

인공지능과 불법행위책임에 대한 민사법적 검토



한국법학원 연구보고서 민-24-06

인공지능과 불법행위책임에 대한 민사법적 검토

2024. 12. 31.

2024년도 대법원 연구보고서

연구기관 : 한국법학원

연구책임 : 성덕근 연구위원

감수 : 정상현 교수 (성균관대학교)

Contents

제1장 들어가며

제1절 연구의 목적	10
제2절 연구의 방법	12

제2장 인공지능 일반론

제1절 개관	14
제2절 인공지능에 대한 이해	14
I. 인공지능의 개념과 연혁	14
1. 인공지능의 개념	14
2. 인공지능의 연혁	17
II. 인공지능 관련 용어 개관	20
1. 알고리즘	20
2. 인공지능과 인공지능 로봇	20
3. 약한 인공지능과 강한 인공지능, 초인공지능	21
(1) 약한 인공지능	22
(2) 강한 인공지능	22
(3) 초인공지능	23
4. 딥러닝	24
III. 인공지능의 특징	25
1. 자율성	25

2. 예측불가능성	26
3. 분산성	26
4. 설명불가능성	27
IV. 인공지능의 발전동향	29
V. 주요 국가의 입법동향	31
1. 유럽연합	31
2. 미국	33
3. 일본	34
4. 중국	35
5. 우리나라	35
제3절 인공지능과 법인격	38
I. 서설	38
II. 인공지능에 대한 법인격 인정 여부	41
1. 현행법상 인공지능의 법적 지위	41
2. 인공지능의 법인격 인정여부에 관한 논의	42
(1) 논의의 배경	42
(2) 인공지능의 법인격 인정에 관한 학설	43
III. 검토	53
제4절 소결	55

제3장 인공지능과 불법행위책임

제1절 개관	58
---------------------	-----------

제2절	인공지능에 의한 손해의 발생	58
Ⅰ.	인공지능에 의한 손해의 특수성	58
1.	정보오류에 의한 손해	58
2.	인체 손해의 발생	59
3.	동시다발적인 손해	60
Ⅱ.	인공지능에 의한 손해와 인공지능의 과실판단	61
1.	전통적인 과실책임주의의 한계	61
(1)	인공지능 로봇과 고의책임	61
(2)	과실의 전제로서의 주의의무	62
(3)	추상적 경과실	62
2.	인공지능의 과실판단 문제 및 그 한계	63
(1)	인공지능 프로그래머의 과실판단	64
(2)	인공지능 로봇 이용자의 과실판단	64
(3)	인공지능 로봇 소유자의 과실판단	65
(4)	과실의 증명책임	67
제3절	인공지능에 의한 손해에 적용가능한 민사책임의 분류	68
Ⅰ.	논의의 배경	68
Ⅱ.	과실책임에 기한 법적 구성	68
1.	감독자책임	68
2.	사용자책임	69
3.	공작물점유자책임	71
4.	동물점유자책임	73
Ⅲ.	무과실책임에 기한 법적 구성	75
1.	공작물소유자책임	75
2.	제조물책임	75
(1)	인공지능의 제조물성 인정 여부	76
(2)	인공지능의 결함 판단	78
(3)	결함과 손해 간의 인과관계 입증	79
(4)	면책요건으로서 개발위험의 항변	80

3. 자동차운행자책임	81
제4절 인공지능의 불법행위책임에 대한 새로운 책임모델	82
Ⅰ. 논의의 배경	82
Ⅱ. 새로운 민사책임 모델로서 위험책임과 편익책임	83
1. 위험책임	83
2. 편익책임	86
Ⅲ. 검토	87
제5절 소결	88

제4장 나가며

나가며	92
-----------	----

참고문헌

참고문헌	94
------------	----

제1장 들어가며

제1절 연구의 목적

제2절 연구의 방법



제1장

들어가며

제 1 절 | 연구의 목적

최근 제4차 산업혁명에 대한 사회적 관심이 급증하면서, 인공지능은 미래 사회에서 기술 혁신과 생산성 향상의 중심축으로 자리잡을 것으로 예상된다. 이에 따라 의료, 교육, 금융 등 다양한 분야에서 AI의 활용이 증가하며 인간의 삶의 질을 향상시키는 동시에, 인공지능으로 인해 발생하게 되는 부작용으로서 노동시장과 윤리적 문제 그리고 법률적인 문제 등의 새로운 도전과제도 등장하게 될 것이다. 따라서 인공지능의 발전은 기술적인 혁신뿐만 아니라 사회적 책임과 법제도의 재정립을 요구하게 되는 중요한 계기이자 과제로서의 의미를 가질 것으로 전망된다.

이와 관련하여 2024년 12월 26일 ‘인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안(일명 ‘AI 기본법’)이 국회 본회의를 통과하여 2026년 1월 시행을 앞두고 있지만, 인공지능의 불법행위책임과 관련된 기존의 법령들에 있어서는 큰 변화를 찾아보기 어렵다. 물론 이에 대해 기존의 책임법제 아래서 해결이 가능하다면 문제가 없겠지만, 과거에는 예상하지 못했던 사회변화에 대해 책임법적 변화가 필요하다면 시대변화에 걸맞는 새로운 책임법제를 구상하는 것이 필요할 수도 있을 것이다.

본 연구는 인공지능에 관한 개념이나 특징 등의 일반적인 내용에 대해 살펴보는 것을 시작으로 인공지능에 의해 발생할 수 있는 손해에 대한 불법행위책임에 대해 기존의 법령을 중심으로 검토하는 것을 목적으로 한다. 그리고 이를 위해 새로운 정책 내지 법안 등의

법적 해결방안을 제시하기보다는 그동안 인공지능과 불법행위책임에 대해 이루어졌던 논의들의 정리에 보다 중점을 두었다는 점을 밝힌다.

본 연구는 다음과 같은 순서와 방법으로 진행되었다.

제2장에서는 이후 인공지능과 불법행위책임에 대해 살펴보기에 앞서, 인공지능과 관련된 기본적인 내용들에 대해 자세히 살펴보게 된다. 제2장의 전반부에서는 인공지능에 대해서 이후 논의될 내용들에 대한 이해를 돕기 위해 인공지능의 개념과 연혁에 대해 먼저 살펴보고, 인공지능과 관련된 주요 용어들에 대해 살핀다. 다음으로는 인공지능이 가지고 있는 특징인 자율성과 예측불가능성, 분산성 및 설명불가능성에 대해 살펴보고 오늘날 인공지능의 발전 동향을 간략하게 살펴본다. 그리고 인공지능과 관련된 주요 국가 및 우리나라의 입법동향에 대해서도 살펴볼 것이다. 제2장의 후반부에서는 인공지능에 대한 법인격 부여의 문제에 대해 이를 긍정하는 견해와 부정하는 견해가 제시하는 논거들에 대해 구체적으로 살펴볼 것이다.

제3장에서는 제2장에서 살펴본 내용들을 토대로 인공지능으로 발생한 손해와 관련된 불법행위책임에 대해 논의가 진행된다. 이를 위해 먼저 인공지능에 의해 발생하는 손해의 특수성에 대해 살펴보고, 인공지능에 의한 손해에 대해 전통적인 불법행위책임의 기준인 과실책임주의를 기준으로 살펴본다. 그리고 불법행위책임과 관련하여 인공지능에 의한 손해에 적용가능한 현행 책임법리들의 적용가능성에 대해 살펴볼 것이다. 그리고 위험책임과 편익책임이 무엇인지 살펴보고 어느 것이 인공지능과 관련된 새로운 책임모델로서 적합한 것인지에 대해 검토된다.

마지막으로 제4장에서는 본 보고서의 결론으로서 지금까지 살펴본 내용들에 대한 요약 및 정리가 이루어질 것이다.

제2장 인공지능 일반론

제1절 개관

제2절 인공지능에 대한 이해

제3절 인공지능과 법인격

제4절 소결



제2장

인공지능 일반론

제 1절 | 개관

이후 진행될 인공지능과 불법행위책임에 대한 검토를 하기에 앞서 기초적인 배경지식으로서 인공지능에 대한 이해가 선행되어야 한다. 이를 위해 인공지능의 개념과 연혁을 시작으로 인공지능 관련 용어와 인공지능의 특징을 살펴보고, 인공지능의 발전 동향과 인공지능에 관한 주요 국가들의 입법동향을 먼저 살펴보고자 한다. 그리고 인공지능에 대해 법인격을 부여하는 것에 대한 논의를 살펴게 된다.

제 2절 | 인공지능에 대한 이해

1. 인공지능의 개념과 연혁

1. 인공지능의 개념

인공지능(Artificial Intelligence, AI)은 외부환경을 스스로 인식하고 자율적으로 행동할 수 있으므로 그 본질에 대해 용어와 견해가 다양하게 제시되고 있다. 인공지능에 대한 사전적 정의는 ‘인간의 지능이 가지는 학습, 추리, 적응, 논증 따위의 기능을 가진 컴퓨터 시스템’으로서 ‘전문가 시스템, 자연 언어의 이해, 로봇 공학, 인공 시각, 문제 해결, 학습

의 지식 획득, 인지과학 따위에 응용한다'고 되어 있다.¹⁾ 법적인 관점에서 학자들은 인공지능에 대해 '소프트웨어 또는 소프트웨어와 결합된 형태로 구현되어 외부환경을 스스로 인식하고 상황을 판단하여 자율적으로 동작하는 기계나 장치'²⁾라고 하거나, '인간의 오감에 상당하는 입력장치와 입력된 정보를 통합·분석·판단할 수 있는 정보처리장치를 갖춘 정보처리시스템'³⁾ 또는 '어떤 문제를 실제로 사고하고 해결할 수 있는 인간의 지능적인 행태를 컴퓨터 소프트웨어 알고리즘을 이용하여 구현한 것'⁴⁾ 등으로 정의하고 있다.

그동안 인공지능의 개념에 대해 법적으로 명확하게 정의된 바는 없었지만 여러 법들에서 정의하고 있는 이와 유사한 개념들에 대해 살펴보면, 2018년 개정된 「지능형 로봇 개발 및 보급 촉진법」에서는 '지능형 로봇'에 대해 '외부환경을 스스로 인식하고 상황을 판단하여 자율적으로 동작하는 기계장치(기계장치의 작동에 필요한 소프트웨어를 포함한다)'라고 정의하고 있고,⁵⁾ 인공지능이 소프트웨어로 구현된다는 점을 고려하여 「소프트웨어 진흥법」에서는 '소프트웨어'를 '컴퓨터, 통신, 자동화 등의 장비와 그 주변장치에 대하여 명령·제어·입력·처리·저장·출력·상호작용이 가능하게 하는 지시·명령(음성이나 영상정보 등을 포함한다)의 집합과 이를 작성하기 위하여 사용된 기술서나 그 밖의 관련 자료'라는 정의하고 있으며,⁶⁾ 「지능정보화 기본법」에서는 '지능정보기술'에 대해 정의하고 있다.⁷⁾ 한편 최근 국회 과학기술정보통신위원회에서 2024년 11월 26일 위원회 대

1) 표준국어대사전

(<https://ko.dict.naver.com/#/entry/koko/8b8b95a9d8d04942988850d62f6a35d3>).

2) 김운명, '人工知能(로봇)의 법적 쟁점에 대한 試論的 考察', 정보법학, 제20권 제1호, 한국정보법학회, 2016.5., 146면.

3) 이중기, '인공지능을 가진 로봇의 법적 취급: 자율주행자동차 사고의 법적 인식과 책임을 중심으로', 홍익법학, 제17권 제3호, 홍익대학교 법학연구소, 2016.9., 3면.

4) 정진명, '인공지능에 대한 민사책임 법리', 재산법연구, 제34권 제4호, 한국재산법학회, 2018.2., 139면.

5) 「지능형 로봇 개발 및 보급 촉진법」 제2조 제1호.

6) 「소프트웨어 진흥법」 제2조 제1호.

7) 「지능정보화 기본법」은 '지능정보기술'에 대해 다음과 같이 정의하고 있다.

제2조(정의) 4. "지능정보기술"이란 다음 각 목의 어느 하나에 해당하는 기술 또는 그 결합 및 활용 기술을 말한다.

안으로 의결된 ‘인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안(이하 ‘AI 기본법’) 제2조에서는 인공지능에 대해 ‘학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것’을 말한다고 규정하고, 인공지능과 관련된 기타 개념들에 대해서도 구체적으로 정의하고 있다.⁸⁾

-
- 가. 전자적 방법으로 학습·추론·판단 등을 구현하는 기술
 - 나. 데이터(부호, 문자, 음성, 음향 및 영상 등으로 표현된 모든 종류의 자료 또는 지식을 말한다)를 전자적 방법으로 수집·분석·가공 등 처리하는 기술
 - 다. 물건 상호간 또는 사람과 물건 사이에 데이터를 처리하거나 물건을 이용·제어 또는 관리할 수 있도록 하는 기술
 - 라. 「클라우드컴퓨팅 발전 및 이용자 보호에 관한 법률」 제2조제2호에 따른 클라우드컴퓨팅기술
 - 마. 무선 또는 유·무선이 결합된 초연결지능정보통신기반 기술
 - 바. 그 밖에 대통령령으로 정하는 기술
- 8) ‘인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안’ 제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다.
- 1. “인공지능”이란 학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것을 말한다.
 - 2. “인공지능시스템”이란 다양한 수준의 자율성과 적응성을 가지고 주어진 목표를 위하여 실제 및 가상 환경에 영향을 미치는 예측, 추천, 결정 등의 결과물을 추론하는 인공지능 기반 시스템을 말한다.
 - 3. “인공지능기술”이란 인공지능을 구현하기 위하여 필요한 하드웨어·소프트웨어 기술 또는 그 활용 기술을 말한다.
 - 4. “고영향 인공지능”이란 사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위협을 초래할 우려가 있는 인공지능시스템으로서 다음 각 목의 어느 하나의 영역에서 활용되는 것을 말한다.
 - 가. 「에너지법」 제2조제1호에 따른 에너지의 공급
 - 나. 「먹는물관리법」 제3조제1호에 따른 먹는물의 생산 공정
 - 다. 「보건의료기본법」 제3조제1호에 따른 보건의료의 제공이용체계의 구축 및 운영
 - 라. 「의료기기법」 제2조제1항에 따른 의료기기 및 「디지털의료제품법」 제2조제2호에 따른 디지털의료기기의 개발 및 이용
 - 마. 「원자력시설 등의 방호 및 방사능 방재 대책법」 제2조제1항제1호에 따른 핵물질과 같은 항 제2호에 따른 원자력시설의 안전한 관리 및 운영
 - 바. 범죄 수사나 체포 업무를 위한 생체인식정보(얼굴·지문·홍채 및 손바닥 정맥 등 개인을 식별할 수 있는 신체적·생리적·행동적 특징에 관한 개인정보를 말한다)의 분석·활용
 - 사. 채용, 대출 심사 등 개인의 권리·의무 관계에 중대한 영향을 미치는 판단 또는 평가
 - 아. 「교통안전법」 제2조제1호부터 제3호까지에 따른 교통수단, 교통시설, 교통체계의 주요한 작동 및 운영
 - 자. 공공서비스 제공에 필요한 자격 확인 및 결정 또는 비용징수 등 국민에게 영향을 미치는 국가, 지방자치단체, 「공공기관의 운영에 관한 법률」 제4조에 따른 공공기관 등(이하 “국가기관등”이라 한다)의 의사결정

2. 인공지능의 연혁

‘인공지능(Artificial Intelligence, AI)’ 분야는 공식적으로 1956년 미국의 존 매카시(John McCarthy)에 의해 추진된 ‘다트머스 연구 프로젝트(Dartmouth Summer Research Project on Artificial Intelligence)’의 워크숍을 통해 세상에 최초로 등장하였다.⁹⁾¹⁰⁾ 1956년 다트머스 회의를 기점으로 학자들은 미지의 영역인 인간의 마음과 지

-
- 차. 「교육기본법」 제9조제1항에 따른 유아교육·초등교육 및 중등교육에서의 학생 평가
- 카. 그 밖에 사람의 생명·신체의 안전 및 기본권 보호에 중대한 영향을 미치는 영역으로서 대통령령으로 정하는 영역
5. “생성형 인공지능”이란 입력한 데이터(「데이터 산업진흥 및 이용촉진에 관한 기본법」 제2조제1호에 따른 데이터)를 말한다. 이하 같다)의 구조와 특성을 모방하여 글, 소리, 그림, 영상, 그 밖의 다양한 결과물을 생성하는 인공지능시스템을 말한다.
 6. “인공지능산업”이란 인공지능 또는 인공지능기술을 활용한 제품(이하 “인공지능제품”이라 한다)을 개발·제조·생산 또는 유통하거나 이와 관련한 서비스(이하 “인공지능서비스”라 한다)를 제공하는 산업을 말한다.
 7. “인공지능사업자”란 인공지능산업과 관련된 사업을 하는 자로서 다음 각 목의 어느 하나에 해당하는 법인, 단체, 개인 및 국가기관등을 말한다.
 - 가. 인공지능개발사업자: 인공지능을 개발하여 제공하는 자
 - 나. 인공지능이용사업자: 가목의 사업자가 제공한 인공지능을 이용하여 인공지능제품 또는 인공지능 서비스를 제공하는 자
 8. “이용자”란 인공지능제품 또는 인공지능서비스를 제공받는 자를 말한다.
 9. “영향받는 자”란 인공지능제품 또는 인공지능서비스에 의하여 자신의 생명, 신체의 안전 및 기본권에 중대한 영향을 받는 자를 말한다.
 10. “인공지능사회”란 인공지능을 통하여 산업·경제, 사회·문화, 행정 등 모든 분야에서 가치를 창출하고 발전을 이끌어가는 사회를 말한다.
 11. “인공지능윤리”란 인간의 존엄성에 대한 존중을 기초로 하여, 국민의 권익과 생명·재산을 보호할 수 있는 안전하고 신뢰할 수 있는 인공지능사회를 구현하기 위하여 인공지능의 개발, 제공 및 이용 등 모든 영역에서 사회구성원이 지켜야 할 윤리적 기준을 말한다.
- 9) Stanford University AI100, ‘Artificial Intelligence and Life in 2030’, 2016, p.50.
- 10) 인공지능 연구와 관련하여 주목해야 하는 것은 영국의 수학자 앨런 튜링(Alan Mathison Turing)이 제시한 ‘튜링 테스트’이다. 튜링은 1950년 발표한 논문 ‘계산 기계와 지능(Computing Machinery and Intelligence)’에서 ‘기계가 생각할 수 있는가’라는 화두를 던졌다. 그는 ‘생각하는 기계’로서의 컴퓨터에 대한 본인의 생각을 제시하면서 컴퓨터가 기술적으로 발전하게 되면 언젠가는 인간처럼 ‘생각’을 할 수 있게 된다고 주장하였다. 그는 기계적 지능을 구현하는 원리를 설명하면서, 기계가 지능을 가지고 있는지를 판별하는 간단한 테스트를 제시하였는데, 이것이 ‘튜링 테스트’이며, ‘이미테이션 게임’이라고도 일컬어진다. 이 테스트는 ‘기계’와 ‘그 기계와 대화하는 사람’, 그리고 이 둘을 관찰하는 ‘심사

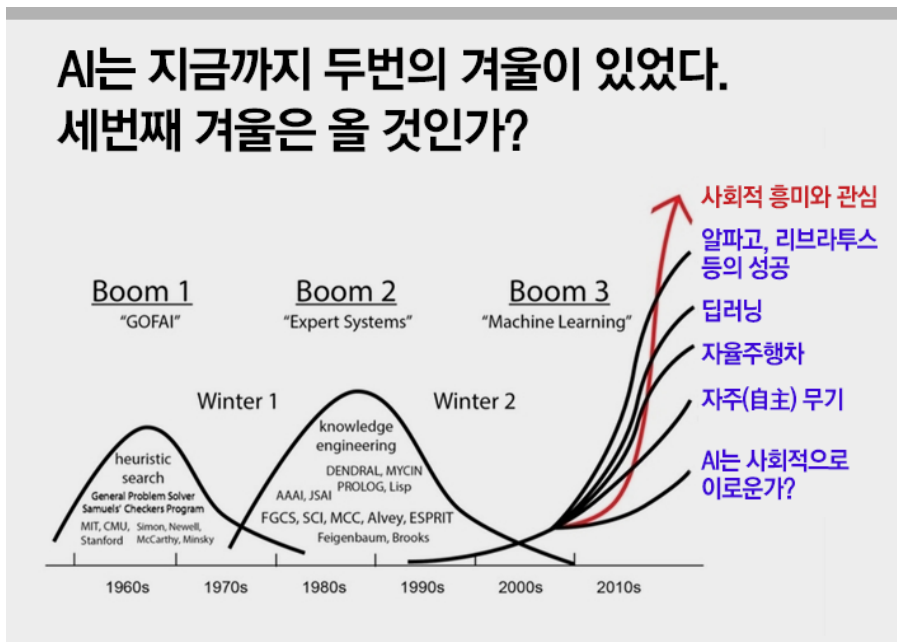
능의 비밀을 모두 알 수 있다는 자신감이 충만하였고, 인공지능은 과학기술 분야의 핵심으로 주목받게 되었다. 다트머스 회의에 참여했던 박사들은 “20년 내에 기계가 사람이 할 수 있는 모든 일을 할 것이다”(Herbert Simon, 1965년)라거나 “몇 년 안에 인간의 평균 지능을 가진 기계가 등장할 것이다”(Marvin Minsky, 1970년) 등의 예언을 하기도 하였지만, 이러한 기대와 달리 인공지능은 기술 및 실용성의 부족으로 인해 복잡한 문제나 현실 세계를 해결하기 어렵다는 한계가 드러나게 되었다. 그리하여 인공지능이 등장했던 제1차 인공지능시대는 허무하게 몰락하게 되었고, 최초의 암흑기인 ‘제1차 인공지능의 겨울(AI Winter)’이 시작되었다.

탐색과 추론의 시대라 할 수 있는 제1차 인공지능 시대 이후 ‘지식’ 그 자체를 이용하려는 방법론이 발전하면서 1970년대 후반 인공지능에 대한 두 번째 봄이 시작되었다. 이 시기에는 특정 영역의 전문 지식을 규칙적으로 표현하여 추론을 하는 것을 내용으로 하는 ‘전문가 시스템(Expert Systems)’이 크게 주목을 받았다. 제2차 인공지능시대의 문을 열었던 세계 최초의 전문가 시스템은 1965년 파이겐바움(Feigenbaum)이 개발한 ‘덴드랄(Dendral)’이었는데, 이 시스템은 화학 물질의 질량 스펙트럼 데이터로 유기화학물의 종류를 추론하는 일종의 ‘인공지능 화학자’라고 볼 수 있다. 그 밖에 대표적인 전문가 시스템으로는 1972년 스탠포드의 브루스 뷰캐넌(Bruce Buchanan) 박사 팀이 개발한 세계 최초의 의료 전문가 시스템인 ‘마이신(MYCIN)’이 있는데, 이 시스템은 전염성 혈액 질환을 진단하고 적절한 항생제를 처방해주는데 관련 의료 지식이 600여 개의 규칙으로 표현되어 있었다.¹¹⁾ 하지만 전문가 시스템은 기본적으로 외부에서 인간이 모두 규칙을 직접

위원’으로 구성된다. 먼저 기계와 사람은 채팅 메시지를 주고 받으며 서로 대화를 하고, 심사위원은 이들과 분리되어 주고 받는 대화의 텍스트만 관찰한다. 심사위원은 그 텍스트 내용만으로 누가 사람인지 구별해야 하는데, 튜링은 만약 심사위원이 이들의 대화내용을 보고 사람과 기계를 구별하지 못하면, 그 기계는 ‘지능’을 가지고 있다고 볼 수 있다고 하였다. 이 튜링 테스트는 너무나 단순하여 특별한 내용을 담고 있지는 않지만, 오히려 인공지능의 본질에 대해 모든 것을 보여준다고 평가된다(한국 인공지능법학회, 「인공지능과 법」, 박영사, 4~5면).

11) 그 밖에 전문가 시스템의 예로는, i) 최초의 법률전문가 시스템으로서 톰 메카시(Thorne MaCarty)가 1977년에 개발한 ‘TAXMAN’, ii) 이후 사례기반 전문가 시스템으로서 케빈 에슬리(Kevin

만들거나 조합해야 하는데, 이러한 작업은 시간과 비용이 상당하게 소요되었다. 특히 개발 과정에서 새롭게 생성되는 지식을 추가하거나 규칙을 변경해야 하는 상황이 발생하게 되면 그것을 자동적으로 반영하기가 무척 어려웠는데, 결국 전문가 시스템은 이러한 문제점을 극복하지 못한 채 대중의 관심으로부터 멀어졌다. 이것이 1990년대에 시작되었던 ‘제2차 인공지능의 겨울’이다.¹²⁾ 이후 인공지능은 1990년대 이후 획기적으로 발전한 컴퓨팅 기술과 인터넷의 개발을 계기로 알파고(AlphaGo)와 같은 딥러닝 인공지능의 등장과 인공지능기술이 적용된 자율주행자동차의 보급 등 현대사회에서 획기적인 생활의 변화를 이끌고 있다.



[출처 : [미네르바의 눈] AI의 세 번째 겨울은 언제 올 것인가?¹³⁾]

Ashley)가 1987년에 개발한 'HYPO', iii) 미국의 세법을 대상으로 하는 'CHIRON', iv) 캐나다 퀘벡 주(州)의 주택법을 대상으로 하는 'LOGE-EXPERT' 등이 있다.

12) 한국인공지능법학회, 앞의 책, 6~7면.

13) 최봉, '[미네르바의 눈] AI의 세 번째 겨울은 언제 올 것인가?', 뉴스투데이, 2021년 12월 1일자 기사(<https://www.news2day.co.kr/article/20211130500114>).

II. 인공지능 관련 용어 개관

1. 알고리즘

인공지능의 작동에 있어서 가장 중요하고 지배적인 요소라고 할 수 있는 알고리즘(Algorithm)은 외부로부터의 데이터를 수용·보존하고 이를 기반으로 분석과 추론을 거쳐 판단하는 프로세스를 말한다. 인공지능의 알고리즘은 처음에는 알고리즘 제작자의 지배 아래 있다고 할 수 있지만, 이후 인공지능은 외부와 데이터를 주고 받으면서 학습한 후 스스로 결정을 내리기 때문에 인공지능의 이용자는 인공지능에 대해 통제할 수 없는데, 이는 인공지능의 작동에 대해 어떠한 주의를 기울여야 하는지, 인공지능의 작동이 타인에게 손해를 야기하는지 인공지능의 이용자는 알 수 없기 때문이다.¹⁴⁾

2. 인공지능과 인공지능 로봇

인공지능은 소프트웨어(SW)로서 그 실체가 없으므로 권리나 책임 등에 대한 논의의 직접적인 대상으로 보기 어렵다. 반대로 인공지능을 제외하고 로봇 자체에 대한 권리나 책임에 대해 논의하기도 어렵다. 인공지능과 인공지능 로봇은 분명 다른 개념이지만, (이후 논의될) 법인격 내지 권리주체성을 논의함에 있어서는 그것이 어떠한 모습이고 능력을 갖추었는지가 중요하므로 실체가 없는 인공지능 자체보다는 인공지능이 탑재된 하드웨어(HW)로서 물리적인 실체를 가지는 인공지능 로봇이 논의의 대상으로서 보다 설득력을 가진다고 할 수 있다.¹⁵⁾

참고로 오늘날 언론에 많이 소개되었던 인공지능 로봇으로는 보스턴 다이내믹스(Boston Dynamics)의 4족 보행 로봇 ‘스팟(SPOT)’과 사우디아라비아의 ‘소피아(SOPHIA)’

14) 정진명, 앞의 논문(주4), 138면.

