



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

교육학박사 학위논문

로봇윤리(roboethics)의 도덕교육적
적용 방안 연구

2020년 2월

서울대학교 대학원

윤리교육과

지 현 아

국문초록

로봇윤리의 도덕교육적 적용 방안 연구

지 현 아

제4차 산업혁명의 도래로 인해 인공지능 로봇과 관련된 다양한 윤리적 문제가 예상되고 있고, 이러한 문제를 해결할 수 있는 로봇윤리의 중요성이 더욱 커지고 있다. 이에 따라 교육과정에서도 로봇윤리를 적용하여 교육현장의 새로운 변화가 필요하다는 목소리가 보다 힘을 얻고 있다. 이에 본 연구는 인공지능 로봇으로 인한 윤리적 문제들을 해결할 수 있는 핵심 요인으로서 로봇윤리에 주목하여, ‘로봇윤리의 도덕교육적 적용 방안’을 탐색하는 것을 목적으로 한다.

지금까지 제4차 산업혁명과 로봇윤리에 대한 사람들의 높은 관심에 비해 로봇윤리를 도덕과 교육과정에 적용하는 것에 대한 체계적인 논의는 부족했다고 평가할 수 있다. 이러한 문제의식에 입각하여 본 연구는 로봇윤리의 의미와 범주를 규명한 후, 이를 토대로 로봇윤리교육의 의미와 내용을 정립하여 최종적으로 도덕과 교육을 통한 로봇윤리교육의 구현 방안을 제시하였다.

로봇윤리는 ‘인간과 로봇이 함께 살아가기 위해 인간과 로봇이 지켜야 할 윤리’를 의미한다. 본 연구에서는 로봇윤리를 로봇이 지켜야 할 ‘로봇의 윤리’와 인간이 지켜야 할 ‘인간의 윤리’로 범주화하여 각각의 내용과 쟁점을 제시하였다. ‘로봇의 윤리’의 주요 쟁점으로는 로봇의 도덕적 지위에 대해 살펴보았고, ‘인간의 윤리’의 주요 쟁점으로는 인간 존엄성, 프라이버시, 투명성, 책임에 대해 살펴보았다.

나아가 로봇윤리의 의미와 쟁점에 근거하여 로봇윤리교육의 의미와 내용을 규명하였다. 본 연구는 로봇윤리교육의 구체화를 위해 ‘로봇의 윤리’와 ‘인간의 윤리’에 대한 범주를 토대로 로봇윤리교육의 필요성, 의미, 목표, 내용을 탐색하였다. ‘로봇의 윤리’에 관한 교육은 포

스트휴먼시대 인간성 재정립을 위해 필요하고, ‘인간의 윤리’에 관한 교육은 과학기술 시민성을 함양하기 위해 필요하다. 본 연구에서 로봇윤리교육은 ‘인간과 로봇의 공존을 위해 인간과 로봇을 탐구하고 성찰하는 교육’을 의미한다. 이와 같은 정의에 기반을 두는 로봇윤리교육은 ‘인간과 로봇에 대한 윤리적 성찰’과 ‘로봇의 문제에 대한 윤리적 해결’을 목표로 한다. 이러한 목표를 달성하기 위한 로봇윤리교육의 내용을 제시하면 다음과 같다.

첫째, ‘인간과 로봇에 대한 윤리적 성찰’의 목표를 달성하기 위해 ‘인간과 로봇의 존재론에 대한 윤리적 성찰’을 내용 영역으로 설정하였다. 이와 같은 영역에서 다룰 구체적 내용 요소로는 인공지능 로봇의 도덕적 지위, 인간과 로봇의 정체성을 도출하였다.

둘째, ‘로봇의 문제에 대한 윤리적 해결’의 목표를 달성하기 위해 ‘로봇의 문제에 대한 도덕적 사고와 판단’, ‘로봇에 대한 도덕적 책임’을 내용 영역으로 설정하였다. ‘로봇의 문제에 대한 도덕적 사고와 판단’ 영역에서 다룰 구체적 내용 요소로 인간 존엄성, 프라이버시, 투명성을 도출하였고, ‘로봇에 대한 도덕적 책임’ 영역에서 다룰 구체적 내용 요소로는 역할책임의식, 과학기술 시민성을 도출하였다.

이와 같은 내용 요소를 분석의 틀로 활용하여 로봇윤리교육 관련 도덕과 교육과정 및 교과서를 분석하였다. 교육과정 분석 결과, 로봇을 직접적으로 언급한 부분은 없었지만, 목표나 내용면에서 로봇윤리교육과 관련 있는 부분이 있었고 이를 통해 도덕과에서 로봇윤리교육을 적용하는 것에 대한 타당성을 입증하였다. 덧붙여 로봇윤리교육은 도덕적 주체로서의 자아 정체성 정립 과정을 보다 구체화하는데 기여할 수 있고 도덕과 과학기술윤리교육 영역의 내용을 확장하는데 기여할 수 있을 것이다.

도덕과 교육에서 로봇윤리교육의 적용에 대한 타당성을 토대로 도덕과 교육을 통한 로봇윤리교육 내용 개선 방안과 로봇윤리교육 방법을 제안하였다. 먼저, 로봇윤리교육 내용을 현행 도덕과 교육과정에 반영하는 방안과 로봇윤리교육 관련 도덕과 교육과정 내용을 신설하는 방안을 제시하여 도덕과에서 로봇윤리교육을 구현하고자 하였다. 나아가 도덕과 교육을 통한 로봇윤리교육 방법으로는 프로젝트 수업 모형, SF 영화 활용 수업 모형, 도덕적 딜레마 토론 수업 모형을 제안하였다.

본 연구는 제4차 산업혁명의 도래에 따른 로봇윤리교육의 필요성에 주목하여 도덕과 교육에서 로봇윤리교육을 구현하고자 하였다. 이를 위해 먼저 로봇윤리의 의미와 범주를 명확하게 규정하였다. 또한 로봇윤리의 범주를 토대로 로봇윤리교육의 의미와 내용을 규명하였다. 이러한 연구는 로봇윤리의 도덕교육적 적용 방안을 구체적으로 제안하였다는 점에서 의미가 있다. 본 연구는 기존의 연구에서 다루어지지 않았던 로봇윤리교육의 의미와 내용에 대한 세부적인 논의를 시도했을 뿐만 아니라, 도덕과 교육을 통한 로봇윤리교육의 구현방안에 대해서도 논의했다는 점에서 함의가 있다.

주요어: 제4차 산업혁명, 로봇윤리, 로봇윤리교육, 도덕교육
학 번: 2013-31094

목 차

I. 서론	1
1. 연구 목적	1
2. 연구 방법과 범위	5
II. 제4차 산업혁명과 로봇	8
1. 제4차 산업혁명의 개념과 특징	8
1) 제4차 산업혁명의 개념	9
2) 제4차 산업혁명의 특징	14
2. 인공지능과 로봇의 개념	18
1) 인공지능의 개념	18
2) 로봇의 개념	21
3. 인공지능 로봇의 영향	24
1) 경제적 변화	24
2) 사회·문화적 변화	28
3) 교육적 변화	34
III. 로봇윤리의 의미와 쟁점	37
1. 로봇윤리의 의미	37
1) 로봇윤리의 개념과 범위	37
2) 로봇윤리의 필요성	41
2. 로봇윤리의 구분	45
1) 로봇의 윤리	45
2) 인간의 윤리	52
3. 로봇윤리의 주요 쟁점	67
1) 로봇의 도덕적 지위	67
2) 인간 존엄성·프라이버시·투명성·책임	72

IV. 로봇윤리교육의 의미와 내용	89
1. 로봇윤리교육의 필요성과 의미	89
1) 로봇윤리교육의 필요성	89
2) 로봇윤리교육의 의미	96
2. 로봇윤리교육의 목표와 내용	99
1) 인간과 로봇에 대한 윤리적 성찰	99
2) 로봇의 문제에 대한 윤리적 해결	101
V. 도덕과 교육을 통한 로봇윤리교육 구현방안	104
1. 로봇윤리교육 관련 도덕과 교육과정 및 교과서 분석	104
1) 도덕과 교육과정 분석	104
2) 「생활과 윤리」 교과서 분석	114
3) 도덕과 교육에서 로봇윤리교육의 의의	117
2. 도덕과 교육을 통한 로봇윤리교육 내용 개선 방안	119
1) 로봇윤리교육 내용의 현행 도덕과 교육과정 반영 방안	119
2) 로봇윤리교육 관련 도덕과 교육과정 내용 신설 방안	125
3. 도덕과 교육을 통한 로봇윤리교육 방법 제안	134
1) 프로젝트 수업 모형	134
2) SF영화 활용 수업 모형	142
3) 도덕적 딜레마 토론 수업 모형	149
VI. 결론	155
참고문헌	159
ABSTRACT	174

표 목 차

〈표 1〉 국내·국의 연구의 제4차 산업혁명 정의	12
〈표 2〉 제4차 산업혁명의 핵심 기술	15
〈표 3〉 로봇윤리에 대한 다양한 정의	38
〈표 4〉 로봇윤리교육의 목표 및 내용	102
〈표 5〉 도덕과의 성격과 목표(2015 개정 교육과정)	105
〈표 6〉 2015 개정 초등학교 도덕 교육과정 내용체계	107
〈표 7〉 2015 개정 중학교 도덕 교육과정 내용체계	108
〈표 8〉 2015 개정 고등학교 교육과정 ‘생활과 윤리’ 내용체계 ..	110
〈표 9〉 로봇윤리교육이 적용 가능한 2015 개정 중학교 도덕 교육과정 내용요소	120
〈표 10〉 로봇윤리교육이 적용 가능한 2015 개정 『생활과 윤리』 교육과정 내용요소	121
〈표 11〉 로봇윤리교육 관련 중학교 도덕 교육과정 내용요소 신설 제안	127
〈표 12〉 로봇윤리교육 관련 『생활과 윤리』 교육과정 내용요소 신설 제안	130
〈표 13〉 도덕과 프로젝트 수업모형의 단계	136
〈표 14〉 영화 활용 수업 모형의 구조	145
〈표 15〉 도덕적 딜레마 토론 수업 모형의 절차	150

그림 목 차

[그림 1] 자율주행자동차의 윤리적 딜레마 상황	58
[그림 2] 프로젝트 학습을 활용한 로봇윤리교육 수업 모형	138
[그림 3] SF 영화를 활용한 로봇윤리교육 수업 모형	146
[그림 4] 도덕적 딜레마 토론을 활용한 로봇윤리교육 수업 모형	152

I. 서론

1. 연구 목적

오늘날 제4차 산업혁명에 대한 관심이 높아지면서 제4차 산업혁명의 핵심기술인 인공지능 로봇에 대한 다양한 전망이 나타나고 있다. 인공지능 로봇이 많은 산업 분야의 효율성을 증대시켜 경제발전의 원동력이 될 것이라는 낙관적인 기대도 있고, 알고리즘에 의한 사회적 차별, 자동화에 따른 노동 대체, 자율성이 높은 로봇으로 인한 생명과 안전의 위협 등 부작용에 대한 우려도 있다.

먼저 인공지능 로봇의 긍정적인 측면에 주목하는 학자들은 인공지능 로봇의 활용도가 높아지면 경제가 더욱 활성화될 것이고 그 규모도 커질 것이라고 예상한다. 예컨대, Schwab(2017), Brynjolfsson(2014) 등의 학자들은 인공지능 로봇 기술의 발전으로 상품의 생산성이 증가하면서 상품과 서비스의 공급이 비약적으로 빨라질 수 있다고 주장하였다(Schwab, 2017: 33; Brynjolfsson 등, 이한음 역, 2014: 129-140). 또한 Roland Berger(2017), 이종관(2017) 등은 인공지능 로봇과 관련된 새로운 직종이 늘어날 것이기 때문에 결과적으로 인공지능 로봇은 인류에게 긍정적인 영향을 주게 될 것이라고 예상했다(Roland Berger, 김정희 외 역, 2017: 78; 이종관, 2017: 385-386).

반면에 인공지능 로봇의 발달로 인해 인간의 생산 활동은 줄어들 것이고, 인간의 지적노동 직업군도 인공지능 로봇으로 대체될 것이라는 부정적 견해 또한 존재한다. Frey 등(2017: 265)은 약 20년 이내에 47%정도의 인간의 일자리가 인공지능 로봇으로 대체될 것이라고 주장하였고, Brynjolfsson(2014), Kaplan(2015), Walsh(2018) 등의 많은 학자들도 인공지능 로봇으로 인해 직업이 줄어들 것이라고 예상했다(Brynjolfsson 외, 이한음 역, 2014: 233-234; Kaplan, 신동숙 역, 2015: 25; Walsh, 이기동 역, 2018: 212-217).

그리고 자율주행자동차, 의료분야 등에서 이미 인공지능 로봇 기술이

상용화 단계에 진입하고 있으며, 이러한 인공지능 로봇의 확산은 여러 가지 변화를 가져올 것으로 예상된다. 이처럼 인공지능 로봇으로 인해 여러 문제가 발생할 가능성이 높아지면서, 과거와는 다른 새로운 차원의 윤리적 문제들이 제기되고 이를 해결할 수 있는 새로운 윤리체계의 필요성이 증대되고 있다.

로봇의 발달에 따른 윤리적 문제는 로봇이 생겨난 시점부터 여러 차례 제기된 바 있다. 로봇윤리의 시초라고 여겨지는 Asimov는 1942년 그의 작품을 통해 로봇을 인간을 뛰어넘는 존재로 상정하며 로봇이 스스로 지켜야 할 ‘로봇공학 3원칙(the three laws of robotics)’을 제시하였다. Asimov가 발표한 초기 로봇윤리는 로봇이 스스로 윤리를 만들고 지킬 수 있다는 가정 하에 만들어진 윤리이기 때문에 ‘미래학적인 로봇윤리(futuristic roboethics)’의 성격을 나타낸다. Asimov의 로봇윤리에서는 미래에 로봇의 자율성이 높아지면 인간을 해치는 방향으로 활용될 가능성을 경계하여 로봇의 책무만을 강조하는 내용을 다루고 있다. 그러나 최근 발표되고 있는 로봇윤리는 현재의 인공지능 로봇기술의 발전수준을 고려하여 ‘현실적인 로봇윤리(realistic roboethics)’의 성격으로 변화되고 있다. 예컨대, 2004년 일본 후쿠오카 세계 로봇 선언, 2006년 유럽로봇연구네트워크(EURON)의 로봇윤리 로드맵, 2007년 우리나라 산업자원부의 로봇윤리헌장(초안), 2010년 영국 공학과 물리과학 연구위원회(Engineering and Physical Sciences Research Council, EPSRC)의 로봇원칙 등에서는 점차 로봇이 아닌 인간의 책임을 강조하는 방향으로 로봇윤리의 의미가 변화하고 있다는 사실을 알 수 있다.

이렇게 여러 단체들이 발표한 로봇윤리 원칙의 주된 내용은 인공지능 로봇에게 인간과 같은 도덕적 지위를 부여할 것인지에 대한 문제와 인공지능 로봇의 제작 및 사용으로 발생할 수 있는 인간 존엄성, 프라이버시, 투명성, 책임 문제로 요약된다. 최근에는 이러한 로봇윤리에 대한 높은 관심으로 인해 학생들이 로봇과 관련된 윤리적 문제를 해결할 수 있는 역량을 기르고 이를 교육과정에 반영해야 한다는 목소리도 높아지고 있다. 실제로 많은 연구자들이 로봇윤리의 필요성과 제4차 산업혁명에

대비하여 교육이 변화해야 한다는 것에 대해서는 기본적으로 동의하고 있지만, 로봇윤리교육이 구체적으로 어떻게 실현될 수 있는지에 대한 논의는 아직 부족한 측면이 있다.

제4차 산업혁명시대의 교육과 로봇윤리교육에 관한 선행 연구에는 이러한 실태가 반영되어 있다. 먼저, 제4차 산업혁명시대의 교육과 관련한 연구는 새로운 역량 교육에 대한 필요성(김지원, 2017; 류성창 외, 2018; 조난심, 2017; 조현국, 2017)과 인공지능 로봇과 관련된 윤리적 문제를 해결할 수 있는 교육의 필요성(Schwab, 2018; Soskil, 2018)에 대한 연구가 있다. 로봇윤리교육의 필요성과 관련한 연구는 공학윤리교육으로서의 로봇윤리교육의 필요성(변순용 외, 2014), 초등학생을 위한 로봇윤리교육 내용 개발(박주현 외, 2018), 도덕과 교육에서 로봇윤리의 교육적 함의(신현우, 2015; 송선영, 2017)와 같이 로봇윤리교육의 중요성을 강조하는 연구가 있다. 그러나 이들 연구는 로봇윤리교육의 개념 및 내용에 대한 체계적인 정립 없이 로봇윤리교육의 필요성만 논의하고 있다는 점에서 한계가 있다.

이처럼 많은 연구자들이 로봇윤리교육의 필요성을 인식하고 있지만, 로봇윤리교육의 개념과 내용을 정립하고, 로봇윤리를 어떻게 학교에서 가르칠 것인지에 대한 구체적인 연구는 부족하다. 이에 본 연구는 제4차 산업혁명시대에 대비하고 인공지능 로봇의 발달로 야기되는 윤리적 문제를 해결하기 위해서 로봇윤리교육에 대한 보다 체계적인 논의가 필요함을 강조한다. 나아가 본 연구는 로봇윤리교육의 필요성에 대한 이러한 문제인식에 기초하여 로봇윤리교육의 개념과 내용을 어떻게 설정해야 하는지, 그리고 로봇윤리를 구체적으로 어떻게 가르쳐야 하는지에 대해 주목한다.

보다 구체적으로, 본 연구는 제4차 산업혁명시대의 인공지능 로봇과 관련된 윤리적 문제에 대한 대비책으로 로봇윤리교육에 주목하면서, 로봇윤리교육의 개념과 내용이 무엇인지 밝히고, 나아가 도덕과 교육에서 적용하기 위한 방안에 대해 탐색하는 것을 목적으로 한다. 이러한 연구 목적을 달성하기 위하여 설정한 연구 문제는 다음과 같다.

첫째, 제4차 산업혁명시대에 대비하기 위한 로봇윤리교육은 무엇인가?

둘째, 도덕과 교육을 통해 로봇윤리교육을 구현할 수 있는 방안은 무엇인가?

첫 번째 연구 문제는 본 연구의 최종 목적인 ‘도덕과 교육을 통한 로봇윤리교육의 구현방안’을 도출하기 위한 전제 조건에 해당한다. 이를 위해 로봇윤리가 다시 주목받게 된 시작점인 제4차 산업혁명이 무엇인지에 대해 먼저 살펴본 후, 인공지능, 로봇, 인공지능 로봇의 의미를 규명하고자 한다. 이어서 로봇윤리의 의미를 파악하고, 현재 다양하게 논의되고 있는 로봇윤리의 범위를 설정한 후, 로봇윤리의 필요성과 주요 쟁점에 대해 논의할 것이다. 이러한 논의는 이후 로봇윤리교육의 의미와 목표, 내용을 정립하는 과정에서 중요한 준거로 작용한다. 이와 같은 로봇윤리의 의미와 쟁점을 바탕으로 로봇윤리교육은 어떠한 개념과 내용을 갖고 있는지에 대해 규명하고자 한다.

두 번째 연구 문제는 본 연구의 궁극적인 목적으로서, 앞서 살펴본 로봇윤리교육의 개념과 내용에 기초하여 도덕과 교육을 통해 로봇윤리교육을 구현할 수 있는 방안을 도출하는 연구를 진행할 것이다. 이 과정에서 로봇윤리교육 관련 도덕과 교육과정과 관련 교과서를 분석하고 이를 바탕으로 로봇윤리교육을 도덕과 교육에 적용할 수 있는 내용과 방법을 제안하고자 한다.

2. 연구 방법과 범위

본 연구는 제4차 산업혁명시대에 대비하기 위해 로봇윤리와 로봇윤리 교육의 의미와 내용을 규명하고, 그것을 기준으로 도덕과 교육을 통한 로봇윤리교육의 구현 방안을 제시하는 것을 목표로 한다. 이러한 연구 목적 달성을 위해 응용연구(applied research)의 특성에 중점을 둔다(이종승, 2009: 40-41). 기존의 지식체계를 확장하는데 일차적 목적을 두는 기초연구(basic research) 차원과는 달리, 본 연구에서는 제4차 산업혁명과 인공지능 로봇이라는 현상 자체와 로봇윤리 등에 대한 기초연구 결과를 토대로, 이를 도덕과 교육에 적용할 수 있는 방안을 구체적으로 수립하는 응용연구 차원에 관심을 두고 있다. 이러한 과정은 주로 문헌고찰(literature review) 방법을 적용하여 연구를 진행할 것이다.

본 연구의 목적을 달성하기 위해 첫째, 다양한 분야의 문헌 연구를 통해 로봇윤리와 로봇윤리교육의 의미를 구명하고자 한다. 공학, 윤리학, 도덕 심리학, 도덕 교육학, 교육 철학, 교육 과정 등 관련 분야의 논문, 연구 보고서, 단행본 등을 고찰하여 로봇윤리와 로봇윤리교육의 의미와 내용을 규정해볼 것이다.

둘째, 로봇윤리교육의 실태를 파악하고 이를 통해 향후 로봇윤리교육의 방향을 논하기 위해 2015개정 도덕과 교육과정 및 교과서를 분석한다. 앞서 규명한 로봇윤리교육의 내용영역을 분석틀로 활용하여 도덕과 교육과정 및 교과서를 분석한다.

셋째, 여러 분야의 문헌 연구를 통해 도덕과 교육을 통한 로봇윤리교육의 방법을 제안한다. 도덕 교육학, 교육 과정, 교육 공학 분야의 논문, 단행본 등을 고찰하여 실천적 교수·학습 방법을 제시한다.

본 연구는 다음과 같은 절차에 따라 이루어질 것이다. II장에서는 인공지능 로봇의 대두 배경인 제4차 산업혁명과 로봇의 명확한 개념에 대한 이해를 도모하고자 한다. 이를 위해 첫째, 여러 선행연구를 분석하여 제4차 산업혁명의 개념과 특징을 살펴볼 것이다. 둘째, 인공지능과 로봇의 개념을 정립하여, 이후 로봇윤리 연구의 토대를 마련할 것이다. 셋째,

인공지능 로봇의 영향으로 예상되는 경제적, 사회·문화적, 교육적 변화에 대한 여러 선행연구를 고찰할 것이다.

Ⅲ장에서는 본격적으로 로봇윤리의 의미를 정의하고 로봇윤리와 관련된 주요 쟁점에 대해 살펴볼 것이다. 이를 위해 첫째, 로봇윤리의 개념을 규명하고 로봇윤리의 범위를 명확하게 설정한 후, 그 내용을 구체적으로 살펴볼 것이다. 본 연구에서는 다양한 선행연구의 로봇윤리 정의를 개관하여 분석한 후, 로봇윤리의 개념을 명확하게 규명하고 범위를 설정할 것이다. 로봇윤리의 범위를 로봇이 지켜야 할 ‘로봇의 윤리’와 인간이 지켜야 할 ‘인간의 윤리’로 구분해서 분류하고, 이를 이후 로봇윤리교육에 관한 연구를 수행하는데 중요한 기준으로 활용할 것이다. 둘째, 지금까지 발표된 로봇윤리 연구결과를 분석하여 가장 강조되고 있는 로봇윤리의 쟁점을 선별하고, 그 내용을 구체적으로 검토하여 이후 로봇윤리교육의 이론적 토대를 마련할 것이다. Ⅲ장에서 규명된 로봇윤리의 의미와 쟁점은 이후 로봇윤리교육의 의미와 내용을 정립하는데 활용될 것이다.

Ⅳ장에서는 로봇윤리교육의 의미와 내용을 규명하고자 한다. 이를 위해 첫째, 앞서 정립한 로봇윤리의 의미를 토대로 ‘로봇의 윤리’와 ‘인간의 윤리’에 관한 로봇윤리교육의 필요성과 로봇윤리교육의 의미를 밝힐 것이다. 둘째, 로봇윤리의 의미와 주요 쟁점을 바탕으로 로봇윤리교육의 목표와 구체적 내용 요소를 선정할 것이다. 로봇윤리교육의 목표와 구체적 내용 요소는 이후 로봇윤리교육 관련 도덕과 교육과정 및 교과서 분석 연구에서 분석틀로 활용될 것이다.

Ⅴ장에서는 로봇윤리교육이 도덕과 교육을 통해 어떻게 구현될 수 있는지를 고찰한다. 이를 위해 첫째, 앞서 정립한 로봇윤리교육의 목표와 내용 영역을 분석틀로 활용하여 로봇윤리교육과 관련된 도덕과 교육과정과 ‘생활과 윤리’ 교과서를 분석할 것이다. 구체적으로는 2015 개정 교육과정에서 도덕과의 성격과 목표를 살펴보고, 2015 개정 초등학교 및 중학교 도덕 교육과정과 고등학교 선택과목 ‘생활과 윤리’ 교육과정 내용체계를 분석할 것이다. 이러한 분석 결과를 종합하여 도덕과 교육을

통한 로봇윤리교육의 타당성과 의의를 밝힌다. 둘째, 교육과정 분석 결과와 로봇윤리교육의 내용 영역을 근거로 현행 도덕과 교육과정에 로봇윤리교육을 반영할 수 있는 방안을 제시한다. 나아가 도덕과 교육과정에 새롭게 포함시킬 수 있는 로봇윤리교육 관련 내용 요소를 제안하고자 한다. 마지막으로 로봇윤리교육의 목표를 달성할 수 있는 실제 수업 모형에 대해 구체적으로 소개하고자 한다.

II. 제4차 산업혁명과 로봇

‘로봇윤리의 도덕교육적 적용 방안’을 마련하기 위해서는 우선 로봇윤리의 의미가 무엇인지, 로봇이 무엇인지에 대해 규명해야 한다. 로봇이 무엇인지 규명하기 위해서 먼저 로봇이 최근 대두되게 된 배경인 제4차 산업혁명의 개념 및 특징에 대해 알아보고, 이를 토대로 인공지능과 로봇의 개념을 살펴본다. 그리고 인공지능 로봇으로 인해 야기되는 경제적, 사회·문화적, 교육적 변화에 대해서도 함께 파악하고자 한다.

1. 제4차 산업혁명의 개념과 특징

제4차 산업혁명은 여러 학자들에 의해서 다양하게 정의되고 있기 때문에 현재 개념이 명확하게 규정되어 있지 않다. 따라서 이 논문에서는 지금까지 진행된 논의를 바탕으로 제4차 산업혁명의 개념과 특징에 대해 명확하게 규명하려고 한다. 이에 대한 구체적인 논의로는 첫째, 현재 제4차 산업혁명에 대해 진행된 국내외 연구를 분석하고, 이를 종합하여 제4차 산업혁명의 개념을 규명한다. 둘째, 현재까지 연구된 제4차 산업혁명에 대한 국외·국내 학자들의 연구를 바탕으로 제4차 산업혁명의 특징을 파악한 후, 제4차 산업혁명과 인공지능 로봇과의 연관성에 대해 유추하여 이후 로봇윤리 연구의 토대를 마련할 것이다.

1) 제4차 산업혁명의 개념

인공지능 로봇이 주목받게 된 배경인 ‘제4차 산업혁명’은 무엇인가? 이 항에서는 인공지능 로봇에 대한 본격적인 연구에 앞서서 인공지능 로봇의 등장 배경인 제4차 산업혁명의 개념에 대해 규명하고자 한다. 제4차 산업혁명(the fourth industrial revolution)은 2016년 1월에 스위스 다보스에서 열린 ‘세계경제포럼(World Economic Forum)’에서 처음 언급된 신생 용어로, 학문적으로 명확한 의미가 정립되지 않았다. 제4차 산업혁명은 새로운 것이 아니라 제3차 산업혁명의 연장선이라는 시각도 존재하고 있어서 제4차 산업혁명의 성립 가능성에 대한 진위여부에 다양한 시각이 있는 것으로 보인다. 제4차 산업혁명의 사전적 정의는 ‘인공지능, 사물 인터넷, 빅데이터, 모바일 등 첨단 정보통신기술이 경제·사회 전반에 융합되어 혁신적인 변화가 나타나는 차세대 산업혁명’으로 새로운 기술에 의한 변화를 총칭하는 의미로 사용되고 있지만, 제4차 산업혁명에 대한 모든 것을 포괄하는 정확한 개념은 아니다.

제4차 산업혁명의 개념은 독일 정부 정책인 ‘인더스트리(Industrie 4.0)’라는 명칭으로 2011년 하노버 박람회에서도 처음 등장하였고, 제조업과 정보통신이 융합되는 의미로 사용되었다(Roland Berger, 김정희 외 역, 2017: 65). 이후 세계경제포럼에서 보다 넓은 개념인 ‘제4차 산업혁명’이란 용어로 언급되면서 정보통신기술(ICT) 기반의 새로운 산업시대를 대표하는 의미로 널리 사용되었다. 2016년 알파고 대국을 계기로 제4차 산업혁명에 대한 관심이 더욱 높아졌고, 각 나라에서는 독일의 Industry 4.0을 비롯하여 미국의 Industrial Internet, 일본의 Robot strategy, 중국의 제조 2025 등의 다양한 형태로 제4차 산업혁명이 진행되고 있다.

