비매품>

\* 본 보고서의 내용은 본원의 공식견해가 아닙니다.