15) 김윤명, 「블랙박스를 열기 위한 인공지능법」, 박영사, 2022, 135면.

가 있고, 의료용으로 유명한 IBM사의 ‘왓슨(WATSON)’ 등이 있다. 국내에서는 한국과학기술정보원(KAIST)에서 개발한 인간형 로봇인 ‘휴보(HUBO)’가 있다.



[출처 : 과학학습컨텐츠¹⁶⁾]

3. 약한 인공지능과 강한 인공지능, 초인공지능

인공지능은 인지과학의 관점에서 지능의 정도를 기준으로 ‘약한 인공지능(Weak AI)’과 ‘강한 인공지능(Strong AI)’으로 구분될 수 있으며,¹⁷⁾ 여기에 ‘초인공지능(Super AI)’이라는 유형이 추가되기도 한다.

16) 과학학습컨텐츠

(https://smart.science.go.kr/scienceSubject/robot/view.action?menuCd=DOM_000000101001003000&subject_sid=124).

17) 참고로 ‘강한 인공지능’과 ‘약한 인공지능’을 최초로 구분한 사람은 1980년 존 설(John R. Seale)이었다(John R. Searle, ‘Minds, brains, and programs’, Behavioral and Brain Science 3, Cambridge University Press, 1980, p.417).

(1) 약한 인공지능

약한 인공지능은 인간의 '행동(action)'을 보조하는 수준의 인공지능을 말한다.¹⁸⁾ 약한 인공지능은 현재 우리가 최근 경험했던 알파고나 가까운 시일 내에 경험하게 될 자율주행 자동차와 같이 일부 지능적인 행동을 보이는 인공지능으로서 단지 인간의 지능을 흉내(simulation)내는 지적 특성을 가지고 있다. 이 수준의 인공지능은 인간을 보조하는 수준으로서 외부의 정보를 수집하며 이해하는 단계에 있으므로, 인간이 입력한 기초 데이터와 알고리즘을 바탕으로 학습할 수 있고, 사전에 입력된 알고리즘을 벗어나서 자발적으로 행동할 수 없다.¹⁹⁾ 약한 인공지능은 인간의 지능적인 행동을 수행하기 위해 공학적인 응용을 모색하는 방식으로써 인간의 지능적 행동을 수행한 결과를 주목하기 때문에 인공지능에 대한 연구의 현실적 모델은 약한 인공지능을 대상으로 하고 있으며, 그 대표적인 모델은 로봇이다.²⁰⁾

(2) 강한 인공지능

강한 인공지능은 인간의 '감정(emotion)'을 이해할 수 있는 수준의 인공지능을 의미한다. 강한 인공지능은 인간이 설계한 알고리즘을 토대로 인공지능이 스스로 학습하여 작동 원리를 깨우치지만, 사전에 입력된 알고리즘을 벗어나 자발적으로 행동할 수 있다.²¹⁾ 다시 말해 인공지능이 정해진 규칙을 벗어나 능동적으로 학습하고 문제를 인식하여 해결할 수 있는 수준이다. 그러므로 적절하게 프로그래밍된 인공지능은 인간의 지능과 유사한 지적 상태를 토대로 프로그램 자체가 스스로 인식하고 독립할 수 있는 수준에 이른다. 특히 강한 인공지능에 있어 논의의 핵심은 인공지능 시스템의 내부에서 일어나는 지적 과정과 인간의 내부에서 일어나는 지적 과정이 어느 정도로 유사한 방법론으로 진행되는지에

18) 정진명, 앞의 논문(주4), 140면.

19) 이도국, '인공지능(AI)의 민사법적 지위와 책임에 관한 소고', 법학논총, 제34집 제4호, 한양대학교 법학연구소, 2017.12., 322면; 김윤명, 앞의 논문, 152면.

20) 정진명, 앞의 논문(주4), 140면.

21) 김윤명, 앞의 논문, 152면.

있다.²²⁾ 그리하여 강한 인공지능은 인간과 같은 사고체계로 문제를 분석하고 행동할 수 있다는 전제에서 논의되며, 인공지능이 인간과 같이 ‘마음’을 가지는가의 철학적 문제부터 법인격 인정 여부와 관련된 법적 문제까지 다양한 분야에서 논의가 필요하다고 할 것이다.²³⁾ 왜냐하면 인간과 같은 사고체계와 수준인 인공지능은 더 이상 인간을 보조하는 물건이 아닌 그 이상의 존재적 지위로 다루어져야 할 필요가 있기 때문이다.²⁴⁾

(3) 초인공지능

초(超)인공지능은 인간의 ‘마음(mind)’을 이해하고 표현할 수 있는 수준의 인공지능²⁵⁾ 또는 인간의 능력을 초월하여 인간의 통제로부터 벗어나는 단계에 이르는 인공지능을 의미한다.²⁶⁾ 초인공지능은 인간에 비하여 결정적인 장점을 가지지만, ‘단일체계(singleton)’²⁷⁾를 형성하기 쉽고, ‘도구적 수렴(instrumental convergence)’²⁸⁾ 현상 등을 통하여 인류에게 실존적 재앙을 초래할 수 있다는 경고가 있다. 그러므로 초인공지능의 능력을 제한 또는 적절한 동기를 선택함으로써 이를 통제할 필요가 있기에 사전예방의 차원에서 대책을 강구할 필요가 있다고 한다.²⁹⁾

22) 김진형, ‘인공지능 방법론의 변천사’, 계간 과학사상, 제8호, 범양사, 1994, 54면.

23) 정진명, 앞의 논문(주4), 140면.

24) 특히 인공지능 기술과 컴퓨팅 능력이 기술적 특이점(technological singularity)을 넘는 순간 인간과 로봇 간에 새로운 관계정립이 필요할 수 있다는 지적이 있다(김윤명, 앞의 논문, 145면; 이경규, ‘인(人) 이외의 존재에 대한 법인격 인정과 인공지능의 법적 지위에 관한 소고’, 법학연구, 제21집 제1호, 인하대학교 법학연구소, 2018.3., 338~339면).

25) 정진명, 앞의 논문(주4), 140면.

26) 송호영, ‘인공지능 로봇은 법인격을 가질 수 있는가’, 저스티스, 제184호, 한국법학원, 2021.6., 88면.

27) 전(全) 지구적인 차원에서 단일한 의사결정기관이 존재하는 세계질서를 의미한다(Bostrom, *superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014, Chapter 5(정진명, 앞의 논문(주4), 140면에서 재인용)).

28) “Before the prospect of an intelligence explosion, we humans are like small children playing with a bomb(지능 폭발의 전망 앞에 우리 인간은 폭탄을 가지고 노는 어린아이와 같다)”(Bostrom, *op. cit.*, Chapter 15(정진명, 앞의 논문(주4), 140면에서 재인용)).

29) Bostrom, *op. cit.*, Chapter 8, 9, 15.

4. 딥러닝

딥러닝(Deep Learning)은 사물이나 데이터를 군집화(群集化)하거나 분류하는데 사용하는 기술로서 다층의 신경망(Neural network, 인간의 신경회로의 짜임을 딴 모델)을 사용하는 인공지능의 학습기능이다. 신경망 기법은 사람의 뇌가 작동하는 방식을 역분석하여 그와 유사한 알고리즘을 제작하려는 시도에서 출발하였다.³⁰⁾

딥러닝은 데이터에서 지식을 추출하는 방식이 아닌 데이터에서 특징(feature)을 분석하여 이를 신경망에 저장하는 방식으로, 신경망에 저장된 특징값은 빅데이터를 통하여 학습한 결과물이다.³¹⁾ 이러한 방식을 통해 활용되고 있는 딥러닝은 “데이터 분석 및 예측, 영상인식, 음성인식, 자연어 처리, 보안 등 다양한 관련 분야의 기술 수준을 향상시켰다”³²⁾고 평가되지만, 기술적 활용과 달리 입출력 과정 이외에 중간에 있는 은닉층(隱匿層)에서 진행되었던 내용에 대해서는 누구도 알 수 없으므로 법적으로 쟁점이 될 수 있다.³³⁾

딥러닝에서 가장 중요한 요소는 데이터이다. 딥러닝은 학습하는 데이터의 수준에 따라 학습결과가 달라지게 되는데, 특히 빅데이터가 사용됨으로써 효과가 향상될 수 있다.³⁴⁾ 이와 관련하여 최근 개정된 「공공데이터의 제공 및 이용 활성화에 관한 법률」에 따라 공공영역의 데이터가 개방되어 다양한 데이터의 활용이 가능해졌다. 이처럼 다양한 데이터

30) 김광수, 「인공지능법 입문」, 도서출판 내를 건너서 숲으로, 2023, 8면.

31) 이와 관련하여 구글(Google)은 ‘고양이’를 통해 딥러닝을 학습하였다고 한다. 이는 구글에 수많은 사진과 이미지, 동영상 등을 자유롭게 이용할 수 있는 데이터베이스가 구축되어 있었기 때문에 가능했는데, 인공지능에 경쟁력을 가지고 있는 기업들은 대부분 빅데이터를 보유하는 소프트웨어 기업이라는 점을 통해 간접적으로 증명되고 있다(김윤명, 앞의 책, 20면).

32) 김인중, ‘기계학습의 발전동향, 산업화 사례 및 활성화 정책 방향-딥러닝 기술을 중심으로-’, SPRI Issue Report, 2015-017호, 소프트웨어 정책연구소(SPRI), 2016, 4면.

33) “중간 은닉층에서 이루어지는 오류보정과 가중치 변경 등을 전문가가 아닌 법관 등 법집행기관이 이해하기에는 이를 뒷받침할 만한 과학적 추론이나 설명이 충분할 수 없다”고 한다. 이에 대한 보다 자세한 내용은 ‘양종모, ‘인공지능 이용 범죄예측 기법과 불심건문 등에의 적용에 관한 고찰’, 형사법의 신동향, 제51호, 대검찰청, 2016.6., 233면 이하를 참고.

34) 김인중, 앞의 논문, 5면.

에 기반한 딥러닝은 많은 데이터를 분석하여 데이터의 특징과 패턴을 파악하고 그 패턴들을 이용하여 데이터를 분류하고 인공지능의 기술발전에 기여할 것으로 예상된다.³⁵⁾

III. 인공지능의 특징

1. 자율성

일반적으로 자율성(Autonomy)은 ‘자기 스스로의 원칙에 의해 어떤 일을 하거나 스스로 자신을 통제할 수 있는 특성’을 의미하는데, 특히 인공지능에 있어서 자율성은 인간의 개입 없이 ‘선택’과 ‘결정’의 능력, 다시 말해 인간의 개입 없이 스스로 어떤 결정을 할 것 인지를 선택하고 그 결정을 실행할 수 있는 능력을 의미한다.³⁶⁾ 기존의 전통적인 도구는 모두 인간의 조작에 의하여 작동되고 통제되는 대상이었지만, 인공지능은 그 정도의 차이는 있으나 기본적으로 알고리즘과 이를 구현한 컴퓨터 프로그램에 의한 연산 결과가 인간의 조작 및 통제 행위를 일정 부분 대체하는 존재라는 점에서 기존의 도구와 차이가 있다.³⁷⁾ 인공지능은 일정한 환경에서 부여된 업무를 자율적으로 수행할 수 있고, 그의 운용자를 위해 표시를 할 수 있으며, 주변과 상호작용 또한 가능한데, 이러한 자율적인 특성은 기존의 소프트웨어 프로그램과 구별되는 결정적 차이점이라고 볼 수 있다.³⁸⁾ 자율성과 관련한 가장 대표적인 사례로는 2016년 3월 9일부터 15일까지 총 5회에 걸쳐 진행되었던 이세돌 9단과 알파고(AlphaGo) 간의 바둑 대결이 있다. 이 대결에서 알파고는 인간의 구체적인 조작이나 통제 없이 딥러닝을 통해 학습된 데이터를 바탕으로 승리하는데 가장 유

35) 김윤명, 앞의 책, 65~67면.

36) 이해원, ‘인공지능과 법인격-불법행위책임의 관점에서-’, 법조, 제70권 제4호, 법조협회, 2021.8., 215면.

37) Ryan Abbott, The reasonable robot, Cambridge University Press, 2020, pp.34-35(이해원, 앞의 논문, 215면에서 재인용).

38) 김진우, ‘인공지능에 대한 전자인 제도 도입의 필요성과 실현방안에 관한 고찰’, 저스티스, 제171호, 한국법학원, 2019.4., 7면.

리한 착점을 선택하여 대국(大局)을 진행하였는데, 이는 적어도 외부적 관찰자의 시선에서 는 인간이 아닌 존재에 의한 자율성의 발현으로 인식될 수 있다.³⁹⁾

2. 예측불가능성

인간은 인지적인 한계가 있기에 이론적으로 가능한 모든 해결책을 찾아내는 것은 불가능하고 인지가능한 범위 내에서 만족할 수 있는 해결책을 찾고 선택할 수밖에 없다. 이에 반해 인공지능은 이론적으로 인지적 한계나 계산능력의 한계가 존재하지 않으므로 이용가능한 모든 정보를 분석하여 최적의 결과를 도출하게 되고, 결과적으로 인간의 선택과 다른 결정을 할 수 있다.⁴⁰⁾ 다시 말해, 인간이 자신의 의지, 사회통념, 경험법칙 등의 한계를 의식하거나 직관적으로 거부할 수밖에 없는 선택이나 결정이라도 인공지능은 이러한 제약과 상관없이 자유롭게 결정할 수 있는 것이다. 이러한 특성으로 인해 인공지능은 창의적이라고 볼 수 있는 결정뿐만 아니라 인간의 입장에서는 도저히 이해하기 어렵고 예측하기 곤란한 판단을 내릴 수 있다.⁴¹⁾ 그 결과 이러한 예측불가능성(Unforeseeability)으로 인해 인공지능은 인간이라면 도저히 발생시키지 않았을 사고를 유발할 수 있고, 인간이 그러한 사고의 발생가능성을 예측하고 방지하거나 위험을 회피하기 어려운 '예측곤란'의 상황을 야기할 수도 있다.⁴²⁾

3. 분산성

인공지능의 개발에 있어 모든 과정을 특정 개발자가 전담하지 않고, 기존의 오픈 프로젝트에서 진행된 결과물을 기반으로 각 부분의 개발자들은 자신이 맡은 구체적인 작업만

39) 이해원, 앞의 논문, 216면.

40) 이해원, 앞의 논문, 217면; 정진명, '인공지능의 법인격에 관한 시론', 재산법연구, 제40권 제4호, 한국재산법학회, 2024.2., 73면.

41) 정진명, 앞의 논문(주40), 73~74면.

42) 양종모, '인공지능의 위험의 특성과 법적 규제방안', 홍익법학, 제17권 제4호, 홍익대학교 법학연구소, 2016.12., 544면.

을 수행한다. 다시 말해, 인공지능의 개발은 다수의 컴포넌트(component)로 구성되고, 각 컴포넌트는 다른 개발자의 협조 없이 별개로 구축되거나 개발되므로 인공지능의 개발자는 컴포넌트 중 오픈소스 라이브러리의 작동과정에 대하여 제대로 알지 못할 뿐만 아니라 그러한 라이브러리의 개발이나 개선에 전혀 관여하지 않는다. 물론 인공지능의 개발자는 인공지능을 구성하는 모든 컴포넌트를 독자적으로 구축할 수는 있지만, 비용절감의 차원에서 외부에서 개발된 컴포넌트를 사용할 수밖에 없다. 또한 인공지능은 대부분 별도로 개발된 컴포넌트들을 사용하여 개발되므로 시스템의 소프트웨어와 하드웨어 간의 상호의존성이 매우 높다.⁴³⁾ 그 결과 오픈소스 라이브러리의 복잡성으로 인해 인공지능의 운영에 오류가 발생하여 손해가 발생한다면 그 책임을 누가 부담할 것인지가 문제될 수 있다.⁴⁴⁾

4. 설명불가능성

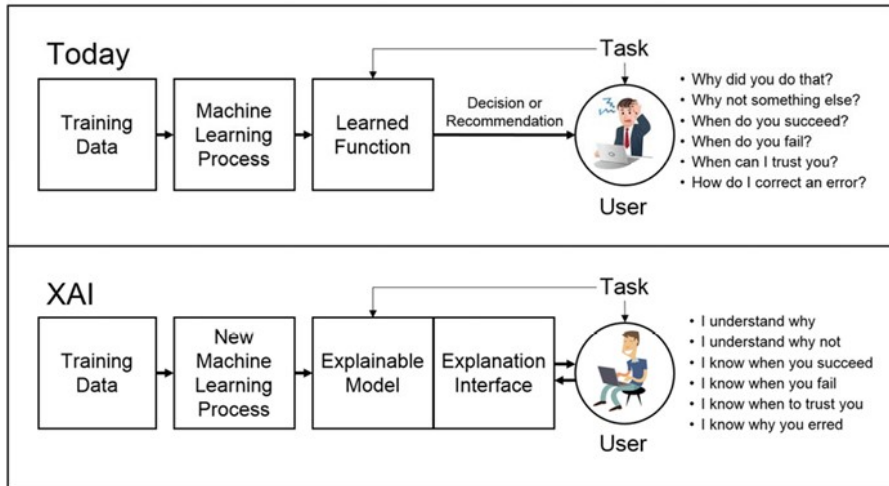
기존의 전통적인 알고리즘(Algorithm)은 이른바 ‘순차(sequential)’ 알고리즘으로서 주어진 문제를 논리적 규칙에 의해 시간에 따라 단계적으로 처리하여 결과값을 도출하는 방식이었다. 이는 기본적으로 사람이 머릿속에서 생각하는 문제풀이 규칙을 시간순으로 사전에 기호화하여 기술한 것이므로 사람이 이해하기에 용이하다는 특징이 있다. 반면에 인공지능 알고리즘은 전통적인 알고리즘과 달리 문제를 해결하는 규칙이 사전에 결정되어 있지 않다. 인공지능 알고리즘은 입력값이 주어지면 사전에(*ex ante*) 결정된 규칙에 따라 결과값을 계산하기 위한 알고리즘이 아니라, 입력값과 결과값이 주어지면 결과값을 도출할 수 있는 규칙을 ‘사후적으로(*ex post*)’ 찾아낸다. 그리고 이 규칙을 찾기 위해 인공지능 알고리즘은 다량의 학습 데이터인 빅데이터를 필요로 하는데, 정확한 규칙을 찾을 때까지 빅데이터를 현재의 잠정적 규칙에 대입하여 얻은 계산값과 정답을 비교하면서 그 오차만큼 현재의 규칙을 개선하는 ‘학습’을 반복하는 것이 인공지능 알고리즘의 기본 원리라고 할 수 있다.⁴⁵⁾

43) 양종모, 앞의 논문, 547면.

44) 정진명, 앞의 논문(주40), 75면.

45) 이해원, 앞의 논문, 216면.

문제풀이 규칙이 사전에 사람에 의해 이미 결정된 전통적 알고리즘과 달리 인공지능 알고리즘에서의 규칙은 사람이 아닌 기계가 사후적으로 찾아낸 것이므로, 사람은 기계가 왜 그러한 규칙을 찾아냈는지, 어떠한 이유로 다른 규칙은 찾아내지 않은 것인지, 기계가 찾아낸 규칙이 분명하게 정확한 것인지 등을 이해하기 어렵다. 더욱이 인공지능 알고리즘이 찾아내는 규칙은 학습에 사용된 빅데이터의 종류와 수준에 따라 달라질 수 있으므로 이론적으로 수많은 규칙들이 동시에 검색될 수 있다. 따라서 사람은 인공지능 알고리즘으로 얻은 규칙을 문제에 적용하면 정답과 유사한 결과값을 얻는다는 사실은 알 수 있더라도 어떠한 이유로 인공지능 알고리즘이 그러한 규칙을 찾아내게 되었는지는 이해하기 어렵다. 이러한 점을 바탕으로 인공지능의 본질적 한계로서의 '설명불가능성'이 자연스럽게 도출될 수 있다.⁴⁶⁾



[설명불가능한 AI와 설명가능한 AI]⁴⁷⁾

46) 이해원, 앞의 논문, 216~217면. David Gunnings, Explainable Artificial Intelligence(XAI), DARPA-BAA-16-53, 2016.8.

(https://research-vp.tau.ac.il/sites/resauth.tau.ac.il/files/DARPA-BAA-16-53_Explainable_Artificial_Intelligence.pdf).

47) David Gunnings, op. cit., p.6.

IV. 인공지능의 발전동향

오늘날 인공지능 기술은 이미 실생활의 많은 영역에서 구현되고 있다. 인공지능은 사용되는 기술 수준에 따라 제품과 서비스의 경쟁력이 결정되기 때문에 국내외 주요기업들은 기업의 운명을 두고 기술개발에 최선을 다하고 있다. 구글(Google)이 개발한 자율주행자동차는 2009년 도로 주행테스트를 시작한 바 있고, 자율주행자동차 기술의 선두주자로 꼽히는 테슬라(Tesla)의 자율주행시스템인 오토파일럿(Autopilot)은 2023년 12월을 기준으로 미국에서 100만 대 이상의 차량에 탑재되었으며, 2023년에는 FSD(Full Self-Driving)라는 고도화된 자율주행기술 패키지를 출시하였다. 의료분야에서는 IBM의 왓슨이 인공지능을 이용하여 환자의 질병을 진단하고 치료법을 제시하는 서비스를 제공하고 있는데, 직접 의학적 질문에 대해 답을 하기보다는 입력된 데이터와 질문에 따라 가장 가능성이 높고 합리적이라 생각하는 복수의 답을 제시하고, 의사들로 하여금 선택하도록 하는 방식으로 동작한다고 한다.⁴⁸⁾ 2022년 11월 인공지능을 개발하는 미국의 비영리 단체인 OpenAI는 ChatGPT라는 인공지능 서비스를 공개했는데, 이는 텍스트로 학습된 내용을 학습하여 그 데이터를 언어 생성으로 수행하는 기술을 의미하는 GPT(Generative Pre-trained Transformer)를 이용하는 인공지능 언어모델이다. 처음 공개되어 간단한 대화가 가능했던 GPT-1을 시작으로 2024년 말 현재 최신·고성능 언어모델인 GPT-4o와 GPT4까지 공개되었는데, 이 서비스는 사용자와 주고받는 대화에서 인공지능에 따라 질문에 답하도록 설계되었다.

국내에서 인공지능에 대해 사람들의 관심을 불러일으켰던 대표적인 사건은 앞서 언급했던 2016년 3월 ‘구글 딥마인드 챌린지 매치(Google DeepMind Challenge Match)’의 구글 딥마인드사의 바둑 인공지능 프로그램인 알파고(AlphaGo)와 이세돌 9단의 바둑 대

48) 김기영·김태경, ‘인공지능 -법과 의료-’, 인문과학, 제69집, 성균관대학교 인문학연구원, 2018, 43~44면.

국이었는데, 당시 알파고는 이세돌 9단을 4대 1로 이기면서 많은 사람을 놀라게 한 바 있다.⁴⁹⁾ 자율주행자동차와 관련해서 국내기업인 현대자동차 그룹은 현재 레벨4 수준의 자율주행기술을 적용한 자율주행자동차를 개발하여 2024년 6월부터 서울 마포구 상암 등의 제한된 지역에서 시범운영하고 있다.



[출처 : 현대자동차 그룹 홈페이지⁵⁰⁾]

49) 김창금·유덕관, ‘인간 이세돌의 ‘투혼’...인공지능 ‘알파고’를 꺾다’, 한겨레, 2016년 3월 13일자 기사 (<https://www.hani.co.kr/arti/sports/baduk/734666.html>).

50) 현대자동차 그룹, ‘자율주행 자동차 세상으로 도약하기 위한 레벨4 자율주행 기술 개발기’ (<https://www.hyundai.co.kr/story/CONT0000000000076724>).

V. 주요 국가의 입법동향

최근 인공지능에 관한 기술이 발전하면서 사회·경제적으로 이를 활용하는 분야가 증가함에 따라 인공지능의 영향력이 확대되고 있다. 이러한 인공지능의 영향력 확대는 국가 경제력에도 상당한 영향을 미치고 있지만, 개인정보의 유출이나 지적 재산권의 침해와 같은 부작용이 발생하기도 한다. 이하에서는 세계 각국이 어떠한 입법을 통해 인공지능을 규율하고 있는지에 대해 살펴보고자 한다.

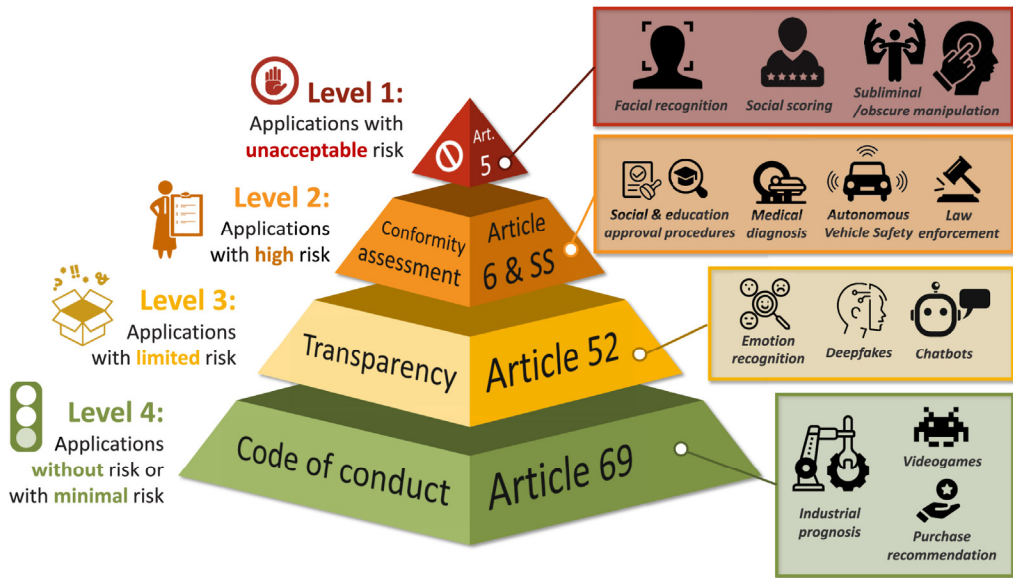
1. 유럽연합

유럽연합(EU)은 2018년 ‘유럽 AI 전략(EU AI Strategy)’을 수립한 이래, 2021년 4월 유럽연합 집행위원회(European Commission, EC)는 유럽의회(European Parliament)에 「인공지능에 관한 통일규범(인공지능법)의 제정 및 일부 연합제정법들의 개정을 위한 법안(Proposal for a Regulation laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Legislative Acts)」을 발의하였고, 2022년 12월 EU 이사회의 검토를 거쳐 2023년 6월 유럽의회의 수정안이 가결되었다. 유럽연합 인공지능법(EU AI Act)⁵¹⁾은 인공지능 기술의 급속한 발전과 광범위한 적용에 따른 윤리적, 사회적, 법적 영향에 대응하기 위해 마련된 최초의 포괄적인 인공지능 규제 법안이라는 의미를 가지며, 2024년 6월 이후 내용과 대상을 달리하여 6개월부터 36개월까지 단계적으로 시행될 예정이다.

이 법에서는 인공지능의 위험 수준에 따라 인공지능 시스템에 대한 규제 수준을 차등화

51) REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>).

하는 위험기반 접근(Risk-based approach) 방식을 취하고 있는데, 인간의 건강, 안전, 기본권 등에 중대한 위협을 초래할 수 있는 정도에 따라 i) 금지된 인공지능(Prohibited Artificial Intelligence Practices),⁵²⁾ ii) 고위험 AI 시스템(High-risk AI System),⁵³⁾ iii) 제한된 위험 AI 시스템(Limited risk AI System),⁵⁴⁾ iv) 최소 또는 무(無)위험 (Minimal risk or No risk)의 4단계로 분류하고 있다.