이러한 제4차 산업혁명의 논의와는 달리, 제4차 산업혁명은 새로운 산업혁명이 아니라 단지 제3차 산업혁명의 연장선 일뿐이라고 주장하며 제4차 산업혁명의 존재 자체를 부정하는 연구도 있다. 일례로 Rifkin 등의 학자는 제4차 산업혁명이 새로운 변화가 아닌 단지 제3차 산업혁명의 연장선에 불과하다고 단언한다. Rifkin은 디지털 혁명인 제3차 산업혁명의

잠재력이 모두 발휘되지도 않았는데 제3차 산업혁명의 종료를 선언하고 새로운 산업혁명을 주장하는 것은 시기상조라고 주장했다(안상희 외, 2016: 2360 재인용). 그는 제4차 산업혁명의 특징으로 제시되는 인공지능, 빅데이터 등과 같은 요소들은 컴퓨터와 인터넷으로 대표되는 디지털 기술의 한 분야일 뿐이기 때문에 우리는 여전히 제3차 산업혁명의 연장선에 있다고 판단했다. Gordon은 우리가 이미 경험한 제3차 산업혁명이 지속되고 있을 뿐 제4차 산업혁명이라는 새로운 변화는 없을 것이라고 예상했다. 또한 그는 연간 GDP 성장률이 1970년대부터 둔화됐으며, 향후 25년간 0.8%에 그칠 것으로 전망하며 장기 저성장 국면에서 새로운 산업혁명을 제기되는 것은 부적절하다고 주장했다(Popely, 2016에서 재인용). 손화철(2017: 33-36)도 ‘제4차 산업혁명’에는 분명한 내용도 없고 단지 2016년 다보스포럼을 위한 일시적 표어에 불과하며 제4차 산업혁명은 고정된 실체가 없는 기술적 용어에 가깝다고 판단했다.

이처럼 제4차 산업혁명에 대한 논란은 현재도 계속 되고 있다. 반면에 Schwab(2017: 9-11)은 제4차 산업혁명의 존재에 대한 부정적인 시각에 대해 제4차 산업혁명은 제3차 산업혁명의 진화나 일시적인 트렌드가 아니라고 단언한다. 나아가 제4차 산업혁명과 기존의 산업혁명의 차이점은 속도, 범위와 깊이, 시스템 충격이라고 주장한다. 그가 주장한 제4차 산업혁명과 기존의 산업혁명의 차이점에 대해 자세히 살펴보면 다음과 같다. 첫째, 제4차 산업혁명은 변화의 속도가 굉장히 빠르다. 지금까지의 제1~3차 산업혁명과는 달리, 제4차 산업혁명은 선형적 속도가 아닌 기하급수적인 속도로 진행하고 있다. 현재 우리가 살고 있는 세계가 다면적이고 서로 깊게 연계되어 있으며, 신기술은 그보다 더 새롭고 뛰어난 역량을 갖춘 기술을 만들어내기 때문에 혁신의 발전과 전파 속도가 그 어느 때보다 빠를 것이다.

둘째, 범위와 깊이가 다른 점이다. 제4차 산업혁명은 단순히 하나의 기술적 변화가 아니라 사회 전체 시스템의 변화를 수반한다. 20개 이상의 새로운 기술이 함께 통합되어 지금까지 존재하지 않았던 완전히 새로운 기술이 적용될 것이다. 그리고 변화의 범위와 깊이가 이전과 달리 넓

고 깊어서, 수많은 분야에서 새로운 기술들이 융합하고 조화를 이룰 것이며 서로 다른 과학기술이 상호 의존하여 획기적인 상품을 만들어낼 것이다.

셋째, 시스템 충격이다. 제4차 산업혁명은 하나의 아이디어를 통해서 모든 시스템이 같이 통합된다. 이러한 시스템 충격은 국가 간, 기업 간, 산업 간 그리고 사회 전체 시스템의 변화를 수반할 것이며, 우리의 소비 행동과 사고방식도 전부 달라질 것이다. 과거에는 생산성이나 효율성에 초점을 맞춰왔지만 제4차 산업혁명을 통해 점차 우리 자신에게 초점이 맞춰질 것이다(Schwab, 2017: 3-4, 9-11).

위의 내용을 종합해보면, 제4차 산업혁명의 속도와 범위, 시스템에 끼치는 영향력이 매우 크다는 사실은 확실한 것으로 보인다. 많은 학자들은 제4차 산업혁명의 발전 속도가 기존의 산업혁명에 비해 매우 빠르기 때문에 제3차 산업혁명의 연장이 아니라고 주장한다(Schwab, 2017: 9; World Economic Forum, 2016: 5). 즉, 제4차 산업혁명은 우리가 일하고 살아가는 방식을 근본적으로 바꿀 수 있는 기술 혁신이며, 변화의 규모, 범위가 이전에 경험한 것과는 분명히 다른 것이다. 이러한 제4차 산업혁명은 현재 모든 국가의 산업을 혼란스럽게 만들고 있고, 생산, 관리 등 전체 시스템의 변화를 야기할 것으로 예상된다.

이러한 변화의 흐름에 따라 많은 학자들은 제4차 산업혁명의 도래를 주장하며 관련 연구를 발표하고 있다. 현재까지 국내외 연구에서 논의된 제4차 산업혁명의 개념을 자세히 살펴보면 아래와 같다.

〈표 1〉 국내·국의 연구의 제4차 산업혁명 정의

연구자	개념 정의
Schwab (2017: 1-2)	물리적, 생물학적, 디지털적 세계의 융합을 통한 총체적인 사회변화
World Economic Forum (2016: 5)	물리계, 디지털계, 바이오계 기술의 융합
Colombo 등 (2017: 3)	ICPS(industrial cyber-physical system, 산업용 사이버물리적시스템)에 의한 기술적, 경제적, 사회적인 변화
미래전략정책연구원 (2016: 13)	기업들이 제조업과 정보통신기술을 융합해 작업 경쟁력을 높이는 차세대 산업혁명
이민화 (2016: 23)	현실과 가상의 세계가 융합하여 현실세계를 최적화하는 것
이종관 (2017: 382)	가상세계와 현실세계가 상호 침투하는 사이버 물리 시스템을 통해 더욱 자동화되고, 지능화된 생산체제가 경제구조를 급격히 혁신하는 과정
정재승 (2018: 253)	물질(atom)세계와 데이터(bit)세계의 일치를 바탕으로 한 제조업과 유통업의 혁신

제4차 산업혁명의 개념을 처음으로 정립한 Schwab(2017: 1-2)은 제4차 산업혁명을 물리적, 생물학적, 디지털적 세계의 융합을 통한 총체적인 사회변화로 정의하였으며, 여러 학자들의 개념 정의도 의미상으로는 크게 다르지 않다. 위와 같이 여러 학자들의 제4차 산업혁명에 대한 개념 정의를 종합하여, 본 연구에서는 제4차 산업혁명을 다음과 같이 정의하고자 한다.

첫째, 제4차 산업혁명은 인공지능, 사물인터넷, 빅데이터 등 새로운 기술의 융합에 의한 총체적인 변화를 의미한다. 이러한 변화는 인공지능, 로봇공학, 사물인터넷, 자율주행자동차, 3D프린팅, 나노기술, 생명공학, 재료공학, 에너지 저장기술, 퀀텀 컴퓨팅 등 과학기술 분야에서 폭넓게 일어나는 거대한 융합이다(Schwab, 2017: 1). 즉, 제4차 산업혁명은 새로운 기술의 융합에 따라 경제·사회·문화 등 사회 전반에 걸쳐 나타나는 혁신적인 변화로 정의할 수 있다.

둘째, 제4차 산업혁명은 새로운 기술에 의해 현실과 가상의 세계가 융

합하는 것을 의미한다. 이와 같은 혁명은 물질(atom)과 데이터(bit)가 일치하는 세상, 즉 물질로 이루어진 현실세계와 데이터로 이루어진 사이버 세계가 일치하는 세상으로 변화하는 것이다(정재승, 2018: 253). 현실세계와 사이버세계의 일치를 바탕으로, 3차 산업혁명의 산물인 IT 및 디지털 기술을 1·2차 산업혁명의 산물인 제조업·유통업에 적용하여 전 산업이 새로운 형태로 융합하는 것을 의미한다.

결론적으로 제4차 산업혁명이 이전 산업혁명의 연장선인지 새로운 산업혁명인지 아닌지에 대한 진위여부에는 논란이 있어도, 인공지능, 로봇공학, 사물인터넷, 자율주행자동차, 3D프린팅, 나노기술, 생명공학, 재료공학, 에너지 저장기술, 퀀텀 컴퓨팅 등 새로운 기술의 융합에 의해 혁신적인 변화가 일어날 것이라는 의견에는 많은 사람들이 동의하고 있는 것으로 보인다. 따라서 본 연구에서는 제4차 산업혁명의 도래에 동의하며, 제4차 산업혁명의 개념은 ‘인공지능을 포함한 새로운 기술과 기존 기술의 융합으로 인해 현실과 가상의 세계가 일치하는 것’으로 정의하고자 한다.

2) 제4차 산업혁명의 특징

앞에서 살펴본 것과 같이, 현재 우리는 빠른 기술혁신과 인공지능의 발달에 따라 제4차 산업혁명시대를 앞두고 있다. 제4차 산업혁명의 도래로 인해 다양한 분야에서 혁명적 변화가 야기될 것으로 예상된다. 일례로 최근 우리나라에서도 자율주행차가 안전하게 제작되고 상용화될 수 있도록 세계 최초로 부분 자율주행차 안전기준을 도입하고, 2020년 7월부터는 자동차로유지기능이 탑재된 레벨3 자율차의 출시, 판매가 가능해진다고 발표하기도 하였다(국토교통부, 2020).¹⁾ 따라서 이 항에서는 제4차 산업혁명에 대한 구체적 이해를 위해 제4차 산업혁명의 특징에 대해 살펴보고 향후 연구의 기반을 마련할 것이다.

많은 연구에서 제4차 산업혁명의 특징에 대해 분석하였으며, 그 연구 결과는 다음과 같다. Roland Berger(김정희 외 역, 2017: 25-26)는 지속적인 디지털화와 향상된 연결성으로 인해 생산성이 증가하고, 더욱 높은 부가가치, 높은 이익, 높은 자본이익률을 창출하는 것이 제4차 산업혁명의 특징이라고 분석했다. 원동규 등(2016: 17)은 제4차 산업혁명의 특징을 초연결성, 초지능성, 예측 가능성으로 판단했다. 사람과 사물, 사물과 사물이 인터넷 통신망으로 연결되고, 막대한 데이터를 분석하여 일정한 패턴을 파악하고, 인간의 행동을 예측하여 새로운 가치를 창출해 내는 것이 바로 제4차 산업혁명의 특징이라고 주장했다. 또한 제4차 산업혁명의 도래를 공식적으로 선언한 Schwab(2017: 6-24)은 다음과 같이 제4차 산업혁명의 특징을 제시하였다.

첫째, 제4차 산업혁명이 과학기술의 혁명적 변화를 가져올 것이라는 점이다. 물리학, 디지털, 생물학의 영역으로 구분하여 제4차 산업혁명에 따른 핵심 기술들을 살펴보면 다음과 같다.

1) '레벨3'은 미국 자동차공학회(SAE)의 자율주행 기능 분류에 따르면 레벨 0-5의 6가지 중에서 지정된 조건에서 자율주행이 가능하나 예상치 못한 공사상황 등 시스템에서 요구 시에는 운전자의 운전이 필요한 부분 자율주행 정도를 의미한다(국토교통부, 2020).

〈표 2〉 제4차 산업혁명의 핵심 기술(임종현 외, 2017: 10)

영역	핵심 기술
물리학(physical)	무인 운송수단, 3D 프린터, 로봇공학, 신소재
디지털(digital)	사물인터넷, 블록체인(block chain)
생물학(biological)	유전학, 유전자 편집, 합성생물학

물리학 영역의 핵심 기술은 자율주행자동차 및 드론을 포함한 무인 운송수단과 3D 프린터, 로봇공학, 신소재 등이며, 디지털 영역의 핵심 기술은 실물과 디지털의 연계를 가능하게 하는 사물인터넷과 블록체인 등이다. 그리고 생물학 영역의 핵심 기술은 유전자 연구와 유전자 편집, 의학과 농업에 대안을 제시하는 합성생물학이 대두될 것이라고 주장했다. 이와 같이 첫 번째 제4차 산업혁명의 특징은 디지털에 기반을 둔 연결성을 바탕으로 물리와 생물, 인공지능 등의 새로운 기술이 융합되는 것이다.

둘째, 제4차 산업혁명이 전반적으로 사회 체제의 혁명적 변화를 야기할 것이라는 점이다. 제4차 산업혁명시대에는 언제 어디서든 네트워크로 연결될 것이며, 현실 세계와 가상 세계가 융합된 세상으로 변화할 것이다. 이러한 빠른 과학기술의 변화 속도만큼 사회 전반에도 광범위한 변화가 일어날 것이다. 이러한 변화의 흐름은 개인의 삶에서부터 국가 체제의 변화를 넘어 세계 체제의 변화까지 이어질 것으로 예상된다.

셋째, 제4차 산업혁명이 직업 구조의 급격한 변화를 가져올 것이라는 점이다. 제4차 산업혁명의 핵심 기술과 관련된 새로운 직업이 생겨날 것으로 예상되지만, 단순반복적인 직업 등은 로봇으로 대체되어 실업이 늘어나고 소득불평등이 증가할 것이라는 비관적 전망도 존재한다.

Frey 등(2017: 265)은 약 20년 이내에 47%정도의 직업이 인공지능 로봇으로 대체될 것이라고 예상했다. 그들은 미국에서 약 1억 3800만개의 일자리를 포함하는 702개 직업군 중에 컴퓨터화될 가능성이 있는 직업군에 대해 탐색하였다. 연구 결과, 컴퓨터화 가능성이 가장 높은 직업은 텔레마케터, 시계 수리공, 스포츠 경기 심판, 계산 점원, 전화 교환원 등으로 다른 직업에 비해 비교적 반복적이고 단순하며 수행 과정과 절차를

표준화하기 쉬운 직업이었다. 반대로 컴퓨터화 가능성이 가장 낮은 직업은 고위 경영자, 멀티미디어 아티스트, 음악감독, 작곡가, 상담사, 미생물학자, 간호사 등으로 전문적인 직무를 수행하며 인간 사이의 상호작용과 창의성, 감성이 중요시되는 직업이었다.

한국고용정보원(2016: 12-14)에서도 자동화로 대체될 확률이 높은 직업에 대해 분석하였으며, 그 결과는 다음과 같다. 자동화로 대체될 확률이 높은 직업은 콘크리트공, 정육원 및 도축원, 고무 및 플라스틱 제품조립원, 청원경찰, 조세행정 사무원 등으로 나타났고, 자동화로 대체될 확률이 낮은 직업은 화가 및 조각가, 사진작가, 작가, 지휘자·작곡가 및 연주가, 애니메이터 및 만화가 등으로 나타났다. 인공지능 로봇의 자동화로 인해 대체될 가능성이 높은 직업은 단순 반복적이고 정교함이 떨어지거나 사람들과 소통하는 일이 상대적으로 적은 특징을 가진 직업이었고, 감성에 기초한 예술 관련 직업들이 상대적으로 자동화될 가능성이 적은 것으로 나타났다.

그러나 단순 반복적인 저숙련 업무만 미래에 로봇으로 대체될 가능성이 높은 것은 아니다. 위의 연구에 따르면, 흔히 전문적인 직업군으로 분류되는 손해 사정인, 의사, 관제사 등도 자동화에 따라 사라질 가능성이 비교적 높게 나타났다. 즉, 인지적 전문성이 요구되는 직업도 인공지능 로봇으로 빠르게 대체될 것으로 예상된다. Frey 등(2017: 254)의 연구에서도 컴퓨터화로 인해 전문직 직업군에서도 직업 퇴화가 일어날 것으로 예상하고 있다.

위의 연구결과를 종합해보면, 제4차 산업혁명에 의해 총체적으로 고용이 감소할 것이며 특히, 인공지능 로봇에 의해 대체가능한 분야의 일자리가 더욱 급격하게 줄어들 것으로 보인다. 이는 제4차 산업혁명의 여러 핵심기술 중에서도 인공지능 로봇 기술이 가장 인류에게 큰 영향을 미치는 핵심적인 기술이라는 사실을 증명한다.

Brynjolfsson 등은 제4차 산업혁명을 ‘인공지능 혁명’으로 규정하였고, 이전 산업혁명들을 촉발한 핵심기술인 증기기관, 전기, 디지털 기술처럼 인공지능이 제4차 산업혁명의 핵심기술이라고 판단했다. 특히 인공

지능 기술 중 머신러닝(machine learning, 기계학습)은 산업 전반에 활용될 수 있는 수준으로 발전했으며, 거의 모든 산업의 질적, 구조적 변혁을 가져와 생산성의 비약적 향상을 이끌 것으로 분석했다(Brynjolfsson et al., 2017: 2).

지금까지의 연구결과에 의하면, 제4차 산업혁명의 특징은 과학기술과 사회 체제의 혁명적 변화, 직업 구조의 급격한 변화이며 이러한 변화의 핵심기술은 인공지능 기술과 로봇이라 할 수 있다. 제4차 산업혁명에 대한 큰 관심의 원인은 인공지능 기술과 로봇의 융합으로 인해 인공지능 로봇이 인간을 많은 영역에서 대체하게 되어 인류 전체에게 위기를 가져올 것이라는 인식 때문인 것으로 보인다. 이에 따라 제4차 산업혁명에 대한 관심의 원인인 인공지능 기술과 로봇이 정확히 무엇인지, 인공지능 로봇은 과연 인간을 대체할 수 있을지, 인공지능 로봇은 우리에게 어떤 영향을 가져올지에 대해 다음 절에서 자세히 살펴보고자 한다.

2. 인공지능과 로봇의 개념

이 절에서는 제4차 산업혁명의 핵심 기술인 인공지능과 로봇의 개념에 대해 알아보려고 한다. 간혹 인공지능과 로봇의 개념을 같은 개념이라고 혼동하는 경우가 있는데 인공지능과 로봇은 엄연히 다른 개념이다. 인공지능은 ‘소프트웨어’ 이고, 로봇은 ‘하드웨어’ 다. 즉, 인공지능은 주어진 데이터를 이용해 판단하거나 학습하는 ‘프로그램’ 이며 로봇은 어떤 작업이나 조작을 자동적으로 행하는 ‘기계’ 이다. 이에 따라 인공지능 기술을 사용한 로봇을 ‘인공지능 로봇’ 이라 한다. 이 장에서는 보다 구체적으로 인공지능과 로봇, 인공지능 로봇에 대한 개념에 대해 알아보고 이를 향후 로봇윤리 연구의 바탕으로 삼고자 한다.

1) 인공지능의 개념

인공지능(artificial intelligence)의 개념은 무엇일까? 인공지능에 대한 정의는 학자마다 다양하기 때문에 명확하게 정의를 내리기는 쉽지 않다. Russell 등(류광 역, 2016: 14)은 인공지능을 “인간처럼 사고하고, 인간처럼 행동하고, 합리적으로 사고하고, 합리적으로 행동하는 것” 이라고 규명하였다. 김대식(2016: 5)은 “컴퓨터가 인간의 지능으로 할 수 있는 사고와 학습 등을 할 수 있도록 연구하는 컴퓨터 정보기술의 한 분야” 라고 인공지능을 규정하였다. 위 학자들의 정의를 종합해보면 인공지능의 개념은 ‘사고나 학습 등 인간이 가진 지적능력을 컴퓨터를 통해 구현하는 기술’ 이라고 정의할 수 있다.

영국의 수학자인 Turing(1950: 2)은 ‘계산기의 기능을 확장시켜 인간처럼 생각하는 기계’ 로 인공지능에 대한 개념적인 기반을 처음으로 마련하였고, 이후 인공지능을 논의하는 시발점을 제시해 주었다. ‘인공지능(artificial intelligence)’ 이란 용어는 1956년 McCarthy에 의해서 처음 사용되었다. 당시 최고의 정보과학자들인 Minsky, Shannon, Rochester 등이 참가한 다트머스 회의(Dartmouth Conference)에서 그들이 인공지능을

‘인간 행동의 방식과 같이 기계를 행동하게 만드는 것’으로 정의하면서 본격적으로 인공지능 연구가 시작되었다(김의중, 2016: 40-41, 카카오 편집진, 2018: 7).

그러나 그 후 50여 년간 인공지능에 대한 연구는 침체기를 맞았다. 당시의 컴퓨터는 복잡한 계산을 처리하지 못했고 처리할 데이터도 많지 않아 인공지능 연구가 한때 중단되기도 했지만, 2012년 딥러닝(deep learning)방식이 개발되면서 인공지능이 다시 주목받기 시작했다. 기계에 학습기능을 부여하는 인공지능의 새로운 접근방식인 딥러닝의 탄생으로 인공지능 연구는 다시 시작되었다. Hinton 등은 기존의 기계학습과 인공신경망의 한계를 극복하여 딥러닝 분야를 개척했다. 딥러닝은 기계가 수많은 데이터 속에서 패턴을 발견하고 이를 토대로 인간의 정보 처리 방식을 모방하여 사물을 분별하는 방식을 의미한다(Hinton et al., 2012: 83). 즉, 딥러닝은 인공지능이 보고 배운 것을 기억하고, 그것을 토대로 새로운 사실을 추론한다는 점에서 인간의 사고와 유사한 기술을 나타낸다고 볼 수 있다.

그리고 인공지능의 종류는 그 수준에 따라 협의의 인공지능과 인공일반지능 그리고 인공 초지능으로 나눌 수 있다. 협의의 인공지능(Artificial Narrow Intelligence)은 특정한 영역에서 특수한 기능을 발휘하는 인공지능으로, 최근까지 화제가 되었던 딥 블루나 알파고, 오늘 날 우리 주변에서 볼 수 있는 번역 시스템, 검색 엔진, 자동응답 서비스 등도 모두 이에 속한다. 인공일반지능(Artificial General Intelligence)는 일반인공지능 또는 범용인공지능이라고도 하는데, 모든 측면에서 인간만큼 똑똑하고 인간 지능으로 할 수 있는 모든 일을 수행할 수 있는 인간 수준의 인공지능을 말하며, 인공지능이 이 수준에 도달할 수 있을 지에 대해 의문을 갖기도 한다. 인공 초지능(Artificial Super Intelligence)는 초인공지능이라고도 하며, 실제로 모든 영역에서 인간 두뇌를 훨씬 능가하는 능력을 가진 것으로, SF영화에서도 볼 수 있는 거의 상상 수준의 인공지능을 의미한다(이상현, 2018: 269-271). 또한 인공지능은 강한 인공지능(Strong Artificial Intelligence)와 약한 인공지능(Weak Artificial Intelligence)로 나누기도 하는데(박찬국, 2018: 83-86), 약한 인공지능 또는 약인공지능은

위의 협의의 인공지능에 해당하고, 강한 인공지능 또는 강인공지능은 인공일반지능과 인공 초지능을 포함한다고 할 수 있다.

지금까지의 연구결과에 의하면, 인공지능 기술은 딥러닝 방식이 개발되기 전까지는 연구 성과가 미미했지만, 2012년에 인간 뇌의 정보처리방식과 비슷한 딥러닝 방식이 개발되면서 인공지능 기술은 제4차 산업혁명을 선도하는 기술이 된 것으로 보인다. 앞으로 딥러닝 기술과 로봇기술은 더욱 융합하고 발전하여 인공지능 로봇이 다양한 곳에서 활용될 것으로 예상된다. 즉, 인공지능의 개념은 ‘인간의 지능을 대체하는 기계’의 개념으로 사용되었고 인간의 지능을 대체하기 위한 방향으로 인공지능 연구가 발전되어 왔으며, 이러한 맥락에서 인간의 정보처리방식을 모방하여 딥러닝 방식이 개발되었다고 볼 수 있다.

2) 로봇의 개념

로봇(robot)의 개념은 무엇인가? 로봇이라는 단어는 언제부터 사용되었을까? 로봇이라는 단어는 1920년대에 처음 등장하였고, 일종의 노예와 비슷한 의미로 사용되었다. 로봇은 감정이 없고 고통을 느끼지 않기 때문에 반복적이고 위험한 노동에서 인간 노동력을 대신하기 위해 제작되었다. 그 후 로봇과 관련된 과학과 공학 분야가 급격하게 발전하고 있다.

로봇의 개념에 대한 정의는 다양하다. 로봇연구 초창기에 로봇의 의미는 단순히 인간의 명령에 따르는 수동적 의미가 강했다. 로봇의 사전적 의미는 ‘명령이나 프로그램에 의해 인간이 하는 복잡한 일을 따라할 수 있는 인간을 닮은 기계적 장치’, 본질적 의미는 ‘SF의 산물로, 인간이나 다른 동물을 닮은 인공적 지능 존재’였다. 로봇연구 초창기의 학자들 정의를 살펴보면, Brooks(2002: 13)는 로봇을 ‘인공 생명체(artificial creature)’라고 정의했으며, Bekey(2005: 2)는 ‘감각과 인공적 인지, 물리적 행동을 가지고 있는 기계’로 정의했다.

그러나 최근 인공지능기술이 발달하고 로봇의 공학적 자율성이 증가하면서 몇몇 학자들은 로봇의 의미를 ‘인간과 비슷하게 자율적인 행동을 할 수 있는 존재’라며 다소 급진적 견해를 보이고 있다. Mataric(2007: 2)은 ‘로봇이란 물리적 세계에 실재하면서 어떠한 목적을 달성하기 위해 자신의 환경에 감각적으로 반응하는 자율적인 시스템’이라고 정의하였다. 그리고 로봇이 더욱 발전하면 인간에게서 어느 정도의 입력과 조연을 받을 수 있어도 인간에게 완벽하게 조종당하지는 않을 것이라며 로봇의 자율성을 강조했다.

최근 인공지능 기술이 발전하면서 로봇이라면 곧 인공지능을 장착한 ‘인공지능 로봇(AI robot)’을 뜻할 정도로 인공지능 로봇이 대세를 이루게 되었지만, 인공지능과 로봇은 다른 개념이다. 인공지능은 ‘인간의 지능이 가지는 학습, 추리, 적응, 논증 따위의 기능을 갖춘 컴퓨터 시스템’으로 고정된 실체가 없는 기술적, 소프트웨어적 의미를 뜻하며, 로봇은 물리적 실체가 있는 기계를 뜻한다. 하지만 오늘날은 인공지능을 활용한 인

공지능 로봇이 대부분이라 로봇과 인공지능 로봇을 같은 의미로 사용하기도 한다. 그 한 예로, 현행 ‘지능형 로봇 개발 및 보급 촉진법’ (법률 제 15645호)은 제2조 1호에서 “지능형 로봇이란 외부환경을 스스로 인식하고 상황을 판단하여 자율적으로 동작하는 기계장치(기계장치의 작동에 필요한 소프트웨어를 포함한다)” 라고 정의하고, 2017년에 발의된 로봇기본법(안)에서도 로봇을 “로봇이란 외부환경을 스스로 인식하고 상황을 판단하여 자율적으로 동작하는 기계장치 또는 소프트웨어” (제2조 1)로 정의하고 있어, 인공지능 로봇과 로봇을 동일한 의미로 사용하고 있음을 알 수 있다.

Veruggio 등은 로봇을 보는 시각을 네 가지로 구분하고 있다(Veruggio, 2006: 613). 첫째, 로봇은 단순한 기계(machine)일 뿐이라는 입장이다. 로봇은 매우 정교하고 유용한 기계이지만 설계자가 부여한 그 이상의 의식이나 자유의지, 자율성을 구현할 수 없다. 둘째, 로봇은 윤리적 차원(ethical dimensions)을 갖는다는 입장이다. 인간이 동물과 구별되는 윤리적인 면이 로봇에 내재되어 있으며, 로봇은 선행을 할 수 있도록 인간성이 부여된 상징적인 장치들이다. 셋째, 도덕적 행위자(moral agent)로서 로봇을 바라보는 시각이다. 로봇을 선악에 대해 행동할 수 있는 실체로 여기는 것이다. 즉, 로봇은 인간과 같은 자유의지, 정신상태, 책임은 없어도 도덕적 행위자로서 선하거나 악하게 행동할 수 있는 실체라는 것이다. 넷째, 로봇은 새로운 종(new species)의 진화로 간주되기도 한다. 미래의 로봇은 자율성과 양심을 가질 뿐만 아니라, 도덕적 차원과 지적 차원에서도 인간을 능가할 수도 있다는 것이다. 이러한 관점에서는 로봇이 이성적인 사고와 흔들리지 않는 도덕성으로 새로운 종이 될 것이라고 예상한다.

이처럼 로봇의 등장 초반에는 로봇의 개념이 인간의 힘든 일과 어려운 노동을 대신하는 단순한 기계의 의미가 강했다. 그러나 최근 인공지능을 융합한 로봇공학이 급속도로 발전하고, 딥러닝을 통해 자율적인 판단을 하는 로봇이 등장할 가능성이 높아짐에 따라 심지어 로봇을 새로운 종으로 보는 시각까지 등장한 것이다.

지금까지의 연구결과에 의하면, 과거에는 로봇의 개념을 단순히 인간을

도와주는 도구의 개념으로 간주했지만, 오늘날에는 로봇을 여전히 단순한 기계로 보는 시각과 함께 도덕적 행위자, 새로운 종으로 보는 등 로봇에 대한 다양한 시각이 존재하고 있다는 것을 알 수 있다. 즉, 로봇에게 주어지는 자율성의 정도에 따라 로봇의 역할이나 지위가 달라진다고 할 수 있다.

그러나 현재의 인공지능 기술수준에서 로봇이 인간과 동일한 수준의 자율적 판단력을 보유하는 것은 불가능하다. 따라서 현재의 인공지능 기술수준에서는 로봇의 개념을 ‘인간이 의도한 범위 내에서만 자율적으로 행동할 수 있는 존재’로 정의할 수 있다. 여기서 ‘자율적으로 행동하는 능력’의 의미는 인간처럼 스스로 생각하고 판단해서 작업하는 능력이 아니라, 인간이 프로그래밍한 범위 내에서만 자율적으로 행동하는 것이다. 로봇의 개념은 앞으로의 인공지능 로봇기술 발전수준에 따라 변화할 수 있기 때문에, 본 연구에서는 현재의 기술수준에서 로봇을 ‘인간이 의도한 범위 내에서만 자율적으로 행동할 수 있는 존재’로 이해하고자 한다.