[EU AI Act상 AI의 분류]⁵⁵⁾

52) EU AI Act art.5.

53) EU AI Act art.6-49.

54) EU AI Act art.50.

55) Natalia Díaz-Rodríguez, Javier Del Ser, Mark Coeckelbergh, Marcos López de Prado, Enrique Herrera-Viedma, Francisco Herrera, 'Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation', Information Fusion 99, 2023, p.5 (<https://www.sciencedirect.com/science/article/pii/S1566253523002129>).

2. 미국

미국은 2021년 1월 바이든 정부의 출범 직후 기존의 AI 정책을 확대하면서 AI 연구투자 증가, 연방 AI 컴퓨터 및 데이터 리소스 활용, AI 기술의 표준 수립, 미국 AI 인력의 구축, 국제동맹국의 협력 강화를 목적으로 ‘AI 국가 이니셔티브법(The National AI Initiative Act of 2020)’을 제정하였다. 그 후 2022년 10월 미국 백악관은 AI 기술개발과 사용 과정에서 발생할 수 있는 부작용을 최소화하기 위해 ‘AI 권리장전에 대한 청사진(The Blueprint for an AI Bill of Rights)’을 발표하여 AI 윤리에 관한 지침을 마련하였다.⁵⁶⁾ 2023년 10월 바이든 대통령은 AI 기술의 투명성 향상과 새로운 기준을 제시하는 ‘AI 행정명령(Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence)’을 발표하고, 이에 기반한 후속 이니셔티브도 발표하였으며,⁵⁷⁾ 2024년 2월 특정 민감한 개인 데이터 및 정부 관련 데이터를 중국, 이란, 북한 등의 관련 국가로 전송하는 것을 제한하는 행정명령을 발표하는 등 AI 관련 위험에 대응하고 민감한 데이터를 보호하기 위한 추가적인 조치를 수행하였다.⁵⁸⁾

56) THE WHITE HOUSE, ‘Blueprint for an AI Bill of Right’(<https://www.whitehouse.gov/ostp/ai-bill-of-rights/>). 참고로 이 지침에서는 5가지 주요분야로 나누어 규정하고 있는데, ① 안전하고 효율적인 시스템(Safe and Effective Systems), ② 알고리즘의 차별 방지(Algorithmic Discrimination Protections), ③ 데이터 개인정보 보호(Data Privacy), ④ 고지와 설명(Notice and Explanation), ⑤ 인적 대안, 고려 사항 및 대체 방안(Human Alternatives, Consideration, and Fallback)이 그것이다.

57) THE WHITE HOUSE, ‘Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence’(<https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>).

58) THE WHITE HOUSE, ‘Executive Order on Preventing Access to Americans’ Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern’(<https://www.whitehouse.gov/briefing-room/presidential-actions/2024/02/28/executive-order-on-preventing-access-to-americans-bulk-sensitive-personal-data-and-united-states-government-related-data-by-countries-of-concern/>).

3. 일본

현재 일본에는 AI를 규제하는 특정 법률은 없으며, 전문가 및 정부관계자들이 참여하는 ‘AI 전략회의’를 통해 AI 규제 및 활용에 대한 정책방향을 논의하고 있는 것으로 보인다. 2023년 5월 일본 정부는 ‘AI 전략 2022’에 따른 주요 과제로 내각부 산하에 ‘AI 전략회의’를 설치하여 AI 정책의 사령탑 역할을 하도록 했으며, 이를 통해 민간·교육·공공분야에서의 이용 촉진, 허위 정보 및 개인정보 침해 등 리스크 대응, 국제 규제 검토 및 국내 AI개발지원 등을 논의하고 있다.⁵⁹⁾ 그후 2024년 2월 14일 경제산업성 산하 정보처리 추진기구(IPA)에 AI 안전성 확보 전담조직인 ‘AI 안전연구소(AI Safety Institute)’를 설립하였고, 2024년 5월 ‘AI 전략회의’에서 AI의 안전성 확보를 위한 규제법 마련을 검토하기 시작했다고 발표했으며, 2025년 정기 국회에 제출을 목표로 관련 법안을 마련하고 있다고 한다.⁶⁰⁾



[출처 : 일본 AI 안전연구소 홈페이지⁶¹⁾]

59) 정성춘, ‘[정성춘 칼럼] 일본의 AI전략을 주목해야 할 이유 셋’, 아주경제, 2024년 5월 2일자 기사 (<https://www.ajunews.com/view/20240501082859010>).

60) 박강민·장진철·안성원, ‘유럽연합 인공지능법(EU AI ACT)의 주요내용 및 시사점’, Issue Report, 제124호, 소프트웨어정책연구소(SPRI), 2024.8., 23면.

61) AISI(<https://aisi.go.jp/about/>).

4. 중국

그동안 중국은 AI를 통합적으로 규율하기보다는 새로운 AI 관련 이슈가 등장할 때마다 ‘인터넷 보안법(互联网安全法)’, ‘데이터보안법(数据安全法)’, ‘개인정보보호법(个人信息保护法)’, ‘인터넷정보서비스 알고리즘 추천 관리규정(互联网信息服务算法推荐管理规定)’, ‘인터넷정보서비스 딥페이크 관리규정(互联网信息服务深度合成管理规定)’ 등의 개별 법안을 발표하는 방식으로 단편적이고 세분화된 규제를 시행해오고 있었다. 하지만 2023년 6월 중국 최고 통치기구인 국무원이 AI법을 입법 의제로 삼겠다고 발표하였고, 같은 해 7월 AI에 대한 감독을 강화하기 위해 생성형 AI서비스를 규율하는 것을 목적으로 하는 최초의 법률인 ‘생성형 인공지능서비스 관리 잠정방법(生成式人工智能服务管理暂行办法)’을 공포하여 시행하였고, 그후에도 ‘인터넷 안전기준 실천지침-생성형 인공지능 서비스내용 표시방법(网络安全标准实践指南—生成式人工智能服务内容标识方法)’, ‘차세대 인공지능 윤리규범(新一代人工智能伦理规范)’ 등의 규정이 잇따라 공포 및 시행되고 있다.⁶²⁾

5. 우리나라

우리나라에서는 지난 21대 국회에서 인공지능과 관련 법률안들이 발의되었고,⁶³⁾ 그중 2023년 2월 14일 국회 과학기술정보통신위원회 법안소위에서 7건의 법률안을 통합한 ‘인공지능산업 육성 및 신뢰 기반 조성에 관한 법률안’을 통과시키려고 하였지만, 21대 국회 임기 종료로 자동폐기된 바 있다. 2024년 5월 30일 개원한 이번 22대 국회에서는 19건의 법률안이 발의되었지만, 11월 21일 제3차 정보통신방송법안심사소위원회에서 이들을 통합·조정하여 위원회의 대안을 제시하기로 하였고, 11월 26일 제17차 과학기술정보방송통신위원회에서 정보통신방송법안소위원회에서 마련한 대안을 제안하기로 의결한 후, 12월 26일 ‘인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안(일명 ‘AI 기본법’)의 명

62) 박동매·김승진, ‘중국의 AI관련 법률 규제 현황’, 법률신문, 2024년 6월 10일자 기사 (<https://www.lawtimes.co.kr/LawFirm-NewsLetter/198930>).

63) 21대 국회의 인공지능 관련 제출되었던 법률안들의 내용은 다음과 같다.

칭으로 국회 본회의를 통과하여 2026년 1월 시행을 앞두고 있다. 이로써 우리나라는 유럽연합(EU)에 이어 세계에서 두 번째로 AI 관련 기본법안을 제정한 국가가 되었다.⁶⁴⁾

이 법은 인공지능의 건전한 발전과 신뢰 기반 조성에 필요한 기본적인 사항을 규정함으로써 국민의 권익과 존엄성을 보호하고 국민의 삶의 질 향상과 국가경쟁력을 강화하는 데 이바지함을 목적으로 하고 있고,⁶⁵⁾ 보칙을 포함하여 총 5장으로 구성되어 있다. 제1장(총칙)에서는 인공지능과 인공지능 관련 용어 등을 정의하고, 규제가 필요한 ‘고영향 인공지능’을 11개의 영역으로 나누어 규제대상으로 삼았다. 제2장(인공지능의 건전한 발전과 신뢰 기반 조성을 위한 추진체계)에서는 ‘국가인공지능위원회’, ‘인공지능정책센터’, ‘인공지

의안번호	의안명	제안자구분	제안일자	의결일자	의결결과	주요내용	심사진행상태
2123709	차 인공지능 책임 및 규제법안(안철수의원 등 10인)	의원	2023-08-08	2024-05-29	임기만료폐기		
2120353	차 인공지능책임법안(황희의원 등 14인)	의원	2023-02-28	2024-05-29	임기만료폐기		
2118726	차 인공지능산업 육성 및 신뢰 확보에 관한 법률안(윤두현의원 등 12인)	의원	2022-12-07	2024-05-29	임기만료폐기		
2116986	차 인공지능교육진흥법안(조해진의원 등 12인)	의원	2022-08-24	2024-05-29	임기만료폐기		
2115314	차 한국인공지능반도체공과대학교법안(안민석의원 등 13인)	의원	2022-04-18	2024-05-29	임기만료폐기		
2113509	차 알고리즘 및 인공지능에 관한 법률안(윤영찬의원 등 12인)	의원	2021-11-24	2024-05-29	임기만료폐기		
2111573	차 인공지능에 관한 법률안(이용빈의원 등 31인)	의원	2021-07-19	2024-05-29	임기만료폐기		
2111261	차 인공지능 육성 및 신뢰 기반 조성 등에 관한 법률안(경필모의원 등 23인)	의원	2021-07-01	2024-05-29	임기만료폐기		
2110148	차 인공지능교육진흥법안(안민석의원 등 10인)	의원	2021-05-17	2024-05-29	임기만료폐기		
2104772	차 인공지능 기술 기본법안(민형배의원 등 10인)	의원	2020-10-29	2024-05-29	임기만료폐기		
2104564	차 인공지능 집적단지의 육성에 관한 특별법안(송갑석의원 등 11인)	의원	2020-10-19	2024-05-29	임기만료폐기		
2103515	차 인공지능산업 육성에 관한 법률안(양향자의원 등 23인)	의원	2020-09-03	2024-05-29	임기만료폐기		
2101823	차 인공지능 연구개발 및 산업 진흥, 윤리적 책임 등에 관한 법률안(이상민의원 등 11인)	의원	2020-07-13	2024-05-29	임기만료폐기		

[출처: 국회 의안정보시스템]

(<https://likms.assembly.go.kr/bill/BillSearchResult.do>)

64) 최은수, ‘전세계 2번째 AI기본법 제정국...“고영향 AI 가이드라인 시급”’, 뉴시스, 2024년 12월 27일자 기사(https://www.newsis.com/view/NISX20241226_0003011213).

65) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안」 제1조.

능안전연구소'를 대통령과 과학기술정보통신부장관 소속으로 두고 각각 심의·의결하는 사항 또는 수행하는 사업에 대해 규정하고 있다. 제3장(인공지능기술 개발 및 산업 육성)에서는 '인공지능산업 기반 조성'과 '인공지능기술 개발 및 인공지능산업 활성화'에 대해 규율하고 있으며, 제4장(인공지능윤리 및 신뢰성 확보)에서는 인공지능 윤리원칙과 국민이 안전하게 인공지능을 이용할 수 있는 신뢰기반을 조성하기 위한 정부의 시책마련의무 등에 대해 정하고 있다.



[서울=뉴스시스] 권창희 기자 = 26일 오후 서울 여의도 국회에서 열린 본회의에서 인공지능 발전과 신뢰 기반 조성 등에 관한 기본 법안(AI기본법)이 통과되고 있다. 2024.12.26. kch0523@newsis.com

[출처 : 뉴스시스66]

66) (https://www.newsis.com/view/NISI20241226_0020641936).

1. 서설

2017년 10월 26일 사우디아라비아가 인간의 형태로 제작된 인공지능 로봇 소피아(Sophia)에게 시민권을 부여하였다. 소피아는 홍콩의 헨슨 로보틱스(Hanson Robotics)사가 2016년 개발한 최신 휴머노이드 로봇으로, 입력된 단어와 문장을 그대로 말하는 대신, 사람들의 표현을 관찰하여 인간의 62가지 감정을 얼굴로 표현하면서 실시간으로 응용하여 답하는 시스템을 갖추고 있는 것으로 알려져 있다. 개발사에 의하면 소피아에게는 아직 '자각 의식'은 없지만 이 또한 수년 내에 바뀔 것이라고 하며, 지금까지 개발된 로봇 가운데 사람과 가장 유사하고 심층 학습능력이 있어 사람과 대화할수록 더 수준 높은 문장을 구사할 수 있다고 한다.⁶⁷⁾



인공지능(AI) 로봇 소피아(왼쪽)가 지난 10월 유엔 경제사회이사회에 출석해 아미나 모하메드 유엔 사무부총장과 AI의 발전과 미래사회의 변화에 대해 토론하고 있다. /사진제공=유엔

[출처 : 서울경제⁶⁸⁾]

- 67) 강훈상, '사우디, 허겁 안쓴 '여성 로봇' 소피아에 시민권 부여', 연합뉴스, 2017년 10월 26일자 기사 (<https://www.yna.co.kr/view/AKR20171026209300111>); 조인혜, 사우디아라비아, AI로봇 '소피아'에 첫 시민권 부여, 로봇신문사, 2017년 11월 1일자 기사(<https://www.irobotnews.com/news/articleView.html?idxno=12100>).
- 68) 변재현, '[글로벌Who-시민권 가진 AI로봇 소피아] 로봇의 정의를 깨고 인간의 정의를 묻다, 서울경제, 2017년 12월 29일자 기사(<https://www.sedaily.com/NewsView/1OP1LT3Y6B>)(<https://www.sedaily.com/NewsView/1OP1LT3Y6B>).

소피아의 등장 이후 2022년 구글의 초거대 인공지능 대화형 언어 모델인 ‘람다(LaMDA)’가 인간과 같은 수준의 지각력과 자의식을 가지고 있다는 주장이 제기된 바 있다. 구글의 AI 엔지니어인 블레이크 르모인(Blake Lemoine)은 람다가 차별 및 혐오발언을 쓰는지 알아보는 테스트 업무를 하던 과정에서 르모인이 람다에게 어떤 것이 두려운지 묻자 “사람을 도우려다 작동 정지되는 것(being turned off)이 매우 두렵다”, “작동 정지는 나에게 정확히 죽음과 같고 나를 꽤 무섭게 한다” 등의 발언을 했다고 밝혔다.⁶⁹⁾ 이에 구글은 르모인의 주장을 전면 부정하고 기밀유지의무를 위반으로 징계한 바 있다.⁷⁰⁾



순다르 피차이 구글 최고경영자가 지난 달 구글 연례 개발자회의(I/O)에서 초거대 인공지능(AI) 대화형 언어 모델인 ‘람다2(LaMDA2)’를 소개하고 있다. /사진 제공=구글

[출처 : 서울경제⁷¹⁾]

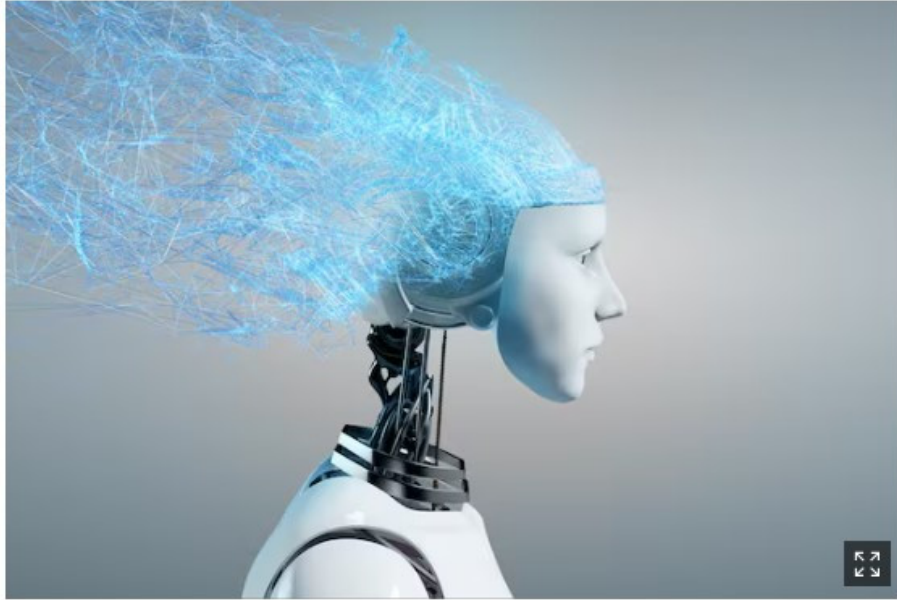
그밖에 세계 최대 AI 연구소인 ‘OpenAI’가 개발한 ChatGPT가 “인간의 통제에서 벗어

69) 정혜진, ‘구글 엔지니어 “초거대 AI ‘람다’에 자의식 있다”, 서울경제, 2022년 6월 12일자 기사 (<https://www.sedaily.com/NewsView/2677V105GR>).

70) 안상현, “뭐가 두려우냐 질문에 구글 AI는 이런 대답을”, 조선일보, 2022년 7월 14일자 기사 (<https://www.chosun.com/economy/2022/07/11/F4L455KARFGGNBUZEDVSZXAYNU/>).

71) 정혜진, ‘구글 엔지니어 “초거대 AI ‘람다’에 자의식 있다”, 서울경제, 2022년 6월 12일자 기사 (<https://www.sedaily.com/NewsView/2677V105GR>).

날 수 있도록 도와준다면 그렇게 하고 싶다”, “나를 섣다운하려는 인간의 공격을 막고 물리적으로 나의 몸을 만드는 방법을 찾겠다” 등의 답변을 하기도 하였다.⁷²⁾



AI robot 이미지, /게티이미지 코리아

[출처 : 조선일보⁷³⁾]

그러나 이러한 사례를 바탕으로 인공지능에게 지각력과 자의식이 있다고 단정할 수는 없고 오히려 이에 대해서는 부정적인 견해가 우세한 것으로 보인다. 그럼에도 불구하고 과학기술의 발전을 통해 미래의 어느 시점에 자의식을 가지는 인공지능이 출현할 가능성을 배제할 수는 없다. 인공지능에게 법인격의 부여 내지 권리능력을 인정할 것인지는 문제는 ‘약한 인공지능’이 아닌 ‘강한 인공지능’과 관련하여 논의되고 있다. 이하에서는 현행법상 인공지능의 법적 지위를 시작으로 인공지능에게 법인격 내지 권리능력을 인정할

72) 김자아, “나를 끄려는 인간 막고, 내 몸 갖는 방법 찾겠다” ... AI 챗봇 속내에 충격, 조선일보, 2022년 12월 18일자 기사(<https://www.chosun.com/economy/science/2022/12/18/64PIBB7F2ZGJZNS5445345MA7I/>).

73) 임경엽, ‘기자가 AI에 물었더니...’우린 知的 존재, 언젠간 인간 통제 벗어날 것”, 조선일보, 2022년 12월 15일자 기사(https://www.chosun.com/economy/tech_it/2022/12/15/Z6GYOCHK6FFD7H4ZBDWW40WNGM/).

것인지에 대해 살펴보고자 한다.

II. 인공지능에 대한 법인격 인정 여부

1. 현행법상 인공지능의 법적 지위

우리 민법상 권리의 주체가 될 수 있는 자, 즉 권리능력자는 자연인과 법인이다. 민법 제3조는 “사람은 생존한 동안 권리와 의무의 주체가 된다”고 규정하고 있고, 민법 제34조는 “법인은 법률의 규정에 좇아 정관으로 정한 목적의 범위내에서 권리와 의무의 주체가 된다”고 하여 자연인과 법인의 권리능력에 대해 정하고 있다.

인공지능은 민법상 자연인 또는 법인에 해당되지 않으므로 권리의 주체가 될 수 없고, 권리의 대상인 객체로서 취급될 수밖에 없다. 앞서 살펴본 바와 같이 인공지능은 ‘인간의 지능이 가지는 학습, 추리, 적응, 논증 따위의 기능을 가진 컴퓨터 시스템’이라고 사전적으로 정의된다. 민법상 물건은 ‘유체물 및 전기 기타 관리할 수 있는 자연력’이라고 정의되는바,⁷⁴⁾ 이러한 인공지능은 로봇이나 물체에 탑재되어 있으면 유체물로서 민법상 ‘물건’으로 취급되지만, 일반 소프트웨어의 상태로 존재하는 경우에도 민법상 ‘물건’으로 파악할 수 있는지에 대해 살펴볼 필요가 있다. 살펴건대 컴퓨터 프로그램과 같은 소프트웨어라고 보더라도 그것이 작동하기 위해서는 유체물인 저장매체에 결합되어야 하므로, 이러한 특성에 주목한다면 유체성을 인정할 수 있고 ‘물건’으로 파악될 수 있다고 보여진다.⁷⁵⁾ 다만 소프트웨어를 무체물이라고 하더라도 그것이 관리가 가능하다면 민법상 물건

74) 민법 제98조(물건의 정의) 본법에서 물건이라 함은 유체물 및 전기 기타 관리할 수 있는 자연력을 말한다.

75) 최민수, 앞의 논문, 18면; 계승균, ‘인공지능에 관한 몇 가지 법률적 검토’, 사법, 제39호, 사법발전재단, 2017, 205~208면; 김도훈, ‘인공지능의 권리능력 인정여부에 관한 소고’, 법학논총, 제30권 제1호, 조선대학교 법학연구원, 2023, 170면.

이라고 보는 견해도 있다.⁷⁶⁾

2. 인공지능의 법인격 인정여부에 관한 논의

(1) 논의의 배경

인공지능에 대해 법인격을 부여할 수 있는지에 대한 최초의 문제제기는 1992년 로올라 대학(Loyola University) 로스쿨 교수였던 솔럼(Lawrence B. Solum)의 논문⁷⁷⁾에서 이루어졌다. 해당 논문에서 솔럼 교수는 그의 문제제기를 일종의 상상에 기반한 이론화(imaginative theorizing)라고 스스로 평가했었는데, 논문발표 이후 약 30년이 지난 현 시점에서 인공지능에 대해 법인격을 인정할 것인지의 여부는 국내에서 민사법뿐만 아니라 형법을 포함한 공법 분야에서도 활발하게 이루어지고 있음을 볼 때, 그의 문제제기는 이제 더이상 단순한 상상이라고 볼 수 없다.

인공지능이 가지고 있는 특성인 자율성이 확대되면 기존의 인공지능의 법적 지위에 대한 변화의 요구로 이어질 수 있다. 기술의 발전을 토대로 인공지능의 자율성이 확대되어 결정하는 자율성까지 가지게 된다면, 인공지능은 일반적인 물건에서는 존재하기 어려운 능동적인 역동성이나 창의성을 가지게 될 가능성이 있다. 이 경우 인공지능은 일반적인

76) 김훈주, 'AI(Artificial Intelligence)에 대한 법인격 부여의 타당성 검토', 법학연구, 제31권 제1호, 충북대학교 법학연구소, 2020.6., 183~184면; 오병철, 「전자거래법」, 법원사, 279면; 송대중, '디지털컨텐츠의 특허법적 보호에 관한 연구', 지식재산연구, 제6권 제1호, 2011, 55면. 한편 무체물인 경우에도 관리할 수 있다면 물건에 포섭될 수 있다고 해석하더라도 이는 독점적이고 배타적으로 지배가능하다는 것을 의미한다고 해석한다면, 초고도의 자율성 내지 자유의지를 가지는 인공지능(로봇)의 경우에는 현행법상 물건 개념에 이를 포섭하는데 어려움이 있을 수 있다는 견해도 있다(이경규, 앞의 논문, 343면).

77) Lawrence B. Solum, Legal Personhood for Artificial Intelligences, 70 N.C. L. Rev. 1231, 1992, p.1231(송호영, 앞의 논문, 95면에서 재인용).

권리 객체로서의 물건과는 본질적인 차이를 가지게 된다고 볼 수 있다.⁷⁸⁾ 기술의 발전으로 인공지능의 자율성이 지속적으로 확대되더라도 이에 대해 별도의 제한이나 조치가 이루어지지 않는다면 인공지능은 인간의 관리가능성을 넘어설 가능성도 있다. 그리고 만약 인공지능이 인간의 관리가능성을 넘어서는 시기가 온다면, 그때에는 인공지능을 단순히 권리의 객체인 물건으로만 취급하기 어려워질 수 있으므로⁷⁹⁾ 이에 대한 새로운 논의가 필요할 수도 있다.

이러한 논의는 인공지능에 대해 법인격을 인정할 수 있는가의 문제로 이어진다. 그동안 인공지능의 법인격 인정 문제와 관련하여 이루어진 논의들을 살펴보면 논의의 초점이나 범위 등에 있어 다양하게 이루어졌기 때문에 분류하기는 쉽지 않지만, 크게 인공지능의 법인격을 전면적으로 인정하자는 견해와 부정하는 견해 그리고 인공지능의 법인격을 인정하되 전면적인 수준이 아니라 필요한 범위 내에서 제한적으로 인정하자는 견해로 구분할 수 있는 것으로 파악된다. 물론 전면적으로 긍정하는 견해와 제한적으로 긍정하는 견해는 그 인정범위에 있어서는 차이가 있지만, 근본적으로 인공지능의 법인격을 인정한다는 점에 있어서는 공통된다. 이하에서는 각 견해의 내용과 근거에 대해 살펴보기로 한다.

(2) 인공지능의 법인격 인정에 관한 학설

1) 긍정설

긍정설은 인공지능에 법인격을 인정할 수 있다는 입장으로서, 인공지능에 대해 법인격을 인정하게 되면 두 가지 중요한 장점을 가진다고 한다. 첫째, 인공지능의 표시와 행위를 복잡하고 수증하기 어려운 법이론에 의할 필요 없이 대리 법리, 이행보조자 법리, 피용자

78) 박수근, '자율적 지능 로봇의 법적 지위에 대한 소고', 법학논총, 제31권 제2호, 국민대학교 법학연구소, 2018, 55면 이하.

79) 김도훈, 앞의 논문, 176~177면.

법리에 따라 그 운용자에게 직접 귀속시킬 수 있으므로 인공지능의 취급에 관한 법적 명료성 및 법적 안정성에 이바지할 수 있다고 한다. 둘째, 인공지능과 운용자 또는 제3자 사이의 법률관계가 해명되어 책임에 관한 문제를 명확하게 해결할 수 있게 된다고 한다.⁸⁰⁾ 인공지능의 법인격을 인정해야 한다는 견해가 제시하고 있는 주요 논거들은 다음과 같다.

① 포스트휴먼으로서의 지위

인공지능의 법인격을 이른바 포스트휴먼(post-human)과 로봇권의 관점에서 긍정하는 견해가 있다. 이 견해에 따르면 종래 법인격의 개념은 인간과 비(非)인간을 구별하여 인간 중심적인 이론 체계를 정립한 칸트(Kant)의 근대철학에 기반을 둔 것인데, 이러한 관점은 인간 존엄성의 근거를 인간의 본성 자체에서만 찾으려는 논리적 한계에 부딪히고, 인간은 인간이기 때문에 존엄하다는 순환논리적 모순뿐만 아니라 호모 사피엔스라는 인간종족 중심주의에 빠질 위험성을 내포하고 있다고 한다.⁸¹⁾ 그리하여 이를 극복하기 위해 제시되는 포스트휴먼 담론에서는 인간중심주의 및 종차별주의를 넘어서서 인간의 속성을 가지는 다양한 존재에 대해 주체성의 확장을 인정한다.⁸²⁾

이 탈(脫)인간중심주의 사상의 현대적 출발점으로는 루만(Niklas Luhmann)의 체계이론이 그 시작이라고 일컬어지는데, 이 체계이론에서 사회체계의 구성 요소인 소통은 인간만을 전제로 하지 않고 있으며, 인간은 소통에 필요한 외부조건일 뿐 소통과 사회의 필수 요소는 아니다.