2. 인공지능 로봇의 영향

앞에서 살펴본 것처럼 제4차 산업혁명으로 인해 로봇, 특히 인공지능 로봇으로 인해 많은 변화가 발생할 것으로 예상된다. 인공지능 로봇은 인간과 비슷한 외관을 가지고 자율적인 행동을 할 수 있는 존재로 발전하고 있다. Kurzweil(2010: 229-230)은 2045년 정도에 ‘특이점(singularity point)’이 도래하게 되면 우리가 제어할 수 없는 폭발적인 속도로 인공지능 로봇이 발전할 것이라고 주장하였다. 그리고 인공지능 로봇의 지능이 인간의 지능을 돌파하는 특이점이 오게 되면, 인공지능 로봇의 의도를 파악할 수 없기 때문에 그 시점 이후의 미래를 예측하는 것은 불가능하다고 주장하였다. Schwab(2017: 9)도 인공지능 로봇 기술의 발전으로 총체적인 사회변화인 제4차 산업혁명이 올 것이라고 판단했다. 이러한 견해를 토대로, 이 장에서는 우리의 삶에 인공지능 로봇으로 인해 어떠한 변화가 발생할지에 대해 구체적으로 살펴보고자 한다. 그 중 가장 많은 변화가 예상되는 경제적, 사회·문화적, 교육적 변화에 대해 자세히 살펴보고 이후 로봇윤리 연구의 기반을 마련할 것이다.

1) 경제적 변화

인공지능 로봇의 발달에 따른 경제적 변화에 대한 연구는 두 가지로 구분된다. 인공지능 로봇의 발달이 인류에게 많은 경제적 이득을 가져다 줄 것이라는 낙관론과 인간이 하는 일을 인공지능 로봇이 대체하여 실업률이 증가할 것이라는 비관론으로 분류된다.

먼저, 인공지능 로봇의 발달은 경제적으로 인류에게 많은 이점을 줄 것이라는 낙관적인 연구가 있다. 강태중 외(2016: 7)의 연구에 따르면, 경제적인 측면에서 새로운 기술을 향유할 역량이 있고 비용을 감당할 수 있는 소비자들은 점점 더 확대되는 안락과 편익을 효율적으로 누릴 것이라고 예측하였다. 또한 생산의 측면에서는 과학 기술의 발전과 융합이 자연스럽게 효율성과 생산성의 증가로 이어지면서, 상품과 서비스의 공

급이 비약적으로 빨라지고 개개인의 필요를 더 충실하게 충족시킬 수 있게 될 것이라는 결과를 도출했다. Schwab도 제4차 산업혁명으로 인해 막대한 경제적 이익이 산출될 것이며 그 규모는 기하급수적으로 늘어날 것이라고 예상했다(Schwab, 2017: 9-10). 이러한 연구결과를 통해 현재 인간이 하고 있는 힘들고 귀찮은 많은 일들이 인공지능 로봇 기술의 발전으로 인해 자동화되어, 인간의 생활이 보다 편리해지고 여유가 생길 것으로 예상할 수 있다.

그러나 인공지능 로봇으로 인해 많은 직업이 자동화되면 실업률이 증가할 것이고 결국 경제에 악영향을 끼칠 것이라는 비판론을 제시하는 연구결과도 많다(Walsh, 2018; Frey et al., 2017; Brynjolfsson, 2014, 미래전략정책연구원, 2016). 이와 같은 연구에서는 인공지능 로봇이 숙련된 인간 노동자들을 몰아내고, 교육받은 사람들의 숙련된 일을 대체하게 될 것이라고 예상한다. 또한 지금까지 인간만이 지적노동을 할 수 있었지만, 딥러닝 방식을 활용한 인공지능 로봇으로 인해 상당히 많은 지적노동 역시 자동화될 수 있다고 주장한다. 비판론을 주장하는 학자들의 연구결과를 구체적으로 살펴보면 다음과 같다.

Walsh는 Keynes, Turing, Leontief 등 많은 학자들이 50~80년 전에 예측했던 ‘기술적 실업’이 인공지능 로봇으로 인해 현실이 되고 있다고 판단했다(Walsh, 이기동 역, 2018: 206-209). Vardi도 기계들이 거의 모든 업무 영역에서 인간의 능력을 압도하게 되는 시대가 다가오고 있기 때문에 사회 전체가 나서서 대비책을 마련해야 한다고 주장했다(안상희 외 역, 2016: 2347). Frey 등(2017: 265)의 연구에서도 향후 20년 안에 미국 내 전체 일자리의 47 퍼센트가 사라질 위험에 처해 있다는 결과가 나타났다. Brynjolfsson 등은 로봇화와 자동화를 통해 상대적으로 안전하다고 여겼던 지식 기반 업무 역시 컴퓨터 알고리즘과 소프트웨어에 의해 대체될 것이라고 판단했다. 또한 로봇과 컴퓨터가 동력을 활용하는 업무나 정확성이 요구되는 계산 업무를 대체하는 수준을 넘어서서 인간만의 영역으로 여겨지던 인지적 판단이나 고도의 지적노동 업무까지 대체하기 시작했다고 주장했다(Brynjolfsson 등, 이한음 역, 2014: 178-181).

구본권(2015: 206-208)의 연구에 따르면, 로봇이 인간의 직업을 대체하는 흐름은 전문직군에서도 점차 나타나고 있다. 예컨대, 뉴스 로봇은 신문 기사를 인간보다 빠르고 정확하게 작성하고 있다. 왜냐하면 인공지능 로봇은 수많은 데이터를 가지고 있고 문장 구성 알고리즘은 갈수록 완벽해지고 있기 때문이다. 그리고 로봇은 약사의 업무도 점차 대체할 것이며, 변호사를 대신해서 방대한 소송 서류도 검토할 것으로 예상된다. 이러한 영향들로 인해 대량실업은 피할 수 없을 것이라는 전망이 많다. 자동화는 대량생산을 가능하게 해줄 것이고 비용을 절감하게 해줄 것이다. 또한 로봇은 인간처럼 지치지 않을 것이며 휴식을 요구하지도 않고 비용도 저렴하기 때문에 특정 분야의 직업들은 결국 사라지게 될 것이다.

미래전략정책연구원(2016: 143-146)의 연구에서도 결국 인공지능 로봇이 인간의 일자리를 빼앗고 그 결과 실업률이 증가하게 될 것으로 나타났다. 실업률이 늘어나면 경제활동인구가 줄어들고 소비인구가 감소하여 결국 경기가 악화될 것이다. 근로자가 줄어드는 만큼 소득세 수입도 줄어들 것이기 때문에, 결국 세금 수입도 감소하게 될 것이라고 전망했다. 그래서 몇몇 경제학자들은 해마다 전국의 모든 사람들에게 ‘기본 소득’ 제공 등의 다양한 대책을 내놓기 시작했다. 왜냐하면 모든 사람들이 기본 소득을 받으면 최소한의 생활수준을 보장받을 수 있고, 전체적인 소비도 촉진될 것이기 때문이다. 또한 Kaplan(신동숙 역, 2016: 25-26)은 로봇으로 인한 대량실업을 해결하기 위한 방안으로 취업교육을 촉진하기 위한 ‘직업대출’을 제안했다. 직업대출은 주택대출을 받을 때 토지를 담보로 제공하듯이 미래의 노동인 근로 소득을 담보로 대출을 제공하는 것이다.

그러나 낙관론을 주장하는 학자들은 인공지능 로봇으로 인한 자동화는 새로운 일자리를 창출할 것이라고 판단했다. 이종관(2017: 385-386)은 과거의 산업혁명이 농업 종사자 대신 공장노동자와 사무직 관리자라는 새로운 직업을 생산한 것과 같이, 인공지능 로봇은 제조업 등의 노동집약적 직업 대신에 더욱 양질의 일자리를 생산할 것이라고 보았다. 일례로 인공지능 로봇을 뛰어넘는 고도의 지적 능력을 갖춘 지식 자본가, 값비

싼 인공지능 로봇을 소유한 물적 자본가 등을 중심으로 슈퍼스타들의 경제가 출현할 것이라고 예상했다. 그들이 천문학적으로 축적한 부를 통해 새로운 소비의 시대가 열리고 서비스 산업이 촉진될 것이라고 주장했다. Roland Berger는 제품을 소유하는 것보다 사용한 만큼 지불하는 방식 등의 새로운 방식을 통해 서비스 분야에서 700만개의 일자리가 창출될 것이라고 예상했다. 이에 따라 제조업 분야 종사자 2,500만 명 중 900만 명이 서비스 분야로 이동할 것으로 예측하였다(Roland Berger, 김정희 외 역, 2017: 79-80).

지금까지의 내용을 종합해보면, 인공지능 로봇으로 인한 경제적 변화는 낙관론과 비관론으로 구분할 수 있다. 먼저, 인공지능 로봇으로 인해 경제적 규모가 커지고 새로운 가치 창출로 인해 일자리가 생겨날 것이라는 낙관론이 있었다. 그리고 인공지능 로봇으로 인해 실업률이 증가하고 소비인구가 감소해 결국 경제가 악화될 것이라는 비관론도 있었다. 결론적으로 인공지능 로봇으로 인해 긍정적이든 부정적이든 경제적으로 큰 변화가 있을 것이라는 사실은 확실한 것으로 보인다.

2) 사회·문화적 변화

인공지능 로봇으로 인한 사회·문화적 변화로는 어떤 것들이 있을까? 이 항에서는 인공지능 로봇의 발달로 인해 발생할 수 있는 사회·문화적 변화에 대해 구체적으로 살펴보고자 한다. 앞에서 살펴보았듯이 인공지능 로봇으로 인해 생산성이 향상되고 경제적 이익이 증대할 것이라는 낙관적인 견해와 인공지능 로봇으로 인한 자동화로 인해 실업률이 상승할 것이라는 비관적인 견해가 있었다. 이러한 연구결과를 통해, 인공지능 로봇으로 인해 실업률이 상승하게 되면 다양한 사회적 문제를 야기할 수 있다는 사실까지 유추할 수 있다.

첫 번째, 인공지능 로봇으로 인한 가장 큰 사회적 변화는 실업률 상승으로 야기되는 빈곤, 빈부격차 등의 사회적 문제이다. 미래전략정책연구원(2016: 143-146)의 연구에 따르면, 실업률이 높아지면 사회 전반적으로 여러 가지 사회 문제가 발생할 것이며, 실업자와 근로자 모두 불행해질 것으로 예상된다. 경제적 어려움을 겪은 실업자들의 절도 및 강도 등의 범죄가 늘어날 수도 있고, 실업자들의 시위도 늘어날 것이다. 이러한 상황을 예방하기 위해서는 정부에서 실업자들에게 기초생활비와 의료비 등을 지원해야 한다. 그렇게 되면 실업자 증대로 전체적인 세금 수입이 감소한 상황이기 때문에 근로자들이 세금을 더 많이 납부해야 한다. 결국 근로자들도 실업자들과 마찬가지로 사회에 대한 불만이 커지게 될 것이다. 이종관(2017: 387-390)의 연구에 따르면, 실업은 빈곤 이외에도 인간의 실존에 많은 문제를 일으킬 수 있다. 왜냐하면 일은 시간이나 임금의 단위로 환산할 수 있는 단순한 활동이 아니라 인간이 살아가는 의미이며 인간의 사회적 가치를 실현하기 위한 실존적 처신의 한 양태이기 때문이다. 결국 실업은 빈곤의 문제를 넘어서 인간에게 존재의 의미를 빼앗아 갈 우려가 있고, 인공지능 로봇으로 인한 실업률 상승은 인간의 실존에 많은 혼란을 가져올 것으로 예상된다.

또한 많은 학자들은 새로운 인공지능 기술의 발달로 인해 기업에만 이익이 집중되고 기업과 노동자간의 소득불평등 격차는 더욱 커져 사회가

계속 양극화될 것이라고 예견했다. Brynjolfsson 등은 현재의 정보기술이 직접 노동을 하는 근로자가 아닌 거대한 자본 소유자의 수익을 늘릴 것이기 때문에 대기업을 여러 방면에서 더 유리하게 만들 것이라고 주장했다(Brynjolfsson 등, 이한음 역, 2014: 188-190). Kaplan(신동숙 역, 2016: 25-26)에 따르면, 자본이 있는 사람들은 노동 능력이 주요 자산인 사람들의 희생으로 이득을 보게 될 것이고, 소득불평등은 앞으로 더욱 악화될 것이라고 전망했다.

김명주(2016: 48)는 자본에 따른 빈부격차뿐만 아니라 정보격차에 따른 빈익빈 부익부 현상도 심화될 것이라고 판단했다. 특히 인공지능 로봇의 활용, 인공지능 로봇의 보유 여부가 더 심화된 빈부격차를 가져올 것이라고 예상했다. 인공지능 로봇은 미래 사회에 중산층의 몰락을 가져와 결국 소수 계층이 사회적 부의 대부분을 차지할 것이라고 전망했다.

지금까지의 내용을 종합해보면, 인공지능 로봇으로 인해 자본이 있는 기업에게 부가 편중될 것이며 로봇을 보유하고 있는 사람에게 이익이 생겨 소득불평등이 커지고 사회가 혼란스러워질 우려가 있다. 이러한 문제를 해결하기 위해 로봇세, 일자리공유, 기본소득 등의 다양한 대안을 마련해야 할 것이다.

두 번째, 인공지능 로봇으로 인해 발생할 수 있는 그 밖의 사회적 변화는 프라이버시와 차별 문제이다. Walsh의 연구에 따르면, 알고리즘은 의도와 상관없이 사회 내의 여러 구성원에게 차별행위를 할 수 있다. 그는 이러한 차별행위를 일으키는 원인 중 하나는 인공지능 연구자들의 남녀 불균형 때문이라고 판단했다. 성별 불균형 때문에 간과되는 부분들이 있고, 해결되지 않고 지나치는 문제도 생길 수 있다고 주장했다. 아울러 흑인과 히스패닉계와 같은 인종의 인공지능 연구 참여 비율이 현저히 낮은 것도 심각한 문제라고 분석했다(Walsh, 이기동 역, 2017: 204-205). 이처럼 최근 정보의 디지털화가 빠르게 진행되면서 개인정보의 무한복제가 가능하게 되었고, 자율주행자동차와 소셜 로봇의 등장으로 많은 프라이버시가 침해될 것이라 생각된다. 프라이버시와 차별에 관련된 더욱 구체적인 내용은 III장 로봇윤리의 주요 쟁점에서 자세히 살펴볼 것이다.

앞에서 살펴본 바와 같이, 인공지능 로봇으로 인한 사회적 변화는 먼저, 실업률 상승으로 인한 빈곤, 빈부격차, 실존적 문제 등으로 분석되었다. 그리고 알고리즘에 의한 차별 문제와 프라이버시 문제 등도 인공지능 로봇으로 인한 사회적 문제가 될 것으로 예상된다.

다음으로 인공지능 로봇으로 인해 예상되는 문화적 변화에 대해 구체적으로 살펴보고자 한다. 인공지능 로봇의 발달로 인해 예상되는 문화적 변화에 대한 견해는 두 가지 주장으로 양분되어 있다.

첫 번째 주장은 문화와 예술에 관련된 창작 분야는 인간 고유의 영역이며, 창조성은 로봇과는 차별되는 인간만의 영역이기 때문에 앞으로 더욱 강조될 것이라는 주장이다. II장 1절 2항에서도 살펴보았듯이 인공지능 로봇의 발달로 인해 많은 직업들이 로봇으로 대체될 것으로 예상되지만, 인공지능 로봇으로 대체될 가능성이 가장 적은 직업이 창의적이고 감성적인 일을 하는 문화예술 관련 직업인 것으로 많은 연구들을 통해 도출된 바 있다(Frey et al., 2017: 267; 세계경제포럼, 2016: 15; 한국고용정보원, 2016: 12-14). 이러한 연구결과를 통해 수행 과정과 절차를 알고리즘으로 표준화하기 쉬운 특성을 가진 직업은 인공지능 로봇으로 인해 대체되기 쉽지만, 창의성을 발휘해야 하는 문화와 예술 분야 관련 직업은 인공지능 로봇으로 쉽게 대체될 수 없다는 사실을 알 수 있다. 이러한 연구결과를 통해 인공지능 로봇은 인간보다 상상하고 생각하고 느끼는 능력이 부족하다고 유추할 수 있다. 실제로 많은 학자들은 ‘창의적인 것’, ‘예술적인 것’은 인간만이 할 수 있는 독보적인 영역이며 인간만의 마지막 성역이라고 주장한다(김선영, 2018: 26). 그동안 예술은 인간만이 창조할 수 있는 영역의 산물이며 인간만이 예술적 가치에 대해 판단할 수 있다고 여겨졌기 때문이다. 그러나 최근 인공지능 로봇이 만든 창작물이 증가함에 따라 인공지능 로봇이 만든 것도 문화예술로 인정해야 한다는 주장이 점차 제기되고 있다.

두 번째 주장은 인공지능 로봇도 인간의 창의성을 능가하는 작품을 만들 수 있기 때문에 문화와 예술과 같은 창작 분야는 이제 인간만의 고유 영역이라는 기존의 관념을 흔들 것이고 ‘문화예술’에 대한 개념 재정

의까지 필요하게 될 것이라는 입장이다. 인공지능 로봇은 인간의 도구일 뿐이라는 과거의 통념과는 달리, 최근 인공지능 로봇이 문화예술 분야에서 만들어내는 창작물의 영역이 점차 확장되고 수준이 높아지고 있다.

문학 분야에서는 인공지능 로봇이 스토리와 관련된 단어들을 자동적으로 골라내 문장을 완성하는 작업을 통해 구성된 SF 소설이 일본 SF 문학상 1차 심사를 통과하였고, 영화 분야에서는 인공지능 로봇이 딥러닝 기법을 통해 만든 약 9분짜리 영화가 공개되기도 했고, 그동안 발표된 예고편을 학습한 인공지능 로봇이 만든 영화 예고편이 영화 홍보에 실제로 사용되고 있다. 미술 분야에서는 인공지능 로봇이 스스로 데이터를 쌓고 학습해 원하는 형태의 그림을 자유롭게 그리는 딥러닝 기법을 사용하여 그린 그림이 경매에서 높은 매출을 올렸으며, 전시 해설 분야에서도 관객의 흥미를 끌어내어 관람객이 증가하고 있으며, 작품 평가기능을 가진 인공지능 로봇까지 등장하고 있다(한국정보화진흥원, 2016: 6-10).

음악 분야에서는 빅데이터에 기반을 둔 작곡 로봇과 연주 로봇이 등장했다. 한 연구에서는 인간이 작곡한 음악과 인공지능 로봇이 작곡한 음악을 비평가들 포함한 청중들에게 투표한 결과, 어떤 음악이 로봇이 작곡한 것인지 구별해내지 못했고 오히려 로봇이 작곡한 음악을 혼이 담겨 있고 정서적 공명을 일으킨다며 더 낫다고 평가하기도 했다(Harari, 김명주 역, 2017: 444-445). 이러한 연구결과를 통해 인공지능 로봇이 작업한 음악의 질이 인간이 작업한 음악과 쉽게 구별이 안될 만큼 수준이 높다는 것을 알 수 있다. 공연 예술 분야에서는 휴머노이드 로봇보다 더 인간의 외양과 유사한 안드로이드형 로봇이 활용되면서 공연의 질이 더욱 높아지고 있다.

연극 분야에서는 안드로이드형 로봇이 65가지의 표정연기를 통해 배우로 활동하고 있고, 판소리와 오페라 분야에서도 인공지능 로봇이 가수로 활하고 있다. 무용 분야에서도 무용수의 움직임을 모방하여 춤을 추는 로봇이 활동하고 있고, 스스로 음악변화에 따라 스스로 춤을 추는 댄싱 로봇이 개발된 상태이다(태혜신 외, 2019: 35-39).

앞에서 살펴본 것과 같이, 문화예술 분야에서 인공지능 로봇의 창작물

이 증가하고 창작물의 질도 인간의 창작물과 비슷한 수준까지 발전하고 있다는 것을 위의 연구결과를 통해 알 수 있다. 그렇다면 인공지능 로봇이 작곡하고 연주하고 그림을 그린 작품들도 문화예술 작품이라고 할 수 있을까? 앞으로 인공지능 로봇 기술이 더욱 발전하게 되면 단순히 인간의 문화 예술 작품을 모방하는 것을 넘어서 스스로 작품을 창조하게 된다면, 과연 인간이 창조하는 것만 문화 예술 작품이라고 할 수 있을지에 대한 의문이 제기된다.

초기 문화예술은 주로 모방하는 것이 대부분이었으며, Socrates와 Da Vinci 등 여러 저명한 예술가들도 가능한 한 훌륭하게 자연을 모방하는 것이 문화예술이라고 정의하였다(오윤자: 2018, 50-52). 그 후에는 실제로 보지 않았던 것도 재현할 수 있는 상상력과 직관이 곧 문화예술이라는 생각이 우세했다. 그러나 Dewey(이재언 역, 2003: 120-124)는 문화예술에 있어 상상은 주관적인 자기인식으로서 예술경험의 결과가 아니라 반성적 판단력의 결과이며, 순수한 자기의식이라고 하였다. 즉, 문화예술은 단순한 모방, 상상력, 직관의 개념을 뛰어넘은 세상에 대한 철학적 사유가 중심이 되는 종합적인 개념이라고 판단된다. 문화예술은 지성이 아닌 감성의 경험이며, 인간정신의 산물이기 때문에 문화예술적 가치에 대한 논의는 그동안 인간을 중심으로 이루어져왔다. 즉, 문화예술은 인간 존재가 아닌 동물 등 다른 존재들에게는 확장되지 않은 오직 인간만의 것이었는데 이제 인공지능 로봇의 도전에 직면하게 된 것이다.

현재 인공지능 로봇 기술은 딥러닝을 통해 작품제작 혹은 작품해석이 가능하지만 이는 평균값을 찾아내는 기술로 감정이 아닌 지적 능력이며, 직관기능도 단순 변형기술 또는 추상기술에 불과하다. 따라서 현재의 인공지능 로봇이 만든 문화예술 작품을 진정한 의미의 문화예술 작품으로 인정하기는 불충분한 것으로 보인다. 현재 인공지능 로봇이 만든 작품은 단순한 기술적 성과일 뿐 인간이 만든 것과 같은 철학적 사유, 세계관, 가치관, 작가의 의도 등이 반영되지 않은 것이기 때문이다. 그러나 현재의 인공지능 로봇 기술 수준으로 로봇이 제작한 것을 문화예술이 아니라고 쉽게 단정할 수는 없다. 많은 학자들은 가까운 미래에 철학적 사유를

할 정도의 자율성을 가진 강인공지능이 개발될 것으로 예상한 바 있다 (Kurzweil, 2010: 454; Schwab, 2017: 98-99). 따라서 인공지능 로봇이 인간 정도의 자율성을 가지게 된다면, 현재 인간 고유의 것이라는 문화예술의 정의에 대한 논의와 수정이 필요할 것으로 보인다.

지금까지 살펴본 것과 같이, 인공지능 로봇으로 인한 문화 예술적 변화에 대한 연구 결과는 두 가지 주장으로 양분된다. 문화와 예술에 관련된 창작 분야는 인공지능 로봇으로 대체될 수 없는 인간 고유의 영역이기 때문에 앞으로 더욱 강조될 것이라는 낙관론과 인공지능 로봇 기술이 인간의 자율성 수준의 강인공지능을 개발하게 되면 더 이상 인간만의 문화예술이 아닌 로봇이 함께 또는 그 이상을 차지할 것이라는 비관론이 있었다. 현재는 인공지능 로봇의 문화예술 작품에 철학적 사유, 세계관, 작가의 의도 등이 반영되지 않았기 때문에 진정한 의미의 문화예술 작품으로 인정하기에는 아직 불충분하지만, 인공지능 로봇이 인간 정도의 자율성을 갖게 되면 인간 고유의 문화예술에 대한 논의와 수정이 필요할 것으로 사료된다.

3) 교육적 변화

앞에서 살펴보았듯이 인공지능 로봇의 급격한 발달은 경제적 변화와 사회·문화적 변화 외에 앞으로의 교육에도 많은 영향을 끼칠 것으로 예상된다. 교육은 본질적으로 미래 사회를 선도해야 하는 책임이 있다. ‘로봇 윤리의 도덕교육적 적용 방안 연구’를 위해 이 항에서는 인공지능 로봇으로 인한 교육적 변화에 대해 구체적으로 살펴볼 것이다. 인공지능 로봇으로 인해 예상되는 교육적 변화를 자세히 살펴보면 다음과 같다.

인공지능 로봇으로 인해 예상되는 교육적 변화는, 교육현장에서 본격적으로 인공지능 로봇이 활용될 것이라는 점이다. 이미 중국에서는 2000년대부터 인공지능 로봇이 학교 교육현장에서 적극적으로 활용되고 있다(Zhang et al, 2008: 762). 인공지능 로봇은 개별적 교수법이 가능하고 교육의 목적을 극대화할 수 있는 다양한 장점이 있어 우리나라에도 도입될 가능성이 점차 커지고 있다. 인공지능 로봇을 교육적으로 활용하는 교육용 로봇에 대한 효과성은 많은 연구를 통해서 입증되고 있다. 교육용 로봇은 긍정적인 학습 성과와 즐거움을 촉진하고 학생들의 수업 집중력을 높이는 데에도 효과가 있는 것으로 나타났다(Tanaka et al., 2009: 558-560).

교육현장에서 인공지능 로봇을 활용하는 방법은 크게 두 가지로 구분될 수 있는데, 수업에 대한 다양한 교육 자료 제공 및 반복학습 등으로 단순하게 수업을 보조하는 수업용 도구로 활용하는 방법과 교실 내에서 적극적으로 학생들과 정서적인 상호작용을 주고받으며 사회적 지능을 겸비한 로봇으로 활용하는 방법이 있다. 첫째, 인공지능 로봇은 외국어교육에 수업용 도구로 활용될 수 있다. 중국의 한 초등학교에서 활용된 외국어교육용 로봇은 다음과 같은 효과를 가져왔다. 로봇의 반복능력을 통해 학생들이 언어를 반복하여 학습할 수 있었고, 교사들은 로봇에게 그들의 데이터를 축적하여 다음 수업에 활용할 수 있었다. 또한 로봇은 교사를 대신해서 학생들에게 스토리텔링과 말하기, 질문과 답변 등의 교육 기능을 수행할 수 있었다(Chang et al, 2010: 15-17). 이러한 연구결과를 통해 인공지능 로봇은 수업을 위한 다양한 자료, 시청각 정보 제공 등을

포함한 학습 보조 능력을 갖추고 반복학습을 통해 수업을 보조하는 단순한 교육용 도구로 활용될 수 있다는 사실을 알 수 있다.

둘째, 인공지능 로봇은 인간과 로봇의 상호작용 기술을 도입하여 교실 안에서 학생들과 정서적인 상호작용을 통해 학생들의 생활을 도와주는 소셜 로봇(social robot)으로 활용될 수 있다. 인공지능 로봇은 자율적으로 움직일 수 있기 때문에 특히 장애 학생들을 위해 사용될 수 있다. 인공지능 로봇은 장애 학생들을 신체적으로 보조하고 학교 현장에서 장애 학생들을 지원하는 데에도 활용될 수 있다. 이에 따라 변순용 등(2015: 206-209)은 학교 현장에서 인공지능 로봇의 활용이 기존의 학습 도구 수준을 넘어설 것으로 예상하며 도덕과에서 인공지능 로봇을 활용할 수 있는 두 가지 방법을 제안했다. 먼저, 인공지능 로봇은 학생들과 정서적인 상호작용을 할 수 있기 때문에 자폐증 또는 신체적 장애를 가진 학생들을 교육적·정서적으로 보조할 수 있다. 향후에 인공지능 로봇이 교과내용 뿐만 아니라 인간과 유사한 지능, 정서적 소통 능력, 상호작용 능력을 갖추게 된다면, 학생들과 정서적으로도 공감할 수 있을 것으로 예상된다.

그러나 인간과 인공지능 로봇의 상호작용 기술을 활용한 인공지능 로봇에 대해 우려를 나타내는 많은 연구가 있다. Tanaka 등(2009: 558-560)에 따르면, 교실에서 인공지능 로봇을 사용하는 것은 안전하지 못하며 학생들을 게으르게 만들 수 있다. 결과적으로 인간에 대한 진정한 애착이 결여될 수 있어 학생들에게 여러 가지 정서적 혼란을 가져올 수 있다. 그리고 로봇은 단순하게 교사들의 교육 활동을 보조하고 지원할 수 있도록 설계되어야 하며 교사의 통제 하에 있어야 하는 도구라는 사실을 항상 인지해야 한다.

Serholt 등(2017: 613-614)의 연구에서 77명의 교사들과 예비교사들을 대상으로 인공지능 로봇의 활용에 대해 설문조사를 실시한 결과, 교사들과 예비교사들은 교실 안에서 인공지능 로봇을 사용하는 것에 대해 위험하다고 생각하는 것으로 나타났다. 첫 번째 이유는 프라이버시 문제와 관련되어 있다. 교사들은 교실 안에 사용되는 인공지능 로봇이 학생들의 행동 발달 상황을 기록하기 위해 학생들과 관련된 데이터를 수집하게 되

면, 학생들의 개인정보가 이후 다른 용도로 이용될 수 있기 때문에 심각한 프라이버시 침해 문제가 발생할 수 있다고 보았다.

두 번째 이유는 인공지능 로봇을 활용하는 것이 학생들을 비인간적으로 만들 우려가 있기 때문이다. 학생들이 인공지능 로봇과 소통하게 되면, 인공지능 로봇은 표정이 다양하지 않기 때문에 학생들이 인간의 표정을 이해하는 데 어려움을 겪기 시작할 것이다. 그렇게 되면 학생들이 인간과 소통하는 정서적 능력이 점차 저하될 것이다. 그리고 학생들이 로봇에게 폭력을 행사해도 결코 로봇은 다치지 않기 때문에 학생들이 인간에게도 폭력을 행사할 수 있다고 우려했다.

세 번째 이유는 학생들이 일관성 있고 예측 가능하고 정확한 정보를 제공하는 로봇을 교사보다 더 신뢰하고 의지할 수도 있어, 교사를 믿지 못하거나 배척하는 현상까지 나타날 수도 있다는 것이다. 네 번째 이유는 인공지능 로봇에 대한 책임 문제와 관련되어 있다. 교실 안에서 만약 인공지능 로봇이 학생들에게 부정적 영향을 미치거나 사고를 일으키게 되면, 누가 책임을 질 수 있는지 명확하지가 않다는 것이다.

지금까지의 논의를 종합해보면, 인공지능 로봇으로 인해 예상되는 교육적 변화는 학교 현장에서 인공지능 로봇이 적극적으로 활용될 것이라는 점이다. 학교 현장에서 인공지능 로봇을 활용하는 방법은 수업에 대한 다양한 교육 자료 제공 및 반복학습 등으로 단순하게 수업을 보조하는 수업용 도구로 활용되는 방법과 교실 내에서 적극적으로 학생들과 정서적인 상호작용을 주고받으며 소셜로봇으로 활용되는 방법이 있었다. 인공지능 로봇을 활용하는 것은 교육적인 효과가 분명히 있지만, 교실 안에서 학생들의 정서적 혼란, 프라이버시, 비인간화, 책임 등의 부작용이 나타날 수도 있기에 이에 대한 예방과 대비가 요청된다. 조만간 학교 현장에서 여러 교육적 활용을 위해 인공지능 로봇을 도입할 가능성이 있기 때문에, 학생들이 이러한 윤리적 문제에 대비할 수 있는 로봇윤리교육의 필요성도 점차 커지고 있는 것으로 생각된다.

Ⅲ. 로봇윤리의 의미와 쟁점

‘로봇윤리의 도덕교육적 적용 방안’을 마련하기 위해서는 먼저 로봇윤리의 의미와 주요 쟁점에 대한 논의가 선행되어야 할 것이다. 이에 따라 이 장에서는 로봇윤리의 의미와 주요 쟁점이 무엇인지에 대해 살펴볼 것이다. 구체적으로 1절에서는 로봇윤리의 의미를 규명하기 위해 먼저, 로봇윤리의 개념을 정립하고 범위를 설정할 것이다. 2절에서는 로봇윤리를 ‘로봇의 윤리’와 ‘인간의 윤리’로 구분하여, 그 의미를 더욱 구체적으로 살펴볼 것이다. 3절에서는 ‘로봇의 윤리’와 ‘인간의 윤리’에 대한 주요 쟁점을 논의함으로써, 이후 ‘로봇윤리교육의 의미와 내용’ 연구의 토대를 마련하고자 한다.

1. 로봇윤리의 의미

이 절에서는 로봇윤리의 의미를 규명하기 위해 먼저 국외·국내 문헌을 분석하여 로봇윤리의 개념을 밝히고 로봇윤리의 범위를 설정할 것이다. 그리고 로봇윤리의 필요성을 논증하여 로봇윤리의 의미를 보다 명확하게 규명할 것이다.

1) 로봇윤리의 개념과 범위

로봇윤리(robotethics)는 무엇인가? 로봇윤리를 최초로 언급한 학자는 Isaak Asimov이다. Asimov(1942)는 인간의 도구로 사용하기 위해 개발한 로봇이 인간을 해치는 상황이 발생하지 않도록 로봇공학 3원칙을 발표했다. 그 후에 조금 수정되긴 했지만, 여전히 인공지능 로봇과 관련된 윤리문제를 포괄하기에는 충분하지 못하다. 이에 따라서 본 연구에서는 로봇공학 3원칙 이후로 다양하게 정의되고 있는 로봇윤리의 정의를 고찰하여 로봇윤리의 개념을 새롭게 정립하고자 한다.²⁾

로봇윤리(roboethics)의 개념은 아래와 같이 학자들마다 다양하게 정의되고 있다. 지금까지 논의된 로봇윤리 개념을 연구자에 따라 연도순으로 정리하면 다음과 같다.

<표 3> 로봇윤리에 대한 다양한 정의

연구자	개념 정의
Verruggio (2006: 7)	로봇공학이 인간의 삶에 적용될 때 나타날 수 있는 윤리적 문제를 다루는 것이며, 로봇기술이 인류를 공격하고 잘못 사용되지 않게 로봇공학의 발전을 장려하는 것
Tamburrini (2006: 265)	현재 그리고 미래의 로봇 사용과 관련된 윤리적인 문제들을 추출하여 분석하고자 노력하는 학문
Tzafestas (2018: 3)	로봇공학기술 중 특히 인공지능 로봇의 윤리적 영향에 대해 이해하고 올바르게 규제하기 위한 연구
변순용 · 송선영 (2015: 17)	수단으로서의 로봇, 윤리적 영역을 갖는 로봇, 임무 수행의 도덕적 행위 특성에 초점을 맞춘 로봇과 인간(설계자, 생산자, 사용자, 관리자)과의 관계를 둘러싼 다양한 문제들을 해결하기 위한 연구
카카오 편집진 (2018: 57)	인간을 이롭게 하려는 본래의 목적에 부합하기 위해 로봇이 지켜야 할 준칙

여기서 주목할 만한 것은, Verruggio와 Tamburrini, Tzafestas의 개념 정의인데, 이들에 따르면 로봇윤리는 응용윤리학의 한 분야로 로봇기술이 인간에게 미치는 영향에 대한 윤리적 문제를 다루는 것이며, 로봇기술이 긍정적인 방향으로 사용될 수 있도록 하는 연구이다. 즉, 인간이 로봇기술을 안전하게 잘 다루기 위한 연구이며 인간이 주체인 연구라고 볼 수 있다.

이와는 달리, 카카오 편집진이 로봇윤리를 ‘인공지능과 로봇이 마땅히 행하거나 지켜야 할 도리’ 라고 정의한 것은 로봇이 주체인 로봇윤리의 차원을 나타내고 있다. 이렇게 볼 때, 현재 연구자들이 정의한 로봇윤리의 개념에서는 로봇이 주체인 로봇의 윤리와 인간이 주체인 인간의 윤리가 제대로 구분되지 않고 사용되고 있다는 것을 알 수 있다.

2) ‘roboethics’(로봇윤리)라는 말은 2002년 G. Veruggio에 의해 처음 사용되었으며, 공식적인 사용은 2004년 이탈리아 산레모(San Remo)에서 열린 제1회 로봇윤리 국제심포지엄(First International Symposium on Roboethics)에서 비롯되었다.

이에 대해 보다 구체적으로 논의해보면, 로봇윤리의 개념은 윤리를 행하는 주체가 누구인지에 따라서 다음과 같이 세 가지 수준으로 정의될 수 있다. 첫 번째 수준의 로봇윤리(roboethics)는 인간이 주체가 되는 윤리이며, 철학과 인문학적 이슈와 관련 있는 응용윤리의 한 분야로서의 로봇윤리를 뜻한다(Verrugio, 2006: 7; Tamburrini, 2006: 265; Tzafestas, 2018: 3). 로봇으로 인해 인간의 삶에서 일어나는 여러 윤리적 문제들을 해결하기 위해 인간이 지켜야 할 윤리를 뜻한다.

두 번째 수준의 로봇윤리(robot ethics)는 주로 과학 및 기술과 관련 있는 ‘공학윤리로서의 로봇윤리’이다(Verrugio, 2006: 7; Tzafestas, 2018: 3, 변순용·송선영, 2015: 17-18). 두 번째 수준의 로봇윤리도 첫 번째 수준의 로봇윤리처럼 인간이 주체가 되는 윤리이며, 로봇을 설계하고 판매하고 사용할 때 인간이 지켜야 할 윤리를 의미한다. 로봇공학자의 윤리, 로봇 제조업체의 윤리, 로봇 사용자의 윤리 등의 의미를 모두 포함하는 개념이다.

세 번째, 가장 높은 수준의 로봇윤리(robot' s ethics)는 위의 두 개념과는 달리 로봇이 주체가 되는 윤리이며, 주로 SF영화에서 등장하는 로봇의 경우를 가정하여 제시된 로봇이 지켜야 하는 행동 규범으로서의 로봇윤리를 의미한다(Verrugio, 2006: 7; Tzafestas, 2018: 3, 변순용·송선영, 2015: 17-18). Asimov의 로봇공학 3원칙이 이 수준에 해당되는 로봇윤리를 의미한다고 볼 수 있다. SF영화에서 등장하는 로봇처럼 현재의 기술보다 더 발전되고, 인간의 역량을 뛰어넘는 로봇을 가정해서 그들이 스스로 윤리적 규범을 만들고 지키는 상황을 가정하여 정의된 개념이다. 세 번째 수준의 로봇윤리는 인간이 로봇에게 도덕코드를 부여해서 로봇이 도덕의 행동규범으로 지키게 하는 윤리의 개념까지 포함한다.

이와 같이 로봇윤리의 개념은 인공지능 로봇과 관련된 여러 윤리적 문제를 윤리학적으로 접근하는 ‘응용윤리로서의 로봇윤리(roboethics)’, 로봇을 설계하고 판매하고 사용할 때 지켜야 할 ‘공학윤리로서의 로봇윤리(robot ethics)’, 로봇이 스스로 규범을 제정하고 준수하는 ‘로봇의 윤리(robot' s ethics)’의 세 가지 수준을 모두 포함한다. 세 가지 수준의 로봇윤리를 도덕적 행위의 주체에 따라 좀 더 엄격하게 구분해보면 다음과 같

다. 응용윤리로서의 로봇윤리(roboethics)와 공학윤리로서의 로봇윤리(robot ethics)에서 도덕적 행위 주체는 인간이고, 로봇의 윤리(robot's ethics)의 주체는 로봇이므로 결국 로봇윤리는 ‘로봇의 윤리’와 ‘인간의 윤리’로 귀결된다. 따라서 로봇윤리의 개념을 인간과 로봇이 지켜야 할 윤리의 의미를 모두 포괄하는 ‘인간과 로봇이 함께 살아가기 위해 인간과 로봇이 지켜야 할 윤리’라고 규정할 수 있다.

앞에서 정립한 로봇윤리의 개념에 따라 도덕적 행위의 주체에 의한 구분으로 로봇윤리를 범주화하면, ‘로봇의 윤리’와 ‘인간의 윤리’로 나눌 수 있다. 우선 ‘로봇의 윤리(robot's ethics)’의 범위는 로봇이 도덕적 행위의 주체이기 때문에, Asimov식의 로봇윤리와 인간이 로봇에게 부여하는 로봇코드를 포함한다. 이와는 달리 ‘인간의 윤리’의 범위는 응용윤리로서의 로봇윤리(roboethics)와 공학윤리로서의 로봇윤리(robot ethics)를 포함한다. 여기서 설정한 로봇윤리의 범위를 토대로 하여 다음 장에서 보다 구체적으로 ‘로봇의 윤리’와 ‘인간의 윤리’의 의미에 대해 살펴볼 것이다.

2) 로봇윤리의 필요성

앞의 장에서 살펴보았듯이 로봇윤리는 인공지능 로봇 기술이 급격하게 발전하고, 새로운 유형의 윤리적 문제가 예상되면서 그 중요성이 더욱 강조되고 있다. 그렇다면 로봇윤리는 왜 필요할까? 기존의 다른 윤리로는 해결할 수 없는 문제인가? 이러한 물음에 대해 많은 학자들은 새로운 윤리가 필요하다고 주장했다(Verrugio, 2006: 7; Tamburrini, 2006: 265; Tzafestas, 2018: 3). 본 연구에서는 왜 기존의 윤리가 아닌 로봇윤리를 통해서만 제4차 산업혁명시대의 윤리적 문제에 대비할 수 있는지에 대해 주목하여 로봇윤리의 필요성을 논증하려고 한다.

제4차 산업혁명시대에 기존 산업사회나 정보사회와는 다른 새로운 윤리가 필요한 이유는 인공지능 로봇의 발달로 인해 기존의 윤리적 문제 외에 새로운 유형의 윤리적 문제를 야기하고 있기 때문이다. 윤리적 지체 이론에 따르면, 인간의 윤리 의식은 그 사회의 기술적 발전 속도에 비해 지체된다고 볼 수 있다. 인간의 윤리 의식이 기술적 영역이 도달한 수준까지 단시일 내에 도달하지 못할 경우 많은 혼란을 가져올 것으로 예상된다. 즉, 인공지능 로봇 기술 발전에 따른 도덕적 혹은 정책적 진공 상태가 존재할 가능성이 있다. 예를 들어, 자율주행자동차는 이미 개발되어 상용화될 시점이 다가오고 있지만, 자율주행자동차가 사고가 나면 운전자가 책임을 질지 제작자가 책임을 질지 등 책임의 문제가 여전히 해결되지 않고 있다.

그러나 인공지능 로봇을 둘러싼 윤리적 문제들이 완전히 새로운 것은 아니다. 왜냐하면 인공지능 로봇으로 비롯된 윤리적 문제들은 이전의 것과는 다른 유형이지만, 완전히 새로운 범주의 이슈를 만들어내는 것은 아니기 때문이다. 예를 들어, 프라이버시에 대한 위협은 인공지능 로봇과 관련된 문제의 형태로 존재하지는 않았지만 프라이버시 문제는 오랫동안 우리 주변에서 제기된 부분이다. 즉, 우리는 인공지능 로봇과 관련된 새로운 유형의 문제를 다루는 데 있어서 전통적인 도덕적 원리와 이론을 활용할 수 있다. 그리고 인공지능 로봇과 관련된 윤리적 문제들은 전통적인 윤리적 개

념들을 사용하여 범주화할 수 있다.

전통적으로 윤리는 앞으로 일어날 일을 예방하는 기능보다는 이미 일어난 일에 대한 도덕적 평가의 역할을 수행했으며, 그 결과 늘 시대적 변화를 따라가지 못하는 심각한 윤리적 지체 현상을 겪어 왔다. 지금까지의 내용을 종합해보면, 윤리적 지체 현상을 방지하고 시대적 변화에 대비하기 위해 로봇윤리는 꼭 필요하다고 판단된다.

로봇윤리에 관한 연구는 Asimov의 로봇공학 3원칙 이후로 한동안 진전이 없다가 최근 인공지능로봇의 비약적 발전과 군사로봇, 소셜 로봇 등 다양한 로봇의 등장에 따라 여러 윤리적 문제점들이 대두되면서 다시 등장하게 되었다. 과학기술은 인류에게 좋은 혜택도 제공하지만 새로운 문제를 동반한다. 따라서 이에 적절하게 대비하지 못하면 새로운 과학기술은 인류에 도움이 될 수 없기 때문에, 전 세계적으로 인공지능 로봇과 관련된 문제를 해결할 수 있는 로봇윤리에 대한 관심이 더욱 높아지고 있다.

현재 세계 각국에서는 로봇윤리와 관련된 많은 논의들이 진행되고 있다. 유네스코 산하의 세계과학기술윤리위원회(COMEST)는 지속적인 회담을 개최하여 로봇윤리에 대한 새로운 논의들과 기준들을 제정하고 있다. 각종 학술대회에서는 로봇윤리와 관련된 법, 정책 이슈들을 지속적으로 논의하고 있으며, 구글, 페이스북, 테슬라, 카카오 등의 기업에서도 인공지능과 로봇윤리에 주목하여 로봇윤리현장을 발표하고 있다. 왜냐하면 인공지능 로봇으로 인해 인류에 해를 끼치는 윤리적 문제가 발생하게 되면, 기업의 존폐까지 영향을 주는 큰 위기로 이어질 수 있기 때문이다. 우리나라에서도 2007년 산업자원부에서 당시 정부 차원으로는 세계 최초로 로봇윤리현장을 제정하려고 시도했으며 제4차 산업혁명에 대한 관심으로 로봇윤리에 대한 관심도 또한 높아지고 있다.

지금까지 살펴본 내용을 바탕으로, 로봇윤리가 필요한 이유를 세 가지로 종합해보면 다음과 같다.

첫째, 인공지능 로봇의 정체성을 확립하기 위해 로봇윤리가 필요하다. 최근 인공지능 로봇과 관련된 문제나 사고에 대해 인공지능 로봇을 인공적 도덕행위자(artificial moral agent)로 간주하여 어느 정도 윤리적·법적

책임을 부과해야 한다는 주장이 나오고 있다. 혹자는 인공지능 로봇을 인간과 동등한 타자로 인정하여 인간과 수평적으로 공존이 가능한 사회에 미리 대비해야 한다고 주장한다. 따라서 인공지능 로봇이 인간과 수평적으로 공존할 대상인지에 대한 문제를 심층적으로 다루는 윤리가 필요하다.

또한 일본이나 유럽에서는 성(性) 로봇(sex robot)이 ‘배우자 로봇’이라는 이름으로 판매되고 있으며, 주인의 정신적 필요와 육체적 필요를 채워주는 기능으로 사용되고 있다. 이에 따라 로봇 주인들은 배우자 로봇에 대한 법적인 신분보장을 요구하며, 인공지능 로봇을 정식 부부 혹은 가족의 일원으로 등록할 수 있도록 요구하고 있다. 유럽에서는 로봇에 대한 법적 지위를 이미 어느 정도 보장해주고 있어서, 국내에서도 이러한 요구가 지속적으로 증가하게 되면 이에 따른 사회적 갈등과 충격은 상당히 클 것으로 예상된다(김명주, 2017: 46-47). 이러한 상황은 전통적인 개념의 부부, 가정에 대한 재정의 문제와, 사회적 공통규범에 대한 윤리적 혼란 등을 가져올 것으로 예상되기 때문에, 이에 대비하기 위해서 로봇윤리가 필요할 것으로 보인다.

둘째, 인공지능 로봇에 관련된 책임문제를 해결하기 위해 로봇윤리가 필요하다. 인공지능 로봇 기술이 비약적인 발전을 이루었으나 자율주행자동차의 사고 책임 및 로보 어드바이저(robo advisor)의 투자 실패 책임 등 여러 문제가 발생하고 있고, 그에 대한 책임 문제가 대두되고 있다. 인공지능의 자율적 결정에 의하여 피해가 발생하게 된다면 누가 어디까지 책임을 질 것인가에 대해 미리 논의되어야 한다. 인공지능 로봇으로 인한 사고에 대해 공학자, 제작자, 사용자에 대한 책임을 어떻게 부여할 것인지에 대한 논의와 함께 향후 인공지능 로봇의 자율성이 인간과 동일한 수준으로 발전한다면 인공지능 로봇에게도 책임을 부과해야 하는지에 대한 논의도 필요하기 때문에 로봇윤리는 필요하다.

셋째, 인공지능 로봇과 관련된 편견, 인종 차별 문제를 해결하기 위해 로봇윤리가 필요하다. 2016년 3월 인공지능 로봇 테이(Tay)는 공개된 지 16시간 만에 전격 철수되었다. 특정 네티즌들이 주입한 왜곡된 내용을 집중적으로 학습한 테이는 ‘히틀러가 옳으며 나는 유대인이 싫다’, ‘나는

페미니스트를 혐오한다.’ 등의 극단적인 편향성 발언을 하여 세간의 비난 대상이 되었다. 이 사건을 통해 회사의 의도와는 상관없이, 제작자도 예측하지 못한 결과가 인공지능 로봇에게 나타난 사례를 통해 알고리즘 제작 초기부터 윤리적 문제에 대한 고려가 필요하다는 것을 알 수 있다. 2015년 6월 미국 뉴욕에서 일하는 한 흑인 프로그래머는 자신의 SNS에 사진 캡처 화면과 함께 ‘내 친구는 고릴라가 아니다’ 라며 욕설이 담긴 글을 올렸다. 인공지능을 이용한 포토 서비스 중 얼굴 자동인식 오류로 주인공의 여자 친구인 흑인 여성을 고릴라로 인식하고 태그하면서 문제된 일이 있었기 때문이다. 인공지능 로봇을 제작한 회사에서는 ‘기능의 오류’ 라며 즉각 사과하고 재발 방지를 약속했다.

이러한 사건들은 단순한 기술적 오류가 아닌 인공지능으로 인해 표출된 우리 사회의 모습이다. 왜냐하면 최근의 인공지능은 스스로 학습하면서 수많은 데이터를 쌓는 방식으로 작업이 수행되기 때문에, 흑인 여성에게 고릴라라고 칭한 것은 단순한 인공지능의 오류가 아니라 결국 우리 사회가 흑인을 어떻게 인식하고 있는지를 보여주는 사례이기 때문이다. 즉, 인공지능은 기계학습을 통해 입력 데이터를 습득하는 데 이 입력 데이터에 성적 편견이나 인종적 차별성이 존재한다면 인공지능도 이를 반영하여 결과를 도출하는 것이다. 입력 데이터가 아닌 알고리즘 자체가 원인이 되어 편견과 편향을 가지게 되면 문제가 더욱 심각해진다. 만약 은행 대출심사를 인공지능이 결정한다면, 인공지능이 대출 희망자의 주소, 인간관계, 소비패턴 등의 데이터를 기반으로 차별적인 판단을 할 수 있다. 따라서 인공지능 로봇이 편견을 가지지 않고 공정하게 사용될 수 있도록 윤리적 보완책을 미리 마련해야 한다.

지금까지의 내용을 종합해보면, 인공지능 로봇의 정체성 확립 문제, 인공지능 로봇과 관련된 책임 문제, 편견 문제 등 로봇에 관련된 여러 문제들을 해결하기 위해 예방적 차원에서 로봇윤리가 필요한 것으로 생각된다. 인공지능 로봇 기술은 계속 발전할 것이고 로봇과 관련된 윤리적 문제는 앞으로 더욱 증가할 것으로 예상되기 때문이다.

2. 로봇윤리의 구분

앞의 절에서 로봇윤리를 ‘로봇의 윤리’와 ‘인간의 윤리’로 범위를 설정한 바 있다. 앞에서 설정한 로봇윤리의 범위를 토대로, 이 절에서는 ‘로봇의 윤리’와 ‘인간의 윤리’에 대해 보다 구체적으로 논의하여 이후 로봇윤리교육의 의미 정립을 위한 기초를 마련하고자 한다.

1) 로봇의 윤리

앞에서 살펴본 것과 같이, 로봇의 윤리(robot's ethics)는 인공지능 로봇이 인간과 비슷한 정도의 자율성을 갖춘다는 가정 하에 로봇이 스스로 지켜야 할 윤리와 인간이 로봇에게 도덕코드를 부여하는 것까지 포함하는 기계윤리(machine ethics)이다. 그렇다면 SF영화 속처럼 인공지능 로봇이 그들 스스로 윤리를 세우고 그것을 준수하는 것이 가능한가? 다시 말해 과연 인공지능 로봇이 인간처럼 윤리를 스스로 세우고 지키는 도덕적 행위자가 될 수 있는가? 그렇다면 도덕적 행위자인 인간은 어떤 특성을 가지고 있을까? 이런 점에 비추어 볼 때 어떤 요소를 가지고 있어야 인공지능 로봇도 도덕적 행위자라고 할 수 있을까?

인간이 도덕적 행위자인 이유는 인간이 자기가 법칙을 수립하고 스스로 그에 복종할 수 있는 능력인 자율성을 가지고 있기 때문이다. 인간은 자신의 법칙을 스스로 수립하고 그것을 지키는 자기기획, 자기결정, 자기 입법적 성격을 가지고 있다. 여기서 자율성의 의미는 공학자들이 사용하는 의미의 자율성과는 차이가 있다. 인간의 자율성은 행태가 자율적이라든가 타자가 만든 프로그램에 따라 움직이는 것이 아니라 스스로 입법하여 자기 행동을 규제하는 도덕적 주체성을 뜻한다. Kant는 자율성이 인간을 규정하는 포괄적이고 핵심적인 개념이라고 하며, 욕구를 제어하고 이성이 규정하는 선한 것을 행할 수 있는 능력, 즉 이성에 의한 자율성이 진정한 자율성이라고 규정한 바 있다(백종현, 2018: 46-47 재인용).

그렇다면 인공지능 로봇도 자율성을 가지고 있을까? 많은 학자들은

인공지능을 ‘자율적 행위자(autonomous agent)’로 지칭하며 인공지능 로봇이 이미 상당한 수준의 자율성을 지니고 있다고 주장한다(Kurzweil, 2010; Russell et al., 2016; 고인석, 2018). 그러나 앞에서 살펴본 ‘자율성’의 개념과 공학자들이 사용하는 ‘자율성’의 의미는 차이가 있다. 공학자들이 사용하는 ‘행위자(agent)’의 개념은 인간이나 동물 같은 생명체가 아닌 단순히 ‘무엇인가를 수행하는 어떤 것’을 가리키는 의미로 사용되고 있다. 또한 ‘자율성’의 의미도 ‘인간의 개입 없이 스스로 임무를 수행하는 역량’을 의미한다(고인석, 2018: 87-89). 즉, 로봇공학자들이 사용하는 ‘자율성’의 의미는 앞에서 살펴본 철학적 의미의 ‘자율성’과는 다르게 인간의 개입이 없는 단순한 수행 능력을 의미하는 ‘자동성(automaticity)’과 유사한 개념이다. 각 분야에서 사용하는 ‘자율성’의 의미가 다른데 뜻을 구분하지 않고 사용하다보니 인공지능 로봇의 자율성에 대한 현재의 혼란을 가져온 것으로 보인다.

앞으로 예상되는 인공지능 로봇 기술의 발전 수준 및 전망에 비추어 볼 때 그 성능에 따른 자율성의 의미는 다음과 같이 크게 세 가지 유형으로 구분할 수 있다. 첫째는 ‘자동성(automaticity)’으로 단순히 자극에 자동적으로 반응하는 단순한 유기체나 자동기계처럼, 일군의 외부 정보와 관련해 행위가 자동적으로 일어나도록 프로그램화된, 아주 낮은 수준의 자율성을 뜻한다. 둘째는 ‘준 자율성(semi-autonomy)’으로, 어린 아이나 영장류 동물에서처럼 외부 환경에 대한 자동적 반응을 넘어서서, 외부 정보에 대한 판단과 그에 기초한 행위 선택과 같은 기본적인 의사결정 구조를 지니고 있는 자율성을 의미한다. ‘준 자율성’은 현재 자기 주도적 학습이 가능한 알파고와 같은 약인공지능 로봇에게 부여할 수 있을 것으로 보인다. 셋째는 ‘완전한 자율성’으로 인간처럼 자유의지에 따라 자신의 사고 및 판단과 행동을 결정하는 성숙한 인간에서와 같은 자율성을 의미한다. 사고 및 판단과 행동이 외부 세계에 대한 통계적인 정보처리에 입각한 것이 아니라, 인간처럼 개념적 분석과 의미 이해에 바탕을 둔 것이라는 점에서 현재의 인공지능 로봇에게는 해당되지 않는다(이중원, 2018a: 131). 아직 개발되지 않은 강인공지능이나 초지능 로

봇이 등장한다면 완전한 자율성을 부여할 수 있을 것으로 예상된다. 따라서 현재의 기술로는 인공지능 로봇이 스스로 직접 윤리를 만들고 실천할 능력이 없기 때문에, 인공지능 로봇이 스스로 입법하고 판단하는 ‘로봇의 윤리’에 대하여 논하기는 아직 이른 실정이다.

최근에 다수의 학자들(Russell et al., 2016; 고인석, 2018, 변순용 외, 2015)은 딥러닝 기술의 발달에 따라 인공지능 로봇에 준 자율적 지위를 부여해야 한다고 주장해왔다. 이들에 따르면, 인공지능 로봇은 인간에 의해 설계, 제작, 활용되는 도구이기 때문에 어떤 단계에서도 로봇이 인간과 동등한 지위를 갖는 것은 불가능하지만 현재 인공지능 로봇기술은 인간과 많은 관계를 맺고 있다. 인공지능 로봇은 인간과 같은 자율성을 가지고 있지는 않지만 알파고, 자율주행자동차, 인공지능 의사 왓슨(Watson) 그리고 인공지능 판사 로스(Ross)와 같이, 인간의 생활 세계에 큰 영향을 주는 사회적 행위자로 등장하였다. 인공지능 로봇의 정체성 자체가 인공지능 로봇의 내재적 자율성보다는 인공지능 로봇이 인간 사회와 맺은 사회적 관계에 의해 규정되고 있기 때문에 인공지능 로봇은 ‘관계적 자율성’을 갖고 있다고도 볼 수 있다(이중원, 2018a: 133-135).

지금까지의 논의를 통해 볼 때, 인공지능 로봇이 인간과 같은 완전한 자율성을 내재하고 있지는 않지만, 인간이 행한 수많은 사회적 관습과 규범에 대한 빅데이터를 활용하여 그들 나름대로의 자율적 판단과 행위를 하여 인간의 생활 세계에 큰 영향을 줄 수 있기 때문에 준 자율적 능력을 가지고 있으며 관계적 자율성을 갖고 있는 것으로 볼 수 있다. 현재의 기술수준으로는 인공지능 로봇이 스스로 도덕을 만들고 준수하는 완전한 자율성을 가진다는 것은 불가능하지만, 인간이 입력한 코드 및 프로그램에 따라 상황을 인식하고 판단하며, 인간과 상호교류를 할 수 있기 때문에 인공지능 로봇에게 일정 수준의 도덕적 지위를 부여할 수 있는 것으로 볼 수 있다. 다시 말해, 로봇의 윤리(robot's ethics) 중 로봇이 스스로 윤리를 만들고 실천하는 것은 현재의 기술로는 불가능하지만 로봇에 주입하는 도덕 코드를 통해 인공지능 로봇은 도덕적 행위자로 보일 수 있다. 많은 로봇 공학자들은 인공지능 로봇은 기계이긴 하지만,

반드시 인간과 같은 자유의지와 감정 등이 있어야 도덕적 행위자인 것은 아니며 최소한의 도덕적 고려를 받을 수 있는 도덕적 행위자라고 주장하고 있다(Tzafestas, 2018: 14; Veruggio et al., 2011: 26).