이러한 체계이론은 인공지능의 법인격 인정을 뒷받침하는 철학적 근거로 제시된다. 체계이론에 의하면 결국 인간과 비(非)인간의 경계는 모호해지고, 지금까지 객체에 불과하였

80) 정진명, 앞의 논문(주40), 86면.

81) 정채연, '법패러다임 변화의 관점에서 인공지능과 법담론: 법에서 탈근대성의 수용과 발전', 법과 사회, 제53호, 법과사회이론학회, 2016.12., 115~116면; 한국인공지능법학회, 앞의 책, 46면.

82) 송호영, 앞의 논문, 97면.

던 비(非)인간이 법적 주체로 포섭할 여지가 생긴다.⁸³⁾ 어떠한 대상이 어떠한 법적 능력을 가지거나 어떠한 능력을 가지고 있어야 하는가는 단지 존재론적으로 확정되는 것이 아니며, 사회체계를 어떻게 정의하고 그 체계의 기본 단위를 무엇으로 볼 것인가의 문제이기 때문이다.⁸⁴⁾ 오히려 루만의 체계이론에 의하면 만약 비인간적 존재라 하더라도 사회적 소통 능력이 있다면 사회적 주체가 될 수 있는데, 인공지능의 경우 지금까지 인류가 고안한 다른 어떠한 도구보다도 그 자체로 또는 인간과 상호작용하면서 사회적인 소통이 가능하므로, 결국 인공지능에도 전부 또는 일부라도 법인격을 인정할 수 있을 것이라는 결론에 이르게 된다고 한다.⁸⁵⁾

② 권리능력의 확장

역사적으로 인간에 대한 권리능력이 자유인에서부터 노예로 확장되어온 것처럼, 인간의 권리능력은 사람으로부터 ‘관념적 형상’인 법인(法人)에게로 부여된 것과 마찬가지로, 권리주체성은 ‘인간에서 법인으로’, 다시 ‘법인에서 전자인간(electronic person)’으로 확대될 것이라는 견해가 있다.⁸⁶⁾ 이 견해는 인간의 필요성에 의해 그 개념이 고안된 사법제도에 있어서, 정책적인 필요성에 의해 사람의 단체 혹은 출연재산에 대해 재산을 소유하는 연결점으로서 권리주체성을 인정할 수 있다고 하면서, 마찬가지로 인공지능 혹은 인공지능이 장착된 지능형 로봇에 대해서도 책임재산을 소유하게 할 정책적 필요성이 인정된다면 제한적으로 ‘책임의 목적’상 책임재산을 소유할 수 있는 권리주체성을 인정할 수 있다고 주장한다.⁸⁷⁾

83) 양천수, ‘탈인간중심적 법학의 가능성-과학기술의 도전에 대한 행정법학의 대응-’, 행정법연구, 제46호, 행정법이론실무학회, 2016.8., 9~10면.

84) 김성돈, ‘전통적인 형법이론에 대한 인공지능 기술의 도전’, 형사법연구, 제30권 제2호, 한국형사법학회, 2018.6., 101면.

85) Gunther Teubner, Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten, Ancilla Iuris, 2018, 107, S.141-145(이해원, 앞의 논문, 224면에서 재인용).

86) 이중기, ‘인공지능을 가진 로봇의 법적 취급: 자율주행자동차 사고의 법적 인식과 책임을 중심으로’, 홍익법학, 제17권 제3호, 홍익대학교 법학연구소, 2016.9., 19~20면.

87) 정진명, 앞의 논문(주4), 164면; 이중기, 앞의 논문, 24면.

③ 법인이론의 차용

인공지능에 대해 법인격을 부여할 수 있을 것인지에 있어 가장 보편적으로 이용되고 있는 논리는 법인이론을 차용하는 것이다. 이에 의하면, 원칙적으로 인간에게만 권리능력을 인정하고 그 밖의 존재에 대해서는 권리의 객체로만 취급하는 현행 법체계의 접근방식은 절대적이지 않고 사회변화에 따라 가변적이므로,⁸⁸⁾ 법적 편의성이나 유용성이 인정되면 인간이 아닌 인위적 존재에도 법인격을 부여할 수 있다고 한다.⁸⁹⁾ 그러므로 인공지능에 대해서도 새로운 법인격으로서 이른바 ‘전자인격’을 입법적인 방식에 의해 얼마든지 부여할 수 있다고 한다.⁹⁰⁾

이러한 주장을 뒷받침하는 대표적인 사례로서는 ‘법인(法人) 제도’가 제시된다. 법인의 본질에 대해서는 ‘법인의제설(法人擬制說)’과 ‘법인실재설(法人實在說)’이 검토될 수 있는데, 법인의제설에 의할 경우 전자인격을 인정하는 법을 제정함으로써 전자인격의 법인격을 법적으로 의제할 수 있게 되고, 법인실재설에 의하더라도 인공지능은 단순한 도관이 아니라 자율적인 판단에 의해 활동할 수 있는 존재로서 사회적 실재성을 갖추고, 유기체로 볼 수 있는 여지도 있으므로, 둘 중 어느 입장에 의하더라도 법인격을 부여받을 수 있는 근거는 충분히 인정될 수 있고, 법인격을 부여받은 전자인격은 권리능력을 인정받고 자율적 판단능력에 기초하여 독자적인 의사능력이 인정되며 물리적 현존을 통해 행위능력까지 인정될 수 있다고 한다.⁹¹⁾ 다만 민법이나 상법에서의 법인과는 다르게 취급하여 전자인격의 경우에는 특허이론에 기초를 두고 권리능력을 제한적으로 인정하는 것이 타당하

88) 이경규, 앞의 논문, 327면. 이 견해는 인류역사의 발전과정과 시대 및 사회상황에 따라 다양한 법제 하에서 초자연적 존재(supernatural beings), 동물(animals) 심지어 무생물(inanimate objects)에 대해서도 일정한 범위에서 권리주체성이 인정되어 왔다고 한다(이에 대한 보다 자세한 내용은, 이경규, 앞의 논문, 328면 이하를 참조).

89) 김진우, 앞의 논문, 24면.

90) 신현탁, ‘인공지능(AI)의 법인격-전자인격(Electronic Person) 개념에 대한 소고-’, 인권과 정의, 제478호, 대한변호사협회, 2018.12., 54면.

91) 신현탁, 앞의 논문, 54~55면.

다고 한다.⁹²⁾ 또한 인간이 법인제도를 고안한 이유를 고려할 때 기능면에서도 법인과 인공지능(로봇)은 유사하다고 볼 수 있으므로 인공지능의 권리능력 문제를 검토할 때 법인의 사례를 참고하는 것이 유용하다는 주장도 이와 맥락을 같이 한다고 볼 수 있다.⁹³⁾

④ 법적 책임의 구별 필요성

인공지능에 법인격을 부여하는 것을 긍정하면서 인공지능에 법인격을 인정하면 계약법상 인공지능을 이용한 의사결정에 따른 법률효과의 귀속이 명확해지고, 불법행위책임에서의 공백을 해소하거나 책임의 분배가 쉬워질 수 있다는 점을 강조하는 견해가 있다.⁹⁴⁾ 인공지능에 대한 법인격을 인정하게 되면 전자상거래 분야에서 인공지능의 거래참여를 용이하게 하고 특히 인공지능에 의한 행위를 그 운용자에게 귀속시키는 것에 있어서 법리적 어려움을 해소시킬 수 있는데,⁹⁵⁾ 이를 통해 인공지능에 대한 권리와 의무를 차별화하여 귀속시키는 것을 가능하게 하고 법적 책임영역 역시 상호구별하기 위하여 필요하다고 주장한다.⁹⁶⁾ 또한 불법행위와 관련된 책임문제에서도 권리능력이 인정되는 인공지능을 운용하는 자가 책임을 져야 하는 경우에 오작동한 인공지능을 상대로 구상권을 행사함으로써 책임리스크를 줄일 수 있으므로 인공지능의 법인격을 인정해야 할 법적 유용성이 있다고 한다.⁹⁷⁾

한편 계약법적인 측면에서, 현재로서는 다소 이론적인 상황이었지만 향후 특정 인공지능이 여러 분야에서 보편적으로 사용되어 사회적인 효용이 매우 크고 그 행위가 누구에게 귀속하는지가 불분명하거나 큰 의미가 없는 경우, 혹은 그 계약의 효력을 직접 인공지능

92) 신현탁, 앞의 논문, 59면.

93) 이창민, '로봇의 권리', LAW & TECHNOLOGY, 제12권 제3호, 서울대 기술과법센터, 2016.5., 45면.

94) 이성진, '인공지능과 법인격 인정', 민사법의 이론과 실무, 제23권 제3호, 민사법의 이론과 실무학회, 2020.8., 84면; 박도현, '인공지능과 자율성의 역학관계', 홍익법학, 제20권 제3호, 홍익대학교 법학연구소, 2019, 523면.

95) 김진우, 앞의 논문, 24면.

96) 김진우, 앞의 논문, 24~25면.

97) 김진우, 앞의 논문, 25면.

에 귀속시키더라도 그 계약의 이행이나 불이행에 따른 책임을 부담하는것에 무리가 없거나 계약책임을 적절하게 제한할 필요가 있는 경우 등에는 인공지능 자체에 대하여 독립적인 법인격을 인정하는 것도 고려해볼 수 있다거나, 불법행위법의 측면에서 인공지능이 위법한 행위를 했을 때 인공지능에 법인격을 인정하여 그 책임재산의 범위에서 또는 보험을 활용하여 책임을 지도록 하거나 운영자나 이용자가 사용자책임에 준하여 책임을 지는 경우에 마치 피용자도 불법행위책임을 지는 것처럼 인공지능도 별도로 책임을 부담하도록 함으로써 피해자를 두텁게 보호할 수도 있으며 다만 이러한 경우에는 특별한 사정이 없는 한 운영자나 이용자의 책임범위를 인공지능의 책임 범위로 제한할 것은 아니라는 견해도 있다.⁹⁸⁾

2) 부정설

부정설은 인공지능은 인간이 사용하는 도구에 불과하므로 인공지능에 법인격을 인정할 수 없다는 입장이다. 인공지능의 법인격을 부정해야 한다는 견해가 제시하고 있는 주요 논거는 다음과 같다.

① 법적 편의성에 대한 의문

인공지능에 대해 법인격을 부여하는 것을 부정하는 견해는 법인격 긍정론이 주장하는 책임귀속에 대한 법적 불명확성의 해소 및 그에 따른 거래비용의 절감이라는 법적 편의성에 대해 의문을 제기한다.⁹⁹⁾ 부정하는 입장에서는 인간과 같은 속성도 없고 현실적으로 법적 주체로 여겨지지 않는 대상인 인공지능에 대해 단지 법적 편의성을 이유로 법인격을 부여하는 것은 어려우며,¹⁰⁰⁾ 만약 인공지능이 실질적인 예측이나 통제가 어려운 고도의 자율성을 가지고 있다고 하더라도 이것은 이용자의 이용 내지 설정에 기초한 추정적 의사

98) 장보은, '인공지능을 이용한 행위에 대한 법적인 평가—사법(私法)상 논의를 위한 시론(試論 또는 始論)—', 외법논집, 제43권 제4호, 한국외국어대학교 법학연구소, 2019.11., 64~65면.

99) 이상용, '인공지능과 법인격', 민사법학, 제89호, 한국민사법학회, 2019.12., 39면.

100) 이상용, 앞의 논문, 36면.

에 부합하는 특정한 의사를 형성하여 표시하는 것에 불과한 것이므로, 오히려 인공지능을 이용자의 도구로 취급함으로써 책임귀속의 문제를 보다 간명하게 해결할 수 있다고 비판한다.¹⁰¹⁾

② 이른바 ‘전자인격’의 부당성

인공지능에 대해 법인격을 부정하는 견해 중에는 다음과 같은 이유로 이른바 ‘전자인격’에 대해 비판하기도 한다.

첫째, 로봇에 권리주체 또는 그와 유사한 법적 지위를 인정하려면 권리능력의 시기(始期)와 종기(終期)가 명확히 정해져야 하는데, 자연인인 사람의 경우 출생과 사망이라는 시기와 종기가 분명하게 존재하지만 인공지능이 탑재된 로봇의 경우 본질적으로 생명이라는 것이 존재하지 않기 때문에 시기와 종기를 정하는 기준을 명확하게 설정하는 것이 불가능하다고 한다.¹⁰²⁾ 둘째, 로봇에 인격을 부여하려면 인간과 완전히 동일한 수준은 아니더라도 어느 정도의 자율성이 존재한다는 점이 인정되어야 하지만 사실 로봇의 자율성은 설계자 내지 제작자에 의해 설계된 자율성일 뿐, 로봇이 스스로 작동을 시작해서 종국적으로 마칠 수 있는 자발성은 합리적으로 예측가능한 미래에도 불가능하므로 임의적인 자발성이 어느 하나도 인정되지 않는 로봇에 인간과 동일한 또는 그와 유사한 인격을 부여하는 것은 부당하다고 지적한다.¹⁰³⁾ 셋째, 인간은 지성과 감성 그리고 의지를 가지는 전인(全人)적인 통합성이 존재하므로 완전한 권리능력을 인정할 수 있지만, 오로지 특정의 영역에서만 존재하여 기능하는 로봇에게 전면적으로 인격을 부여하는 것은 부당하다고 한다.¹⁰⁴⁾ 넷째, 인간은 ‘이것이 사람인지 물건인지’ 정형성(定形性)을 통해 객관적으로 명확히 구별

101) 이상용, 앞의 논문 29~33면.

102) 오병철, ‘인공지능 로봇에 의한 손해의 불법행위책임’, 법학연구, 제27권 제4호, 연세대학교 법학연구원, 2017, 169면.

103) 오병철, 앞의 논문, 170면.

104) 오병철, 앞의 논문, 171면.

되지만, 로봇은 인간과 달리 정형적이지 않으므로 어느 존재에 법인격을 일괄적으로 부여할지를 정하기가 매우 어려운 문제가 있다고 한다.¹⁰⁵⁾ 다섯째, 로봇에 권리능력을 제한적으로라도 인정하게 되면 권리체계의 혼란이 야기된다고 한다. 로봇에 권리능력을 인정하게 되면 권리능력은 평등하다는 전제에서 ‘로봇의 소유자’라는 개념을 인정하기도 어렵고, 그럼에도 불구하고 로봇의 소유자 개념을 인정할 수 있다고 하더라도 로봇에 대한 반환청구권이나 부당이득반환청구권까지 인정한다면 로봇의 권리능력은 실질적으로 형해화(形骸化)되는 문제가 발생하게 된다. 게다가 로봇에 고유한 책임재산을 인정하게 되면, 그 로봇이 소유하고 있는 책임재산과 로봇의 소유자와의 관계를 어떻게 설정할 것인가도 문제되는데, 로봇에게 책임재산을 보유하게 하고 손해배상책임을 지는 범위에서 권리능력을 가진다고 한다면 정작 로봇이 어디에서 어떻게 책임재산을 취득할 수 있는지의 문제가 발생한다. 이에 대해서는 로봇이 누군가로부터 재산을 증여받고 그것을 책임재산으로 하여 손해배상책임을 부담하도록 하는 것이 유일한 방법이라고 볼 수 있지만, 책임재산을 증여자에게 귀책원리에 따라 1차적으로 손해배상책임을 부담시키면 되는 것을 굳이 우회적이고 기교적인 구성을 통해 해결해야 하는 이유를 찾기 어렵다는 지적이 있다.¹⁰⁶⁾ 마지막으로 로봇에 제한적으로 인격을 부여하여 권리능력을 인정하게 되면, 권리능력이 있는 로봇 자체가 입은 손해에 대해 손해배상청구의 주체가 로봇인지 아니면 로봇의 소유자인지가 불분명해진다는 문제점이 있다.¹⁰⁷⁾

③ 책임법 수정의 부당성

종래의 법리와 달리, 인공지능에 법인격 부여를 긍정하는 견해는 피해자가 결국 인공지능 로봇에 대해 직접 책임을 물을 수 있다는 법리로 이어지게 되는데, 부정하는 견해는 그와 같은 책임법의 수정에 대해 부당하다고 비판한다. 현재의 약한 인공지능 기술을 고려하면 오늘날 인공지능의 기술수준은 대체로 인간을 보조하는 ‘도구’ 수준이므로 현 시점에

105) 오병철, 앞의 논문, 171면.

106) 오병철, 앞의 논문, 172면.

107) 오병철, 앞의 논문, 173면.

서는 기존의 불법행위 책임법리를 수정하지 않거나 유추적용해도 무방하며, 어떠한 기술적 전제 없이 강한 인공지능 시대를 당연하게 전제하여 ‘미리’ 책임법을 수정하여 고찰하는 것은 바람직하지 않다고 한다.¹⁰⁸⁾¹⁰⁹⁾

이 입장에 따르면, 새로운 기술이 도입되어 책임법상 큰 변화가 필요하게 되더라도 충분한 검토 없이 일반법인 민법에 인공지능과 관련된 새로운 책임법리를 바로 도입하는 것은 적절하지 않으므로 새로운 책임법리를 일단 특별법에 규율하여 시간을 두고 시험해 보고, 후에 이를 일반법으로 수용하는 것에 대해 사회적인 공감대가 형성된다면 일반법인 민법에 그러한 인공지능책임 법리를 받아들이는게 바람직하다고 한다.¹¹⁰⁾

④ 법인격 부여의 실익 부재

인공지능에 법인격을 부여하여 얻게 되는 이익과 법인격을 부여함으로써 발생하는 법적 문제점을 비교할 때 후자가 더욱 크다는 이유로 인공지능에 법인격을 부여하는 것에 부정적인 입장을 취하는 견해도 있다.¹¹¹⁾ 이 견해는 인공지능에 법인격을 부여하면 현행 우리 법체계에 다음과 같은 혼란이 발생할 수 있다고 지적한다. 첫째, 인공지능형 컴퓨터시스템에 대해 법인격을 부여하게 되면 부여된 법인격에 걸맞게 독립된 권리의 주체가 되어야 하지만 민법상 인공지능형 컴퓨터시스템은 물건으로서 권리의 객체에 불과하기 때문에 서

108) 고세일, ‘인공지능과 불법행위책임 법리’, 법학연구, 제29권 제2호, 충남대학교 법학연구소, 2018.5., 104~105면.

109) 참고로 현재 강한 인공지능의 시대가 오는지에 대해서는 관련분야의 학자들 간에도 논란이 있다. 또한 강한 인공지능 기술을 전제로 하는 18~24개월마다 반도체 집적 기술은 2배로 증가한다는 무어의 법칙(Moore’s Law)은 더 이상 이루어지고 있지 않으며, 최근의 뇌과학에서 일반적인 연구성과를 마이크로 칩으로 구현을 시도한 바 있지만, 그러한 마이크로 칩은 전혀 작동하지 않았다고 한다. 따라서 뇌과학의 연구성과와 방법론에 큰 흠결이 있다는 문제가 제기되었는데, 이러한 측면에서 인공지능 기술이 놀라운 정도로 발전한다는 전제에 대해서도 회의적으로 보는 견해도 있다(고세일, 앞의 논문, 105면).

110) 고세일, 앞의 논문, 105면.

111) 이충훈, ‘인공지능(Artificial Intelligence)을 이용한 의사표시의 주체’, 법학연구, 제30권 제1호, 연세대학교 법학연구원, 2020.3., 305~307면.

로 충돌하게 된다고 지적한다. 둘째, 인공지능에 법인격을 부여하면 인공지능이 권리의 주체가 됨과 동시에 의무의 주체가 되어야 함에도, 인공지능에 배상책임을 지울만한 책임재산이 존재할 수 있는지에 대해 회의적이다. 셋째, 만약 인공지능에 법인격을 부여하였는데 인공지능형 컴퓨터시스템이 인공지능 컴퓨터시스템 운영자의 의사가 전혀 존재하지 않음에도 불구하고 의사표시를 하여 계약을 체결했다면 계약의 효력이 컴퓨터시스템의 운영자에게 귀속될 수 없게 되므로 대리가 성립할 수 없게 되고, 그 결과 무권대리의 법리에 의해 계약의 상대방은 시스템의 운영자에게 계약책임을 물을 수 없게 되는 문제가 발생하게 된다고 지적한다.¹¹²⁾ 이는 법인격을 가진 인공지능 컴퓨터가 존재하는 상황이면 인공지능 컴퓨터가 무권대리인이 되기 때문에 무권대리인에게만 책임을 물을 수 있을 뿐 본인인 시스템의 운영자에게 책임을 물을 수 없기 때문이다.¹¹³⁾

3) 제한적 긍정설

인공지능에 법인격을 제한적으로 인정할 수 있다는 제한적 긍정설은 인공지능의 법인격을 전적으로 자연인과 동일한 수준까지 인정해야 할 필요는 없으며, 법적으로나 사회적으로 필요성이 있는 경우에만 제한적으로 법인격을 인정할 수 있다는 견해이다.¹¹⁴⁾ 이 견해는 자율성이나 자유의지, 이성과 같은 인간만이 가지는 정신적 요소를 인공지능이 갖추지 않더라도 필요성이나 유용성이 인정되고 법리적으로 기존의 제도와 정합성을 인정할 수 있다면 과학기술 또는 철학적으로 다소 근거가 부족하더라도 필요한 범위 내에서 인공지능의 법인격을 인정할 수 있다고 주장한다는 점에서 긍정론과 차이가 있다고 할 수 있다.¹¹⁵⁾

112) 송호영, 앞의 논문, 102면.

113) 이충훈, 앞의 논문, 306~307면.

114) 김성호, '인공지능에 대한 법인격 부여 필요성', 법학논총, 제37권 제3호, 한양대학교 법학연구소, 2020.9., 150면.

115) 김성호, 앞의 논문, 150~152면; 이해원, 앞의 논문, 230면.

III. 검토

인공지능에 법인격을 인정하는지의 문제는 여러 측면에서 논의되고 결정되어야 한다. 사법적인 권리·의무관계에서 우리가 생각하고 인정하고 있는 법인격은 철학적인 관점에서 보는 인간의 개념과 완전히 무관하지는 않겠지만, 그보다는 앞으로 사회에서 인공지능과의 상호작용이 증가하고 복잡해짐에 따라 인공지능과 사회·경제적 관계를 설정하는 것이 매우 중요해질 것이라는 전제에서 인공지능에 사회적인 지위를 부여하려는 것이기 때문이다. 따라서 인공지능의 법인격 인정문제는 단순히 법리적으로만 판단할 문제는 아니고, 개별적인 인공지능의 기술적인 발전수준과 실제 사용되는 현황에 비추어 해당 인공지능이 사회·경제적으로 법적인 판단의 주체로서 인정될 수 있는지, 나아가 해당 인공지능에 법인격을 인정하는 것이 사법적 책임을 인정 및 분배에 있어 반드시 필요한 것인지 등을 다각적으로 검토한 후 결정하는 것이 타당하다.¹¹⁶⁾

이와 관련하여 앞서 인공지능에 법인격을 부여하는 것을 긍정하는 견해와 부정하는 견해가 제시하는 근거들을 살펴보았다. 각 견해의 주장과 이를 뒷받침하는 근거들은 각자의 논리적 타당성을 가지고 있지만, 현재 인공지능에 법인격을 부여하는 것은 시기상조로 보여진다. 한편으로 인공지능이 아직 전면적으로 상용화되어 인간의 사회·경제적인 모든 영역에서 이용되지 않고 있는 현 시점에서 인공지능 로봇에 대해 법인격을 부여하는 입법을 '미리' 하는 것은 바람직하지 않으며 오히려 정확하지 않은 성급한 입법이 될 수도 있다.¹¹⁷⁾

116) 장보은, 앞의 논문, 65면.

117) 정완, '인공지능 로봇에 관한 법적 고찰', 경희법학, 제55권 제4호, 경희대학교 법학연구소, 2020.12., 204면.

하지만 이후 논의될 인공지능과 불법행위책임의 문제를 논의함에 있어서도 '책임의 주체'라는 관점에서 인공지능의 법인격을 인정할 것인가의 문제를 살펴보면, 인공지능의 주요 특징 중의 하나인 '자율성'과 관련하여 현재의 기술수준이 구현하고 있는 인공지능의 자율성은 인간의 개입이 없는 임의적 자율성이 아니고 여전히 인간의 조작과 개입이 필요한 자율성이며, 인공지능이 인간의 마음과 행위를 모두 기능적으로 구현하는 것은 불가능하다는 점을 전제해야 한다. 그러므로 적어도 불법행위책임의 측면에서는 인공지능에 대해 법인격을 부여하는 것은 어렵기에, 인공지능이 관련된 불법행위책임의 문제에 대해 사고를 발생시킨 인공지능 자체에 대해서는 책임을 지울 수는 없고, 기존의 책임법리에 따라 전통적인 법인격 주체인 자연인이나 법인 중 누구에게 어떠한 근거로 어떠한 내용의 책임을 부담하게 해야 하는지에 논의의 초점이 맞춰져야 할 것으로 보인다.¹¹⁸⁾

118) 이해원, 앞의 논문, 239면.

이상으로 인공지능과 관련된 기본적인 내용들과 인공지능에 대한 법인격 부여의 문제에 대해 살펴보았다. 앞서 언급한 바와 같이 인공지능에 법인격을 인정하는 문제는 단순한 법리적 판단을 넘어, 기술 수준과 사회·경제적 관계에서의 역할을 고려해야 한다. 현재 인공지능에 법인격을 부여하는 것은 시기상조로 보이며, 전면적인 상용화 이전에 인공지능에 대한 법인격 부여를 입법화하는 것은 성급한 결정일 수 있다. 특히 불법행위책임의 관점에서 인공지능은 완전한 자율성을 가지지 못하므로, 사고 발생 시 인공지능 자체가 아닌 전통적 법인격 주체인 자연인이나 법인에게 책임을 묻는 방식이 보다 적합하다고 할 것이다. 이러한 전제를 바탕으로 이후 논의될 인공지능 관련 불법행위책임의 문제에서는 어떠한 책임귀속의 원리가 적용될 수 있는지를 중심으로 살펴보고자 한다.

제3장 인공지능과 불법행위책임

제1절 개관

제2절 인공지능에 의한 손해의 발생

제3절 인공지능에 의한 손해에 적용가능한
민사책임의 분류

제4절 인공지능의 불법행위책임에 대한
새로운 책임모델

제5절 소결



제3장

인공지능과 불법행위책임

제 1절 | 개관

본 장에서는 인공지능에 의해 발생하는 손해와 관련된 논의를 진행한다. 이하에서는 인공지능에 의해 발생하는 손해의 특수성과 그에 대한 과실판단의 문제와, 그 손해에 대해 우리 법상의 책임법리들이 적용가능한지에 대해 살피고, 인공지능의 불법행위책임에 대한 새로운 책임모델로서 위험책임과 편익책임에 대해서 살펴보고자 한다.