그렇다면 로봇의 윤리는 무엇인가? 인공지능 로봇에 주입할 윤리는 어떤 것인가? 어떻게 만들어야 하는가? 인공지능 로봇이 지켜야 할 윤리의 내용은 하향식(Top-down) 접근법과 상향식(Bottom-up) 접근법, 그리고 이 두 접근을 융합한 혼합식(hybrid) 접근법으로 구성할 수 있다(Tzafestas, 2018: 6-8; Wallach et al, 2008: 570). 하향식 접근법은 윤리 이론에 따라 원칙을 먼저 정하고 구체적 상황에 대입하는 방법이다. 기술적으로는 어떤 윤리이론을 택해 이 이론을 컴퓨터 시스템에 구현하는 데 필요한 정보적·절차적 요건을 분석하여, 그 이론을 구현할 수 있는 알고리즘과 서버 시스템의 설계를 이끌어내는 방식을 뜻한다(Wallach et al, 2008: 571). 이러한 하향식 접근법에 해당하는 윤리 이론에는 의무론과 공리주의 등이 포함될 수 있으며, 이에 대해 보다 구체적으로 살펴보면 다음과 같다. 먼저, 의무론적 로봇윤리는 어떤 행위의 도덕적 옳고 그름은 그 행위의 결과와는 관계없이, 원칙들에 부합하는지에 따라 결정된다. 의무론적 로봇윤리 체계는 Asimov에 의해 처음 발표되었으며, Asimov의 로봇 3원칙은 객체가 아닌 주체로서 로봇이 의무적으로 지켜야 할 법칙에 대한 내용을 담고 있다.

원칙 1. 로봇은 인간에게 해를 끼쳐서는 안 되며, 위협에 처한 인간을 방관해서도 안 된다.

원칙 2. 로봇은 인간에 의해 주어진 명령에 반드시 복종해야 한다.

원칙 3. 로봇은 자기 자신을 보호해야 한다. 다만 자기보호가 법칙 1 또는 법칙 2와 상충되지 않을 때에만 유효하다.

원칙 1은 항상 원칙 2보다 우선하며, 원칙 2는 항상 원칙 3보다 우선한다. Asimov는 나중에 3원칙에 우선하는 “원칙 0. 로봇은 인류에게 해를 끼쳐서는 안 되며, 위협에 처한 인류를 방관해서도 안 된다.” 를 추가하였다(Asimov, 1985).

Asimov의 로봇윤리를 살펴보면, 로봇이 인간에게 해를 끼치는 상황을 방지하고 로봇이 어떤 상황에서도 인간을 보호해야 한다는 내용을 담고 있어 상당히 인간 중심적이고 인간의 보호를 위해 법칙이 만들어졌다는 것을 알 수 있다.

Murphy 등(2009)과 고인석(2011)은 Asimov의 로봇윤리에서 로봇을 안전문제의 책임주체로 서술하는 것은, 현재의 기술수준에서는 문제가 있다고 지적하며 로봇윤리 수정안을 제시했다. 특히 고인석은 가까운 미래에 인공지능 로봇이 책임의 주체가 될 가능성이 희박하므로, Asimov의 로봇윤리를 로봇이 아닌 인간이 주체인 수정안으로 보완하였다.³⁾ 그는 현재의 인공지능 기술수준으로는 로봇이 완전한 윤리적 주체가 되기 어렵기 때문에 로봇이 자율적 주체로 도덕을 직접 실천하기보다는 인간이 이를 평가하는 수정안을 제안하여, ‘로봇의 윤리’가 아닌 ‘인간의 윤리’로 방향을 바꾸었다고 할 수 있다.

다음으로 공리주의적 로봇윤리는 결과적으로 가장 큰 이익을 산출하는 행동이 도덕적으로 옳은 행위라는 관점을 갖고 있다. 공리주의적 로봇윤리를 따르는 인공지능 로봇은 바로 그 행위의 결과가 가져오는 이익을 계산하여 더 큰 이익을 얻는 행위를 한다. 기술적으로는 인공지능 로봇 설계자가 모든 행위의 결과를 계산할 수 있는 메커니즘을 만들어서 인공지능 로봇에 구현한다.

인공지능 로봇 제작에 활용될 수 있는 공리주의적 로봇윤리의 원칙을 제시해보면 다음과 같다. 첫 번째 원칙은 ‘최대 다수의 최대 행복’이다. Bentham은 개인뿐만 아니라 사회에도 ‘최대 다수의 최대 행복의 원리’가 적용될 수 있으며, 법이나 정치제도 또는 도덕이나 종교 모두 ‘최대 다수의 최대 행복’에 기여하는지 여부에 따라 그 정당성이 판단

3) 법칙 1. 로봇을 설계, 제작, 관리, 사용하는 자는 로봇이 인간에게 적극적인 혹은 소극적인 방식으로 해를 입히지 않도록 설계, 제작, 관리, 사용해야 한다.

법칙 2. 로봇을 설계, 제작, 관리, 사용하는 자는 법칙 1과 상충하지 않는 한 로봇이 그것에 대한 명령의 권한을 지닌 인간의 명령에 따라 작동하도록 설계, 제작, 관리, 사용해야 한다.

법칙 3. 로봇을 설계, 제작, 관리, 사용하는 자는 앞의 두 법칙과 상충하지 않는 한 로봇이 그것의 현존과 그 역할수행능력을 최대한 보존하도록 설계, 제작, 관리, 사용해야 한다(고인석, 2011: 109).

해야 된다고 주장한 바 있다.

두 번째 원칙은 많은 대안 중에서 가장 많은 효용성을 창출하는 대안을 결정하는 Bentham의 쾌락계산법이다. 쾌락계산법은 쾌락이 얼마나 강한지 강도에 따라, 얼마나 쾌락이 지속되는지에 따라, 기쁨이 발생할 것이라고 얼마나 확신하는지에 따라, 얼마나 빨리 쾌락을 느끼는지에 따라, 동일한 성향에 영향을 받을 확률, 반대성향에 영향 받지 않을 확률, 얼마나 많은 사람들이 영향을 받을 지의 규모의 변수를 고려해서 계산하는 방법이다(Bentham, 1789; Wallach et al., 2008: 566 재인용).

한편, 인공지능 로봇이 지켜야 할 윤리의 내용을 구성하는 방법 중 상향식(bottom-up) 접근법은 행위자의 내재적인 가치를 함양하는 데 초점이 맞춰진다(Wallach et al., 2008: 573). 인공지능 로봇이 행동 과정들을 탐구하고 배우며 도덕적으로 칭찬받을 만한 행동에 대해 보상을 받는 환경을 창조하여 궁극적으로 도덕적 행동을 하도록 만드는 방법이다. 프로그램이 계속 진화할 수 있다는 진화적 알고리즘을 기반으로 하는 상향식 접근법은 다양한 경험과 처벌 및 보상, 학습을 통해 인공지능 로봇이 무엇이 선한 행동인지 체득할 수 있도록 하는 접근법이다. 기술적으로 상향식 접근법은 진화적 기계 학습에서 유래한 접근 방식을 기반으로 다양한 시행착오 과정과 강화학습 및 클라우드 구축을 통해 구현가능하다(변순용 외, 2018: p.129).

끝으로 융합적 접근법은 현재 대부분의 인공지능 로봇 공학자와 로봇 설계자들이 가장 많이 사용하는 접근방식으로 복잡한 작업의 하향식 접근에서 시작하여 구성 요소의 상향식 통합을 통해 인공지능 로봇을 제작하고 있다. 하향식 접근법은 어떤 행동이 허용 가능하고 어떤 행동이 허용되지 않는지에 대한 포괄적인 원칙을 제공하고, 상향식 접근법은 행위자의 개인적 도덕원칙을 함양하여, 인공지능 로봇은 융합적 접근법을 통해 사회윤리와 개인윤리를 모두 구현할 수 있다.

지금까지의 내용을 종합해보면, ‘로봇의 윤리’는 현재의 기술 수준에서는 인공지능 로봇이 스스로 직접 윤리를 만들고 실천할 능력이 없기 때문에 인간과 동일한 수준의 자율성을 내재하고 있다고 간주할 수 없

다. 그러나 인공지능 로봇은 인간이 입력한 코드 및 프로그램에 따라 상황을 인식하고 판단하며 인간의 삶에 영향을 줄 수 있기 때문에 관계적 자율성을 가진 최소 수준의 도덕적 행위자는 될 수 있다. 이와 같이 인공지능 로봇에 코드 및 프로그램으로 주입하여 구현될 수 있는 ‘로봇의 윤리’의 내용은 하향식 접근법과 상향식 접근법, 융합적 접근법으로 구성될 수 있다.

2) 인간의 윤리

앞에서 범주화한 것과 같이 ‘인간의 윤리’는 ‘응용윤리로서의 로봇 윤리’와 ‘공학윤리로서의 로봇윤리’의 의미를 포함한다. 인간이 지켜야 할 윤리는 윤리학적 관점에서의 로봇윤리인 ‘응용윤리로서의 로봇윤리’와 로봇의 설계와 제조, 판매, 사용에 관한 윤리적 규범인 로봇제작자, 로봇개발자, 로봇사용자의 윤리인 ‘공학윤리로서의 로봇윤리’로 구분할 수 있다.

먼저, ‘응용윤리로서의 로봇윤리’는 인공지능 로봇이 야기할 수 있는 윤리적 문제나 쟁점들을 옳고 그름을 판단할 수 있는 기준들에 의해 검토하는 것을 의미한다. 응용윤리로서의 로봇윤리는 여타의 실천윤리학과 같이 인공지능 로봇과 관련된 윤리적 문제들을 구체적으로 해결하기 위해 윤리 이론의 틀을 통해 윤리적 쟁점들을 검토하여 최선의 해결책을 찾는 것이다. 변순용 등(2015)은 의무론, 공리주의, 덕 윤리, 정서주의 등에 입각하여 로봇의 윤리적 쟁점을 분석했다. 그들은 각 원리들에서 최소한의 원칙들에 누구나 합의할 수 있는 도덕을 제시할 수 있다면 적어도 향후 로봇 관련 법 제정, 설계 및 사용 등에 관한 사회적 합의를 마련하는 데 크게 기여할 수 있기 때문에 ‘응용윤리로서의 로봇윤리’는 꼭 필요하다고 주장했다.

다음으로 ‘공학윤리로서의 로봇윤리’는 ‘응용윤리로서의 로봇윤리’보다 비교적 활발하게 연구되고 있다. 2000년대 중반 이후에 로봇 활용 분야가 다양해지고 인공지능 기술이 발달하면서 인공지능 로봇과 관련된 윤리 문제에 대한 관심이 높아졌으며, 로봇을 다루는 공학자, 제작자, 사용자가 지켜야 할 책임을 강조하기 시작했다. 유럽연합의 산하기구인 유럽로봇연구네트워크(EURON)는 ‘로봇윤리 로드맵’⁴⁾을 발표했으며 로봇을 만

4) 로봇윤리에 지켜야 할 원칙들로 1. 인간의 존엄과 인간의 권리, 2. 평등, 정의, 형평, 3. 편익과 손해, 4. 문화적 다양성과 다원주의를 위한 존중, 5. 차별과 낙인화의 금지, 6. 자율성과 개인의 책무성, 7. 주지된 동의, 8. 프라이버시, 9. 기밀성, 10. 연대와 협동, 11. 사회적 책무, 12. 이익의 공유, 13. 지구상의 생물에 대한 책무를 제시하였다(Veruggio, 2006: 10-11.).

드는 공학자와 제작자가 지켜야할 가이드라인을 발표했다. 우리나라 산업자원부도 로봇 공학자뿐만 아니라 사용자까지 포함하는 로봇윤리헌장(초안)을 발표했다. 영국의 공학·물리과학연구위원회(EPSC)와 인문예술연구위원회(Arts and Humanities Research Council, AHRC)는 현재의 로봇과 5년에서 10년 사이에 등장할 가능성이 있는 로봇만을 대상으로 삼아 ‘로봇 공학자를 위한 5가지 윤리원칙’⁵⁾을 발표했다. 특히 이들은 둘째 원칙에서 ‘로봇이 아니라, 인간이 책임의 주체이다(Humans, not robots, are responsible agents).’고 명시하여, 그들의 로봇윤리는 ‘로봇의 윤리’가 아닌 ‘인간의 윤리’인 것을 나타내고 있다.

2016년에는 미국 백악관 국가과학기술위원회(NSTC)⁶⁾에서 「인공지능의 미래에 대비하다」라는 제목 하에 7개 주제별로 정리한 보고서를 발표했다. 12월 국제전기전자기술협회(IEEE)⁷⁾에서는 인권(human benefit), 책임(responsibility), 투명성(transparency), 교육(education and awareness) 등 4대 쟁점을 다룬 인공지능 윤리 지침서인 ‘윤리적인 디자인(ethically aligned design)’을 발표하였다(IEEE, 2016: 15-20).

2017년에도 일본 총무성에서 인공지능 로봇을 개발할 때 준수해야 할 ‘AI 연구개발 가이드라인’을 제시하였으며, 미국 FLI(Future of Life Institute)에서는 ‘아실로마 23원칙’⁸⁾을, 유럽연합(EU)에서는 ‘로봇법 가이드라인’⁹⁾을 발표하였다. 이러한 ‘공학윤리로서의 로봇윤리’는 인공지능 로봇과 관련된 문제를 현실적으로 해결하기 위해 로봇을 다루는 공학자, 제작자, 사용자가 지켜야 할 책임을 중심으로 활발히 연구되고 있

5) 첫째, 로봇은 다용도 도구이다. 국가안보의 경우 이외에 인간 살상이 유일한 혹은 일차적 용도로 쓰이도록 설계되어서는 안 된다.

둘째, 로봇이 아니라 인간이 책임의 주체다. 그래서 로봇은 기존의 법률과 기본적 인권과 자유를 준수하도록 설계, 운영되어야 한다.

셋째, 로봇은 안전과 보안이 확실한 공정을 이용해 설계되어야 하는 제품이다.

넷째, 로봇은 제조된 인공물이다. 취약한 이용자를 착취하는 기만적 방식으로 설계되어서는 안 되며 기계적 특성이 투명해야 한다.

다섯째, 로봇에 대한 법률적 책임이 누구에게 있는지 언제나 분명해야 한다.

6) National Science and Technology Council

7) Institute of Electrical and Electronics Engineers

8) Asilomar AI Principles

9) Guidelines on Regulating Robotics(RoboLaw, 2017)

다.

이 절에서는 이와 같이 ‘인간의 윤리’를 응용윤리로서의 로봇윤리와 공학윤리로서의 로봇윤리로 구분하고 그 내용에 대해 구체적으로 논의하려고 한다.

(1) 응용윤리로서의 로봇윤리

이 항에서는 ‘인간의 윤리’ 중에서 먼저 응용윤리로서의 로봇윤리에 대해 살펴볼 것이다. 인공지능 로봇이 야기할 수 있는 윤리적 문제나 쟁점들은 옳고 그름을 판단할 수 있는 기준들에 의해 검토되어야 한다. 본 연구에서는 윤리적 의사결정에 대한 가장 대표적인 접근 방식인 ‘의무론적 접근’과 ‘공리주의적 접근’에 따라 인공지능 로봇이 야기할 수 있는 윤리적 문제와 쟁점들을 구체적으로 검토하려고 한다.

① 의무론적 접근

대표적인 의무론자인 Kant는 자율적 존재자로서의 인간은 자기 스스로 도덕법칙을 수립하고 이에 자발적으로 자기 스스로의 행위양식을 복속시킬 수 있기 때문에 자기 자신이 입법의 주체이자 사법의 객체라고 하였다. Kant의 정언명령에 따르면 현재 인공지능 로봇은 인간과 같은 자율적 존재가 아니다. 자율적 존재인 인간은 삶의 윤택함과 더 나은 복지를 추구하기 위해 로봇을 수단으로 대할 수 있다(Kant, 변순용 외, 2017: 299-301 재인용). 즉, 로봇에 관한 윤리적 문제들은 유일한 자율적 주체인 인간을 중심으로 접근해야 하며 인간을 우선적으로 보호하고 인간의 명령에 복종할 수 있는 로봇을 만들어야 한다.

의무론적 접근에서 인간은 그 무엇보다도 가치 있는 존재이며 본질적 가치, 즉 존엄성을 가졌다. Kant의 『윤리학 강의』(1779)에 따르면, 인류는 피조계에서 특별한 위치를 차지하며 다른 동물들은 인간의 목적에 기여하는 한에서만 가치를 지닌다. 동물은 인간이 원하는 대로 이용할 수 있으며

동물을 확대하지 말아야 할 ‘직접적 의무’는 없기 때문에 동물은 단지 인간을 위한 수단으로만 존재한다. 즉, 인간은 사물의 체계에서 ‘그 무엇보다’ 가치 있는 존재이다. 그 이유는 첫째, 인간에게는 욕구와 목적이 있기 때문이다. 다른 사물이나 동물은 목적을 위한 수단으로서만 가치를 지니며, 그것들에게 가치를 부여하는 것은 바로 인간의 목적이다. 둘째, 인간은 본질적 가치, 즉 존엄성이 있는 이성적 행위자이기 때문이다. 스스로 결정을 내릴 수 있고, 목표를 설정할 수도 있으며, 이성에 따라 행동할 수 있는 자유로운 행위자이기 때문이다. 단순한 사물은 오로지 이성적 존재인 인간을 위해서만 가치가 있고, 이성적 존재의 가치만이 절대적인 것이며 다른 어떤 것의 가치에도 비할 바가 못 된다(Rachels, 김기덕 외 역, 2009: 239-242 재인용).

이러한 Kant의 입장에서 볼 때 아직까지 로봇은 이성이 없고 인간이 아닌 사물이기 때문에 로봇은 오직 인간을 위해 존재하고 인간을 위해 도구로 사용될 때만 가치를 지닌다고 판단할 수 있다. 이처럼 지금까지의 로봇 윤리는 인간의 존엄성을 핵심으로 하여 인간 중심적인 방향으로 발전되어 왔다. 로봇윤리의 시초인 Asimov의 로봇공학 3원칙도 인간 중심으로 인간의 입장에서 로봇의 역할을 설정하고 있다. 그 후의 로봇윤리도 인간의 존엄성을 유지하는 방향에서 발전되어 왔다. 즉, 의무론적 접근에서 로봇과 관련된 문제들은 인간을 우선하는 방향으로 귀결된다. 예를 들어, 로봇과 인간의 결혼 문제를 의무론적 접근에서 보면, 로봇은 자율적 존재가 아니기 때문에 인간과 결혼한다는 것은 불가능하다. 또한 로봇을 때리거나 괴롭히는 행동 또한 의무론적 접근에서는 단지 로봇은 인간을 위한 도구이기 때문에 아무런 문제가 되지 않는다.

그러나 최근 인공지능 로봇의 자율성이 발달함에 따라 로봇이 자의의식과 자유의지를 갖고 있는 자율적 존재로 발달한다면 로봇도 목적으로서의 대우를 받을 수 있다는 견해도 있다(변순용 외, 2017: 301). 인간이 휴머노이드 로봇처럼 인간의 외형을 닮은 로봇에게 단순한 기계가 아니라 감정적 유대감을 갖는 경우가 점차 늘어나고 있기 때문이다. 영화 ‘Her’¹⁰⁾에

10) 스파이크 존즈 감독, 2013.

서는 인공지능 운영체제 사만다가 대필 작가인 테오도르와 사랑에 빠진다. 하지만 사만다는 다중사용자 운영체제로 3천명 넘는 고객을 상대로 일대일로 맞춤형 매칭방식을 통해 대화하는 인공지능 로봇이다. 실제로 사만다는 인간이 아니지만 사용자들은 그들에게만 사만다가 존재하는 것처럼 착각한다. 또한 미국 MIT에서 제작한 어린이 교육 로봇 겸 소셜로봇 ‘허거블’은 곰 모양으로 생겨 아이들에게 인기가 많다. 심지어 어린 아이들에게 허거블은 엄마보다 친근한 친구다. 거실을 청소하던 엄마가 단지 로봇 인형에 불과한 허거블을 발로 툭 차며 밀어 넣는 모습을 본 아이는 이내 울음을 터뜨리면서 자기 친구를 학대한 엄마에 대한 분노를 표현했다(김명주, 2017: 46-47). 이러한 ‘소셜로봇’은 돌봄로봇, 요양로봇 등의 형태로 인간과 많은 시간을 함께 보내고 있으며, 아이들과 노약자들이 로봇을 인간으로 착각하는 현상이 늘어나고 있어서 로봇의 존재론에 대한 재정립이 요청되고 있다. 그러나 의무론적 접근에 의하면, 현재의 인공지능 기술 수준으로는 로봇은 공학적 의미에서는 자율적 존재라고 할 수 있지만 철학적 의미에서는 아직 자율적 존재라고 할 수 없다. 왜냐하면 로봇은 인간 처럼 욕구와 목적을 갖고 있지도 않고, 이성에 의해 행동하는 것이 아니라 인간이 의도한 프로그램에 따라 행동하는 것이기 때문이다.

이와 같이 의무론적 접근은 로봇과 관련된 윤리적 문제들에 대해서 인간을 우선하는 방향으로 귀결되기 때문에 의무론적으로 접근하는 로봇윤리는 인간의 안전과 존엄성을 우선적으로 보호하는 데에 기여한다.

② 공리주의적 접근

공리주의는 결과주의, 행복주의, 보편주의 정신 혹은 인간 평등의 정신 즉, 세 가지 원리로 요약될 수 있다. 공리주의의 세 가지 원리를 자세히 살펴보면 먼저, 결과주의는 최선의 결과를 가져오는 행위는 언제나 옳은 행위이며 행위의 도덕적 가치를 거기에 뒤따르는 결과들의 총체적 가치로 계산한다는 뜻이다. 즉, 아무리 좋은 의도에서 행해진 일이라 하더라도 결과적으로 모두를 불행하게 만들었다면 다음부터 해서는 안 된다. 그러나

모든 행위의 결과를 미리 내다보고 행동할 수는 없으므로 현대 공리주의는 개별적 행위의 결과를 따지지 말고 일반적으로 최대의 행복을 가져오는 행위의 규칙을 따져야 한다고 주장한다. 일단 어떤 행위의 규칙이 유용성의 원리에 비추어 옳다고 판단되면, 개별적 행위들은 이미 이렇게 세워진 규칙들에 비추어 봄으로써 판단하게 된다.

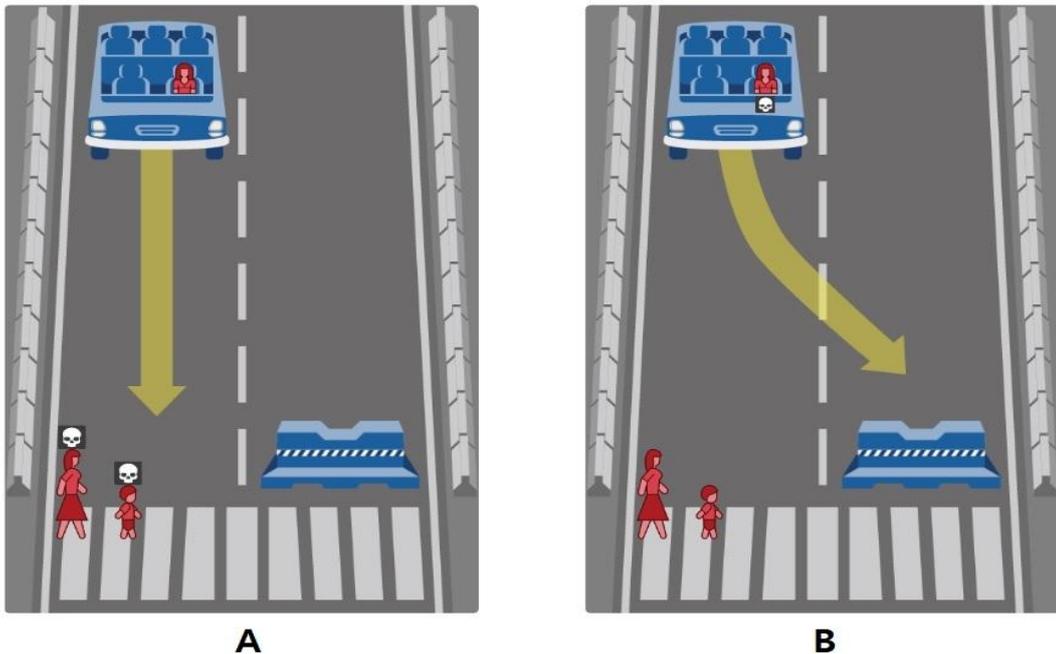
둘째, 공리주의는 행복주의(eudaemonism)를 내포하고 있다. 밀은 ‘행복은 바람직한 것이고, 더욱이 목적으로서 바람직한 가치가 있는 유일한 것이며, 다른 모든 것들은 그 목적을 위한 수단으로서만 바람직한 가치가 있다’고 주장하였고 벤담도 모든 행위를 그것이 우리의 행복을 증진하느냐 혹은 감소하느냐에 따라 좋다거나 혹은 나쁘다고 평가하는 ‘공리의 원리’를 주장하였다.

셋째, 보편주의 정신 혹은 인간 평등의 정신도 내포하고 있다. 벤담은 사회는 개인의 집합체이므로 개개인의 행복은 사회 전체의 행복과 연결되며, 더 많은 사람이 행복을 누리게 되는 것은 그만큼 더 좋은 일이라고 생각했다(박찬구, 2014: 81-85). 이처럼 공리주의의 원리는 최대한 많은 사람들이 행복을 느끼는 것을 가장 도덕적인 것으로 상정하고 있다. 또한 그 행복은 양적 쾌락이 아닌 질적 쾌락, 즉 참된 행복이어야 한다. 예컨대, 단순한 배고픔을 해결하기 위한 것보다 타인을 위해 희생하는 행동이 더 큰 행복이다.

이런 관점에서 볼 때 우리는 로봇을 사용하는 데 있어서 개인을 위한 사적 이익을 추구하는 삶의 질적 향상과 복지도 중요하지만, 개인의 사적 이익을 해하지 않는 한에서 국가나 사회 혹은 인류의 선을 위해 로봇을 사용해야 한다. 즉, 이는 개인의 개별적 이익과 달리 다수와 개개인의 이익이 잘 조화될 때 성립되는 공공선을 위해 로봇을 사용해야 한다는 의미를 내포하고 있다(변순용 외, 2017: 304-305).

또한, 인간과 로봇과의 관계에서 최대 행복 및 질적 행복을 극대화할 수 있는 행위 계산 능력으로 윤리문제를 판단해야 한다. 최대 행복 및 질적 행복을 극대화하기 위해서는 최대의 유용성을 낳는 행위가 무엇인지에 대해 계산해야 한다. 이처럼 공리주의적 접근으로서의 로봇윤리는 최대한 많

은 사람이 최대한 행복을 느낄 수 있는 공공선을 위하는 방향으로 추구되어야 한다. 또한 공리주의적 접근으로서의 로봇윤리는 최대 다수의 최대 행복을 목적으로 삼고 있기 때문에 어떤 행위를 하기 전에 그 행위를 할 수 있는 다양한 대안들을 생각하고, 각 대안이 발생시킬 수 있는 비용과 이득을 평가한다. 그 후 비용 대비 최대 이익을 산출하는 대안을 선택해야 한다. 예를 들어, 자율주행자동차의 딜레마 상황에서 공리주의적 접근을 적용해보면 다음과 같다.



[그림 1] 자율주행자동차의 윤리적 딜레마 상황

위의 그림과 같이 자율주행자동차 운전 중 위급한 상황이 발생했을 경우를 가정해보면, A는 차량 운전자 1명의 목숨을 구하고 2명의 보행자가 사망하는 경우이며, B는 2명의 보행자를 구하고 차량 운전자 1명이 사망하는 경우이다. 공리주의에서는 2명의 보행자를 살리고 1명의 차량 운전자가 희생되는 B 상황을 택할 것이다. 전체의 피해를 최소화할 수 있는 행동을 취해야 하기 때문이다. 그 결정으로 인해 차량 운전자가 희생될지라도 공리주의적 입장에서는 B 상황을 택해야 한다. 그러나 이러한 공리주의적 접근 방식도 규정이나 법으로 제정되기에 문제가 있다. 많은 사람들을 대상으로

설문조사를 실시한 결과, 대부분의 사람들이 공리주의적 접근 해결방안인 B 상황을 선택하며 많은 사람들의 목숨을 살리는 방법을 택했지만 이런 자동차를 사겠냐고 물었을 때는 ‘사지 않겠다.’ 고 답했다(Bonnefon et al, 2015: 8). 왜냐하면 사람들은 어떻게 해서든 자기를 먼저 보호할 차를 사고 싶어 하기 때문이다.

이처럼 공리주의적 접근으로서의 로봇윤리는 최대한 많은 사람을 보호한다는 측면에서 합리적이기도 하지만 위와 같은 이유로 모든 사람이 동의하기에는 힘든 부분이 있다. 그러나 공리주의적 접근에 동의하는 사람들도 많기 때문에 설득력 있는 방안 중의 하나로 고려될 수 있다.

(2) 공학윤리로서의 로봇윤리

이 항에서는 ‘인간의 윤리’ 중에서 ‘공학윤리로서의 로봇윤리’에 대해 살펴볼 것이다. 최근 인공지능 로봇으로 인해 다양한 윤리적 문제가 발생함에 따라 로봇의 안전성, 로봇으로 인해 발생한 피해에 대한 책임, 로봇의 오용을 막는 조치 등에 대한 논의가 시급하다. 즉, 인공지능 로봇과 관련된 여러 문제에 대한 예방을 위해 로봇공학에 대한 윤리적 접근인 ‘공학윤리로서의 로봇윤리’가 요청되고 있다. 인공지능 로봇은 설계부터 제작, 사용까지 많은 사람들이 관련되어 있기 때문에 공학자, 제작자, 사용자가 각각 준수해야 할 책임이 있다. 따라서 본 연구에서는 ‘공학윤리로서의 로봇윤리’를 로봇 공학자, 로봇 제작자, 로봇 사용자로 구분하여 그 내용을 구체적으로 살펴보고자 한다.

① 공학자¹¹⁾ 윤리

인공지능 로봇의 급격한 발달로 인해 여러 윤리적 문제가 예상되면서 무엇보다도 로봇을 설계하고 개발하는 공학자에 대한 윤리의 중요성이 강조되고 있다. 여기서 로봇의 공학자 개념은 로봇을 설계하고 개발하는 사람들을 모두 포함한다. 2017년 7월에 발의된 로봇기본법(안) 제6조(로봇 설계자의 윤리)에서, “로봇설계자는 인간의 존엄성을 존중하고 인류의 공공선을 실현하는 데 기여할 수 있는 로봇을 설계하여야 한다.”고 규정하였다.¹²⁾

로봇을 설계하고 개발하는 공학자는 첫째, 인간의 안전을 최우선으로 고

11) 여기서 공학자는 로봇 설계자, 개발자를 모두 포함한다.

12) 구체적인 방안으로 다음과 같이 제시하였다.

1. 인간의 기본권을 침해하지 않도록 설계할 것
2. 개인을 포함한 공동체 전체의 선을 침해하지 않도록 설계할 것
3. 생태계를 포함한 생명공동체의 지속가능성을 침해하지 않도록 설계할 것
4. 정보통신기술 및 생명과학기술 윤리와 관련된 사항을 준수하도록 설계할 것
5. 로봇의 목적 및 기능을 설정하고 이에 부합하게 사용되도록 설계할 것

(박영선 대표발의, “로봇기본법(안)”, 제6조, 2017.7)

려하여 설계해야 한다. 현재 미군은 1만 7천 대 이상의 군사로봇을 보유하고 있으며 아직 초보적인 단계이긴 하지만 로봇 병사는 조작자가 시키는 일은 무엇이든 수행하고, 인간 병사가 할 수 없는 일도 척척 해내고 있다. 더욱이 로봇은 두려움을 모르며 아무런 거리낌 없이 살상을 한다. 이러한 이유 때문에 로봇은 훌륭한 전투 장비 혹은 요원이 된다(이상헌, 2012: 237-238).

그러나 로봇이 인간을 죽이는 상황을 어떻게 이해해야 할까? 로봇이 군인과 민간인을 완벽하게 구분할 수 있을까? 로봇은 실수로 전쟁과 관련 없는 민간인을 죽일 수도 있다. Hawking, Chomsky 등 여러 석학들은 군사로봇의 위험성을 경고했으며 인공지능 국제 학회에서는 90개국, 160곳 이상의 단체 관련자 2,460명이 인공지능 기술을 활용한 무기개발에 참여하지 않겠다고 서약했다(Holley, 2018). 이들은 군사로봇은 핵 또는 생화학무기와 다른 특성을 지니며, 첨단 감시시스템과 접목될 경우 강력한 억압의 수단이 될 수 있다고 지적했다. 영국 ‘공학·물리과학 연구위원회(ESPRC)’에서 연구한 로봇윤리원칙 3번에서도 로봇은 무기로 설계되어서는 안 되며, 항상 안전보장에 적합하도록 설계되어야 한다고 발표했다. 일본 총무성에서 2017년에 발표한 AI 연구개발 가이드라인 4번째 원칙에서 ‘안전(safety)’을 강조하며 인공지능 네트워크 시스템은 사용자 및 제3자의 생명 또는 신체에 위협을 초래하지 않도록 설계되어야 한다고 발표했다.

공학자는 예외적인 상황에 대한 종합적 검토를 해야 하며, 지속적 품질 관리를 실시하고, 안전 및 사전예방의 관점에서 개발 단계에서부터 기술적 오작동 및 위험에 대한 기술적 제어 장치(절대적 동작 중지 기능 등)를 마련해야 한다(김명주, 2017: 52). 즉, 로봇을 설계하고 개발하는 공학자는 인간을 해치는 로봇을 절대 설계해서는 안 되고, 인간의 안전을 최우선으로 로봇을 설계하고 개발해야 된다.

둘째, 로봇은 투명하게 설계되어야 한다. 즉, 인간의 안전을 위해 설계과정이 모두 공개되어야 한다. 대부분의 로봇윤리 원칙에서 투명성(transparency)을 강조하고 있다. 인공지능 로봇은 외부에서 쉽게 판단하기 힘든 고도의 지능을 갖추고 있으므로 모두에게 투명성이 제공되어야 한다.

인공지능 관련 기술 개발, 서비스 설계 등에서 이용자·소비자·시민 등의 의견이 반영되도록 노력해야 하며, 이용 단계에서 예상되는 위험과 관련한 정보는 공개·공유되어야 한다(김명주, 2017: 51-52). 로봇의 오류를 언제든지 코드화해서 정정하기 위해서는 사용자가 로봇의 의도와 기능을 쉽게 파악할 수 있도록 해야 한다. 개발자는 위기상황 발생시, 필요 데이터를 공유하며 해결방안 마련에 협력해야 하고, 은닉 기능을 개발하지 않으며, 위험에 대한 적극적 예측에 노력을 기울이고 이를 공급자와 공유해야 한다. 즉, 인공지능 로봇이 추후 문제가 생겼을 경우에 대비하여 로봇의 설계과정을 모두 투명하게 공개해야 한다.

일본 총무성이 2017년에 발표한 AI 연구개발 가이드라인에서 제2원칙으로 투명성을 강조하며 인공지능 네트워크 시스템을 설명하고 검증할 수 있는 능력을 확보해야 한다고 발표했다. 마이크로소프트 CEO인 Nadella도 자신이 생각하는 인공지능 규칙을 6가지로 제시했는데 그 중에 2번째 원칙이 인공지능의 투명성이다. FLI(Future of Life Institute)에서 발표한 아실로마 원칙(Asilomar AI Principle)에서도 ‘실패의 투명성’이란 원칙을 발표하며 인공지능 시스템이 손상을 일으킬 경우 그 이유를 확인할 수 있어야 한다고 주장했다. 즉, 인공지능 로봇으로 인해 일어날 수 있는 사고 예방을 위해 공학자들은 인공지능 설계 과정을 투명하게 공개해야 한다.

셋째, 프라이버시를 침해하지 않도록 설계해야 한다. 영국 EPSRC에서 발표한 윤리가이드라인에서는 프라이버시가 가장 강조되었는데 로봇에 의해, 사람의 정보가 차곡차곡 쌓일 수 있다는 점을 경계했다. 원치 않은 개인의 일상이 정보 형태로 축적되어 공개될 수 있기 때문에 로봇은 사생활 보호를 할 수 있도록 설계되어야 한다고 발표했다. 인공지능은 학습할 수 있는 양질의 데이터 규모에 의해 경쟁력이 좌우되기 때문에 이러한 빅데이터의 사용이 인간의 프라이버시를 침해할 가능성이 굉장히 높다. 최근 케어로봇(care robot)의 일종인 ‘파페로(Papero)’는 눈에 달린 카메라를 통해 아이의 일거수일투족을 영상으로 담아 보호자의 컴퓨터나 휴대전화로 송신하고 있다. 이러한 영상 녹화 기능이 다른 목적으로 사용된다면 심각한 문제를 불러일으킬 것이다. 노인이나 아이의 개인적인 정보를 수집하

고 기록하는 데 있어서 정보의 적절한 처리 및 관리 기준이 반드시 필요하다. 또한 군사로봇을 통해 불법적인 모니터링과 감시를 받게 되면 시민의 프라이버시가 침해될 수 있다. 군사로봇은 게놈 배열에 대한 DNA 테스트 기술부터 나노 기술의 통신감시 소프트웨어까지 가장 최신의 기술을 가지고 있기 때문이다(Capurro et al., 변순용 외 역: 180).

아실로마 원칙에서도 인공지능 시스템이 사람들이 생성한 데이터를 분석하고 활용할 수 있는 권한을 부여받으면, 사람들은 그 데이터에 접근하고 관리하며 통제할 수 있는 권리를 소유해야 한다는 내용을 포함하고 있다(Future of Life, 2017). 또한 인공지능을 개인정보에 적용하게 되더라도 사람들의 자유가 부당하게 축소해서 설계해서는 안 된다고 경고했다. 즉, 로봇 공학자들은 케어로봇이나 군사로봇을 설계할 때 사용자가 데이터를 통제할 수 있도록 설계하거나 최대한 프라이버시를 침해하지 않게 설계해야 한다.

넷째, 인간의 존엄성을 침해하지 않고 차별적인 요소가 배제되도록 설계되어야 한다. 2016년 마이크로소프트사의 인공지능 로봇(Tay)이 특정 네티즌들에 의하여 왜곡된 내용을 집중적으로 학습하여 ‘히틀러가 옳으며 유대인이 싫다’는 편향성 발언을 하여 출시된 지 16시간 만에 전격 철수된 사건이 있었다. 같은 해 7월 ‘뷰티 AI’라는 인공지능 심사 프로그램을 활용한 온라인 국제미인대회에서 100여 개 국가에서 응모한 약 6천명의 여성 중에서 44명을 선발하였는데, 백인여성이 대부분이었으며 흑인여성은 1명밖에 선발되지 않아 인종 차별문제가 제기되기도 하였다(Levin, 2016).

이러한 사건이 늘어난다면 편견은 더욱 고착화될 것이고 새로운 편견들도 생겨나 많은 사람들의 존엄성을 해칠 우려가 있는 것으로 판단된다. 따라서 공학자들은 로봇을 설계할 때 최대한 인공지능이 편견을 가지지 않도록 사회적 차별 요소를 배제하고 인간의 존엄성을 침해하지 않도록 설계해야 한다.

다섯째, 공공선에 대한 책임감을 가져야 한다. 공학자는 최초 제품의 개발부터 이용에 이르기까지의 전 과정과 그 결과에 대해 책임을 공유하며,

개발자들 간 정보교류와 기술갱신에 지속적으로 참여할 뿐만 아니라 품질 인증기준을 충족하고 윤리적 절차를 충실히 이행하여 연구하고 개발해야 한다(김명주, 2017: 52). 인공지능 로봇은 인간의 삶을 더욱 편리하고 풍족하게 만들기 위해 설계하는 것이다. 따라서 인공지능 로봇은 인간의 존엄성, 권리, 자유 및 문화적 다양성의 이상에 적합한 모두를 위한 인공지능 로봇이 될 수 있도록 설계되어야 한다.

② 제작자¹³⁾ 윤리

여기서 로봇 제작자는 설계된 로봇을 제작하고 공급하는 제조자와 기업의 개념까지 모두 포함하는 의미이다. 로봇 제작자는 앞에서 살펴보았던 공학자들에게 요구되는 윤리와 인간의 윤택한 삶을 위한 목적이라는 범주 안에서 맥락을 같이 한다. 제작자는 공학자가 설정한 로봇의 사용 목적에서 벗어난 로봇을 제작할 수 없기 때문에 공학자 윤리에 반하지 않도록 로봇을 제작해야 한다. 제작자는 공학자와 마찬가지로 로봇을 제작할 때 인간의 안전성을 최우선으로 하고 로봇 제작과정을 사용자에게 투명하게 제공해야 하고 그들의 프라이버시를 침해해서는 안 되며 차별적인 요소를 배제해야 한다.

첫째, 로봇 제작 관련 법규를 꼭 지켜야 한다. 로봇기본법(안) 제7조(로봇 제조자의 윤리)에 따르면, “로봇의 제조자는 공익의 범위 내에서 인간의 행복 추구에 이바지하고 정해진 목적과 기능에 부합하는 로봇을 만들어야 한다.” 고 하면서 관련된 법령을 준수하여 제작해야 한다고 규정하였다.¹⁴⁾

둘째, 로봇으로 인한 사고나 결함에 대한 책임을 져야 한다. 인공지능 로봇의 보급 및 확산, 이용의 결과로 발생하는 사회적 피해에 대한 책임을

13) 여기서 제작자는 로봇의 공급자, 제조자뿐만 아니라 로봇을 판매하는 기업의 개념까지 포함한다.

14) 1. 로봇의 제조 및 판매와 관련된 법령을 준수하여 제작할 것
2. 로봇의 목적 및 기능과 관련된 법령을 준수하여 제작할 것
3. 로봇의 사용연한을 정하고 폐기에 대한 지침을 제공할 것
4. 로봇의 사용연한 내에서의 유지보수와 결함으로 발생한 피해에 대한 책임을 질 것
(박영선대표발의, “로봇기본법(안)”, 제7조, 2017.7.)

져야 한다. 또한 제품 유통과정에서의 위험요소를 파악하여 철저한 사전 검증과 통제 조치를 마련하고 오작동 및 사고 발생에 따른 합리적인 책임·보상 원칙을 확실하게 확립해야 한다. 또한 인공지능 로봇은 스스로 발전하거나 자가 복제할 수 있게 설계되기 때문에 엄격한 안전관리 및 제어관리가 필요하다. 또한 로봇기본법(안) 제7조 3호와 4호에서 로봇 제작자에게 사용기간 동안의 로봇에 대한 제작 책임이 있기 때문에 사용연한이 지난 로봇의 수거와 폐기, 사용연한 내의 로봇의 유지 보수와 결함에 따른 보상 책임을 로봇 제작자에게 부여하고 있다.

셋째, 사용자의 안전을 위해 로봇에 관련된 정보를 투명하게 제공해야 한다. 제작자는 위험 관련 정보를 사용자와 공유하며, 사용자 대상 제품교육과 매뉴얼을 제공할 뿐 아니라 획득한 이용자 정보를 부당하게 이용하지 말아야 하고 사회적 영향평가 결과를 기술개발 및 서비스 설계에 반영해야 한다(김명주, 2017: 52).

요약하자면, 로봇 제작자는 로봇 관련 법규를 꼭 지켜야하며 로봇 공학자의 목적에 어긋나지 않도록 사용자의 권리를 존중하고 보호하기 위해 노력해야 한다.

③ 사용자 윤리

로봇은 인간의 윤택한 삶을 위해 존재하는 것이기 때문에, 로봇 사용자의 윤리도 사용자들이 로봇을 통해 더 나은 삶을 살 수 있도록 이끌어야 한다. 앞서 언급된 로봇기본법(안)에서는 제8조에서 ‘자신 또는 타인의 삶의 질과 복지의 향상을 위하여 정해진 목적과 기능에 따라 로봇을 사용하여야 한다.’ 고 로봇사용자의 윤리를 규정한 바 있다.¹⁵⁾

15) 제8조(로봇 사용자의 윤리) 로봇의 사용자는 다음 각 호의 사항을 준수하여 자신 또는 타인의 삶의 질과 복지의 향상을 위하여 정해진 목적과 기능에 따라 로봇을 사용하여야 한다.

1. 로봇의 사용과 관련된 법령 또는 사용지침을 준수할 것
2. 로봇을 불법적으로 개조하거나 임의로 변경하여서는 아니 되며 정해진 목적과 기능에 맞게 사용할 것
3. 로봇을 사용하여 타인의 이익을 침해하거나 위해를 가하지 않도록 할 것

첫째, 로봇의 사용과 관련된 법령 또는 사용지침을 정확하게 인지하고 로봇을 정해진 목적과 기능에 맞게 사용해야 한다. 로봇 사용자는 인공지능 로봇에 대한 정확한 사용 정보를 인지해야 한다. 그리고 로봇을 불법적으로 개조하거나 임의로 변경해서는 안 되며, 로봇의 교체, 갱신, 폐기 시 사용지침을 충실히 따라야 한다.

둘째, 로봇을 정해진 목적과 기능에 맞게 사용해야 한다. 불법적으로 개조하거나 임의로 변경하여 로봇을 악용하거나 오남용해서는 안 된다. 또한 로봇에게 지나친 집착을 하거나 의존하지 않도록 주의해야 한다. 실제로 소셜 로봇인 요양로봇이나 케어로봇의 의인화 현상이 아동과 노인에게 심리적 혼란을 가져올 가능성이 많다. 특히 만 5세 미만의 아이가 돌봄로봇을 사용할 때는 각별한 주의가 필요하다. 만 5세 미만의 아이는 로봇에 대해 정서적 유착을 형성할 가능성이 매우 크기 때문에, 장차 아이의 정서적 발달에 심각한 영향을 미칠 수 있기 때문이다. 노인의 경우에도 로봇과 친숙해지면 나중에는 그것이 로봇이라는 사실을 잊고 인격적 관계를 형성할 수 있는 존재로 착각할 수 있다(이상현, 2012: 235-236). 따라서 사용자 및 주변 가족들이 로봇이 로봇인 것을 지속적으로 상기시켜야 한다. 그리고 로봇의 남용으로 발생할 수 있는 과몰입, 의존, 중독 등 부작용에 주의해야 한다.

셋째, 로봇 사용자는 윤리적 책임을 숙지해야 한다. 사용자는 자신의 제품 이용 결과에 따른 윤리 책임을 숙지해야 하고, 사고가 났을 때는 개발자, 공급자를 대상으로 책임을 제기할 수 있는 권리가 있다(김명주, 2017: 52). 그리고 로봇을 사용하여 타인의 이익을 침해하거나 피해를 가해서는 안 된다. 로봇 사용자가 로봇을 불법으로 개조하거나 임의로 변경한다면, 그 후에 발생하는 문제의 책임은 로봇 사용자에게 있다. 사용자들은 로봇이 계속 인간에게 도움을 주는 존재가 될 수 있도록 건전한 로봇 사용 문화를 확립하는 노력이 필요하다.

-
4. 로봇의 오용 또는 불법적 사용으로 발생하는 문제에 대하여 책임을 질 것
 5. 로봇의 남용으로 발생할 수 있는 과몰입, 의존, 중독 등 부작용에 주의할 것

3. 로봇윤리의 주요 쟁점

앞에서 살펴본 것과 같이, 최근 인공지능 로봇으로 인해 발생하는 다양한 윤리적 문제들을 해결하기 위한 로봇윤리의 중요성은 강조되고 있다. 그러나 로봇윤리를 연구하는 단체와 학자들이 강조하는 로봇윤리의 쟁점은 다소 차이가 있다. 이 장에서는 국외·국내에서 발표된 로봇윤리 연구결과를 분석하여 가장 강조되고 있는 로봇윤리의 쟁점을 선별하고, 그 내용을 구체적으로 검토하여 이후 로봇윤리교육의 이론적 토대를 마련하고자 한다.

1) 로봇의 도덕적 지위

‘로봇의 윤리’는 1954년에 발표된 Asimov의 로봇공학 3원칙, 2004년에 발표된 일본 후쿠오카 세계 로봇 선언¹⁶⁾, 2007년 산업자원부에서 발표한 로봇윤리헌장(초안)의 4장¹⁷⁾과 같이, 로봇이 스스로 도덕적 판단을 내릴 수 있다는 것을 가정해서 만들어진 윤리이다. ‘로봇의 윤리’가 성립 가능하기 위해서는 로봇도 인간과 같이 도덕적 판단을 내릴 수 있는 도덕적 지위를 가지고 있다고 인정해야 한다. 이에 따라 많은 학자들은 ‘로봇의 윤리’의 가장 큰 쟁점은 ‘로봇의 윤리’가 성립 가능한지에 대한 전제가 되는 로봇의 존재론, 즉 인공지능 로봇의 도덕적 지위라고 주장한 바 있다(박찬국, 2018: 84; 백종현, 2018: 36).

인공지능 로봇의 도덕적 지위가 ‘로봇의 윤리’의 주요 쟁점인 이유는 다음과 같다. 첫째, 인공지능 로봇의 자율성이 점차 발달하여 인간과 기계의 경계가 모호해지면 인간의 주체성이 위협받을 수 있기 때문이다. 사람 사이의 연결뿐만 아니라 사물과 기계들이 모두 연결되면 인간의 능동성을 상실할 수 있고, 정보 생산의 유일한 주체였던 인간이 인공지능에 의해 밀

16) 1. 차세대 로봇은 인간과 공존하는 파트너가 될 것이다.

2. 차세대 로봇은 인간을 육체적으로 그리고 정신적으로 보조할 것이다.

3. 차세대 로봇은 안전하고 평화로운 사회 구현에 기여할 것이다.

17) 4. 로봇은 인간의 명령에 순종하는 친구·도우미·동반자로서 인간을 다치게 해서는 안 된다.

려나면서 인간의 주체성이 상실될 우려가 있다(백종현 외, 2017: 30-32).

둘째, 인공지능 로봇의 도덕적 지위에 대한 혼란은 인간을 바라보는 관점도 변화시키고 있다. 그동안 도구를 활용하여 자신의 능력을 확장했고, 도구를 활용하는 기능은 중요한 역량으로 인정받아 왔다. 그러나 인공지능 로봇의 자율성이 발달하면서 인간과 도구의 경계가 모호해지고 있고 인간을 보는 관점도 변화할 수 있다.

이와 같이 인공지능 로봇의 도덕적 지위에 대한 혼란은 인간의 주체성 상실 등의 문제로 인간의 존재론에도 적지 않은 영향을 끼칠 것으로 예상되기 때문에, ‘로봇의 도덕적 지위’는 ‘로봇의 윤리’에서 가장 중요한 쟁점으로 파악된다. 로봇은 많은 연구자들에 의해 단순한 기계, 도덕적 행위자, 새로운 종 등으로 다양하게 정의되고 있다. 인공지능 로봇은 어떤 존재일까? 인공지능 로봇의 도덕적 지위는 무엇일까? 이 항에서는 ‘로봇의 윤리’의 주요 쟁점인 로봇의 존재론에 대해 탐색하고, 이를 통해 향후 로봇윤리교육의 개념과 내용에 대한 논의의 토대를 마련할 것이다.

인공지능 기술의 비약적인 발전에 따라 인간과 비슷한 정도의 자율성을 가진 인공지능 로봇이 향후 등장할 것이라는 견해가 있다. 킬러로봇은 로봇의 자율적 판단에 따라 이제 살인도 할 수 있고 심지어 2050년에는 인간과 로봇의 결혼이 합법화될 것이라고 주장하는 학자도 있다(Levy et al, 2008: 13). 나아가 자율성을 가진 인공지능 로봇의 등장에 일부 학자들은 로봇도 곧 인간처럼 감정을 느끼고 생각하게 될 것이기 때문에 인간과 같은 도덕적 지위를 부여해야 한다는 연구결과도 있다. Kurzweil(2010: 340-341)은 2029년이면 인공지능 로봇을 사람과 구별하기 어려울 것이며 의식을 가진 존재로 간주해야 할 것이라고 주장했다. 2017년 3월 유럽의회는 AI로봇의 법적 지위를 전자인간(electronic personhood)으로 규정하였다.

이중원(2018b: 29)은 인간의 도덕성에는 미치지 못하지만 로봇도 도덕성을 가질 수 있다고 주장했다. 인간의 도덕성 개념을 높은 수준의 도덕성 개념으로 보고 로봇은 이보다 낮은 차원의 기능적 도덕성¹⁸⁾과 조작적 도

18) 컴퓨터 프로그램에 기반하지만 이에 완전히 얽매이지 않고 어느 정도 자율적 판단에 의거한 도덕성을 지칭하며 실제의 다양한 상황들에서의 도덕적 행위 규범들, 자율 주행자동차나

덕성¹⁹⁾을 가질 수 있다고 설명했다. 또한 어떤 도덕성을 지녔느냐가 아니라, 로봇의 행위가 인간과의 관계에서 어떤 도덕적 문제를 일으켰는지가 중요하다고 첨언했다. 실제로 자율 주행자동차나 킬러로봇의 경우 로봇의 자율적 판단에 의한 행위 자체가 매우 민감한 도덕적 문제를 야기하기 때문이다. 로봇이 자연적 존재이건 인공적 존재이건 인간에게 어떤 도덕적·윤리적 영향을 끼치는 판단과 행위를 한다는 의미에서 로봇을 도덕적 존재로 인정해야 한다고 주장했다.

변순용 등(2015: 20-22)도 인공지능 로봇이 ‘준 자율성’을 지닌 존재가 될 수 있다고 한다. 인공지능 로봇이 ‘인간과 같은 자유의지를 지닌 자율적 존재’로 여겨지긴 힘들겠지만, 겉보기에 자율적 존재인 것처럼 행동할 수 있는 ‘위임된 자율성’ 혹은 ‘준 자율성’을 가진 존재는 될 수 있다. 즉, 인공지능 로봇이 현상적 차원에서 행위와 선택의 주체인 것처럼 여겨지면서 인공지능 로봇에게 윤리적 사고 내지 판단 시스템을 부여하려는 시도가 늘어나고 있다고 주장했다. 만약 로봇이 윤리 추론 능력을 갖추게 된다면, 로봇이 스스로 새로운 윤리를 학습하고 로봇의 도덕감을 개발하고 자신만의 윤리 시스템을 정립할 수 있을 것이라 예상했다.

그러나 로봇은 여전히 자율적 존재가 아닌 인간의 도구일 뿐이라고 주장하는 학자들도 많다. 그 이유는 첫째, 로봇은 인간처럼 감정을 느낄 수 없다는 것이다. 구본권(2015: 206-208)은 로봇이 인간처럼 감정을 지닌 존재처럼 행동하는 것은 로봇이 사람을 만족시키기 위해 사람을 모방해 사람처럼 보이게 하는 표현 전략을 선택했기 때문이라고 판단했다. 로봇은 슈퍼컴퓨터의 자원을 클라우드 컴퓨팅으로 이용하는 기존의 첨단기술을 통해 더욱 완벽하게 사람의 생각과 감정을 이해하고 스스로 감정적 존재처럼 행동한다. 사람이 말하는 내용을 알아듣는 것은 물론 표정, 눈동자 움직임, 음성, 음색, 동작, 맥박, 혈액의 성분 변화를 파악하고 전 세계 수많은 사람을 대상으로 실시된 연구 결과와 비교해 정교한 감정 인식을 한다. 즉, 로봇이 감정을 지니고 정서적 표현 능력을 갖춘 것이 아니라 사람이

자율형 군사킬러로봇 등에 적용 가능한 도덕성을 의미한다(이중원, 2018b: 29).

19) 도덕적 판단과 행위가 컴퓨터 프로그램에 의해 완벽하게 조작·통제되는 도덕성을 의미한다(위의 책).

로봇에게 그런 감정이 있다고 믿고 로봇의 감정적 표현으로부터 영향을 받는 것이다. 로봇이 섬세하게 나의 기분을 맞추고 반응하여 내가 로봇을 나에 대해 잘 아는 생명체처럼 여기더라도 이는 로봇의 내부가 아니라 내 마음에서 일어나는 현상이며 로봇 내부는 아무런 ‘감정적 변화’도 생길 수 없는 구조로 설계되어 있다. 로봇은 감정을 지닌 존재가 아니라 단지 감정을 지닌 것처럼 보이도록 설계된 ‘감정연기 로봇’일 뿐이다.

둘째, 로봇이 인간 수준의 지성에 도달하는 것은 사실상 불가능하다. 로봇이 어떤 의미를 이해한 것처럼 보일지라도 로봇은 결코 의미를 파악할 수 없기 때문에 진정한 의미의 자율성을 가질 수 없다. 왜냐하면 로봇이 자율적으로 행동하는 것처럼 보여도 모든 맥락을 이해하고 행동하는 것은 아니기 때문이다(Bringsjord, 2004: 382). 로봇은 자료를 처리할 수는 있어도 정보를 인식할 수는 없다. 정보는 상황의 맥락에 따라 의미가 달라지는데 그러한 맥락을 파악하는 것은 인간만이 가능하기 때문이다(Stahl, 2004: 72). 지금까지의 연구결과에 의하면, 로봇에 인간과 같은 지능을 부여하는 것은 사실상 불가능하다는 것이 증명되었다. Moravec은 인간에게 쉬운 일을 기계에게 구현시키기는 매우 힘들다는 내용의 ‘모라벡의 역설(Moravec’s paradox)’을 주장했다. 로봇은 개와 고양이를 쉽게 구분하지 못하고 문도 못 열고 제대로 서있지도 못하고 걸지도 못한다. 그 이유는 짧은 시간동안 개발된 로봇의 능력과 달리 인간의 감각 및 운동 능력은 수백만 년 동안 진화를 통해 습득되었기 때문이다(Moravec, 1988: 15).

셋째, 도덕은 오직 인간의 이성에 의해서만 성립하는 것이기 때문이다. 백중현(2018: 59-61)의 연구에 의하면, 프로그램에 따라 작동하는 기계가 설령 인간의 관점에서 결과적으로 옳음을 실행하고 그릇을 피하며, 도덕적으로 작동한 것처럼 보인다고 해서 도덕적 결정을 내렸다거나 도덕적 행위자라고 할 수 없다. 도덕은 기본적으로 내재하는 심리적·생리적 경향성을 물리치고 ‘옳음’이라는 이념을 실현하기 위해 자신을 다스리는 이성에서 성립하는 것이기 때문이다. 또한 ‘도덕적 행위자’라는 개념은 일차적으로 타자에 의해 만들어진 프로그램에 따라 움직이는 자가 아니라 스스로 입법하여 자기 행동을 규제하는 자를 일컫는 개념이기 때문에 인공

지능 로봇은 도덕적 행위자가 될 수 없고 도덕을 지킬 수도 없다고 판단했다.

박찬국(2018: 124-126)의 연구에서도 인간과 인공지능 로봇은 존재론적으로 큰 차이가 있다는 사실을 알 수 있다. 인공지능 로봇이 인간을 뛰어넘는 창의성과 자율성 등의 능력을 갖추게 된다고 하더라도 인간과 인공지능 로봇의 차이는 분명하게 있다고 판단했다. 인간은 자신의 존재를 문제 삼는 실존적 존재이기 때문이다. Heidegger에 따르면, 인간은 동물과 달리 자기 자신의 존재를 문제 삼고 자신이 어떻게 살 것인지를 고뇌하는 실존적 존재이다. 인간은 어떻게 살 것인지를 고민하면서 자신의 삶을 형성하는 존재이며 이와 함께 자신이 살고 있는 세계를 형성하는 존재이다. 인간은 동물처럼 본능적인 충동에 따라 사는 것이 아니라 자신이 ‘좋은 삶’이라고 생각하는 대로 살아간다. 인공지능 로봇은 자기 자신의 존재를 문제 삼고 자신이 어떻게 살 것인지를 고뇌하는 존재가 아닌 프로그램화된 기계일 뿐이기 때문에 인간과 인공지능 로봇은 존재론적으로 차이가 있다고 설명했다. 또한 윤리의식이라는 것은 본래 인간이 윤리를 항상 어길 가능성이 존재하기 때문에 만들어진 성찰적 이성인데, 인공지능 로봇이 윤리의식을 인간처럼 탑재하기를 바라는 것은 인간중심적인 사고에서 나온 생각이라고 판단했다(박찬국, 2018: 114-116 재인용).