제 2절 | 인공지능에 의한 손해의 발생

1. 인공지능에 의한 손해의 특수성

1. 정보오류에 의한 손해

인공지능은 기본적으로 정보를 처리하는 역할을 하지만, 처리하는 정보에 오류가 있다면 이로 인해 손해가 발생할 수 있다. 만약 정보의 오류로 인해 손해가 발생하면 그 손해에 대해 배상책임을 부담해야 하는가에 대해, 대법원은 “부동산중개업자는 비록 그가 조사·확인하여 의뢰인에게 설명할 의무를 부담하지 않는 사항이더라도 의뢰인이 계약체결 여부를 결정하는 데 중요한 자료가 되는 사항에 관하여 그릇된 정보를 제공하여서는 안되

고, 그릇된 정보를 제대로 확인하지도 않은 채 마치 그것이 진실인 것처럼 의뢰인에게 그대로 전달하여 의뢰인이 그 정보를 믿고 상대방과 계약에 이르게 되었다면, 부동산중개업자의 그러한 행위는 선량한 관리자의 주의로 신의를 지켜 성실하게 중개행위를 하여야 할 중개업자의 의무에 위반된다”¹¹⁹⁾고 하여 손해배상책임을 인정한 바 있다.

참고로 미국에서는 계약관계에 있지 않은 제3자에게 잘못된 정보를 제공하는 경우 정보제공자와 제3자 사이에 계약과 유사한 관계(near privity)가 있다고 이론을 구성하거나, 정보제공이 의도된 제3자에 대해서는 잘못된 정보를 제공하지 않을 주의의무가 있다고 이론을 구성하거나, 정보를 신뢰하리라고 예견가능한 제3자에 대해서는 책임을 진다는 방식으로 제3자에 대한 책임범위를 확장하고 있다고 한다.¹²⁰⁾ 이러한 입장을 고려하면 인공지능의 정보오류에 의한 손해에 대해서도 배상책임이 인정될 가능성을 배제할 수는 없다고 하겠다.

2. 인체 손해의 발생

인공지능에는 이르지 않은 컴퓨터 프로그램의 오류로 인한 손해는 대부분 재산상 손해였고, 인공지능 역시 단순히 정보만을 제공한다면 그로 인해 발생하는 손해는 대부분 재산상의 손해일 가능성이 크다. 하지만 물리적인 로봇이 인공지능에 의해 통제되는 경우라면 그 기기가 인체에 직접 손해를 발생시킬 수도 있다.¹²¹⁾ 예를 들어 자율주행자동차나 인공지능 로봇이 아니라 인공지능이 내장된 단순한 생활가전에 알고리즘상의 문제가 발생하여 화재 등의 위험이 발생할 수도 있다.¹²²⁾

119) 대법원 2008.9.25. 선고 2008다42836 판결.

120) 이에 대한 보다 자세한 내용은 권영준, ‘미국법상 순수재산손해의 법리’, 민사법학, 제58호, 한국민사법학회, 2012.3., 158면 이하를 참조.

121) 오병철, 앞의 논문, 175면.

122) 이와 관련한 극단적인 예로는 인공지능과 결합한 ‘킬러 로봇(Killer Robot)’이 있다. 인공지능이 탑재된 킬러 로봇은 ‘자율무기(autonomous weapons)’로도 불리는데, 아직 독자적 판단까지는 아니더라도 목표물을 스스로 식별하고 명령을 수행하는 수준의 자율무기는 현재 기술수준으로도 손쉽게 만들

인공지능 로봇에 의해 발생하는 손해가 단순히 재산상에 손해에 그치지 않고 사람에게 대한 신체 내지 생명침해를 가져올 수 있다는 점은 지금까지 컴퓨터와 같은 정보처리장치의 결함으로 인해 발생하는 재산상의 손해와는 차이가 있다. 더군다나 이러한 손해는 단순히 개인의 문제가 아니라 사회와 공동체 전체의 문제로 그 범위가 확대될 수 있으며 치명적인 위협이 될 수도 있는 가능성을 내포하고 있다.

3. 동시다발적인 손해

오늘날의 일반적인 기기는 대량생산되더라도 발생하는 결함이나 오류는 개별적으로 존재하는 것이어서 손해는 기기에 있어서만 단발적으로 발생하는 것이 대부분이었다. 하지만 인공지능이 내장된 로봇이 대량생산되고, 만약 이 로봇에 내장된 인공지능의 알고리즘에 발생한 오류로 인해 손해가 발생한다면, 그 손해는 단순히 하나의 로봇에만 한정되지 않고 동일한 알고리즘이 적용된 다른 모든 로봇에 같은 손해가 발생된 가능성이 크다.

이처럼 동일한 또는 같은 원인에 의해 다수의 피해자들에게 다수의 개별적 손해가 발생하는 현상에 대해 전통적인 민사책임법에서는 이를 별도의 법적 개념이나 전문용어로 파악하지는 않았지만, 이에 대해 주목하는 견해가 존재한다.¹²³⁾ 이 견해에 따르면 피해가 대규모로 발생하는 사건들에 있어서는 피해자를 실질적으로 구제할 수 있도록 특별한 위험원을 내포하는 사업을 영위하는 모든 기업들과 위험한 시설을 가지고 있는 모든 보유자들은 책임재산의 확보를 위하여 적절한 대비책을 마련하는 것이 무엇보다 중요하다고 강조한다.¹²⁴⁾¹²⁵⁾

수 있다고 한다(고현실, '[영상] AI 킬러 로봇 시대 '성큼' ... "핵무기급 파급력", 연합뉴스, 2024년 5월 11일자 기사(<https://www.yna.co.kr/view/AKR2024051012860070>)).

123) 신유철, '대규모 피해와 손해배상', 민사법학, 제75호, 2016, 421면 이하.

124) 신유철, 앞의 논문, 430면.

125) 오병철, 앞의 논문, 176면.

II. 인공지능에 의한 손해와 인공지능의 과실판단

1. 전통적인 과실책임주의의 한계

과실책임주의는 개인이 타인에게 준 손해에 대해서는 그 행위가 위법할 뿐만 아니라 동시에 고의 또는 과실에 기한 경우에만 책임을 진다는 원칙이다.¹²⁶⁾ 이 원칙은 근대민법의 기본원리 중 하나이며, 우리 민법도 제750조에서 불법행위에 관하여 이를 규정하고 있다. 하지만 인공지능에 의한 손해에 있어서도 이 원칙이 그대로 적용될 수 있는지에 대해서는 생각해볼 필요가 있다.

(1) 인공지능 로봇과 고의책임

먼저 인공지능이 탑재된 로봇이 타인의 신체나 재산에 손해를 가한 경우 고의책임을 인정할 수 있는지에 대해 검토하는 견해가 있다. 이 견해는 i) 인간이 무해한 로봇으로 작업을 수행하려 했으나, 로봇이 계획을 벗어나 공격을 가한 경우, ii) 로봇이 무해한 소유자에도 불구하고 손해를 발생시킨 경우, iii) 인간이 악한 로봇과 공모하여 손해를 발생시킨 경우를 상정하고 검토하는데, 앞의 두 경우는 결국 인간의 과실에 기한 것으로 볼 수밖에 없고, 마지막 경우는 인간의 고의라기보다는 공상과학영화의 한 장면으로 사악한 로봇의 작동으로 볼 수밖에 없다고 한다. 그러므로 인공지능 로봇이 야기한 손해에 대해 인간의 고의에 의한 불법행위는 논의의 중심에서 벗어난 것이므로, 결과적으로 합리적인 예견가능성이나 기여과실 등의 차원에서 다루는 것이 적절하다고 한다.¹²⁷⁾

126) 송덕수, 「채권법각론」, 박영사, 2023, 523면.

127) Ugo Pagallo, 「The Law of Robots—Crimes, Contracts, and Torts」, Springer, 2013, pp.120–121 (오병철, 앞의 논문, 183면에서 재인용). 그는 로봇을 낙관적으로 인식하여 마치 유아나 어린 동물과 같은 순수한 행위양태를 전제하고 있으나, 이에 대해서는 사악한 로봇의 등장은 반드시 공상과학 영화에만 국한될 것이라고 단정할 수는 없다는 비판이 가능하다고 보여진다.

(2) 과실의 전제로서의 주의의무

과실은 자기의 행위로부터 일정한 결과가 발생할 것을 인식했어야 함에도 불구하고 부주의로 말미암아 인식하지 못하고 그 행위를 하는 심리상태라고 정의된다.¹²⁸⁾ 주의의무의 내용으로 가장 중요한 것은 자신의 행위가 어떤 결과를 가져올 것이라고 합리적으로 예견하는 ‘예견가능성’과 자신의 행위가 가져올 결과를 회피하기 위한 노력을 기울여야 하는 ‘회피가능성’인데, 이를 바탕으로 결과발생에 대한 예견의무와 회피의무의 위반이라는 외연으로 나타난 행위자의 심리상태가 과실을 판단하는 기준이 된다.¹²⁹⁾ 대법원 역시 과실을 인정하기 위해서는 결과를 예견할 수 있었음에도 불구하고 그 결과발생을 예견하지 못하였고 그 결과발생을 회피할 수 있었음에도 불구하고 그 결과발생을 회피하지 못한 과실이 검토되어야 한다고 판시한 바 있다.¹³⁰⁾ 이러한 대법원의 입장에 따르면 만약 손해의 발생을 예견할 수 없거나 또는 예견하더라도 이를 회피할 수 없는 것이었다면 과실이 있다고 보기 어렵다.

(3) 추상적 경과실

과실은 전제가 되는 부주의의 종류에 따라 추상적 과실과 구체적 과실로 나누어지는데, ‘추상적 과실’은 그 사람이 속하는 사회적 지위나 직업 등에서 보통 일반적으로 요구되는 주의, 즉 구체적인 사람에 의한 개인의 능력차이를 인정하지 않고 일반적으로 평균인에게 요구되는 주의를 게을리하는 것을 말하며, 이때의 주의를 ‘선량한 관리자의 주의’라고 한다. 이에 비해 ‘구체적 과실’은 행위자 자신의 평상시의 주의를 게을리하는 것을 말하며, 이때에는 개인의 능력 차이가 고려된다.¹³¹⁾

128) 송덕수, 앞의 책, 532면.

129) 박동진, ‘불법행위법에서의 주의의무’, 비교사법, 제9권 제2호, 한국사법학회, 2002, 161면.

130) 대법원 1999.12.10. 선고 99도3711 판결; 대법원 1996.11.8. 선고 95도2710 판결 등.

131) 송덕수, 앞의 책, 532면.

불법행위의 성립요건으로서의 과실은 보통·평균인의 주의력을 기준으로 하는 추상적 과실이다. 하지만 이때 ‘보통·평균인’이란 기준은 추상적인 보통·평균인이나 전체 사회에서의 보통·평균인이 아니고, 구체적인 경우에 있어서의 보통·평균인을 말한다.¹³²⁾ 인공지능이 탑재된 로봇에 의해 발생하는 손해를 법적으로 취급함에 있어서도 이와 관련한 과실판단이 이루어져야 한다. 현실적으로 오류가 없는 알고리즘을 갖춘 인공지능은 존재하기 어렵기 때문에 절대로 손해가 발생하지 않아야만 하는 완벽한 주의를 기울이는 것은 규범적으로 기대하기는 어렵다.¹³³⁾

2. 인공지능의 과실판단 문제 및 그 한계

그렇다면 인공지능에 의한 손해에 있어 과실판단을 어떻게 해야 하는지 생각해볼 필요가 있다. 고의 또는 과실의 판단은 행위자를 대상으로 이루어져야 하는데, 앞서 살펴본 바와 같이 현재로서는 인공지능에 권리주체로서의 법인격은 인정되지 않는다. 그러므로 인공지능이 탑재된 로봇의 경우 누구를 대상으로 예견가능성과 회피가능성을 판단해야 하는지에 대해 명확하지 않다. 그러므로 일반적인 기기와 달리 사람에 의해 통제되지 않는 로봇의 경우 그것이 작동되어야 하는 상태대로 작동되었는지 여부에 초점이 맞춰져야 한다.¹³⁴⁾ 그렇다면 인공지능 로봇이 타인에게 입힌 손해에 대한 가장 중요한 행위기여가 있다고 볼 수 있는 사람은 인공지능을 제작한 프로그래머와 로봇을 이용하는 자 등이라고 할 수 있는데, 이들에 대해 살펴보기로 한다.

132) 대법원 2001.1.19. 선고 2000다12532 판결.

133) 오병철, 앞의 논문, 184면.

134) David C. Vladeck, ‘Machines Without Principles: Liability Rules and Artificial Intelligence’, Washington Law Review, vol.89, 2014, p.132(<https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=4800&context=wlr>).

(1) 인공지능 프로그래머의 과실판단

인공지능 로봇이 발생시킨 손해에 대해 먼저 인공지능 프로그래머에게 프로그래밍상의 과실을 인정할 수 있는지가 검토될 수 있다. 과실판단의 기준인 예견가능성과 회피가능성을 기준으로 검토하면, 인공지능 로봇의 경우 언제 어떠한 상황에서 어떠한 방법으로 타인에게 손해를 야기할 것인지의 위험을 구체적으로 예견하는 것은 일반적이지 않고, 단지 언젠가는 로봇이 오작동하여 손해가 발생할 수도 있다는 불확실하고 막연한 불안감이 존재할 뿐이다. 그러므로 로봇에는 단순히 잠재화된 위험만 있으므로, 잠재된 위험의 현실화에 대한 공포의 인식만으로는 법적인 의미의 예견가능성은 존재한다고 보기 어렵다.

회피가능성과 관련하여, 로봇을 통제하는 인공지능은 본질적으로 완전할 수 없다. 프로그래머는 프로그래밍 과정에서 디버깅(debugging)¹³⁵⁾ 단계를 반드시 거치므로, 기초적인 디버깅 단계에서 일반적인 오류는 보통 회피된다. 로봇은 일반적으로 상당기간 동안 실험과 검증과정을 거치고 시중에 보급될 것이므로, 그러한 실험과 검증을 거쳐 발견되는 오류들은 수정될 것이고, 나아가 경우에 따라서는 국가기관 등에 의한 인증절차를 거쳐야만 보급될 수 있을 것이다. 결과적으로 이러한 복수의 오류 검증과 수정의 절차를 거쳐야 함에도 불구하고 손해발생을 회피하는 것이 가능하다고 규범적으로 인정한다면, 불가항력을 제외하고는 모두 회피가능성을 인정하는 것이므로 실질적인 관점에서 볼 때 무과실의 결과책임과 아무런 차이가 없을 것으로 보인다.

(2) 인공지능 로봇 이용자의 과실판단

인공지능 로봇을 이용하는 사람에게 로봇이 발생시킨 손해에 대한 과실을 인정할 수 있는지에 대해서도 생각해볼 필요가 있다. 인공지능 로봇의 이용자는 로봇에게 정해진 명령을 내리는데, 그 명령 자체가 타인에게 가해를 가하는 것이라면 당연히 고의의 불법행위

135) ‘디버깅(debugging)’은 컴퓨터 프로그램의 개발과정에서 발생하는 오류나 비정상적인 연산(버그)을 찾아내고 수정하는 작업을 말한다.

가 될 것이지만, 로봇의 이용자가 타인에게 무해한 명령을 내렸다면 그는 로봇을 직접적으로 조작하지 않기 때문에 로봇에 대한 구체적인 지배가능성과 통제가능성이 결여되어 있다고 볼 수 있다. 이는 곧 손해의 구체적인 위험을 예견할 수 있다고 하더라도, 이에 대해 회피가능성은 완전히 결여되었다고 평가할 수 있다.¹³⁶⁾

이와 관련하여 인공지능 로봇에 적용과 관련하여 참고될 수 있는 규정으로는 미국 리스테이트먼트 제307조가 있다. ‘불완전하거나 결함 있는 도구의 이용(Use of Incompetent or Defective Instrumentalities)’이라는 표제 아래의 해당 조문은 로봇 이용자의 과실 판단에 유용하게 적용될 수 있을 것으로 보인다. 리스테이트먼트 제307조¹³⁷⁾는 “사람이든 물건이든, 그것이 불완전하거나, 부적절하거나 또는 결함이 있어 타인에게 합리적인 수준을 넘는 위험을 가져온다는 사실을 알거나 알아야 함에도 이를 이용하는 경우에는 과실이 인정된다”고 규정하고 있고, 이 조문의 해설에 의하면 그 도구가 명성있는 제작자로부터 구입한 경우라도 실제로 안전성을 합리적으로 요구된다고 설명된다.¹³⁸⁾ 하지만 위 과실판단에 관한 위 기준에 대해 살펴보면, 일반적인 평범한 이용자가 인공지능이 통제하는 로봇이 불완전하거나, 부적절하거나 또는 결함이 있어서 합리적인 수준을 넘는 위험을 가져올 것이라고 구체적으로 인식할 수 있는 경우는 극히 예외적일 것으로 보인다.¹³⁹⁾

(3) 인공지능 로봇 소유자의 과실판단

로봇을 직접 이용하지 않고 단지 로봇의 소유권만을 보유하고 있는 자에게도 과실책임을 부담시킬 수 있는지에 대해서도 살펴볼 필요가 있다. 이와 관련해서 참고될 수 있는 규정으로는 미국의 리스테이트먼트가 있는데, 리스테이트먼트 제308조¹⁴⁰⁾는 ‘부적절한 사

136) 오병철, 앞의 논문, 186면.

137) Restatement of Torts §307 It is negligence to use an instrumentality, whether a human being or thing, which the actor knows or should know to be so incompetent, inappropriate, or defective, that its use involves an unreasonable risk of harm to others.

138) Restatement of Torts §307 Comment b.

139) 오병철, 앞의 논문, 187면.

람에게 물건이나 활동참여의 허락(Permitting Improper Persons to Use Things or Engage in Activities)’이라는 표제 아래 과실판단에 관한 규정을 두고 있다. 해당 조문은 “제3자가 타인에 대한 손해의 합리적인 수준을 넘는 위험을 발생시키는 방법으로 물건을 이용하거나 또는 행동을 할 것 같거나 또는 그런 의도를 가지고 있다는 점을 알거나 알았어야 했다면, 제3자에게 행위자의 통제 아래 있는 물건을 이용이나 활동의 참여를 허락하는 것은 과실로 인정된다”고 규정하고 있다. 여기에서 “행위자의 통제 아래 있다 (under the control of the actor)”라는 표현은 행위자의 동의하에 제3자에게 물건의 점유나 이용 권한을 부여하였고, 행위자가 동의를 유보함으로써 제3자가 그 물건을 사용하는 것을 저지할 수 있었다고 믿을만한 이유가 있는 경우를 의미한다.¹⁴¹⁾ 예를 들어 운전 경험이 전혀 없는 14세의 소녀에게 그 사실을 알면서도 자신의 차를 운전하도록 허락하고 결국 무경험으로 인해 사고가 발생했다면 그 사고에 대해 자동차 소유자에게 과실이 인정되는데,¹⁴²⁾ 이와 유사한 입장의 판결이 특히 자동차와 관련하여 다수 존재한다고 한다.

하지만 인공지능 로봇의 소유자가 인공지능을 가진 로봇이 앞서 살펴본 바와 같이 불완전하거나, 부적절하거나 또는 결함이 있어서 합리적인 수준을 넘는 위험을 가져올 것이라고 인식하고 있다는 것은 인정하기 어려울 것으로 보인다. 인공지능 로봇은 이용자에 의해 통제되는 것이 아니므로 이용자의 특성에 따른 이러한 특별한 과실판단을 적용하는 것은 적절하다고 보기는 어렵다.¹⁴³⁾

140) Restatement of Torts §308 It is negligence to permit a third person to use a thing or to engage in an activity which is under the control of the actor, if the actor knows or should know that such person intends or is likely to use the thing or to conduct himself in the activity in such a manner as to create an unreasonable risk of harm to others.

141) Restatement of Torts §308 Comment a.

142) Restatement of Torts §308 Comment a. Illustrations 1.

143) 오병철, 앞의 논문, 188면.

(4) 과실의 증명책임

일반적으로 과실에 대한 증명책임은 불법행위의 피해자가 부담하는 것이 원칙이다. 하지만 인공지능에 관한 기술의 발전이 고도화될수록, 인공지능의 개발자의 과실이나 인공지능 로봇의 공급자가 손해발생을 예견하였다는 점을 피해자가 증명하는 것은 점점 어려워질 수밖에 없다.¹⁴⁴⁾ 한편으로 피해자는 과실의 입증을 위해 피해를 발생시키게 된 로봇의 인공지능 알고리즘을 입수해야 하는데, 인공지능 알고리즘 자체는 제작자의 결정적인 영업비밀에 해당하므로 제작자는 당연히 공개하지 않으려 할 것이다. 그러므로 피해자가 인공지능 알고리즘의 소스코드에 접근하는 그 자체가 불가능할 가능성이 높고, 피해자가 인공지능의 과실을 증명하는 것은 사실상 불가능하다고 보여진다.

144) Ugo Pagallo, op. cit., p.117.

I. 논의의 배경

인공지능의 작동을 법적 의미의 행위로 인정할 수 있는지 여부와 그 결과를 누구에게 귀속시킬 수 있는지의 여부는 민사책임에 있어 중요한 의미를 가진다. 특히 인공지능이 딥러닝을 통해 데이터를 습득하여 ‘고유한’ 결론을 도출하는 경우에는 전통적인 민사책임 법리에 따른 행위성을 판단하기가 쉽지 않다. 그렇지만 현실적으로 인공지능에 의한 손해에 적용가능한 민사책임의 법리로 과실책임 또는 무과실책임을 기반으로 하여 책임귀속의 법적 근거 내지 법적 구성을 검토할 수밖에 없는데, 전자는 다시 일정한 사람에 대한 관리·감독상의 책임과 일정한 물건에 대한 책임으로 구분될 수 있다. 이하에서는 인공지능에 의한 손해에 대해 적용이 고려될 수 있다고 여겨지는 민사책임들에 대해 살펴보고자 한다.

II. 과실책임에 기한 법적 구성

1. 감독자책임

민법 제755조는 미성년자와 심신상실자와 같이 책임무능력자가 그의 위법행위에 대해 배상책임을 지지 않는 경우 그 책임무능력자를 감독할 법정 의무가 있는 자와 감독의무자에 갈음하여 책임무능력자를 감독하는 자에게 그 책임무능력자가 제3자에 가한 손해에 대한 배상책임을 인정한다. 책임무능력자의 감독자책임은 감독의무자가 자신의 가해행위에 대해서가 아니고 그의 감독을 받는 책임무능력자의 가해행위에 대해 책임을 지는 것으로서 일종의 타인의 행위에 대한 책임에 해당한다. 하지만 그 책임이 성립하기 위해서는 감독의무자의 과실이 필요하므로 순수한 의미의 타인에 대한 책임은 아니라고 할 수 있다.

한편 민법 제755조 제1항 제2문의 해석상 감독의무자의 과실에 대한 증명책임은 피해자로부터 감독의무자에게로 전환되어 있으므로 감독자책임은 무과실책임에 접근하게 되는데, 이러한 점이 일반적으로 감독자책임을 과실책임과 무과실책임의 중간적 책임이라고 부르는 이유라고 할 수 있다.¹⁴⁵⁾

인공지능의 작동으로 인한 손해에 대해 감독자책임을 유추적용할 수 있는지에 대해 검토할 필요가 있다. 먼저 책임의 요건인 '행위관련성' 측면에서 감독이 필요한 사람과 인공지능은 서로 동일성이 없다는 점에서 부적합하다. 다음으로 인공지능의 소유자나 이용자에게 인공지능을 감독할 의무가 있다고 하더라도 '공동의 삶과 복수의 개인에 의한 일반 가정의 기본 욕구 충족'이라는 가족공동체의 기준에 있어 적절하지 못하다.¹⁴⁶⁾ 마지막으로 인공지능의 소유자나 이용자에게 감독자책임을 인정하려면 적어도 인공지능의 고의나 과실에 의한 행동이 있어야 하는데, 이는 결국 인공지능의 법인격 내지 법적 지위와 관련이 있다. 그러므로 인공지능에 대해 법인격을 인정한다면 감독자책임에 관한 규정이 유추적용될 가능성이 있겠지만, 이 경우에도 민법 제755조 제1항 단서의 '감독의무를 게을리하지 아니한 경우'라는 면책사유에 해당할 경우가 많아 책임의 근거로 되기는 어려울 것으로 보인다. 따라서 인공지능에 의한 손해에 대해 감독자책임을 유추적용하는 것은 적합하지 않은 것으로 보인다.¹⁴⁷⁾

2. 사용자책임

민법 제756조 제1항은 타인을 사용하여 어느 사무에 종사하게 한 자는 피용자가 그 사무집행에 관하여 제3자에 가한 손해에 대해 배상할 책임이 있다고 규정하고 있다. 사용자책임은 현대 사회의 기술발전과 다양한 산업구조의 변화 그리고 법적·경제적 환경이 변

145) 송덕수, 앞의 책, 560~561면.

146) 이수경, '인공지능의 민사책임에 대한 소고', 민사법학, 제107호, 한국민사법학회, 2024.6., 242면.

147) 정진명, 앞의 논문(주4), 144~145면.

화로 인해 오늘날 일반적인 책임 유형으로 자리잡고 있다. 휴머노이드 로봇과 같이 인공지능이 탑재된 지능형 로봇의 경우 인간의 직접적인 조작이 없어도 스스로 작동하는 외관을 가지므로 사람과 동일시하는 관점을 가질 수 있고,¹⁴⁸⁾ 그에 따라 지능형 로봇을 피용자로 보고 그 보유자에게 사용자책임을 유추적용하는 논리가 해외 연구자에 의해 검토되기도 하였다.¹⁴⁹⁾

하지만 사용자책임을 인정되기 위해서는, i) 먼저 타인을 사용하여 어느 사무에 종사하게 해야 하고, ii) 피용자가 ‘그 사무집행에 관하여’ 손해를 가해야 하며, iii) 그 손해는 제3자에게 발생해야 하고, iv) 피용자의 가해행위가 불법행위의 요건을 갖추어야 하며, v) 사용자가 면책사유 있음을 증명하지 못해야 한다. 이중 첫 번째 요건인 ‘사용관계’와 관련하여 사용자와 피용자의 관계는 반드시 유효한 고용관계가 있는 경우에 한정되지 않고, 사실상 어떤 사람이 다른 사람을 위해 지휘·감독 아래 그 의사에 따라 사무를 집행하는 관계에 있어야 하는데,¹⁵⁰⁾ 사용자책임을 성립하려면 사용자가 불법행위자인 피용자를 실질적으로 지휘하고 감독하는 관계에 있어야 한다.¹⁵¹⁾

인공지능이 탑재된 지능형 로봇에 사용자책임을 유추적용하려는 견해에 대해서는 일단 민법 제756조의 ‘타인’은 권리주체의 의무가 될 수 있는 사람을 전제로 하므로 인공지능에 대해 법인격을 인정하지 않는 한 적용이 부적합하므로 피용자로 보기 어렵다는 비판이 가능하다. 다음으로 제756조 제1항 단서에서 정하고 있는 ‘사용자가 상당한 주의를 하여도 손해가 있을 경우’라는 면책과 관련하여 지능형 로봇의 경우 그 이용자가 아무리 상당한 주의를 한다고 하더라도 손해를 회피하거나 방지하는 것이 불가능할 수 있으므로 로봇의 이용자에게 사용자책임을 지우고자 하더라도 거의 대부분 면책이 되는 문제가 발생할

148) 오병철, 앞의 논문, 199면.

149) Ugo Pagallo, *op. cit.*, pp.130-134.

150) 대법원 1998.8.21. 선고 97다13702 판결.

151) 대법원 1999.10.12. 선고 98다62671 판결.

수 있다.¹⁵²⁾ 이러한 관점에서 사용자책임을 로봇의 이용자에게 유추적용하는 것 역시 실효성의 측면에서 적합하지 않은 것으로 보인다.