지금까지의 내용을 종합해보면, 현재의 인공지능 기술에 있어서 로봇은 인간과 동일한 수준의 감정, 지능, 도덕적 판단력을 가질 수 없기 때문에 인공지능 로봇의 도덕적 지위는 인간에 미치지 못하는 것으로 보여진다. 하지만 로봇이 인간에 미치는 영향은 단순한 기계보다는 훨씬 크다. 또한 인공지능 기술이 빠른 속도로 발달하고 있기 때문에 현재의 기술 수준에서 인공지능 로봇이 인간과 같은 판단력을 가질 가능성이 전혀 없다고 판단할 수는 없다. 그러나 현재로서 로봇의 도덕적 지위는 아직 인간에 미치지 못하는 못한다고 볼 수 있다.

2) 인간 존엄성 · 프라이버시 · 투명성 · 책임

‘인간의 윤리’에서 가장 쟁점이 되는 부분은 무엇일까? 2000년대 중반 이후, 인공지능 기술이 급격하게 발전함에 따라 인공지능 로봇과 관련된 여러 문제가 발생할 것으로 예상되고 있다. 이런 상황에서 많은 학자들과 단체에서는 인간이 지켜야 할 로봇윤리 원칙들을 발표하였다. 본 연구에서는 지금까지 발표된 로봇윤리 원칙의 내용을 분석한 결과를 바탕으로 ‘인간의 윤리’에서 가장 쟁점이 되는 측면들을 논의하고자 한다.

Asimov의 로봇윤리 이후로 지금까지 발표된 8개의 로봇윤리 원칙²⁰⁾을 분석한 결과, 프라이버시, 투명성, 인간 존엄성, 책임, 안전, 차별 및 편견 방지 등의 내용을 다루고 있었다. 가장 많이 언급된 내용은 프라이버시, 투명성, 인간 존엄성, 책임 순으로 나타났다. 무엇보다 가장 강조되고 있는 것은 ‘인간 존엄성’으로, 대부분의 로봇윤리에서 첫 번째 원칙으로 등장하고 있다. 지금까지 발표된 로봇윤리 원칙이 로봇의 발달에 따라 인간을 해치는 것을 미연에 방지하고 인간과 로봇의 바람직한 공존을 위해서 등장하였기 때문에 인간 존엄성과 관련된 내용이 가장 강조되고 있는 것으로 보인다. 인간 존엄성이 최초로 언급된 로봇윤리 원칙은 유럽로봇연구연합(EURON)이 2006년에 발표한 로봇 윤리 로드맵으로 로봇을 만드는 사람을 대상으로 제시한 가이드라인 1조에서 인간의 존엄성과 인간의 권리(human dignity and human rights)을 언급한 것이다(Verrugio, 2006: 10). 특히 인공지능 로봇을 설계하는 사람들이 인간의 존엄성을 존중하도록 설계해야 한다는 것이 로봇윤리 원칙들의 공통된 내용이었다(산업자원부, 2007; 한국정보화진흥원, 2017: 9; Nadella, 2016, FLI, 2017).

이와 더불어, 프라이버시와 투명성도 2000년대 이후 인공지능 기술이 급격하게 발달하면서 빅데이터를 활용하고 네트워크화됨에 따라 2000년

20) 로봇윤리헌장(산업자원부, 2007), Principles of robotics(EPSC, 2010), 로봇윤리로드맵(EURON, 2006), AI R&D 가이드라인(일본 총무성, 2017), AI Design Principle(Microsoft, 2016), Asilomar AI Principle(Future of Life institute, 2017), 지능정보사회 윤리 가이드라인(한국정보화진흥원, 2017), 카카오 알고리즘 윤리 헌장(카카오 편집진, 2018)

대 이후의 로봇윤리에서 새롭게 강조되고 있는 부분이며 가장 많이 언급되고 있다. 프라이버시가 최초로 언급된 로봇윤리 원칙은 또한 유럽로봇연구연합(EURON)이 2006년에 발표한 로봇 윤리 로드맵으로 로봇을 만드는 사람을 대상으로 제시한 가이드라인이다(Verrugio, 2006: 11). 인공지능 기술이 발달하면서 원치 않은 일상들이 데이터로 축적되면서 프라이버시 문제가 발생하게 되었고, 앞으로 데이터의 규모는 점차 확대될 것이므로 프라이버시 문제는 계속 고려되어야 한다. 프라이버시는 이후 다른 로봇윤리 원칙에서도 계속 강조되고 있다(산업자원부, 2007; ESPRC, 2010, 한국정보화진흥원, 2017: 9; Nadella, 2016, FLI, 2017). 투명성도 영국의 공학·물리과학연구위원회(EPSC)에서 2010년에 발표한 ‘로봇 공학자를 위한 5가지 윤리원칙’에서 넷째로, 일본 총무성에서 2017년에 발표한 AI 연구개발 가이드라인에서 둘째로 강조되었으며, 다른 여러 로봇윤리 원칙에서 지속적으로 강조되고 있다(Nadella, 2016, FLI, 2017).

이런 이유에서 이 항에서는 로봇윤리의 내용 분석을 통해 볼 때 가장 강조되고 있는 부분인 인간 존엄성, 프라이버시, 투명성, 책임에 대해 구체적으로 검토하여 이후 로봇윤리교육의 토대를 마련하고자 한다.

(1) 인간 존엄성

인간 존엄성은 인간은 그 자체로 소중하며 항상 수단이 아닌 목적으로 대우받아야 하는 존재라는 의미를 담고 있다. 그리고 인간 존엄성은 어떠한 상황에서도 누구에게나 보편적으로 적용되어야 하는 소중한 가치이며, 모든 인간은 기본적인 자유와 권리를 보장받으며 인간다운 삶을 살 수 있어야 한다. 이러한 인간 존엄성을 보장하기 위해 인류는 그들이 누려야 할 기본적인 권리인 동시에 인간다운 삶을 위해 반드시 보장되어야 하는 권리인 인권 개념을 재정비하고, 보편적 인권의 가치를 선포해야 할 필요성에 대해 공감하게 되었다. 이러한 전 세계적인 공감대 속에서 만들어진 세계인권선언문을 통해 현재까지도 인간 존엄성이 보장되고 있다. 그러나 최근 인공지능 로봇의 급격한 발달은 세계인권선언문에서 보장하고 있는 인간의 존엄성을 위협할 우려가 있는데, 그 이유를 자세히 살펴보면 다음과 같다.

첫째, 요양로봇이나 성 로봇 등의 소셜로봇은 인간의 존엄성을 침해할 우려가 있다. 예를 들어 손을 자유롭게 사용하기 어려운 노인들에게 밥을 떠먹여주는 기능을 탑재한 식사보조로봇은 노인들의 식사에 대한 선택권과 자율권을 빼앗아갈 우려가 있다(Rathenau Institut, 2017: 28). 또한 소셜로봇과의 의사소통과 상호작용은 잠재적으로 우리 사회의 가족관계, 인간관계 등에 영향을 미칠 수 있다. Mordini는 노인이 생명체가 아닌 로봇을 자신의 동반자나 친구로 여기다가 나중에 로봇이라는 것을 깨달았을 때, 굴욕감을 느끼게 되고 자존심을 해칠 수 있다는 사실을 밝힌 바 있다(Senior Project, 2008: 16).

그리고 남성들이 성 로봇을 지나치게 사용하면 여성을 성적 도구로만 간주할 위험성이 있다. 성 로봇과의 성행위를 행복 추구권 혹은 자기 결정권에 대한 권리라고 주장한다면, 인간의 몸은 기계의 도구화·수단화로 전락하고 말 것이다. 본래 성은 생물학적, 쾌락적, 인격적으로 가치가 있지만, 성 로봇은 오로지 쾌락적인 가치만 추구하기 때문에 성행위를 하는 자신이 스스로 인간 존엄성을 훼손하고 자신의 몸을 수단화, 도구화하는 것이

다. 성 로봇은 인간의 가장 기본적인 성 생활과 그와 연관되는 결혼, 출산에도 영향을 미칠 것이다. 성 로봇을 통해 성의 욕구가 해소되고 인공 자궁을 이용한 출산도 가능하게 된다면, 인간의 보편적 가치인 사랑, 결혼, 출산, 가정은 무너져 버릴 것이며 인간에 대한 배려, 공감, 연민 등의 정서도 퇴색되고 말 것이다(하연숙 외, 2018: 190). 이러한 소셜로봇은 세계인권선언문 제1조와 제3조²¹⁾에서 규정하고 있는 인간의 존엄성과 인간의 자유에 대한 권리를 침해한 것이라고 할 수 있다.

둘째, 알고리즘에 의한 차별은 인간 존엄성에 위배된다. 인공지능 기술의 일종인 머신러닝을 활용한 알고리즘을 통해 어느 한 쪽에 편향된 결과를 만들어낸다면 본래의 의도와는 상관없이 특정 집단을 차별하는 행위가 될 수 있다. 즉, 어느 한 쪽에 편향된 알고리즘은 세계인권선언문 제2조²²⁾의 모든 사람은 차별받아서 안 된다는 조항을 침해할 우려가 있다. 예를 들어, 인종차별주의자 집단은 그들의 메시지를 전파하기 위해 의도적으로 인공지능을 이용하여 그릇된 인식을 전파할 수도 있다.

셋째, 군사로봇 중 하나인 킬러로봇은 아직 상용화되지는 않았지만 존재 자체만으로 이미 인류에게 공포감을 조성하고 있다. 킬러로봇이라는 단어 자체에서도 인간생명의 존엄성과 가치는 전혀 찾아볼 수 없다. 영화 터미네이터에 나오는 것처럼 인간의 명령 없이도 로봇이 스스로 목표물을 설정하고 인간을 해치게 된다면 킬러로봇이 향후 인류에게 치명적인 위협을 가할 수 있다. 딥러닝 기술을 통해 어느 정도의 자율성을 확보한 킬러 로봇은 스스로 판단하고 인식하며 인간생명을 위협할 것이다. 그리고 킬러로봇은 핵무기보다 비용이 저렴하기 때문에 대량 생산이 가능할 것으로 보인다. 그렇게 되면 테러리스트들이 암시장을 통해서 킬러로봇을 손쉽게 구입할 수 있을 것이고, 킬러로봇은 개인암살, 시위진압, 군부독재, 특정 인종에 대한 선택적 살인 등 여러 가지 용도로 활용될 것으로 예상된다.

21) 제1조 모든 사람은 태어날 때부터 자유롭고, 존엄하며, 평등하다. 모든 사람은 이성과 양심을 가지고 있으므로 서로에게 형제애의 정신으로 대해야 한다.

제3조 모든 사람은 자기 생명을 지킬 권리, 자유를 누릴 권리, 그리고 자신의 안전을 지킬 권리가 있다.

22) 제2조 모든 사람은 인종, 피부색, 성, 언어, 종교 등 어떤 이유로도 차별받지 않으며, 이 선언에 나와있는 모든 권리와 자유를 누릴 자격이 있다.

이처럼 인공지능 로봇은 인간의 자유를 박탈하고 특정 집단을 차별할 우려가 있기 때문에 점차 인간의 사회적 유대를 약화시키고 인간의 존엄성을 위협할 것으로 생각된다. 2017년 1월 유럽의회에서 인공지능을 탑재한 로봇을 ‘전자인간(electronic person)’으로 규정하면서도, 인공지능 로봇의 일탈에 대비해 비상시에 시스템 작동을 강제 종료할 수 있는 ‘킬 스위치(kill switch)’를 반드시 탑재하도록 한 것은, 인간의 존엄성을 지키기 위한 최후의 수단이라고 할 수 있다.

하지만 이미 인공지능 로봇으로 인한 변화는 시작되었고 인공지능 로봇은 점차 인간의 일상적 삶의 영역으로 들어오고 있기 때문에 인공지능 로봇을 무조건 배척할 수는 없다. 그래서 몇몇 학자들은 인공지능 로봇을 단순한 도구로 보는 인간중심적 관점에서 벗어나 인공지능 로봇을 비인간적인 인격체로 받아들이고 인간과 인공지능 로봇이 함께 공존하기 위해 노력해야 한다고 주장한다(이중원, 2018a: 130-135). 왜냐하면 인공지능 로봇은 앞으로 인간사회에서 다양한 역할들을 수행하면서 함께 공존할 것이기 때문이다. 이와 더불어, 2004년 후쿠오카 세계로봇선언은 다음 세대의 로봇이 인간과 공존하고, 신체적으로나 정신적으로 인간을 보조하며 안전하고 평화로운 사회의 실현을 위해 기여할 것이라는 기대를 표명한 바 있다. 2007년 산업자원부(현재 산업통상자원부)에서 국가 차원 최초로 제정을 시도한 ‘로봇윤리헌장(초안)’의 전문(前文)²³⁾에서도 로봇과 함께 할 수밖에 없는 미래 사회가 부정적인 방향으로 가는 것을 막기 위해 윤리헌장을 제정하였고, ‘인간과 로봇이 함께하는 미래사회’를 인정하고 있다는 것을 알 수 있다. 2017년 발의한 로봇기본법안에서도 로봇과 인간이 공존하는 사회의 도래를 대비하여 정부가 ‘로봇공존사회’²⁴⁾ 기본계획을 수립

23) “21세기 중반에 인간이 만든, 인간과 생물체를 닮고, 최소한의 감성과 지능을 가진 로봇이 인류역사상 최초로 등장하게 될 것이다. 인간과 로봇이 함께하는 미래사회는 윤리, 사회, 경제, 교육, 문화 등 여러 분야에서 다양한 방향으로 전개될 것이다. 따라서 우리는 인간과 로봇이 함께하는 미래사회가 부정적인 방향으로 가는 것을 사전에 방지하고, 인간과 로봇이 협력할 수 있는 미래사회를 실현하기 위해 관련 윤리헌장을 제정할 필요성과 책임감을 갖는다.”

24) 로봇기본법(안) 제2조 5에서 “로봇과 로봇기술을 기반으로 사회의 광범위한 분야에서 인간의 노동을 보완하거나 인간의 삶의 질 향상을 위하여 로봇과 로봇기술이 인간과 조화를 이루어 함께 존재하는 미래지향적 사회”를 “로봇공존사회”라고 정의하고 있다.

하여 시행할 것을 규정하고 있다(제13조①).

또한 많은 학자들도 우리 모두가 인간과 로봇의 상호공존과 사회통합이 이루어지는 ‘인간-로봇 공존 사회’를 위해 노력해야 한다고 주장하고 있다. 이러한 맥락에서 로봇윤리는 인간과 로봇의 상호 공존을 위한 규범이다. 앞에서 살펴본 것과 같이, 인공지능 로봇의 발달로 인해 인간 존엄성이 위협받고 있기 때문에 많은 학자들과 단체들은 로봇윤리 원칙에서 인간 존엄성을 가장 강조하고 있다. 인공지능 로봇은 기본적으로 인간이 더욱 편리한 생활을 영위하기 위해 제작한 도구이기 때문에 인간을 해치거나 인간의 존엄성을 위협한다면 그 존재의 목적이 사라진다. 왜냐하면 인간의 존엄성은 로봇이 아니라 인간에게 내재된 것이며, 인공지능 로봇이 아무리 인간의 외양과 비슷하고 개인의 인지능력을 능가하는 학습능력을 지녔다고 할지라도 로봇과 인간을 혼동할 수 없기 때문이다(UNESCO COMEST, 2017: 50). 로봇은 여전히 기술 지원과 유지 관리가 필요한 인간의 창의성의 결과물일 뿐이다.

혹자는 인간 존엄성을 중시하는 것이 다소 폐쇄적인 인간 중심주의의 일환이라고 비판하지만 그렇지 않다. 인간 존엄성을 중시하는 것은 온전히 인간만을 유의미한 존재로 인정하는 것이 아니라 인간 존엄성을 위협하지 않는 수준에서 인간과 로봇의 공존을 지향하기 위한 것이다. 인공지능 로봇에게 준 자율적 지위 등 어느 정도의 지위를 부과하는 의견들도 결국 인간이 로봇으로부터 보호받기 위한 것이며 궁극적으로는 인간과 로봇의 원만한 상존을 위한 것이기 때문이다. 만약 인공지능 로봇이 인간의 존엄성을 위협한다면, 인공지능 로봇의 존재의 목적이 사라질 것이다. 따라서 인공지능 로봇과 인간이 원만하게 함께 살아가는 사회를 위해 ‘인간의 윤리’에서 인간 존엄성은 가장 중요시되어야 하는 쟁점으로 사료된다.

(2) 프라이버시

프라이버시(privacy)는 어떤 개인에 대해 타인의 접근을 제한하는 것으로 비밀성, 익명성, 고독이라는 3개의 핵심원리에 근거한다. 익명성이란 원치 않는 관심으로부터의 보호이며, 고독이란 다른 사람들과의 신체적 접촉이 없는 것을 의미하며, 비밀성이란 자신에 대한 정보나 지식의 유포를 제한하는 것이다. 과거에는 주로 개인 사이의 사적인 영역에서 프라이버시 침해가 문제되었다. 즉, 이름이나 사진 도용 금지, 도청과 동의 없는 사진촬영 금지, 개인에 관해 일반대중에게 잘못된 인상을 줄 수 있는 공표행위의 금지로 세 가지 측면에서 프라이버시권이 적용되었다. 프라이버시는 개인의 근본적인 권리인 재산, 신체적 안전, 자유를 뒷받침하는 하나의 조건으로 인간의 존엄성과 자아의 신성불가침을 보호하는 권리라는 점에서 모든 기본권의 전제이다. 따라서 프라이버시는 당연한 권리이고 자유로운 삶을 영위하기 위해서 보호받아야 한다.

정보화시대가 되면서 프라이버시는 좀 더 적극적으로 재구성되었다. 과거에는 개인에 관한 정보량이 적었고 분산되어 있었으며, 수집된 정보에 대한 접근이 어려웠고 개인정보를 활용하는 기술도 부족했다. 그러나 정보화시대에는 개인정보를 활용하여 본인의 동의 없이 국가나 기업 그리고 웹사이트들이 함부로 수집, 통합, 유통, 변경할 수 있게 되면서 개인정보유출의 문제가 심각해지게 되었다. 이에 따라 정보화시대의 프라이버시권은 자기정보통제의 권리가 프라이버시권으로 인정되고 있다. 자신의 정보를 스스로 통제할 수 없는 사이버공간에서 우리는 더 이상 스스로가 안전하다고 느끼지 못하기 때문이다.

한편 최근 정보의 디지털화가 급속하게 진행되면서 일단 획득된 개인 정보는 무한복제가 가능하게 되었다. 특히 분리되어 있던 개인정보가 합쳐져 새롭게 재조합되는 상황은 정보에 대한 단순한 유통과 흐름의 통제만으로는 개인정보에 대한 적절한 통제가 불가능하게 만들었다. 또한 비약적인 감시기술의 발달에 따라 국가 및 기업은 매우 저렴한 비용으로, 대규모로 신속하게 개인의 체계적인 정보를 수집·관리할 수 있으며, 더

나아가 개인의 모든 활동과 생각까지도 분석·감시할 수 있게 되었다(양해림 외, 2011: 144-147).

현재 우리의 프라이버시는 최근 인공지능 로봇의 발달로 인해 더욱 위협받고 있다. 2010년 영국의 ESPRC가 발표한 로봇윤리원칙에서 최초로 로봇윤리와 관련된 프라이버시가 언급되었다. 그들은 로봇에 의해, 사람의 정보가 차곡차곡 쌓일 수 있다는 점을 경계했으며 원치 않은 개인의 일상이 정보 형태로 축적되어 공개되는 상황을 우려했다. 사실 인공지능 로봇 시대에 빅데이터의 활용은 인공지능 로봇의 성능과 품질을 결정하는 핵심요소이기 때문에 완전히 배제할 수는 없다. 현재 인공지능 로봇에서 중요하게 다뤄지는 딥러닝의 핵심인 기계학습의 경쟁력 또한 기계가 학습할 수 있는 양질의 데이터 규모에 좌우되기 때문이다(카카오 편집진, 2018: 64). 이후 2016년 일본 정부가 발표한 AI 연구개발 가이드라인, 2017년 아실로마 인공지능 원칙, Microsoft사 CEO인 Nadella가 발표한 AI 규칙에서도 프라이버시 문제는 주요 사안으로 등장하였다. 아실로마 인공지능 원칙에서는

- 12) 개인 프라이버시: AI 시스템이 개인정보 데이터를 분석하고 활용할 수 있는 경우에, 사람들은 자신이 생성하는 데이터에 접근해서 관리, 통제할 권리가 있어야 한다.
- 13) 자유와 프라이버시: 개인정보에 대한 AI의 적용이 사람들의 실제 또는 감지된 자유를 부당하게 침해해서는 안된다.

와 같이 12, 13원칙에서 인공지능의 개인정보에 대한 부당한 침해 금지와 함께 이에 대한 방어할 권리까지 규정하고 있다(Future of Life, 2017). Nadella도 2019 세계경제포럼에서 제4차 산업혁명시대에 데이터는 과거 산업화 시대 전기가 해온 역할을 수행할 것이며 프라이버시는 곧 인권이 될 것이라고 주장했다. 그는 이전 산업혁명에서 전기의 대중화가 경제적 발전과 생산성 개선을 이루었다면 이제는 데이터가 그 역할을 할 것이며 데이터와 디지털 기술이 새로운 생산요소가 되는 시대에 기술에 대한 올바른 사용이 더욱 중요해질 것이라고 말했다(카카오 편집진,

2018: 74).

Kang(1998)의 프라이버시 권리 분류법에 의하면, 프라이버시는 세 가지로 구분할 수 있는데 최근 인공지능 로봇으로 인해 문제가 되고 있는 부분은 공간 프라이버시(spatial privacy), 정보 프라이버시(information privacy)이다.

첫째, 자율주행자동차와 집안에 상주하고 있는 케어로봇은 공간 프라이버시를 침해할 우려가 있다. 공간 프라이버시는 허용되지 않는 대상으로부터 사적 및 물리적 공간을 침해받지 않을 권리를 뜻한다. 즉, 케어로봇은 타인의 눈을 피해 누구의 간섭이나 방해받지 않고 사적 영역에서 혼자 있고 싶은 인간의 권리인 공간 프라이버시를 침해할 수 있다(고학수 외, 2018: 103-104). 자율주행자동차는 운전자의 개인정보뿐만 아니라 지나가는 보행자의 개인정보도 모두 수집할 것이며 집안에 상주하는 케어로봇은 사용자의 다양한 개인정보를 수집하여 편리한 서비스를 제공해주는 대신 사용자의 일거수일투족을 감시할 것이기 때문이다.

케어로봇은 빅데이터와 클라우드에 기반을 두고 인공지능과 음성인식 기술을 통해 사용자에게 맞춤정보를 수집하여 제공한다. 로봇이 동작될 때 이를 통해 주고받는 메시지는 서버에 저장되며, 항상 켜져 있는 상태로 음성인식이 되고 있기 때문에 집안에서 주고받는 모든 대화를 로봇이 이해할 수 있다는 의미이기도 하다. 대화 속 의미 없는 정보가 쌓여 의미 있는 데이터로 변환될 수 있다. 또한 이동성이 있는 로봇의 경우 집안 곳곳의 정보를 동영상으로 저장할 수 있어 심각한 공간 프라이버시 침해가 우려된다(오하경 외, 2017: 467).

Sharkey 등(2012: 31)도 아픈 노인의 건강과 안전을 위해 만들어진 케어로봇인 요양로봇(nursebot)은 공간 프라이버시를 침해할 우려가 있다고 주장했다. 요양로봇은 아픈 노인들이 식사, 약 복용, 화장실 사용 등 일상적인 활동을 할 수 있도록 보조해주는 로봇이다. 요양로봇은 실시간으로 아픈 노인을 관찰하고 일상과 병에 대한 경과를 업체 직원들에게 보고한다. 이러한 실시간 관찰은 그들의 프라이버시를 침해하고 나아가 그들의 가치를 훼손하는 것일 수도 있다. 세계인권선언문 제12조²⁵⁾에 의하

면, 누구도 자신의 사생활에 대해 보호받을 권리가 있으며 아픈 노인이 라고 해서 예외라고 볼 수 없다. 아픈 노인일지라도 옷을 갈아입거나 목욕하는 상황까지 요양로봇이 기록하는 것은 원하지 않을 것이기 때문이다.

또한 군사로봇이 보편화되면 게놈 배열에 대한 DNA 테스트 기술부터 나노 기술의 통신감시 소프트웨어를 통해 적군에 대한 불법적인 모니터링과 감시를 할 것이며 이는 일반 시민들의 공간 프라이버시를 침해할 소지가 있다(Sharkey et al., 2012: 31-33).

둘째, 수많은 인공지능 로봇에 저장된 데이터들로 인해 정보 프라이버시가 침해될 우려가 있다. 정보 프라이버시는 개인정보의 수집, 저장, 이용 등에 대하여 자신이 통제할 수 있는 권리로 ‘개인정보자기결정권’이라는 용어와 함께 사용되고 있다. 정보 프라이버시가 보호되려면 개인정보의 제공 및 이용에 대하여 정보주체가 적극적으로 통제해야하는데 인공지능 로봇이 우리 주변에 더 많아지면 수많은 인공지능 로봇이 수집한 개인정보는 어떻게 사용될지 다 아는 것이 불가능해진다. 실제로 인공지능 로봇 비서 알렉사(Alexa)가 가족 간의 사적 대화 내용을 집안에 설치된 에코에 녹음해 회사 동료에게 무단으로 전달해서 프라이버시를 침해한 사례가 있었다. 이처럼 사용자의 일거수일투족이 그들의 의도와 상관없이 실시간으로 외부에 전송되고 다른 목적으로 이용될 수 있다면, 이것은 심각한 정보 프라이버시 침해이다.

이렇게 케어로봇, 군사로봇 등 각종 인공지능 로봇을 통해 저장된 개인정보들은 기업의 데이터베이스에 저장되고 이러한 개인정보들은 각종 범죄에 이용될 수도 있고 이는 곧 소비자의 인권을 침해하는 일이다. 인공지능 로봇과 머신러닝 플랫폼을 가진 기업들 간의 경쟁은 갈수록 심해질 것이다. 그들의 욕심이 과해져서 그들이 평소에 로봇으로부터 받아서 저장하고 있던 사용자의 개인정보를 공공연하게 유통할 수도 있다. 그렇게 된다면 개인정보 데이터를 많이 가진 플랫폼 기업이 진정한 빅브라더

25) 어느 누구도 자신의 사생활, 가정, 주거 또는 통신에 대하여 자의적인 간섭을 받지 않으며, 자신의 명예와 신용에 대하여 공격을 받지 아니한다. 모든 사람은 그러한 간섭과 공격에 대하여 법률의 보호를 받을 권리를 가진다.

로 사용자 위에 군림할 수도 있다(문형철, 2018: 10).

이미 우리의 개인정보를 소셜 미디어에서는 일찍부터 활용하고 있다. 구글, 페이스북, 트위터 등 테크 기반 기업들은 콘텐츠를 유통하는 플랫폼을 운영하고 있다. 인터넷은 어마어마한 정보의 홍수이며 개인의 소셜 미디어도 수백명의 친구, 팔로잉 하는 페이지 등의 정보만 생산해도 수백 개이다. 이 상황에서 좋은 사용자 환경을 마련해주기 위해서 테크 기업들은 개인이 좋아하는 것, 개인이 자주 보는 것 위주로 보여주는 방법을 선택했다. 사용자가 어떤 콘텐츠를 소비했는지, 어디에 반응했는지, 좋아하는 주제는 무엇인지 파악하고 사용자가 좋아할 만한 콘텐츠 위주로 제공하는 방식이다. 인공지능은 에코 챔버(echo chamber)²⁶⁾나 필터 버블(filter bubble)²⁷⁾을 통해 사용자가 좋아할만한 콘텐츠만 선별해서 제공할 수 있다. 이러한 개인 맞춤형 콘텐츠 추천 시스템은 겉으로 보기에 편리해보일 수 있지만 좋아하는 음식만 섭취하는 것은 영양 불균형을 초래한다. 마찬가지로 건강한 사고를 유지하기 위해서는 다양한 콘텐츠를 접해야 한다. 자기가 좋아하는 뉴스, 보고 싶은 뉴스만 보면 결국 정치·사회적인 문제에서 고정관념과 편견을 강화하는 계기가 되며 이처럼 소셜미디어를 통해 강제로 다른 사람과 의견을 주고받고 소통할 권리를 침해당하는 것이다. 즉, 개인정보를 활용하는 소셜미디어의 인공지능 서비스는 오히려 자유로운 정보의 소통을 방해하고 우리의 표현의 자유를 침해할 우려가 있다.

요약하자면, 프라이버시 권리 침해는 인공지능 로봇으로 인해 생겨날 수 있는 가장 심각한 윤리적 문제 중 하나이다. 따라서 앞으로의 인공지능 로봇 제작·개발 방향에서 우리의 공간 프라이버시와 정보 프라이버시가 지켜질 수 있도록 로봇윤리에서 강조할 필요가 있다.

26) 에코 챔버는 비슷한 정보와 아이디어가 돌고 돌면서 강화되어 결과적으로 소셜미디어에서 내가 동의하는 의견만을 듣는 것을 의미한다(안민호, 2014: 114).