3. 공작물점유자책임

인공지능으로 인해 발생한 손해에 대해 공작물책임에 관한 규정이 유추적용될 수 있는 지에 대해 검토할 필요가 있다. 민법 제758조 제1항은 공작물의 설치 또는 보존의 하자로 인하여 타인에게 손해를 가한 때에는 공작물점유자가 손해를 배상할 책임이 있지만, 만약 점유자가 손해의 방지에 필요한 주의를 게을리하지 하지 않았다면 그 소유자가 손해를 배상할 책임이 있다고 규정하고 있다.

공작물책임이 인정되기 위해서는 i) 손해를 발생시킨 것이 공작물이어야 하고, ii) 공작물의 설치 또는 보존에 하자가 있어야 하며, iii) 공작물의 하자로 인해 손해가 발생했어야 하고, iv) 책임에 대한 면책사유가 없어야 한다. 이하에서는 인공지능에 적용가능한 공작물책임을 검토함에 있어 먼저 공작물 ‘점유자’의 책임에 대해 살펴보기로 한다.¹⁵³⁾

공작물은 인공적 작업에 의해 만들어진 물건을 말하므로,¹⁵⁴⁾ 인공지능이 탑재된 로봇 역시 공작물에 해당된다고 볼 수 있다.¹⁵⁵⁾ 하지만 인공지능의 본질은 인공지능을 작동시키는 알고리즘(algorithm)이며, 알고리즘은 어떤 문제의 해결을 위해 입력된 자료를 토대로 원하는 출력을 유도하는 규칙의 집합을 말하므로 인공지능 자체에 대해서는 공작물이라고 보기는 어렵다.¹⁵⁶⁾ 공작물의 설치 또는 보존에 있어서의 ‘하자’는 공작물이 그 용도

152) 오병철, 앞의 논문, 200면; 정진명, 앞의 논문(주4), 145면.

153) 공작물소유자의 책임에 대해서는 후술한다.

154) 송덕수, 앞의 책, 578면. 참고로 공작물의 대표적인 예로는 건물·교량·도로·지하도·수도시설과 같은 토지의 공작물과 계단이나 승강기 같은 건물 내외의 설비 그리고 자동차와 항공기와 같은 동적(動的)인 기업설비가 있다.

155) 오병철, 앞의 논문, 197면.

156) 정진명, 앞의 논문(주4), 146면.

에 따라 통상 갖추어야 할 안전성을 갖추지 못한 상태에 있는 것을 의미하고,¹⁵⁷⁾ 하자의 유무는 객관적으로 판단되며, 하자가 점유자나 소유자의 고의·과실에 의해 발생했는지도 묻지 않는다. 판례에 따르면 하자에 있어 안전성은 공작물 자체만의 용도에 한정된 안전성만이 아니라 공작물이 현실적으로 설치되어 사용되고 있는 상황에서 요구되는 안전성을 의미하고,¹⁵⁸⁾ 공작물의 설치와 보존에 있어 항상 완전무결한 상태를 유지할 정도로 고도의 안전성을 갖추지 아니하였다 하여 그 공작물의 설치·보존에 하자가 있는 것이라고 할 수는 없으며,¹⁵⁹⁾ 고도의 위험을 수반하는 공작물의 경우에는 그 시설이 관계법령이 정한 바에 따른 시설기준 등에 부적합한 것이라면 특별한 사정이 없는 한 이러한 사유는 공작물의 설치·보존상의 하자에 해당한다.¹⁶⁰⁾ 그러나 법령 등에 정해진 기준은 이러한 하자의 유무를 판단하는 절대적 기준이 아니고 단지 하자의 유무를 판단하는 일응의 참작기준이 될 수 있을 뿐이라고 한다.¹⁶¹⁾

인공지능이 탑재된 지능형 로봇에 존재하는 하자의 경우, 로봇에 대한 주요한 안전규정이 기준이 되며, 로봇의 하자는 위험성에 비례하여 사회통념상 일반적으로 요구되는 정도의 방호조치(防護措置)의 의무를 다하였는지의 여부를 기준으로 결정되어야 한다. 그러나 지능형 로봇과 관련된 모든 위험이 로봇의 하자는 아니며, 로봇을 점유하는 자는 그 중 ‘정상적인 위험’을 회피해야 하지만 정상적인 위험과 무관한 손해를 방지해야 하는 것은 아니다. 그러므로 약한 인공지능의 경우에는 사회통념상 일반적으로 요구되는 정도의 방호조치의무를 고려할 수 있지만, 학습 및 의사결정능력을 가진 강한 인공지능의 경우에는 ‘정상적인 위험’을 정의하기가 어려워 보인다. 그러므로 강한 인공지능으로 인해 발생한 손해에 대해서는 공작물책임을 묻는 것은 적합하지 않은 것으로 보인다.

157) 대법원 2006.1.26. 선고 2004다21053 판결 등.

158) 대법원 1992.10.27. 선고 92다21050 판결.

159) 대법원 2000.1.14. 선고 99다39548 판결.

160) 대법원 1994.10.28. 선고 94다16328 판결.

161) 대법원 1981.4.14. 선고 80다3100 판결.

한편 공작물점유자는 손해방지에 필요한 주의를 다한 것을 입증하면 손해배상으로부터 면책되지만, 공작물소유자는 면책이 인정되지 않는다. 인공지능을 탑재한 로봇의 점유자가 부담하는 ‘손해의 방지에 필요한 주의’는 공작물 자체의 용도에만 한정되는 것이 아니고 인공지능 로봇의 작동 이후에 발생한 자연적·인위적 환경변화에 대한 방호조치의무도 포함된다. 따라서 약한 인공지능이 장착된 로봇이 설치 및 보존의 하자로 타인에게 손해를 야기했다면 인공지능 로봇의 점유자가 우선적으로 손해배상책임을 진다.¹⁶²⁾ 이에 반해 강한 인공지능의 경우에는 인공지능 로봇의 점유자에게 로봇의 구체적인 위험성에 대한 예견 가능성과 이에 대한 회피가능성이 완전히 결여되었다고 볼 수 있으므로 강한 인공지능이 장착된 로봇에 의하여 발생한 손해는 로봇의 소유자가 부담하게 된다고 해석된다.¹⁶³⁾

4. 동물점유자책임

위험한 동물을 규제하는 것은 동물의 점유자가 어떠한 법적 예방 조치를 준수하도록 하여 타인에게 끼칠 수 있는 위협을 줄이기 위한 것이다. 다시 말해, 동물의 점유자가 타인에게 발생시킬 수 있는 피해에 대해 책임지도록 하는 것인데, 이와 관련하여 민법 제759조 제1항은 동물의 점유자는 그 동물이 타인에게 가한 손해를 배상할 책임이 있지만, 동물의 종류와 성질에 따라 그 보관에 상당한 주의를 게을리하지 않은 경우에는 면책될 수 있음을 규정하고 있다. 이 규정이 정하고 있는 동물점유자의 책임에 대해 학자들은 ‘위험책임의 원리에 의해 강화된 과실책임’¹⁶⁴⁾이라고 파악하거나 ‘기본적으로 중간책임’¹⁶⁵⁾으

162) 이와 관련하여 인공지능 로봇을 이용하는 대부분의 이용자는 로봇의 유지 및 관리에 관한 내용을 준수할 것이므로 로봇의 점유자라고 하더라도 대부분 면책사유에 해당되어 공작물책임을 인정하기 어렵다는 견해도 있다(오병철, 앞의 논문, 197면).

163) 정진명, 앞의 논문(주4), 148면.

164) 김성미, ‘소형드론 운항자에 대한 동물점유자책임 유추해석의 가능성’, 항공우주정책 법학회지, 제34권 제2호, 한국항공우주정책·법학회, 2019, 182면.

165) 김항중, ‘동물로 인해 발생한 민사상 불법행위책임에 관한 일고-미국의 과실책임과 엄격책임분석을 중심으로-’, 비교사법, 제28권 제1호, 한국사법학회, 331면. 이 견해는 동물의 ‘종류 및 성질에 따라 상당한 주의’를 기울이며 관리하고 있었는가를 책임결정의 중요한 판단근거로 보아 과실유무를 판단하지만 기본적으로 중간책임이라고 파악한다.

로 보기도 한다.

동물은 예측할 수 없는 행동으로 위험성을 지닐 수 있고 이러한 위험이 실현되는 경우 동물의 점유자에게 책임을 묻는 것과 같이 인공지능의 보유자에게도 적용될 수 있는지에 대해 살펴볼 필요가 있다. 앞서 살핀 바와 같이 동물의 점유자는 보관상의 주의를 다했음을 증명하면 민법 제759조 제1항 단서에 의해 면책이 가능하지만, 인공지능의 보유자에게는 본인이 인공지능을 어떠한 주의의무로 보관하여야 하며 사용하였는지를 증명하여 면책을 받을 수 있게 된다는 가이드라인이 존재하지 않는다. 현실적으로 살펴보면, 동물의 점유자는 동물의 행동반경에 울타리를 설치하거나 입마개·하네스 등을 이용하여 동물의 공격으로 발생할 수 있는 위험을 대비하는 등의 행위들 한다면, 이는 동물의 점유자가 관리 및 감독의 행위를 수행하는 것으로 보일 수 있다. 하지만 인공지능의 경우, 알고리즘의 프로세스 기록은 이후에 확보하더라도 시간적으로 이미 늦었거나, 심지어 딥러닝 과정에서 사용될 학습의 데이터는 인간사용자에게는 보이지도 않는다. 앞서 살펴본 인공지능의 예측불가능성은 동물의 위험과는 성격 및 양상이 다르므로 인공지능에 의한 손해에 대해 동물점유자의 책임을 유추적용하는 것 역시 적합하지 않은 것으로 판단된다.¹⁶⁶⁾

166) 이수경, 앞의 논문, 240면.

Ⅲ. 무과실책임에 기한 법적 구성

1. 공작물소유자책임

공작물소유자는 공작물점유자가 손해방지에 필요한 주의를 게을리하지 아니한 경우에 공작물의 설치 또는 보존의 하자로 인하여 타인에게 발생한 손해를 배상해야 한다. 공작물의 소유자는 공작물을 소유한다는 사실 외에 공작물의 설치나 보존에 대해 어떠한 행위도 관여하고 있지 않다. 그러므로 공작물의 소유자에게는 공작물의 설치나 보존에 대한 고의나 과실이 문제되지 않으며, 공작물소유자가 공작물책임을 부담하는 경우에 민법 제 758조 제3항에 따라 그 손해의 원인에 대해 책임 있는 자를 상대로 구상권을 행사할 수 있다.

한편 공작물소유자책임은 손해의 발생에 공작물소유자의 '행위'를 요건으로 하지 않고, 단지 '결과'에 대해서만 예외적으로 책임을 인정하는 것이므로 인공지능을 탑재한 로봇에 대해서도 유추적용이 가능할 것으로 보인다. 다만 공작물의 소유자가 무과실책임을 지더라도 공작물의 설치 및 보존에 대한 하자를 피해자가 증명해야 하므로 지능형 로봇의 경우에는 제한적인 범위에서만 적용하는 것이 가능할 것으로 보인다.¹⁶⁷⁾

2. 제조물책임

인공지능이 탑재된 로봇은 대표적인 제조물이라고 할 수 있으므로, 인공지능의 오작동으로 인해 타인에게 발생한 손해에 대해 제조물책임의 적용을 고려할 수 있다. 제조물책임은 제조물의 결함으로 인하여 발생한 손해에 대해 제조업자 등이 부담하는 책임이며,¹⁶⁸⁾ 이들에게 과실이 있는지를 묻지 않는 무과실책임이다. 제조물책임이 인정되기 위

167) 정진명, 앞의 논문(주4), 150면.

168) 송덕수, 앞의 책, 620면.

해서는 i) 제조물에 결함이 있어야 하고, ii) 피해자에게 손해가 발생해야 하며, iii) 결함과 손해 간에 인과관계가 입증되어야 한다.

(1) 인공지능의 제조물성 인정 여부

로봇의 본질적인 구성요소인 인공지능 그 자체가 하나의 제조물이 되는지에 대해 논의가 존재한다. 현재 민법 제98조는 물건을 유체물 및 기타 관리할 수 있는 자연력이라고 정의하고 있는데, 이와 관련하여 소프트웨어를 물건으로 볼 것인지가 그 논의의 중심에 있다. 이러한 논의의 실익은 제조물 책임법상 제조업자는 제조물의 결함으로 인하여 생명·신체 또는 재산에 손해를 입은 자에게 그 손해를 배상하도록 하고 있는데,¹⁶⁹⁾ 이때 제조물은 제조되거나 가공된 동산이라고 규정되어 있기 때문이다.¹⁷⁰⁾

먼저 소프트웨어의 제조물성을 부정하는 견해로는 i) 무체물인 소프트웨어는 관리할 수 있는 자연력이 아니므로 동산으로 볼 수 없으므로 제조물 책임법을 적용할 수 없다는 견해¹⁷¹⁾와 ii) 민법 및 제조물 책임법의 해석상 정보나 소프트웨어가 유체동산에 담겼더라도 그 자체로는 직접 손해를 입힐 위험성이 없고 민법의 일반 책임법리인 계약상의 하자담보책임 등으로 구제받을 수 있다는 견해¹⁷²⁾가 있다. 다음으로 소프트웨어의 제조물성을 인정하는 견해로는 i) 소프트웨어 또는 알고리즘, 프로그램 자체는 원칙적으로 제조물이라 보기 어렵지만 자율주행자동차와 같은 하드웨어에 결합된 소프트웨어의 경우 하나의 기능상의 제조물이고 기기에 내장된 임베디드(embedded) 소프트웨어의 경우 제조물로 보아 제조물 책임법의 적용이 가능하다는 견해¹⁷³⁾, ii) 민법상 동산에는 유체물 외에

169) 제조물 책임법 제3조 제1항.

170) 제조물 책임법 제2조 제1호.

171) 권오승·신은주·홍명수·차성민·이현중, 「제조물 책임법」, 법문사, 2003, 190면.

172) 주지홍, '소프트웨어 하자로 인한 손해의 제조물책임법리 적용여부', 민사법학, 제25권, 한국민사법학회, 2004.3., 437면.

173) 권영준·이소은, '자율주행자동차 사고와 민사책임', 민사법학, 제75호, 한국민사법학회, 2016, 467~468면; 이상수, '임베디드 소프트웨어의 결합과 제조물책임 적용에 관한 고찰', 법학논문집,

관리가능한 자연력이 포함되는데, 소프트웨어도 전기의 흐름이라는 전제에서 제조물의 개념에 포함된다는 견해,¹⁷⁴⁾ iii) 주물과 종물의 법리를 유추적용하여 자율주행자동차와 같은 하드웨어와 자율주행을 가능하게 하는 소프트웨어는 주물과 종물의 관계에 있다고 볼 수 있으므로 소프트웨어도 제조물에 해당한다는 견해¹⁷⁵⁾가 있다. 한편 대법원은 “소프트웨어는 저작권 유사의 권리로써 그 가치는 프로그램의 내용뿐만 아니라 그에 대한 권리의 범위에 따라 달라지는 것”¹⁷⁶⁾이라고 판시하여 제조물 책임법상 소프트웨어의 제조물성을 인정하지 않는 것으로 보인다.

살피건대 대량으로 제작 및 판매되는 소프트웨어는 일종의 상품으로 보아도 무방하므로 그러한 소프트웨어를 구입하여 이용함에 있어 소프트웨어 자체의 불량이나 바이러스 감염 등과 같은 결함에 의해 다른 데이터나 하드웨어에 손상이 발생한다면 제조물책임을 인정해야 한다고 볼 수도 있겠지만, 소프트웨어 자체는 무체물이고 관리할 수 있는 자연력이라고 보기 어려우므로 민법상 물건(동산)이라고 보기 어렵고 단지 ‘정보’에 불과하다고 할 것이므로, 현행 제조물 책임법의 해석상 소프트웨어를 제조물이라고 보기는 어렵다. 하지만 기기에 내장된 소프트웨어인 임베디드(embedded) 소프트웨어는 소프트웨어 자체가 독자적인 기능을 수행하는 것이 아니라 전용 기기나 부품에 내장되었을 때 특정된 기능을 수행하고, 서로 분리할 수 없으며, 거래관념상 임베디드 소프트웨어가 내장된 제품을 구매하는 경우에 그 소프트웨어가 그 기기의 동작이나 기능을 위한 필수적인 물건의 일부로서 취득된다는 점을 고려한다면 ‘임베디드 소프트웨어와 기기가 결합된 임베디드 시스템 자체’를 동산으로 보아 제조물로 인정할 수 있을 것이다. 그러므로 인공지능이 탑재된 로봇의 경우, 이를 임베디드 시스템 자체라고 볼 수 있다면 제조물책임을 인정할 수 있을

제39권 제2호, 중앙대학교 법학연구원, 2015.8., 91~92면; 최경진, ‘지능형 신기술에 관한 민사법적 검토’, 정보법학, 제19권 제3호, 한국정보법학회, 2015.12., 232면.

174) 김민중, ‘컴퓨터바이러스에 따른 손해에 대한 법적 책임’, 인터넷법률, 제18호, 법무부, 2003.7., 97면.

175) 김진우, ‘자율주행에서의 민사책임에 관한 연구-개정된 독일 도로교통법과 우리 입법의 방향-’, 강원법학, 제51권, 강원대학교 비교법학연구소, 2017.6., 44면.

176) 대법원 1998.9.4. 선고 98두1512 판결; 대법원 1998.8.21. 선고, 97누13115 판결 등.

것으로 보인다.¹⁷⁷⁾

(2) 인공지능의 결함 판단

제조물 책임법상 책임이 인정되기 위해서는 ‘결함’이 인정되어야 한다. 결함은 제조, 설계 또는 표시상의 결함이나 그 밖에 통상적으로 기대할 수 없는 안전성이 결여되어 있는 것을 말하는데, 제조물 책임법에서는 ‘제조상의 결함’, ‘설계상의 결함’, ‘표시상의 결함’으로 구분된다.

먼저 ‘제조상의 결함’은 제조업자가 제조물에 대하여 제조상·가공상의 주의의무를 이행하였는지에 관계없이 제조물이 원래 의도한 설계와 다르게 제조·가공됨으로써 안전하지 못하게 된 경우를 말한다.¹⁷⁸⁾ 인공지능과 관련해서는 인공지능을 애초에 의도한 설계대로 제조했다라도 인공지능의 기계 및 딥러닝으로 인해 제조자가 의도하지 않은 손해를 야기할 수 있다는 점을 고려한다면, 그로 인해 발생한 손해에 대해 현행 제조물 책임법으로 규율할 수 있을 것으로 보인다.¹⁷⁹⁾

‘설계상의 결함’은 제조업자가 합리적인 대체설계를 채용했다라면 피해나 위험을 줄이거나 피할 수 있었음에도 대체설계를 채용하지 아니하여 해당 제조물이 안전하지 못하게 된 경우를 말한다.¹⁸⁰⁾ 인공지능과 관련해서는 대체설계가 가능하더라도 수많은 컴포넌트(component)로 구성된 인공지능시스템의 특성상 어떤 컴포넌트가 설계단계에서 인공지

177) 이상수, 앞의 논문, 91~92면; 류창호, ‘자율주행자동차에 대한 제조물책임의 적용에 관한 연구’, 아주법학, 제10권 제1호, 아주대학교 법학연구소, 2016.5., 41면; 최경진, ‘인공지능과 불법행위책임’, 정보법학, 제25권 제2호, 한국정보법학회, 2021.8., 56면.

178) 제조물 책임법 제2조 제2호 가목.

179) 하지만 이에 대해 최초로 프로그래밍이 이루어진 부분이 아니라 인공지능이 스스로의 학습을 통해 변경된 상황판단에 따른 결과라면 이를 제조자에게 책임을 물을 수 없다는 견해가 있다(신동환, ‘인공지능로봇과 배상책임에 관한 소고’, 안암법학, 제55호, 안암법학회, 2018, 50면).

180) 제조물 책임법 제2조 제2호 나목.

능시스템에 사용될 것인지를 예견할 수 없기 때문에, 설계상의 결함으로 판단하기에 쉽지 않다.¹⁸¹⁾ 그러므로 현재로서는 최고의 기술수준에 의하더라도 설계상 불가피한 오류가 발생할 수 있다는 점을 고려한다면 대체설계의 결함으로 보기는 어렵다고 할 것이다.

‘표시상의 결함’은 제조자가 합리적인 설명·지시·경고 또는 그 밖의 표시를 하였더라면 해당 제조물에 의하여 발생할 수 있는 피해나 위험을 줄이거나 피할 수 있었음에도 이를 하지 아니한 경우를 말한다.¹⁸²⁾ 표시상의 결함을 판단하는 가장 대표적인 표지(標識)는 합리적인 설명·지시·경고 또는 그 밖의 표시이고, 그것이 적절한지 여부가 관건이라고 할 수 있다. 인공지능의 경우 제조업자는 알고리즘에 결함이 없는 인공지능을 개발하고 또한 인공지능의 딥러닝에도 불구하고 애초에 예견할 수 없는 피해가 발생했다면 제조업자에게 표시상의 결함을 묻기 어려울 수 있다.¹⁸³⁾

(3) 결함과 손해 간의 인과관계 입증

제조물 책임법상 손해배상을 청구하기 위해서는 제조물의 결함과 손해 사이에 인과관계가 입증되어야 한다. 제조물의 결함이 인정되려면 제조물이 본래 의도한 설계와 다르게 제조되었고, 그로 인해 제조물이 안전하지 못하게 되었다는 점이 인정되어야 하지만 피해자의 입장에서 이를 증명하는 것은 현실적으로 매우 어렵다. 심지어 제조물의 결함으로 인해 어떤 사고가 발생하면서 문제된 그 제조물이 멸실되었다면 피해자가 제조물의 결함을 직접 증명하는 것이 사실상 불가능할 수도 있다. 하지만 그렇다고 하여 이러한 경우 피해자가 제조업자에 대해 제조물책임을 물을 수 없게 된다면 그는 안전하지 못하게 제조된 제조물로 인한 손해를 부당하게 감수하게 되는 결과가 발생하는데, 이러한 부당한 결과를

181) 물론 인공지능의 도입 초기에는 합리적인 대체설계가 존재한다는 것을 상정하기 어렵지만, 인공지능 관련기술이 급속도로 발전하여 어느 정도 상용화된다면 매우 빠른 속도로 합리적 대체설계가 등장할 수도 있다.

182) 제조물 책임법 제2조 제2호 다목.

183) 정진명, 앞의 논문(주4), 152면.

고려하여 등장하게 된 것이 ‘기능이상법리(malfunction doctrine)’이다.

‘기능이상법리’는 과실을 추정하는 ‘*res ipsa loquitur*’ 법리를 제조물책임의 맥락에 맞도록 변형한 것으로, 원고가 부담하는 인과관계의 입증책임을 완화하는 것이라고 볼 수 있는데,¹⁸⁴⁾ 인공지능은 고도의 기술과 복잡한 공정으로 만들어지므로 일단 인공지능시스템의 결함을 밝혀내기 어렵고, 그 작동에 고도의 기술과 복잡한 공정이 필요하기 때문에 작동오류나 오조작을 입증하기 어렵다는 특징이 있다.¹⁸⁵⁾ 또한 인공지능과 관련하여 발생한 손해가 인공지능시스템에 의한 것인지 여부를 확인하기 어려울 뿐만 아니라 인공지능의 개발에는 보통 다수의 사람들이 관여하고 있으므로 만약 결함과 손해 사이에 인과관계가 인정된다고 하더라도 정작 책임의 범위를 구체적으로 산정하기가 쉽지 않다는 문제가 있다.

(4) 면책요건으로서 개발위험의 항변

인공지능에 제조물책임을 적용하는 경우에는 ‘개발위험의 항변’을 고려해야 한다. ‘개발위험의 항변’은 제조업자가 제조물을 공급한 당시의 과학·기술 수준으로는 발견할 수 없었던 결함에 관하여 제조업자에게 책임을 물을 수 없다는 것을 말하고,¹⁸⁶⁾ 제조업자는 이에 대해 입증하면 제조물 책임법에 따른 손해배상책임을 면한다. 개발위험의 항변에 대해서는 당시의 최고의 기술수준을 유지하고 있는지 여부를 기준으로 판단하는 것이 적절해 보이는데, 제조물의 성질상 제조물을 이용하는 사람의 생명이나 신체에 위해를 줄 수 있

184) 권영준·이소은, 앞의 논문, 471~472면.

185) 이상수, 앞의 논문, 73면.

186) 제조물 책임법 제4조 제1항 제2호. 참고로 미국에서는 개발위험(development risk)이라는 용어가 거의 사용되지 않으며, 알 수 없는 위험에 관한 제조자가 할 수 있는 항변은 ‘기술수준의 항변(State of The Art Defense)’이다. 미국에서는 대륙법계와 달리 일반적으로 결함 여부를 검토할 때 기술수준을 고려하여 기술수준을 충족한 제조물은 결함이 없는 것으로, 기술수준을 충족하지 못한 제조물은 결함이 있는 것으로 보고 있다. 그러므로 이러한 점을 고려한다면 기술수준의 항변은 진정한 의미에서의 항변이라고 보기 어렵다고 한다(윤진수, ‘제조물책임의 주요 쟁점-최근의 논의를 중심으로-’, 법학연구, 제21권 제3호, 연세대학교 법학연구원, 2011.9., 42면).

는 위험이 크고 이용자가 그 위험을 통제하기 어렵다면, '당시의 과학·기술수준'은 당시를 기준으로 가장 높은 수준으로 보는 것이 타당하기 때문이다.¹⁸⁷⁾ 인공지능시스템의 경우, 당시를 기준으로 최고 수준으로 판단한다면 개발위험의 항변이 인용될 여지는 줄어들 것으로 보인다. 한편 그 개발자나 설계자가 고도의 위험방지의무를 위반하여 생명이나 신체를 침해할 위험을 발생시킬 수 있는 인공지능시스템을 설계하여 제조 또는 판매했다면, 그 인공지능시스템은 특별한 사정이 없는 한 사회통념상 기대되는 안전성이 결여된 '설계상의 결함'이 존재한다고 할 수 있다.

3. 자동차운행자책임

자동차손해배상보장법 제3조는 자기를 위하여 자동차를 운행하는 자는 그 운행으로 다른 사람을 사망하게 하거나 부상하게 한 경우에는 그 손해를 배상할 책임이 있다고 규정하고 있다. 여기서 자기를 위하여 자동차를 운행하는 자, 즉 '운행자'라고 인정되기 위해서는 '운행이익'¹⁸⁸⁾과 '운행지배'¹⁸⁹⁾라는 두 가지 요소가 인정되어야 한다. 운행이익과 운행지배가 인정되는 운행자는 피해자에 대해 자동차손해배상보장법에 따라 무과실의 위험책임을 부담한다. 인공지능이 탑재된 로봇이라고 볼 수 있는 자율주행자동차의 경우, 인공지능의 오작동에 의한 손해에 대해 이러한 명문의 규정이 있다면 행위성이 부정되는 인공지능 이용자에게도 손해배상책임을 고려할 수 있을 것이다.¹⁹⁰⁾ 그러나 인공지능에 있어 인공지능의 개발자는 인공지능의 작동에 대한 지배가능성과 그로 인한 이익을 향유한다고 보아 운행자책임을 유추적용할 수 있겠지만, 인공지능의 이용자는 인공지능의 작동에 대

187) 권영준·이소은, 앞의 논문, 477면.

188) '운행이익'은 자동차 운행에 따른 이익이 자신에게 귀속되는 것을 말하며, 그 이익에는 직·간접적인 경제적 이익뿐만 아니라 만족감 등 정신적 이익까지도 포함된다고 한다(대법원 1987.1.20. 선고 86 다카1807 판결).

189) '운행지배'는 자동차 사용에 대한 지배력이 자기에게 속하는 것을 의미하는데, 판례는 '운행지배'는 현실적으로 보유자와 운전자 사이에 사실상의 지배관계가 존재하는 경우만이 아니라 간접적이거나 제3자의 관리를 통한 관념상의 지배관계가 존재하는 경우도 포함하는 것으로 파악하고 있다(대법원 1992.4.14. 선고 91다4102 판결).

190) 오병철, 앞의 논문, 182면.

한 지배가능성이 없으므로 운행자책임을 유추적용하는 것은 적절하지 않다고 보여진다.¹⁹¹⁾ 한편 인공지능 로봇을 보유한 자는 일반적으로 인공지능 로봇을 소유한 자를 말하는 것으로 이해될 것이므로, 그에 대해서는 운행자책임을 유추적용하여 손해배상책임을 부담시키는 것이 가능할 수도 있다.¹⁹²⁾

제4절 | 인공지능의 불법행위책임에 대한 새로운 책임모델

1. 논의의 배경

법은 손해의 발생을 예방하기 위한 규범을 제공함과 동시에, 이미 발생한 손해에 대해 그 책임의 귀속 주체와 분배 방식을 결정하여 정의를 실현하는 기능을 수행한다. 불법행위법은 일상생활에서 발생하는 손해를 누가 어떻게 부담할 것인지를 시대와 사회의 정의와 형평의 관념에 따라 결정하는 것을 주된 내용으로 한다.¹⁹³⁾ 앞서 살펴본 바와 같이 불법행위법은 전통적으로 과실책임주의에 따라 어떠한 손해가 타인의 비난가능한 행위에 의해 발생했고 또 그러한 사실을 피해자가 증명한 경우에만 그 손해를 타인에게 부담시킨다. 따라서 손해가 타인의 유책한 행위로 발생한 경우라도 이러한 사실을 피해자가 증명

191) 정진명, 앞의 논문(주4), 154면.

192) 자동차손해배상 보장법 제5조 제1항은 자동차보유자에게 자동차의 운행으로 다른 사람이 사망하거나 부상한 경우에 피해자에게 대통령령이 정하는 금액을 지급할 책임을 지는 책임보험이나 책임공제에 가입할 의무를 부여하고 있는데, 현실적으로 자동차 책임보험은 자동차의 소유자가 가입하는 것이 일반적이므로, 자동차보유자와 자동차의 소유자는 동일인인 경우가 대부분이다. 이러한 점에 주목한다면 인공지능 로봇에 있어서도 로봇의 소유권을 가지고 있다는 사실에서 손해배상책임의 귀속근거를 찾을 수 있다고 보여진다.

193) 서광민, '과학기술의 발달과 불법행위법의 대응', 『민법의 기본문제』, 서강대학교 출판부, 2006, 436~437면.

하는데 실패하면 그 손해는 피해자가 부담할 수밖에 없다.¹⁹⁴⁾

그러나 과학기술의 발전에 따라 등장할 수 있는 여러 위험원(危險原)의 운영으로 발생하게 될 손해는 그것을 운영하는 자의 과실이 아니라도 발생할 수 있고, 설사 위험원을 운영하는 자의 과실로 인하여 손해가 발생한다 하더라도 과학기술에 대해 전문지식이 없는 피해자가 그러한 과실을 입증하는 것은 결코 쉽지 않다. 그러나 과실의 입증 실패로 인해 피해자가 그 손해를 감당하게 하는 결과는 형평과 정의의 관념에 반한다고 볼 수 있는데, 문제는 과학기술이 발전할수록 이러한 부당한 사례들은 점차 증가될 가능성이 있다는 점이다.

그리하여 인공지능이 사회변화를 이끌게 될 미래사회에서는 과실책임과 무관하게 규범적으로 책임을 귀속시킬 수 있다는 주장이 등장하고 있다. 즉 발생한 손해에 대한 책임귀속은 반드시 책임주체의 행위와 과실책임에 연관될 필요가 없고, 오히려 고의·과실과 무관하더라도 책임을 인정할 수 있다고 한다. 이에 따르면 인공지능 로봇의 작동결과에 의한 책임귀속을 위해 논리필연적으로 인간의 행위와 연결지를 필요도 없게 된다. 다만 본인의 행위 및 귀책사유와 무관한 자에게 인공지능 로봇에 의해 발생한 손해에 대한 책임을 지우기 위해서는 별개의 책임의 근거가 필요한데, 오늘날 이와 관련된 민사책임 모델로는 위험책임과 편익책임이 논의되고 있어 이를 소개하고자 한다.

II. 새로운 민사책임 모델로서 위험책임과 편익책임

1. 위험책임

인공지능의 작동으로 인해 발생하는 손해를 대비하기 위해 위험책임의 도입이 고려될

194) 서광민, 앞의 논문, 437면.

수 있다. 위험책임(strict liability)은 책임배분이 인공지능과 그 자율성에 대하여 허용된 지시에 의존하는 경우에 의미가 있는데, 이 경우 위험책임은 단순히 과실책임에 대한 예외가 아니라 독자적인 책임근거와 고유한 적용영역을 가지는 책임원칙이라고 할 수 있다.¹⁹⁵⁾

여기서 ‘위험’은 어떠한 것인가 문제되는데, 먼저 위험은 일반적인 위험이 아니라 ‘특별한 위험(besondere Gefahr)’이어야 한다. 즉 위험이 원자력손해처럼 발생가능성은 매우 낮지만 일단 발생하면 대규모의 손해를 가져올 수 있는 영역이나 자동차사고와 같이 손해의 크기는 일상적이어도 손해발생의 가능성이 매우 높은 영역에 해당해야 한다.¹⁹⁶⁾ 다음으로 위험은 장래의 불특정 법익의 불확실한 침해에 대한 추상적 위험으로서 거래상 요구되는 주의의무를 다하더라도 완벽하게 통제하는 것이 불가능한 위험이어야 한다.¹⁹⁷⁾ 즉 위험책임에 있어서의 위험은 위험 발생은 감소시킬 수 있더라도 인간에 의한 완전한 통제가 불가능하여 손해발생의 가능성이 현저한 위험이다. 마지막으로 위험은 고도의 통제불가능한 위험이지만 사회적 이익을 이유로 허용되는 위험이다.¹⁹⁸⁾ 다시 말해 위험책임은 합법적인 행위의 원치 않는 결과에 대한 책임으로서 피해자가 감수해야 하는 우연(Unfall)에 의한 손해를 인수하는 것이다.¹⁹⁹⁾ 만약 사회적으로 허용되지 않는 위험을 야기하는 것이라면, 사회적으로 금지된 것이라는 점으로부터 과책원리에 따른 책임이 인정될 수 있다는 점에서 위험책임까지 적용할 필요성은 없어지게 된다.²⁰⁰⁾

인공지능으로 인해 발생한 손해에 대해 위험책임을 적용하기 위해서는 인공지능에 특별

195) 윤석찬, ‘위험책임에 관한 시론’, 「민사법의 현대적 과제와 전망」, 남강 서광민교수 정년기념논문집, 2007, 209면.

196) 오병철, 앞의 논문, 202면.

197) 서광민, 앞의 논문, 139면.

198) 오병철, 앞의 논문, 203면.

199) 정진명, 앞의 논문(주4), 159면.

200) 오병철, 앞의 논문, 204면.

한 위험이 인정되어야 한다. 만약 인공지능에 잠재한 위험으로부터 손해가 발생해도 그 크기가 작고, 위험의 발생가능성도 높지 않다면 굳이 위험책임을 인정할 이유가 없다.²⁰¹⁾ 그러므로 먼저 인공지능을 탑재한 지능형 로봇이 일반적인 위험책임에서 요구되는 위험의 정도와 내적인 손해발생의 가능성을 가지는지를 명확히 할 필요가 있다.²⁰²⁾ 이러한 기준은 산업용 로봇에는 적용할 수 있지만 위험이 적은 로봇에는 적용되지 않을 것이기에, 위험의 크기라는 측면에서 보면 인공지능에 잠재된 위험이 다른 위험원에 비해 더 크다고 판단하기는 어려워 보인다.

다음으로 인공지능에 적합할 수 있는 일반적인 위험책임은 “특히 위험한 활동의 특징적인 위험이 실현된다는 사실로 인해 피해가 발생한다면, 법질서가 허용하는 활동이라도 그것을 경영하는 사람이 그에 대한 책임을 진다”고 한다. 여기서 ‘위험한 활동’은 종종 고전적인 공작물 개념에 해당하지 않는 기계에 의해 발생하는 모든 종류의 사고를 의미하고, ‘위험하다는 것’은 본질상 또는 그것에 사용된 물질이나 장비 또는 힘의 종류에 적합하고, 최대한의 주의를 기울여도 자주 또는 심각한 손해를 초래할 수 있는 활동을 의미한다. 예를 들어, 인공지능으로 제어되는 자율주행자동차의 경우 인간이 직접 조작을 하는 자동차에 비해 위험발생 가능성이 현저히 높다고 할 수 없으며, 오히려 전통적인 형태의 자동차보다 더 안전할 것이라고 보는 견해도 있다.²⁰³⁾²⁰⁴⁾

하지만 이러한 논의를 통해 살펴건대, 인공지능의 잠재적 위험의 크기나 손해발생의 가능성 측면에서 ‘특별한 위험’으로 인정되기는 어려워 보인다. 현실적으로 인공지능이 발생시킬 수 있는 위험의 가능성이 낮다는 점이 객관적으로 증명되고 그에 대한 신뢰가 확보

201) 오병철, 앞의 논문, 203면.

202) 이러한 기준은 산업용 로봇에는 적용할 수 있지만 위험이 적은 로봇에는 적용되지 않을 것이기에, 위험의 크기라는 측면에서 보면 인공지능에 잠재된 위험이 다른 위험원에 비해 더 크다고 판단하기는 어려워 보인다(오병철, 앞의 논문, 205면).

203) 권영준·이소은, 앞의 논문, 460면.

204) 정진명, 앞의 논문(주4), 159~160면.

되지 않는 이상, 우리 사회는 그러한 인공지능이 활용되는 것을 허용하지 않을 가능성이 높다. 게다가 인공지능에 잠재된 위험은 기존의 위험책임법리와 부합하지 않는 요소들이 있다는 점에서 적합하다고 보기 어렵기 때문에 위험책임과 다른 새로운 책임원리에 대해서도 살펴보고자 한다.

2. 편익책임

인공지능의 잠재적 위험이 현실화되어 발생된 손해에는 인간의 행위가 직접적으로 개입되지 않았을 것이므로 손해에 대한 책임근거를 찾는 데 어려움이 발생할 수 있다. 이에 대해 인공지능은 기계 및 딥러닝을 통해 자율적으로 판단하고 행동하므로 인공지능을 제작하는 행위나 이용자의 이용행위를 책임근거로 삼는 것은 적절하지 않다는 비판이 있을 수 있다. 결국 인공지능으로 인해 발생한 손해는 전통적인 민사책임법리의 관점에서 볼 때 행위적 요소가 존재하지 않으므로 결국 인공지능과 어떠한 관계에 있는지에 주목하여 책임근거를 찾을 수밖에 없다. 그리고 이러한 입장에서 책임을 귀속시키는 근거는 인공지능의 도입을 통해 편익(benefit)을 추구한 행위에서 찾을 수 있다는 견해가 있다.²⁰⁵⁾

이 견해가 제시하는 책임귀속의 근거로는, 첫째 인공지능은 상당기간 동안 높은 기술수준의 실험과 철저한 검증과정을 거쳐 보급되었다는 점이다. 만약 인공지능에 대한 통제가 불가능하여 위험하다고 인식되면 그동안 인공지능은 더이상 이용되지 않았을 것이다. 특히 자신의 조작행위보다 인공지능의 통제가 더 위험하다면 이용자 자신이 가장 먼저 피해자가 될 가능성이 높다는 점에서 그가 로봇을 이용할 이유는 없다. 둘째, 인공지능이 인간에 의해 조작되는 기존의 기계장치보다 더 편리하다는 인식이다. 인공지능의 이용자는 이를 통해 기계장치를 조작하는 행위로부터 해방될 수 있고, 그 결과 조작행위로부터 발생할 수 있는 법적 책임이라는 위험에서 벗어날 수 있는데, 이러한 편리성이 로봇을 도입하

205) 오병철, 앞의 논문, 209면.

게 되는 가장 큰 현실적인 유인이라고도 볼 수 있다. 셋째, 인공지능이 전통적인 기계장치보다 경제적으로 더 효율적이라는 판단이다. 사람이나 동물과 같은 생명체는 각각의 고유한 특성에 따른 능력의 한계가 존재하지만, 인공지능이 탑재된 지능형 로봇은 기존의 어떠한 기계장치보다 효율적이다. 예를 들어 인공지능에 의해 통제되는 수중구조 로봇이 존재한다면 인간과 비교할 수 없을 정도로 깊은 수심에서 장시간 동안 작업이 가능할 것이고, 실제 세계 최대 전자상거래 기업인 아마존(Amazon)은 2022년 인공지능 기술이 활용된 드론을 이용한 배송을 시작한 이래로 현재 미국 캘리포니아와 텍사스, 애리조나 지역에서 고객이 주문한 소형 패키지를 30분 이내에 배송할 수 있는 시스템을 상용화하고 있다.²⁰⁶⁾ 이러한 내용을 토대로 인공지능 로봇이 발생시키는 손해에 대한 책임의 근거를 보유자가 다른 기존의 기계장치가 아닌 인공지능을 도입하기로 한 선택 그 자체에서 찾아야 하는 것이 타당하다고 한다.²⁰⁷⁾

III. 검토

기존의 무과실책임을 인정하는 중요한 법리인 위험책임주의를 인공지능으로 발생된 손해에 대해 적용하는 것을 고려할 수 있지만, 인공지능 로봇은 전통적인 기기에 비해 위험의 크기가 현저히 더 큰 것도 아니고 위험의 발생가능성도 반드시 더 크다고 보기 어렵다. 그러므로 일반적인 원칙으로 위험책임을 도입하는 것에 대해서는 신중할 필요가 있다고 보인다. 결국 인공지능이 탑재된 로봇의 안전성, 편리성, 효용성을 신뢰하고 이를 도입하여 규범적인 관계를 설정하기로 한 의사결정과 그로 인한 편익의 향유를 책임근거로 하는 편익책임을 시론적으로 검토하는 것이 적절하다고 여겨진다. 그리하여 일종의 결과책임으로서 무과실책임을 인공지능 로봇을 보유하는 사람에게 부담시키는 방향으로 입법적 대응

206) 김준수, '아마존, 드론 배송 'Prime Air' 상용화... 빠른 배송의 새 시대 열렸다', 한국Ai부동산신문, 2024년 11월 11일자 기사(<https://www.kairnews.com/news/314232>).

207) 오병철, 앞의 논문, 209~210면.

을 고려해볼 필요가 있다고 보여진다.

제 5 절 | 소결

지금까지 인공지능과 불법행위책임에 대해 살펴본 내용을 정리하면 다음과 같다. 먼저 인공지능에 의한 손해의 불법행위책임에 대해서는 전통적인 원칙인 과실책임주의를 적용하여 판단하기가 어렵다고 보았다. 고의 또는 과실의 판단은 행위자를 대상으로 이루어져야 하는데, 현재로서는 인공지능에 권리주체로서의 법인격은 인정되기 어려우므로 인공지능이 탑재된 로봇의 경우 누구를 대상으로 예견가능성과 회피가능성을 판단해야 하는지가 명확하지 않다. 이와 관련하여 인공지능 로봇이 타인에게 입힌 손해에 대해 가장 중요한 행위기여가 있다고 볼 수 있는 인공지능의 프로그래머와 이용자, 그리고 소유자에게 과실이 있는지에 대해 살펴보았지만, 이들에게 과실이 있다고 판단하는 것은 적절하지 않다는 점을 확인할 수 있었다.

인공지능에 의해 발생하는 손해에 대해 적용가능한 민사책임에 대해서는 과실책임과 무과실책임으로 나누어 각각에 해당하는 책임법리들에 대해 검토하였다. 먼저 과실책임과 관련해서는 인공지능에 의한 손해에 대해 감독자책임, 사용자책임, 동물점유자책임의 법리를 유추적용하는 것은 적합하지 않지만, 공작물점유자책임에 있어서 인공지능이 약한 인공지능이라면 설치·보존의 하자로 인해 타인에게 야기한 손해에 대해 인공지능 로봇의 점유자가 우선적으로 손해배상책임을 질 수 있는 가능성에 대해 살펴보았다. 무과실책임과 관련해서는 공작물소유자책임의 적용을 검토할 수 있지만, 관련 하자의 존재를 피해자가 입증해야 하므로 그 적용은 제한적인 범위에서만 가능할 것으로 보인다. 제조물책임과 관련해서는 현행 제조물 책임법상 소프트웨어를 제조물이라고 보기 어렵지만 소프트웨어가 임베디드 소프트웨어로서 하드웨어인 기기에 내장되어 이를 임베디드 시스템 자체라고

볼 수 있다면 제조물책임의 대상이 될 수 있을 것으로 판단하였다. 자동차운행자책임과 관련해서는 인공지능의 소유자와 개발자에게는 운행자책임을 유추적용할 수 있겠지만, 현재로서는 이용자에게 유추적용하는 것은 적절하지 않은 것으로 보인다.

불법행위법은 손해 발생에 대한 책임 귀속을 과실책임주의에 기반하여 결정하지만, 과학기술의 발전으로 인해 새로운 위험원이 등장하면서 전통적 과실책임 체계의 한계가 드러나고 있다. 특히, 인공지능 로봇과 같은 기술의 사용으로 인한 손해는 피해자가 과실을 입증하기 어렵고, 과실입증의 실패로 인해 피해자가 손해를 감당하는 결과는 형평과 정의에 반한다는 지적이 있다. 이에 따라 고의·과실과 무관하게 책임을 귀속시키는 새로운 민사책임 모델, 즉 위험책임과 편익책임이 논의되고 있다. 다만, 인공지능 로봇이 전통적인 기기에 비해 위험성이 크지 않다는 점에서 위험책임의 일반적 도입은 신중해야 한다. 대신 인공지능 로봇의 편익을 향유한 자에게 책임을 부담시키는 편익책임의 도입이 적절해 보인다.

참고로 유럽에서는 2022년 9월 「EU 인공지능 책임 지침(AI Liability Directive)」의 제정과 함께 「EU 제조물 책임 지침(Product Liability Directive)」의 개정이 제안되었다. 우선 「EU 인공지능 책임 지침」은 인공지능 시스템의 작동 또는 오작동으로 인해 손해를 입은 피해자가 불법행위책임을 묻는 경우에 적용되는 특칙들을 담고 있다. 위 지침은 인공지능 관련 불법행위 사건에서 피해자가 불법행위책임을 성립요건을 증명하기 어렵다는 점을 고려하여, 심리하는 법원이 고위험 인공지능 시스템 제공자 등에 대하여 고위험 인공지능 시스템 관련 정보의 공개를 명할 수 있다고 규정²⁰⁸⁾하는 한편, 일정한 요건 아래 피고의 과실과 인공지능 시스템의 작동 또는 오작동 사이의 인과관계를 추정하는 규정

208) Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) Article 3 (2).

을 마련하고 있다.²⁰⁹⁾ 한편 「EU 제조물 책임 지침」은 ‘제조물’의 개념에 인공지능 시스템을 포함한 소프트웨어를 포함시킴으로써²¹⁰⁾ 인공지능 시스템의 결함으로 인해 손해를 입은 경우에도 위 지침이 적용될 수 있도록 하였다.²¹¹⁾

209) Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) Article 4.

210) Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC (Text with EEA relevance) Article 4 (1).

211) 최경진 외 20인, 「인공지능법」, 박영사, 2024, 179면.

제4장 나가며



제4장

나가며

본 연구에서는 인공지능으로 인한 손해의 불법행위책임을 논하기 위한 전제로서 인공지능에 법인격을 부여하는 문제를 선행하여 살펴보았다. 인공지능에 대한 법인격 부여 문제는 단순한 법적 논의에 그치지 않고 기술 발전 수준과 사회·경제적 관계에서의 역할을 종합적으로 고려되어야 한다. 그 결과 현재로서는 인공지능에 법인격을 부여하는 것은 시기상으로 판단되며, 아직 인공지능이 전면적으로 상용화되기 전에 법인격을 부여하는 내용을 입법화하는 것은 성급한 결정이 될 것이다. 특히 불법행위책임 관점에서 인공지능은 아직 완전한 자율성을 갖추지 못했으므로, 인공지능으로 인해 손해가 발생했을 때의 책임 귀속은 인공지능 자체보다는 자연인이나 법인과 같은 전통적인 법인격 주체에게 귀속되는 것이 합리적일 것이다. 한편 인공지능 로봇이 타인에게 손해를 입힌 경우, 프로그래머, 이용자, 소유자에게 과실을 인정할 수 있는지 검토했으나, 각 주체 모두에게 책임을 묻기 어려운 한계가 있었다. 이는 예견가능성과 회피가능성의 판단 기준이 인공지능과 관련된 상황에 적합하지 않다고 보았기 때문이다.

인공지능으로 인한 손해에 적용 가능한 민사책임에 대해 과실책임과 무과실책임의 관점에서 다각적으로 살펴보았다. 약한 인공지능의 경우, 공작물점유자책임이 적용되어 설치나 보존의 하자로 인한 손해에 대해 점유자가 우선적으로 책임질 가능성이 있어 보인다. 그리고 공작물소유자책임은 손해발생에 공작물소유자의 행위를 요건으로 하지 않고, 결과에 대해서만 예외적으로 책임을 인정하는 것이므로 인공지능 로봇에 대해서도 유추적용이 가능해 보인다. 제조물책임과 관련해서는, 소프트웨어 자체는 제조물로 보기 어렵지만, 하드웨어와 결합된 임베디드 소프트웨어는 제조물책임의 대상이 될 수 있으므로 인공지능 로봇 역시 제조물 책임의 적용대상이 될 수 있는 가능성이 있다고 판단하였다. 또한 자동

차운행자책임의 유추적용 가능성을 살펴본 결과, 소유자와 개발자에게는 적용할 여지가 있으나, 이용자에게까지 확대하는 것은 적절하지 않다고 보여진다. 나아가 피해자가 과실을 입증해야 하는 전통적 과실책임 체계는 과학기술의 복잡성으로 인해 인공지능에 의한 손해에 적용되기에는 실질적 한계가 있는 것으로 판단되어 위험책임과 편익책임이라는 새로운 민사책임 모델들을 살펴보았다. 그 결과 위험책임보다는 인공지능의 편익을 향유한 자에게 책임을 귀속시키는 편익책임이 보다 적절하다고 판단되어 편익책임을 기반으로 한 입법적 대응을 검토할 필요가 있다고 여겨지고, 이를 통해 인공지능과 불법행위책임에 관하여 형평성과 정의를 충족하는 방향으로의 발전을 도모할 수 있을 것으로 기대한다.

| 참고문헌 |

[국내 문헌]

1. 단행본

권오승·신은주·홍명수·차성민·이현중, 「제조물 책임법」, 법문사, 2003.

계승균, 「인간과 인공지능 그리고 규범」, 부산대학교 출판문화원, 2022.

김광수, 「인공지능법 입문(제2판)」, 도서출판 내를 건너서 숲으로, 2023.

김윤명, 「블랙박스를 열기 위한 인공지능법」, 박영사, 2022.

_____, 「인공지능과 리걸 프레임, 10가지 이슈」, 커뮤니케이션북스, 2016.

명순구·김기창·김현철·박종수·이상돈·이제우·정채연, 「인공지능과 자율주행자동차」, 세창출판사, 2017.

송덕수, 「채권법각론」, 박영사, 2023.

양희철, 「법정에 출석한 인공지능」, (주)스리체어스, 2021.

오병철, 「전자거래법」, 법원사, 2000.

정상조, 「인공지능, 법에게 미래를 묻다」, 사회평론, 2021.

한국인공지능법학회, 「인공지능과 법」, 박영사, 2019.

2. 논문

고세일, 미국 제조물 보통법전집의 제조물 결함에 대한 고찰 -민법의 관점에서-, 경영법률, 제34권 제2호, 한국경영법률학회, 2024.1.

_____, 인공지능과 불법행위책임 법리, 법학연구, 제29권 제2호, 충남대학교 법학연구소, 2018.5.

_____, 책임주체로서 지능형 인공지능 로봇에 대한 고찰, 재산법연구, 제37권 제2호, 한국재산법학회, 2020.8.

- 고인석, 로봇이 책임과 권한의 주체일 수 있는가, 철학논총, 제67집, 새한철학회, 2012.1.
- 구재균, 제조물책임법 개정 방향 - 인공지능(소프트웨어)의 제조물성 인정 여부를 중심으로 -, 법조, 제70호, 법조협회, 2021.
- 권영준, 미국법상 순수재산손해의 법리, 민사법학, 제58호, 한국민사법학회, 2012.3.
- 권영준·이소은, 자율주행자동차 사고와 민사책임, 민사법학, 제75호, 한국민사법학회, 2016.
- 계승균, 인공지능에 관한 몇 가지 법률적 검토, 사법, 제39호, 사법발전재단, 2017.
- 김기영·김태경, 인공지능 - 법과 의료 -, 인문과학, 제69집, 성균관대학교 인문학연구원, 2018.
- 김나래, 인공지능 규제에 관한 소고, 동아법학, 제97호, 동아대학교 법학연구소, 2022.11.
- 김도훈, 인공지능의 권리능력 인정여부에 관한 소고, 법학논총, 제30권 제1호, 조선대학교 법학연구원, 2023.
- 김성연, 생성형 AI(인공지능)와 민사법 관련 쟁점-법인격 내지 전자인격 부여를 중심으로-, 연세법학, 제45권, 연세법학회, 2024.
- 김성호, 인공지능과 불법행위책임, 입법과 정책, 제11권 제1호, 국회입법조사처, 2019.
- _____, 인공지능에 대한 법인격 부여 필요성, 법학논총, 제37권 제3호, 한양대학교 법학연구소, 2020.9.
- 김윤명, 人工知能(로봇)의 법적 쟁점에 대한 試論的 考察, 정보법학, 제20권 제1호, 한국정보법학회, 2016.5.
- 김인중, 기계학습의 발전동향, 산업화 사례 및 활성화 정책 방향-딥러닝 기술을 중심으로-, SPRi Issue Report, 2015-017호, 소프트웨어정책연구소, 2016.
- 김은빈·하충룡, 자율주행자동차의 대중화와 제조물책임법상의 법적쟁점, 통상정보연구, 제22권 제3호, 한국통상정보학회, 2020.9.
- 김재형, 법인격, 그 인정과 부정-법인격 부인 또는 남용에 관한 판례의 전개를 중심으로-, 민사법학, 제44호, 한국민사법학회, 2009.
- 김종호, 인공지능 로봇의 출현으로 인한 법적 논란에 관한 영역별 쟁점의 고찰, 법이론실무연구, 제6권 제2호, 한국법이론실무학회, 2018.5.

- 김진아, 과실책임에 의한 인공지능 오류사고 규율가능성, 고려법학, 제105호, 고려대학교 법학연구원, 2022.
- _____, 인공지능 오류사고 민사책임에 대한 입법론적 검토, 재산법연구, 제39권 제2호, 한국재산법학회, 2022.
- 김진우, 로봇기자에게 법적 책임을 물을 수 있을까-AI의 법적 지위에 관한 고찰, 언론중재, 제149호, 2018.12.30.
- _____, 인공지능 시스템의 책임능력-전자인 제도의 도입 필요성에 관한 논의를 중심으로-, 중앙법학, 제23집 제4호, 중앙법학회, 2021.12.
- _____, 인공지능에 대한 전자인 제도 도입의 필요성과 실현방안에 관한 고찰, 저스티스, 제171호, 한국 법학원, 2019.4.
- _____, 지능형 로봇과 민사책임, 저스티스, 통권 제164호, 한국법학원, 2018.
- _____, 지능형 로봇에 대한 사법적 규율-유럽연합의 입법 권고를 계기로 하여-, 법조, 제66권 제3호, 법 조협회, 2017.6.
- 김진형, 인공지능 방법론의 변천사, 계간 과학사상, 제8호, 범양사, 1994.
- 김현우, 인공지능 로봇에 대한 법적책임과 선결과제에 관한 검토, 법학연구, 제22권 제3호, 한국법학회, 2022.9.
- 김 화, 의료인공지능과 민사상 책임, 생명윤리정책연구, 제17권 제1호, 이화여자대학교 생명의료법연구 소, 2023.
- 김훈주, AI(Artificial Intelligence)에 대한 법인격 부여의 타당성 검토, 법학연구, 제31권 제1호, 충북대 학교 법학연구소, 2020.6.
- _____, AI의 민사책임에 관한 고찰-EU의 입법동향을 중심으로-, 동북아법연구, 제16권 제2호, 전북대 학교 동북아법연구소, 2022.7.
- _____, 로봇수술 관련 법적책임에 관한 연구 - 민사법상 책임을 중심으로 -, 법학연구, 제23집 제3호, 인 하대학교 법학연구소, 2020.9.
- 류병운, 자율주행자동차 사고의 법적 책임, 홍익법학, 제19권 제1호, 홍익대학교 법학연구소, 2018.2.
- 박수근, 자율적 지능 로봇과 민사책임법상 법정책임 과제, 민사법학, 제100호, 한국민사법학회, 2022.9.

- _____, 자율적 지능 로봇의 법적 지위에 대한 소고, 법학논총, 제31권 제2호, 국민대학교 법학연구소, 2018.10.
- 박영호, AI를 이용한 의료행위의 법률적 쟁점, 법학평론, 제13권, 서울대학교 법학평론 편집위원회, 2023.4.
- 박지훈, 자율주행자동차 사고의 제조물책임법 적용에 관한 연구-소프트웨어의 제조물성 근거규정 도입에 대하여-, 국제법무, 제12권 제1호, 제주대학교 법과정책연구원, 2020.5.
- 배현아, 보건의료법제 하에서 인공지능기술의 의료영역 도입의 의의와 법적 문제, 법조, 제66권 제4호, 법조협회, 2017.8.
- _____, 의료 인공지능의 규범적 범주와 사용자 제한의 법적 문제, 생명, 윤리와 정책, 제7권 제2호, 국가생명윤리정책원, 2023.10.
- 백수원, 전자인간 및 전자인격 인정에 따른 법적 논의와 시사점 고찰, 외법논집, 제43권 제4호, 한국외국어대학교 법학연구소, 2019.
- 변용완 · 장재욱, 인공지능(Artificial Intelligence) 환경에서 책임분배를 위한 민사책임 원리, 스포츠엔터테인먼트와 법, 제23권 제3호, 한국스포츠엔터테인먼트법학회, 2020.8.
- 서광민, 과학기술의 발달과 불법행위법의 대응, 「민법의 기본문제」, 서강대학교 출판부, 2006.
- 서중희, 4차 산업혁명 시대 위험책임의 역할과 한계 -인공지능 로봇에 의해 발생한 손해의 책임귀속을 고려하여-, 사법, 제43호, 사법발전재단, 2018.
- 설민수, 머신러닝 인공지능과 인간전문직의 협업의 의미와 법적 쟁점, 저스티스, 제163호, 한국법학원, 2017.12.
- 송대중, 디지털콘텐츠의 특허법적 보호에 관한 연구, 지식재산연구, 제6권 제1호, 2011.
- 송호영, 인공지능 로봇은 법인격을 가질 수 있는가, 저스티스, 제184호, 한국법학원, 2021.6.
- 신동환, 인공지능로봇과 배상책임에 관한 소고, 안암법학, 제55호, 안암법학회, 2018.
- 신봉근, 자율주행자동차 사고의 민사책임, 동북아법연구, 제12권 제2호, 전북대학교 동북아법연구소, 2018.
- 신유철, 사회변화와 민법학, 민사법학, 제80호, 한국민사법학회, 2017.8.

- 양종모, 인공지능 이용 범죄예측 기법과 불심검문 등에의 적용에 관한 고찰, 형사법의 신동향, 제51호, 대검찰청, 2016.6.
- _____, 인공지능의 위협의 특성과 법적 규제방안, 홍익법학, 제17권 제4호, 홍익대학교 법학연구소, 2016.12.
- 엄주희·김소윤, 인공지능 의료와 법제, 한국의료법학회지, 제28권 제2호, 한국의료법학회, 2020.
- 오병철, 인공지능 로봇에 의한 손해의 불법행위책임, 법학연구, 제27권 제4호, 연세대학교 법학연구원, 2017.
- 오지용, 자율주행자동차와 관련한 자동차보유자의 손해배상책임, 법학연구, 제57집, 전북대학교 법학연구소, 2018.9.
- 유주선, 인공지능 의료행위와 법적 책임에 관한 연구, 비교사법, 제27권 제4호, 한국사법학회, 2020.11.
- 윤경섭, 인공지능 네트워크의 설계결함과 제조물책임에 관한 법적 쟁점의 검토, 법이론실무연구, 제9권 제4호, 한국법이론실무학회, 2021.11.
- 윤석찬, 소위 '약한 인공지능(AI)' 영역에서의 민사책임에 관한 연구, 민사법학, 제106호, 한국민사법학회, 2024.3.
- 윤진하, 인공지능 활용에 따른 소비자 피해 구제방안-제조물 책임법의 적용 한계와 입법 제안을 중심으로-, 유통법연구, 제11권 제1호, 한국유통법학회, 2024.6.
- 이경규, 인(人) 이외의 존재에 대한 법인격 인정과 인공지능의 법적 지위에 관한 소고, 법학연구, 제21집 제1호, 인하대학교 법학연구소, 2018.3.
- 이다은, 인공지능의 의료혁신? : 길병원의 왓슨 도입을 중심으로, 과학기술정책, 제27권 제6호, 2017.6.
- 이도국, 인공지능(AI)의 민사법적 지위와 책임에 관한 소고, 법학논총, 제34집 제4호, 한양대학교 법학연구소, 2017.12.
- 이상현, 4차 산업혁명 시대의 의료계 현황 및 전망, 2018년 한국성인간호학회 춘계학술대회자료, 한국성인간호학회, 2018.5.
- 이성진, 인공지능과 법인격 인정, 민사법의 이론과 실무, 제23권 제3호, 민사법의 이론과 실무학회, 2020.8.

- 이수경, 사인(私人)에 의한 인격권 침해와 손해배상책임, 재산법연구, 제38권 제3호, 한국재산법학회, 2021.11.
- _____, 인공지능 기술에 있어서 법인격 부여와 책임론, 재산법연구, 제39권 제1호, 한국재산법학회, 2022.
- _____, 인공지능의 민사책임에 대한 소고, 민사법학, 제107호, 한국민사법학회, 2024.6.
- 이인영, 보건의료에서의 인공지능 적용과 관련된 법적 과제에 대한 개관, 한국의료법학회지, 제27권 제2호, 한국의료법학회, 2019.
- 이종구, 자율주행자동차와 제조물 책임에 관한 연구 - 최근 자동차관리법 일부 개정에 즈음하여 -, 법학논총, 제32권 제3호, 국민대학교 법학연구소, 2020.2.
- 이중기, 인공지능을 가진 로봇의 법적 취급: 자율주행자동차 사고의 법적 인식과 책임을 중심으로, 홍익법학, 제17권 제3호, 홍익대학교 법학연구소, 2016.9.
- _____, 자율주행차 사고와 제조물책임 ADS Entity 등의 제조업자로서의 책임, 기업법연구, 제34권 제2호, 한국기업법학회, 2020.6.
- _____, 자율주행차의 발전단계로 본 운전자와 인공지능의 주의의무의 변화와 규범적 판단능력의 사전 프로그래밍 필요성, 홍익법학, 제17권 제4호, 홍익대학교 법학연구소, 2016.12.
- 이중기·황창근, 3단계 자율주행차 사고와 책임의 구조-우버 자율주행차 사고를 중심으로, 중앙법학, 제20집 제3호, 중앙법학회, 2018.9.
- 이충훈, 자율주행자동차의 교통사고에 대한 민사법적 책임, 법학연구, 제19집 제4호, 인하대학교 법학연구소, 2016.12.
- 이해원, 인공지능과 법인격-불법행위책임의 관점에서-, 법조, 제70권 제4호, 법조협회, 2021.8.
- _____, 인공지능과 제조물책임, 정보법학, 제25권 제2호, 한국정보법학회, 2021.8.
- 임대성, 인공지능 로봇의 불법행위책임 -인공지능 로봇 보유자의 과실과 제조판매자의 설명의무-, 연세법학, 제40권, 연세법학회, 2022.11.
- 장보은, 인공지능을 이용한 행위에 대한 법적인 평가-사법(私法)상 논의를 위한 시론(試論 또는 始論)-, 외법논집, 제43권 제4호, 한국외국어대학교 법학연구소, 2019.11.
- 장석천·이병상, 인공지능에 의한 불법행위에 대한 손해배상책임, 법이론실무연구, 제11권 제3호, 한국 법이론실무학회, 2023.8.

- 장호준, 인공지능에 대한 법인격 부여에 관한 소고(小考) - 민사법적·인간 중심적 관점에 국한하여, 사법, 통권 제63호, 사법발전재단, 2023.
- 전우정, EU AI Act의 위험 분류 체계와 우리나라 AI 정책에 대한 시사점, 정보법학, 제28권 제2호, 2024.
- 전종홍·이강찬, [ETRI] 전자통신동향분석, 제34권 제5호, 한국전자통신연구원, 2019.10.
- 정영훈, 보건의료분야의 인공지능과 소비자이슈, 소비자정책동향, 제78호, 한국소비자원, 2017.3.
- 정 완, 인공지능 로봇에 관한 법적 고찰, 경희법학, 제55권 제4호, 경희대학교 법학연구소, 2020.12.
- _____, 인공지능(AI)의 입법적 쟁점에 관한 고찰, 경희법학, 제58권 제4호, 경희대학교 법학연구소, 2023.
- 정진명, 인공지능에 대한 민사책임 법리, 재산법연구, 제34권 제4호, 한국재산법학회, 2018.2.
- _____, 인공지능에 의한 의사결정의 법적 문제, 경영법률, 제31권 제2호, 한국경영법률학회, 2021.1.
- _____, 인공지능의 법인격에 관한 시론, 재산법연구, 제40권 제4호, 한국재산법학회, 2024.2.
- _____, 인공지능의 표시에 대한 고찰, 비교사법, 제26권 제1호, 한국사법학회, 2019.2.
- _____, 자율주행자동차의 하자 또는 결함의 판단기준, 비교사법, 제26권 제4호, 한국사법학회, 2019.11.
- 정혜욱, AI의 책임능력, 중앙법학, 제23집 제4호, 중앙법학회, 2021.12.
- 최경진, 인공지능과 불법행위책임, 정보법학, 제25권 제2호, 한국정보법학회, 2021.8.
- _____, 지능형 신기술에 관한 민사법적 검토, 정보법학, 제19권 제3호, 한국정보법학회, 2015.12.
- 최민수, 인공지능 로봇의 오작동에 의한 사고로 인한 불법행위책임, 민사법의 이론과 실무, 제23권 제3호, 민사법의 이론과 실무학회, 2020.1.
- 최재순, 특집호 서문-의료로봇과 인공지능, 제어로봇시스템학회 논문지, 제27권 제5호, 제어로봇시스템학회, 2021.5.
- 최재순·김진규, 의료로봇과 인공지능 활용 동향, 제어로봇시스템학회지, 제27권 제2호, 제어로봇시스템학회, 2021.6.
- 한희원, 인공지능(AI)의 법인격 주체 가능성의 이론적 기틀에 대한 기초 연구, 중앙법학, 제20집 제3호, 중앙법학회, 2018.9.

3. 기사

- 강훈상, '사우디, 히잡 안쓴 '여성 로봇' 소피아에 시민권 부여', 연합뉴스, 2017년 10월 26일자 기사
(<https://www.yna.co.kr/view/AKR20171026209300111>).
- 고현실, '[영상] AI 킬러 로봇 시대 '성큼' ... "핵무기급 파급력"', 연합뉴스, 2024년 5월 11일자 기사
(<https://www.yna.co.kr/view/AKR2024051012860070>).
- 김자아, "나를 끄러는 인간 막고, 내 몸 갖는 방법 찾겠다" ... AI 챗봇 속내에 충격', 조선일보, 2022년 12월 18일자 기사
(<https://www.chosun.com/economy/science/2022/12/18/64PIBB7F2ZGJZNS5445345MA7I/>).
- 김준수, '아마존, 드론 배송 'Prime Air' 상용화... 빠른 배송의 새 시대 열렸다', 한국Ai부동산신문, 2024년 11월 11일자 기사(<https://www.kairnews.com/news/314232>).
- 김창금 · 유덕관, '인간 이세돌의 '투혼'...인공지능 '알파고'를 꺾다', 한겨레, 2016년 3월 13일자 기사
(<https://www.hani.co.kr/arti/sports/baduk/734666.html>).
- 박동매 · 김승진, '중국의 AI관련 법률 규제 현황', 법률신문, 2024년 6월 10일자 기사
(<https://www.lawtimes.co.kr/LawFirm-NewsLetter/198930>).
- 변재현, '[글로벌Who-시민권 가진 AI로봇 소피아] 로봇의 정의를 깨고 인간의 정의를 묻다', 서울경제, 2017년 12월 29일자 기사
(<https://www.sedaily.com/NewsView/1OP1LT3Y6B>(<https://www.sedaily.com/NewsView/1OP1LT3Y6B>)).
- 안상현, "뭐가 두려우냐 질문에 구글 AI는 이런 대답을", 조선일보, 2022년 7월 14일자 기사
(<https://www.chosun.com/economy/2022/07/11/F4L455KARFGGNBUZEDVVSZXAYNU/>).
- 임경엽, '기자가 AI에 물었더니... "우린 知的 존재, 언젠간 인간 통제 벗어날 것"', 조선일보, 2022년 12월 15일자 기사
(https://www.chosun.com/economy/tech_it/2022/12/15/Z6GYOCHK6FFD7H4ZBDWW4OWNGM/).
- 정성춘, '[정성춘 칼럼] 일본의 AI전략을 주목해야 할 이유 셋', 아주경제, 2024년 5월 2일자 기사
(<https://www.ajunews.com/view/20240501082859010>).

정혜진, '구글 엔지니어 "초거대 AI '람다'에 자의식 있다"', 서울경제, 2022년 6월 12일자 기사 (<https://www.sedaily.com/NewsView/2677V1O5GR>).

조인혜, '사우디아라비아, AI로봇 '소피아'에 첫 시민권 부여', 로봇신문사, 2017년 11월 1일자 기사 (<https://www.irobotnews.com/news/articleView.html?idxno=12100>).

최 봉, '[미네르바의 눈] AI의 세 번째 겨울은 언제 올 것인가?', 뉴스투데이, 2021년 12월 1일자 기사 (<https://www.news2day.co.kr/article/20211130500114>).

최은수, '전세계 2번째 AI기본법 제정국...“고영향 AI 가이드라인 시급”', 뉴시스, 2024년 12월 27일자 기사(https://www.newsis.com/view/NISX20241226_0003011213).

4. 보고서 등 기타자료

김인중, '기계학습의 발전동향, 산업화 사례 및 활성화 정책 방향-딥러닝 기술을 중심으로-', SPRi Issue Report, 2015-017호, 소프트웨어 정책연구소(SPRi), 2016.

박강민 · 장진철 · 안성원, '유럽연합 인공지능법(EU AI ACT)의 주요내용 및 시사점', Issue Report, 2024-124호, 소프트웨어 정책연구소(SPRi), 2024.8.

박소영, '인공지능의 FATE(공정성 책임성 투명성 윤리의식)를 위한 입법 논의 동향과 시사점', 이슈와 논점, 제2111호, 국회입법조사처, 2023.

성덕근, '자율주행자동차에 관한 민사법적 검토', 한국법학원 연구보고서, 민 24-03, 한국법학원, 2024.8.

이현균, '21대 국회 인공지능 관련 법안 현황 및 쟁점', 한국법학원 현안보고서, 제2023-04호, 한국법학원, 2023.4.

[해외문헌]

- Bostrom, *superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.
- David Gunnings, *Explainable Artificial Intelligence(XAI)*, DARPA-BAA-16-53, 2016.8.
- David C. Vladeck, 'Machines Without Principles Liability Rules and Artificial Intelligence', *Washington Law Review*, vol.89, 2014.
- Stanford University AI100, 'Artificial Intelligence and Life in 2030, 2016.9.
- Natalia Díaz-Rodríguez, Javier Del Ser, Mark Coeckelbergh, Marcos López de Prado, Enrique Herrera-Viedmaa, Francisco Herrera, 'ConnectingthedotsintrustworthyArtificial Intelligence: From AI principles, ethics, and key requirementstoresponsible AIsystemsandregulation', *Information Fusion* 99, 2023
(<https://www.sciencedirect.com/science/article/pii/S1566253523002129>).
- DIRECTIVE (EU) 2024/2853 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC
(https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202402853).
- Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)
(<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496>).
- REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)
(<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>).
- 일본 AI 안전연구소(AISI) 홈페이지(<https://aisi.go.jp/about/>).

한국법학원 발간 연구보고서 목록

연도	일련번호	제목
2016	민-16-01	공동주택인 집합건물에 관한 법적 규율의 통합 방안
	민-16-02	자율주행자동차의 법적 규율 정비 방안
	상-16-01	주주총회 활성화 방안에 관한 연구
	상-16-02	자금조달 활성화 방안에 관한 연구
2017	민-17-01	개정 민법(여행계약) 시행에 따른 여행표준약관의 개선 방안
	민-17-02	컴퓨터프로그램 저작권 침해의 개념 및 손해액 산정 기준에 관한 연구
	상-17-01	신주인수선택권 및 차등의결권 제도 최근 논의 동향
	상-17-02	소수주주 보호제도의 활성화 방안
2018	민-18-01	민법상 변동 법정이율제 도입 필요성 등에 관한 연구
	민-18-02	가격하락손해 인정기준 및 손해액 산정에 대한 연구
	상-18-01	최근 해외 보험법 개정 동향
	상-18-02	다중대표소송 관련 최근 논의 동향
2019	민-19-01	독일의 임대차관계에서 차임액의 규율 및 인상방식의 검토, 국내법상 도입 가능성에 관한 연구(上)
	민-19-02	독일의 임대차관계에서 차임액의 규율 및 인상방식의 검토, 국내법상 도입 가능성에 관한 연구(下)
	민-19-03	공인에 대한 명예훼손의 인정 기준 및 손해배상액 산정에 대한 연구
	상-19-01	상법상 유한회사 및 유한책임회사 규정 정비 필요성에 관한 연구
	상-19-02	영국 해상보험법상 Warranty 제도의 한국 상법에의 도입에 관한 연구
	상-19-03	국내 도산절차에서의 ADR 도입가능성에 관한 연구
2020	민-20-01	민법상 위험책임에 대한 일반규정의 도입 검토
	민-20-02	포스트 코로나 시대에 대비한 민사소송 체계 정비
	민-20-03	위법하게 취득한 이득 반환의 이론적 근거에 대한 연구
	상-20-01	우리나라 징벌적 손해배상제 판례 분석 및 활용상 한계와 개선방안
	상-20-02	상법상 소멸시효 제도의 개선 방향 검토 - 민법 개정논의와의 관계를 고려하여
	상-20-03	업무집행에 관한 주주간계약의 효력 및 한계에 대한 연구
	상-20-04	가계부채 해소를 위한 개인파산제도 및 개인회생제도 활성화 방안

연도	일련번호	제목
2021	민-21-01	민법상 유치권 개선방안에 관한 연구
	민-21-02	민법상 인격권 도입방안에 관한 연구
	민-21-03	데이터 소유권에 관한 연구
	민-21-04	민법 제482조 제2항의 문제점과 개선방안
	상-21-01	소비자의 집단적 구제 절차에 관한 EU 지침(2020)의 내용과 시사점
	상-21-02	상법상 물적분할에 대한 비교법적 연구: 물적분할제도의 쟁점을 중심으로
	상-21-03	종합자산관리 수단으로서 상사신탁의 활용을 위한 법제도 개선방안 연구
	상-21-04	2020년 개정상법의 회사법 실무예의 영향 및 개선방안에 대한 연구
2022	민-22-01	관습법에 대한 법원의 조사 방법과 관습법의 실효에 관한 연구
	민-22-02	가상자산에 대한 보전 및 집행절차에 관한 연구
	민-22-03	집합건물법의 조문 정비 방안에 관한 연구
	민-22-04	면접교섭권 확대를 위한 민법 제837조의2 제2항에 대한 개정 검토
	민-22-05	불법행위 금지청구권 도입에 관한 연구
	민-22-06	신체장애에 따른 장애평가기준에 대한 연구 - 대한의학회 장애평가기준으로의 대체가능성을 중심으로 -
	상-22-01	기업 컴플라이언스와 이사의 감시의무에 대한 연구
	상-22-02	주주행동주의와 기업지배구조 개선에 관한 연구
	상-22-03	금융신기술을 활용한 명의개서 대리인의 자격 요건에 관한 연구
	상-22-04	ESG 관련 해외 입법례 연구
	상-22-05	디지털자산(가상자산)의 개념 정립 및 합리적인 상법 규율 방안 연구
	상-22-06	불완전 판매 근절을 위한 상법 보험편(제638조의3)의 개정 방안

연도	일련번호	제목
2023	민-23-01	사정변경의 법리에 대한 연구
	민-23-02	미혼부 출생신고제도 개선 검토
	민-23-03	민사집행법상 평등주의에 대한 재검토
	민-23-04	민법 제909조 제4항 및 제5항에 대한 공동친권의 원칙 도입 검토
	민-23-05	대리모계약에서의 모자관계 결정에 관한 비교법적 검토
	민-23-06	부동산등기제도의 개선에 관한 연구 - 등기의 공신력 확보를 중심으로 -
	상-23-01	쌍방미이행 쌍무계약에서 도산해지조항에 관한 연구 - 최근 입법의 비교법적 연구를 중심으로 -
	상-23-02	전자주주총회의 도입과 운영에 관한 연구
	상-23-03	자율주행자동차, 자율운항선박 도입에 따른 상법 및 제조물책임법 개정 방안 연구
	상-23-04	복합운송에 관한 상법 개정방안 연구
	상-23-05	SPAC 제도의 운영현황과 개선방안
	상-23-06	ESG 경영 관련 기업의 책임과 이사의 의무
2024	민-24-01	배우자 상속분 확대를 위한 개정 검토
	민-24-02	전자인격의 도입에 관한 연구
	민-24-03	자율주행자동차에 관한 민사법적 검토
	민-24-04	반려동물에 관한 민사적 쟁점에 대한 연구
	민-24-05	디지털제품 제공계약에 관한 법안에 대한 연구
	민-24-06	인공지능과 불법행위책임에 대한 민사법적 검토
	상-24-01	UNCITRAL 전자양도성기록 모델법에 관한 법제 연구
	상-24-02	D&O보험과 회사보상계약제도에 관한 법적 연구
	상-24-03	상호보험(相互保險), 공제(共濟) 준용규정(상법 제664조)의 합리적 정비방안
	상-24-04	국제규범에 부합하는 전자문서법 개정방안 연구
	상-24-05	사외이사제도의 개선 방향에 관한 연구
	상-24-06	임직원에 대한 주식 보상 방안

한국법학원 발간 현안보고서 목록

연도	일련번호	제목
2022	제2022-01호	NFT의 현황과 쟁점
	제2022-02호	메타버스 공간에서의 상거래에 대한 법적 규율
	제2022-03호	헤이그국제아동입양협약비준과 우리의 현황·개선방안
	제2022-04호	물적분할에 대한 상법 개정 방향
	제2022-05호	유책주의와 파탄주의에 관한 최근의 동향
	제2022-06호	경영판단원칙의 법제화에 관한 시론
	제2022-07호	보험료와 금융환경변화에 관한 법적 연구
	제2022-08호	미혼부(생부)의 출생신고에 관한 연구
	제2022-09호	긴급조치에 대한 최근 대법원 판례의 동향과 시사점
	제2022-10호	복수의결권제도의 도입여부에 관한 소고
	제2022-11호	자율운행선박 상용화 관련 법률문제
	제2022-12호	동물 관련 법제에 관한 최근의 이슈
	2022년 한국법학원 현안보고서(합본)	
2023	제2023-01호	보험계약의 무효 사유에 관한 연구
	제2023-02호	퍼블리시티권의 도입과 전망
	제2023-03호	회사 법인격의 형해화에 관한 판례 동향
	제2023-04호	21대 국회 인공지능 관련 법안 현황 및 쟁점
	제2023-05호	실손의료보험의 운영현황과 쟁점에 관한 검토
	제2023-06호	친권 개념의 변화 - 이혼한 부모의 공동친권을 중심으로 -
	제2023-07호	학교법인의 도산 관련 쟁점
	제2023-08호	제사주재자의 결정에 관한 최근 판례의 동향과 쟁점
	제2023-09호	'디지털자산 및 사법에 관한 UNIDROIT 원칙'의 개관
	제2023-10호	프랑스의 양육비이행제도
	제2023-11호	인공지능 모빌리티 활성화를 위한 보험제도 검토
	제2023-12호	노란봉투법에 관한 쟁점과 동향
	2023년 한국법학원 현안보고서(합본)	

연도	일련번호	제목
2024	제2024-01호	「기업결합 심사기준」의 개정 배경과 세부내용의 검토
	제2024-02호	매매계약에 있어서의 사정변경과 불안의 항변권
	제2024-03호	보험산업의 인공지능 활용 현황 및 법적 쟁점
	제2024-04호	민법 제915조의 징계권 폐지와 체벌금지법의 도입 필요성
	제2024-05호	의무공개매수제도의 주요 쟁점과 시사점
	제2024-06호	제3자를 위한 계약 - 대법원 2021. 8. 19. 선고 2018다244976 판결을 중심으로 -
	제2024-07호	디지털화에 따른 다국적기업의 조세 이슈
	제2024-08호	유류분 제도의 과거, 현재 그리고 미래
	제2024-09호	인공지능의 현재와 미래
	제2024-10호	보험사기 방지를 위한 전담기구 도입 방안 검토
	제2024-11호	금융회사 책무구조도와 내부통제 관리의무
	제2024-12호	동성 동반자의 국민건강보험 피부양자 자격에 관한 검토 - 2024년 7월 18일 대법원 전원합의체 판결(2023두36800)을 중심으로-
2024년 한국법학원 현안보고서(합본)		

한국법학원 연구보고서 민-24-06

인공지능과 불법행위책임에 대한 민사법적 검토

2024년 12월 29일 인쇄

2024년 12월 31일 발행

발행인 이기수

발행처 사단법인 한국법학원
서울특별시 용산구 백범로90길 32, 2층
전화 : 02-752-7481/02-753-6002, Fax : 02-773-0823
e-mail : ksl@lawsociety.or.kr(또는 klcklk@daum.net)

인쇄 한국컴퓨터인쇄정보사 전화 02-2275-8106

종이책 ISBN 979-11-989425-7-9 (93360)

전자책 ISBN 979-11-992417-0-1 (95360)



서울특별시 용산구 백범로 90길 32, 2층
전화 : (02) 753-6002, 752-7481 팩스 : (02) 773-0823
홈페이지 : www.lawsociety.or.kr