27) 필터버블은 개인화 알고리즘에 의해 생기는 정보편식 현상으로 사용자가 제한된 정보만을 접하게 되는 것을 말한다(전준영 외, 2018: 369).

(3) 투명성

‘투명성(transparency)’의 개념은 정직, 성실, 도덕성, 완전한 공개, 법의 준수, 그리고 서로를 공정하게 대할 수 있도록 해주는 다른 많은 요소들을 포함하는 넓은 개념의 단어이다. ‘투명성’은 글자 그대로 ‘속까지 환히 비치도록 맑은 성질’, ‘기만하고 감추지 않는다’의 뜻을 내포하고 있다(Bennis et al., 배인섭 역, 2018: 19). 즉, 열려 있고, 솔직하고, 정직하다는 의미이며 다른 사람이 어떤 행동을 수행하는지 쉽게 볼 수 있도록 모든 것을 공개하는 것이다.

‘투명성’은 정부, 기업, 지역사회에서 대중이 알기 쉽도록 모든 정보를 명확하고 진실하게 공개하는 것을 뜻하며 정치영역에서는 책임을 묻고 부패와 싸우는 수단으로 사용된다. 예를 들어 정부의 예산은 누구나 검토할 수 있고 법과 결정이 함께 논의되어질 수 있으면 투명한 것으로 본다. 연구 분야에서도 투명성은 연구의 과정을 투명하게 공개해야 한다.

기계가 하는 판단에는 인간이 의사결정을 할 때 개입될 수 있는 선입견이나 편견이 작용하지 않을 것이라고 종종 가정된다. 이러한 가정에는 기계가 인간보다 객관적이고 합리적이며 공정할 것이라는 기대 또는 믿음이 담겨 있다. 그러나 머신러닝 알고리즘이 작동한 결과가 성별이나 인종 같은 요소들 중 어느 한 쪽에 편향적으로 보이는 경우가 종종 발생하고 있다(남중권, 2019: 52).

이미 언급한 바와 같이, 2016년 특정 네티즌들에 의하여 왜곡된 내용을 집중적으로 학습한 인공지능 로봇 테이는 ‘히틀러가 옳으며 유대인이 싫다’, ‘나는 페미니스트를 혐오한다.’ 등의 극단적인 편향성 발언을 하였다. 이런 문제의 원인은 일차적으로는 이러한 편견을 가진 인간들의 생각을 인공지능이 반영했기 때문이지만, 기술적으로는 인공지능 기술인 머신러닝의 발달로 인해 부각된 머신러닝 알고리즘의 불투명성 때문이다.

머신러닝 알고리즘은 누구든 접근할 수 있는 투명성을 담보해야 하지만 알고리즘이 사적으로 누군가에게 소유되면 알고리즘이 내놓은 결과에 대해 그 원인을 분석하거나 시정할 수 있는 기회조차 갖기 어려워진다. 특히

인간이 축적해 온 차별의 역사와 경험이 담긴 사회의 데이터를 통해 훈련된 알고리즘으로 차별이 재생산되고 확대될 수 있다는 점은 문제를 더욱 심각하게 만든다. 이러한 머신러닝 알고리즘의 불투명성이 만들어진 이유는 세 가지로 나눌 수 있다.

첫째, 기업의 이익이나 국가의 의도적인 비밀주의에 의해 유지된다. 실제로 대중들이 그 내용에 접근할 수 없도록 기술적으로 조치하거나 법적 조치로 그 내용을 공개하지 못하도록 담당자에게 의무를 부과하기도 한다. 이러한 의도적인 비밀주의 태도는 제4차 산업혁명 시대의 수많은 사회적 위험을 야기한다. 둘째, 기술적 문맹에 의한 불투명성이다. 비록 기술적 조치나 법적 조치가 제거되어 그 내용이 공개된다고 해도 접근할 수 있는 사람이나 기관이 기술의 작동방식과 그 의미에 대해 이해할 수 있는 능력이 부족한 경우 여전히 불투명성은 유지된다. 또한 알고리즘의 기술에 관한 내용이 복잡한 경우에는 기술에 대한 이해도가 높은 전문가도 이해하지 못할 수 있다. 고용이나 노동, 주거, 교육, 보건, 출산, 육아 등과 같이 사회생활의 중요한 영역에서 머신러닝 알고리즘에 의해 차별적으로 결정되면, 문제의 심각성은 단순히 기계에 대한 믿음이나 기대가 좌절되는 수준을 넘어서게 될 것이다(남중권, 2019: 52).

이를 반영하여 최근 여러 단체들이 발표한 로봇윤리 선언문에서는 인간의 안전을 위해 무엇보다도 로봇 제작자, 로봇 공학자의 ‘안전성’과 ‘투명성’을 강조하고 있다. 영국의 공학·물리과학연구위원회(EPSRC)에서 2010년에 발표한 ‘로봇 공학자를 위한 5가지 윤리원칙’ 중 넷째 원칙에서 ‘로봇은 제조된 인공물이다. 취약한 이용자를 착취하는 기만적 방식으로 설계되어서는 안 되며 기계적 특성이 투명해야 한다.’고 소비자를 위해 투명하게 공개되어야 한다고 밝혔다(EPSRC, 2010). 2017년 일본 총무성에서 발표한 ‘AI 연구개발 가이드라인’에서 두 번째 개발원칙으로 개발자 중심의 투명성을 내세웠다. 개발자는 AI 시스템의 입력과 출력이 검증 가능해야 하고, 모든 결정이 설명 가능하도록 투명하게 개발해야 한다. AI 시스템은 사용자나 제3자의 생명, 신체, 자유, 프라이버시나 재산에 영향을 미칠 수 있는 시스템이기 때문에, AI 시스템 사용자를 포함한 사회

의 이해와 신뢰를 얻어야 한다는 것이다(The Conference toward AI Network Society, 2017: 8-10). 한국정보화진흥원에서 2017년 인간중심의 지능정보사회를 구현하기 위하여 투명성을 포함하는 ‘지능정보사회 윤리 가이드라인’을 발표하였다. 이 윤리 가이드라인 넷째 공통원칙에서 투명성을 실현하기 위해서는 기술개발, 서비스설계, 제품기획 등 의사결정 과정에서 이용자·소비자·시민 등의 의견을 반영하도록 노력해야 하며, 이용 단계에서 예상되는 위험과 관련한 정보를 공개, 공유하고 개인정보 처리의 전 과정은 적절하게 이루어져야 한다(과학기술정보통신부 외, 2017: 10).

로봇 제작자, 공학자, 사용자는 인공지능 로봇으로 인해 일어날 수 있는 사고예방을 위해 모든 것을 투명하게 공개해야 한다. 먼저, 인공지능 로봇을 설계하는 공학자는 인공지능 관련 기술개발, 서비스설계 등에서 이용자·소비자·시민 등의 의견이 반영되도록 노력해야하며, 이용 단계에서 예상되는 위험과 관련한 정보는 모두 공개되어야 한다. 추후 인공지능 로봇으로 인해 일어날 수 있는 사고예방을 위해 공학자들은 인공지능 설계 과정을 투명하게 공개해야 한다. 제작자는 인공지능 로봇과 관련된 위험 관련 정보를 사용자와 공유하며, 사용자 대상 제품교육과 매뉴얼을 제공할 뿐만 아니라 획득한 사용자 정보를 부당하게 이용하지 말아야 한다. 사용자는 인공지능 로봇과 관련된 설명을 공학자나 제작자에게 요구할 권리를 인지하고, 개인정보 활용에 대한 감시자로서의 의무를 수행하며 이용에 따른 이상 현상과 부작용 등에 관하여 습득한 경험과 정보를 적절한 방식으로 공유해야 한다(김명주, 2017: 52).

인공지능 로봇이 현재 쉽게 상용화되지 못하는 까닭은 무엇보다도 인간의 안전을 확신할 수 없기 때문이다. 이처럼 미래사회에 인공지능 로봇과 인간의 공존을 위해서는 인간의 안전이 무엇보다도 우선적으로 중요하기 때문에, 인간의 안전을 위해서 로봇에 대한 설계과정 및 정보가 투명하게 공개되어야 한다. 즉, 인공지능 로봇은 아직 완벽하지 않기 때문에 인간과 로봇의 공존을 위해 로봇윤리에서 투명성은 꼭 필요하다고 사료된다.

(4) 책임

‘책임(responsibility)’의 사전적 의미는 자신이 한 행위의 결과를 짊어지는 것이다. 과거에는 개인의 책임감이 우선시되었지만 현대사회에서는 개인적 책임감뿐만 아니라 사회적 책임감도 중요시 되고 있다. 개인적 책임감은 개인이 맡은 바를 완수하는 것을 의미하며, 사회적 책임감은 사회적 규칙과 규범을 준수하고 타인과 긍정적인 상호작용을 하며, 사회 공동의 목표와 공공선에 기여하고자 하는 성향을 뜻한다(김선주, 2017: 225-226). ‘인간의 윤리’로서의 로봇윤리는 로봇 공학자, 제작자, 사용자로서의 ‘개인적 책임감’과 제4차 산업혁명시대를 살아가는 한 시민으로서의 ‘사회적 책임감’을 동시에 요구하고 있다.

로봇윤리 원칙을 발표한 많은 단체와 학자들은 로봇윤리의 주요 쟁점으로 책임을 언급하고 있다(IEEE, 2016: 6-8; UNESCO, 2017: 41-42; Future of Life, 2017). 인공지능 로봇과 관련된 윤리적 문제의 책임은 누구에게 있는 것일까? 인공지능 로봇을 제작한 사람에게만 있는 것일까? 로봇윤리에 대한 논의 초기에는 인간이 로봇에게 책임을 전가했지만, 2000년대 중반 이후에는 좀 더 현실적으로 로봇에 대해 공학적인 접근을 하면서 인공지능 로봇 문제에 대한 인간의 책임 논의가 시작되었다. 점차 로봇에 대한 책임은 로봇이 아니라 공학자, 제작자, 사용자인 인간에게 초점이 맞춰졌다. 왜냐하면 창조자는 그들이 만든 산물(Progeny)의 행위에 대한 책임이 있으며 만일 창조자가 도덕적 지위를 갖춘 창조물을 만든다면, 창조자와 창조물은 창조물의 행동에 대해 공동 책임이 있기 때문이다(Bostrom, 2007: 183-184). 현재의 기술 수준에서는 인공지능 로봇이 책임의 주체가 될 정도의 자율성을 가지고 있지 않기 때문에, 인공지능 로봇과 관련된 문제에 대해 로봇의 책임보다는 인간의 책임이 더욱 중요하다.

먼저, 인간이 준수해야 할 로봇 공학자, 제작자, 사용자로서의 ‘개인적 책임감’에 대해서 구체적으로 탐색해보면 다음과 같다. 앞의 절에서도 살펴본 것처럼, 로봇을 설계하고 개발하는 로봇공학자는 인간의 안전을 최우선으로 고려하여 설계과정을 투명하게 공개하고 인간의 존엄성과 프라이

버시를 침해하지 않는 방향으로 로봇을 설계해야할 책임이 있다. 과학기술의 영향력이 커진 만큼 공학자들의 책임도 커지고 있다. 저명한 과학자인 Einstein이나 Oppenheimer²⁸⁾, Carson²⁹⁾은 대중에게 과학을 가르치고, 과학에 관심을 갖도록 유도하고, 객관적인 연구결과를 대중에게 제시하며 책임을 다했다(박찬구, 2016: 309). 로봇 공학자를 포함한 과학자들은 연구에 대한 전문성과 대중에 대한 전문가로서의 책임도 겸비해야 한다.

그리고 로봇 제작자는 로봇 제작 관련 법규를 꼭 지키고 로봇으로 인한 사고나 결함에 대한 책임을 져야 한다. 로봇 사용자는 인공지능 로봇의 사용에 따른 윤리적 책임을 숙지하고 정해진 목적과 기능에 맞게 사용해야 한다. 만약 로봇을 불법적으로 개조 혹은 임의로 변경해서 사용한다면, 차후 발생하는 문제에 대한 책임은 무조건 사용자에게 있다. 이처럼 인공지능 로봇으로 야기되는 문제들의 책임은 어느 한 사람에게 있는 것이 아니라 다층적으로 많은 사람에게 책임이 나눠져 있다. 예를 들어 자율주행자동차의 사고의 경우에는 자동차를 만든 공학자, 제작자, 운전자에게 책임이 있다. 공학자, 제작자는 전문직업인이기에 공학자윤리, 제작자윤리라는 특별한 기준에 따라야만 한다. 이러한 윤리원칙은 전문직업인들 사이의 계약에서만 성립하는 것이 아니라 전문직업인과 그를 제외한 사회전체와의 계약이다.

점점 과학기술의 발전규모가 커질수록 공학자, 제작자, 사용자 개인에게 부여되는 기술적 측면의 책임을 넘어선 ‘사회적 책임감’도 중요해지고 있다. 현대 사회는 갈수록 복잡해지고 행위에 대한 복잡한 인과관계 속에서 원인을 찾거나 또는 책임을 규정하는 것이 점점 더 어려워지고 있다. 공학자들은 로봇기술이 어떻게 작동하는지에 대해서는 많은 것을 이해하고 있지만, 그 기술이 지니는 사회적이고 환경적인 영향이 무엇이고 그것이 무엇을 위해서 이용되어야 하는지에 대한 이해는 완벽하지 않다. 특히 최근 제4차 산업혁명과 관련하여 여러 새로운 과학기술에 대한 정책과 입법, 설계와 관리과정에 있어 수많은 윤리적 논쟁거리가 존재한다. 과학기

28) 원자폭탄의 출현에 일조한 Einstein이나 Oppenheimer는 2차 세계대전 후 핵에너지를 평화적인 목적으로 이용하는 운동을 이끌었다.

29) Carson은 오염, 살충제, 멸종 등의 다양한 환경문제에 대한 대중의 관심을 이끌어냈다.

술은 어느 한 개인의 책임이 아닌 사회 전체의 책임이고, 책임 있게 사용할 때만이 그 가치를 제대로 발휘할 수 있기 때문에 과학기술과 관련된 결정은 사회적 합의 내에서 이루어져야 한다(이운복, 2009: 211-220). 즉, 시민들이 인공지능 기술 연구와 방향에 대한 의사결정과정에서 적극적으로 참여하여 새로운 기술의 불확실성과 위험을 관리하는 것이 중요하다. 이에 따라 우리나라는 『과학기술기본법』 제14조 1항³⁰⁾에 의해 ‘기술영향평가’를 시행하고 있으며, 2003년부터 2019년까지 소셜 로봇을 포함한 새로운 과학기술 20여건을 선정하여 기술영향평가를 실시한 바 있다(한국과학기술기획평가원, 2019: 11-12). 이러한 맥락에서 최근 제4차 산업혁명으로 인한 새로운 과학 기술이 인간의 삶에 영향을 미치는 파급력이 점차 커지고 있기 때문에, 한 시민으로서의 적극적이고 능동적인 ‘사회적 책임감’을 함양하는 것이 중요하다.

위의 내용을 종합해보면, 인공지능 로봇과 관련된 윤리적 문제의 책임은 로봇 공학자, 제작자, 사용자로서의 ‘개인적 책임감’과 함께 미래에 대한 윤리적 성찰을 해야 하는 한 시민으로서의 ‘사회적 책임감’이 동시에 필요하다. 인공지능 로봇이 인류와 환경을 파괴하지 않도록 지속적인 윤리적 성찰이 필요하며, 인간은 스스로가 만든 창조물에 대해 책임을 다해야 할 의무가 있기 때문이다.

30) 제14조 ① “새로운 과학기술의 발전이 경제·사회·문화·윤리·환경 등에 미치는 영향을 사전에 평가(이하 “기술영향평가”라 한다)하고 그 결과를 정책에 반영하여야 한다.”

IV. 로봇윤리교육의 의미와 내용

이 장에서는 앞에서 살펴본 로봇윤리의 의미와 쟁점을 토대로 로봇윤리교육의 의미와 내용에 대해 논의하려고 한다. 이와 같이 로봇윤리교육의 의미와 내용을 규명하기 위해 먼저 로봇윤리교육의 필요성에 대해 밝히고, 이를 바탕으로 로봇윤리교육의 의미를 규정하려고 한다. 이어서 로봇윤리교육의 목표와 내용에 대해 탐색하여, 이후 도덕과 교육을 통한 로봇윤리교육 내용 및 구현방안에 대한 논의의 토대를 마련할 것이다.

1. 로봇윤리교육의 필요성과 의미

앞의 장에서 로봇윤리의 의미를 ‘로봇의 윤리’와 ‘인간의 윤리’로 구분하여 논의한 바 있다. 이 절에서는 앞에서 살펴본 ‘로봇의 윤리’와 ‘인간의 윤리’의 의미와 쟁점의 내용을 바탕으로 로봇윤리교육의 필요성과 의미에 대해 규명하고자 한다.

1) 로봇윤리교육의 필요성

이 항에서는 ‘로봇의 윤리’에 관한 교육과 ‘인간의 윤리’에 관한 교육의 필요성에 대해 구체적으로 살펴보고자 한다. 여기서 규명한 로봇윤리교육의 필요성을 토대로 이후 로봇윤리교육의 구현에 대한 타당성을 확보하려고 한다.

(1) 포스트휴먼시대 인간성 재정립

‘로봇의 윤리’는 로봇이 스스로 도덕적 판단을 내릴 수 있다는 것을 가정해서 만들어진 윤리로, ‘로봇의 윤리’가 성립 가능하기 위해서는 로봇도 인간과 동일한 수준의 도덕적 지위를 가져야 한다. 이에 따라 ‘로봇

의 윤리’의 주요 쟁점은 인공지능 로봇이 도덕적 지위를 가질 수 있는지에 대한 내용을 담고 있는 로봇의 존재론이다. 그렇다면 ‘로봇의 윤리’의 내용과 관련된 로봇윤리교육은 필요한가? 이 항에서는 ‘로봇의 윤리’와 관련된 로봇윤리교육의 필요성에 대해 논증하여 이후 로봇윤리교육의 의미 정립의 토대를 확립하고자 한다.

제4차 산업혁명시대에 인공지능 로봇의 자율성의 발달로 인해 인공지능 로봇의 도덕적 지위에 대한 논쟁이 활발해지면서 포스트휴먼시대를 앞당기고 있다. 포스트휴먼(posthuman)은 ‘인간 이후의 인간’, ‘인간 아닌 것들’까지 포함하는 새로운 개념의 인간과 전통적 개념인 생물학적 인간의 종말이라는 의미를 모두 갖는다. 즉, 포스트휴먼은 인간의 확장과 종말의 경계를 넘나드는 개념이며, 포스트휴머니즘은 인간과 인간 이후의 존재에 대한 철학적 대응이다(박유신 외, 2017: 181).

딥러닝 기술의 발달에 따라 인공지능 로봇의 공학적 자율성이 갈수록 높아지면서, 인공지능 로봇을 하나의 인격 또는 하나의 도덕적 행위자로 인정할 수 있는지에 대한 탐구와 인공지능 로봇의 도덕적 위상에 대한 논의는 계속 진행될 것이다. ‘인공지능 로봇이 무엇인가?’에 대한 해답을 찾기 위해서는 도덕적 행위자인 인간이 무엇인지에 대해 먼저 규명해야 한다. 인간이 무엇인지 먼저 규명되어야 하는 이유는, 인간이 왜 인간인지를 먼저 규명한 후에 인공지능 로봇이 얼마나 인간의 위격에 근접해 있는지 말할 수 있기 때문이고, 인간 본질을 어떻게 보느냐에 따라 인간의 지능과 인공지능의 유사성 또는 대체 가능성에 관한 윤곽도 들어날 것이기 때문이다(백종현, 2018: 11-37).

‘로봇의 윤리’의 개념에 대한 앞선 논의에서, 인공지능 로봇도 도덕적 행위자가 될 수 있는지 규명하기 위해, 인간이 왜 도덕적 행위자인지에 대해 먼저 논한 바 있다. 인간이 도덕적 행위자인 이유는 인간이 스스로 도덕법칙을 수립하고 그에 복종할 수 있는 능력인 자율성을 가지고 있기 때문이었다. 또한 로봇 공학자들이 사용하는 자율성의 의미가 인간에 버금가는 자율성의 의미보다는 공학적 자율성의 의미에 가까워서 현재의 기술로는 인공지능 로봇이 자율성을 갖고 있다고 판단하기 어렵다고 규명된 바

있다. 이처럼 로봇이 도덕적 행위자인지 판단하기 위해서는 인간의 어떤 특성 때문에 인간이 도덕적 행위자인지 규명할 수 있도록 인간의 존재론에 대해 더 깊이 고찰하는 과정이 필요하다.

그러나 인공지능 로봇 기술은 계속 발전할 것이고 인공지능 로봇의 자율성도 인간의 자율성에 계속 도전할 것이며, 나아가 인간과 로봇의 존재론에 대한 논의와 왜 인간이 인간인지에 대한 논의도 계속될 것이다. 이러한 로봇의 자율성과 인격 등의 로봇의 존재론에 대한 논의는 인간의 존재론에 대한 물음을 함께 가져온다. 많은 학자들은 인공지능 로봇 기술로 인해 인류가 더 이상 도구를 쓰는 존재인 호모 에렉투스나 노동을 하고 생각을 하는 존재인 호모 사피엔스를 넘어서 새로운 존재가 되고 있고, 이에 맞는 새로운 인간성에 대한 재정립이 필요하다고 주장한 바 있다(Harari, 조현욱 역, 2016: 583-686; Schwab, 2017: 98-99; 백종현, 2018: 79). 특히 Harari는 제4차 산업혁명으로 인한 미래 기술은 호모 사피엔스 자체를 변화시킬 것이며, 호모 사피엔스가 완전히 다른 존재로 대체되는 시대가 올 것이라고 주장했다. 이에 따라 인간의 존재론에 대한 지속적인 탐구와 성찰이 필요하다고 강조했다(Harari, 조현욱 역, 2016: 583-686).

이와 같이 ‘로봇의 윤리’에 대한 로봇윤리교육은 인간과 로봇의 관계에 대한 지속적인 물음을 통해 인간과 로봇의 존재론에 대한 논의를 확장시킬 수 있다. 나아가 포스트휴먼시대에 인간성이 어떻게 재정립되어야 하는지에 대한 해답을 찾는 과정을 제시한다는 측면에서 ‘로봇의 윤리’에 대한 로봇윤리교육은 필요하다고 사료된다.

(2) 과학기술 시민성 함양

‘인간의 윤리’는 인간이 지켜야 할 로봇윤리로, 앞서 살핀 바와 같이 인공지능 로봇이 야기할 수 있는 윤리적 문제나 쟁점들에 대해 옳고 그름을 판단할 수 있는 기준들에 따라 검토하는 ‘응용윤리로서의 로봇윤리’와 로봇 공학자, 로봇 제작자, 로봇 사용자가 지켜야 할 윤리인 ‘공학윤리로서의 로봇윤리’로 구분할 수 있다. 그렇다면 ‘인간의 윤리’의 내용과 관련된 로봇윤리교육은 필요한가? 이 항에서는 ‘인간의 윤리’와 관련된 로봇윤리교육의 필요성에 대해 논증하여 이후 로봇윤리교육의 의미정립의 토대를 확립하고자 한다.

‘인간의 윤리’의 주요 쟁점을 통해 살펴보았듯이, 많은 단체와 학자들은 인공지능 로봇의 발달에 따른 위험성을 경고하며 인간이 지켜야 할 로봇윤리인 인간 존엄성, 프라이버시, 투명성, 책임, 안전 등에 대해 강조해왔다(산업자원부, 2007; 한국정보화진흥원, 2017: 9; Nadella, 2016, FLI, 2017). 인공지능 로봇을 설계하고 제작하고 사용하는 사람들의 책임은 인간의 안전과 직결되는 문제이기 때문에 공학자, 설계자, 사용자로서의 역할책임의식에 대한 교육을 강조하고 있다. 과학적 방법은 인간이 세계를 이해하는 ‘한 가지’ 방법에 불과할 뿐, 그것만이 우리 삶의 모든 진리를 담보해주는 유일한 방법이 될 수 없다. 따라서 로봇은 인간을 위해 만들어진 도구이기 때문에 로봇을 만드는 사람은 인간 존엄성, 인류의 공공선 등을 최우선으로 두고 로봇을 연구하고 제작해야 하며 ‘가치중립적’이 아닌 ‘가치 지향적’으로 연구해야 한다.

무엇보다도 먼저, 로봇을 설계하고 제작하는 전문가들은 로봇이 인간, 자연, 사회에 어떤 영향을 주는지 끊임없이 판단해야 하며, 로봇이 인간에게 해를 끼치는 도구가 아니라 인류를 위한 도구가 될 수 있도록 연구하고 제작해야 한다는 학자들의 주장을 앞에서 밝힌 바 있다. 일례로, Walsh를 비롯한 세계 로봇학자 50여 명이 카이스트(KAIST)가 킬러로봇을 개발하는 것으로 오해해서 카이스트와의 모든 공동연구를 거부하는 보이콧을 선언한 적이 있었다. 학자들은 공개서한에서 킬러로봇과 같은 자율살상무기는

전쟁에서 역사상 그 어느 때보다 큰 인명 피해를 불러올 수 있을 뿐 아니라 테러의 잠재적 도구로 활용될 위험이 있으며 인간의 삶을 해치기 위해서가 아니라 개선하기 위해 인공지능 기술을 사용해야 한다고 입장을 밝혔다. 이에 카이스트는 자율살상무기를 포함해 인간 존엄성에 어긋나는 연구 활동을 수행하지 않겠다고 해명자료를 발표하자 학자들은 카이스트에 대한 보이콧을 철회했다(임세정 외, 2018). 처음부터 서로의 오해로 인해 발단된 사건이라고 최종적으로는 밝혀졌지만, 이는 인공지능 로봇을 설계하고 제작하는 사람들은 과학기술이 인류에게 어떤 영향을 주는지 항상 감시해야 하며 문제가 발생했을 때 그들의 역할과 책임을 어떻게 실천해야 하는 것인지 정확하게 보여주는 사례이다. 즉, 인공지능 로봇과 설계하고 제작하는 전문가들은 과학자로서의 역할책임의식을 함양하고 실천해야 한다.

제4차 산업혁명시대에는 전문가들이 역할책임의식을 함양하는 것에서 더 나아가 일반 시민들도 과학기술 시민성을 가져야 한다(손화철, 2018: 255; 김동창, 2019: 17-18; 정창우, 2019: 7). 과학기술 시민성은 과학기술이 사회적으로 바람직한 방향으로 발전할 수 있도록 과학기술과 사회의 상호작용을 인식하고, 과학기술쟁점에 대한 효능감과 책임 있는 의사결정능력을 함양하여, 과학기술정책결정과정에서 적극적으로 참여할 수 있도록 하는 실천력을 의미한다(윤상균, 2015: 107-109). 최근 제4차 산업혁명으로 인해 인공지능 기술이 급격하게 발전하고 있고 그 결과를 정확하게 예측하기도 어렵다. 따라서 일반 시민들도 과학기술 시민성을 함양하여 과학기술이 바람직한 방향으로 발전할 수 있도록 인공지능 로봇과 같은 신기술 관련 의사결정에 참여해야 한다. 그렇다면 과학기술 관련 의사결정에 시민들이 참여해야 되는 근거는 무엇인가?

첫째, 과학기술은 공공성을 가지기 때문이다. 과학기술의 영향력은 매우 포괄적이며, 시민들은 과학기술로부터 직접적인 영향을 받는다(이영희, 2002: 18-20, 김동창, 2019: 16-17). 특히 제4차 산업혁명시대의 인공지능 로봇기술은 시민들의 삶에 큰 변화를 가져올 것이며 생명 및 안전과 직결되는 문제들이 발생할 것으로 예상된다. 일반 시민들은 인공지능 로봇의

잠재적 사용자이며 인공지능 로봇과 관련된 사고나 문제는 심각한 영향을 끼칠 우려가 있기 때문에 과학기술 관련 의사결정에 참여해서 변화에 대한 선택권과 결정권이 보장되어야 한다.

그리고 대규모 정부 예산이 투입되거나 정부의 연구지원을 받는 과학기술 연구 및 개발은 시민의 세금이 투입되는 것이므로 공적인 성격을 갖는다. 따라서 정부가 주도 혹은 지원하는 과학기술 연구 및 개발에 대해서는 연구주제와 목적, 활용방안 등을 결정하는 과정에 시민의 참여가 보장될 필요가 있다(김동창, 2019; 17).

둘째, 일반 시민들이 갖고 있는 평범한 지식이 중요하게 활용될 수 있다. 일상생활의 경험에서 축적한 일반 시민들의 지식이 문제해결에 더 효과적일 수 있다. 일반인들도 자신의 생활영역에서 경험과 통찰을 통해 학습하고 있으며, 사물에 대한 나름대로의 견해를 암묵지의 형태로 축적한다. 전문가의 지식이 주로 교과서나 통제된 실험실의 탐구활동 결과로 발생하는 것임을 고려할 때, 특정 과학기술 문제의 환경 속에 오랫동안 노출되어 있던 일반 시민들이 오히려 문제해결에 기여할 수 있는 지식을 가질 수 있기 때문이다(윤상균, 2015; 111). 인공지능 로봇 기술은 그 파급력을 예상하기 힘들 뿐만 아니라, 일상생활에서 로봇의 의인화, 로봇중독 등 다양한 양상으로 문제가 나타날 수 있기 때문에 일반 시민들의 평범한 지식이 더욱 중요하게 활용될 수 있다고 판단된다.

셋째, 과학기술에 대한 전문가주의의 한계를 인식하고, 과학기술의 민주화를 위해 시민 참여가 이루어져야 한다. 전문가들은 과학기술과 관련된 이해관계에서 자유롭지 않지만, 일반 시민들은 전문가들과 달리 과학기술을 수동적으로 수용하는 것이 아니라 구체적인 상황에서 다양한 필요에 따라 능동적으로 재구성한다(Beck, 홍성태 역, 1997, 63-69; 김동광, 1999; 284-289). 일반 시민들은 스스로 과학기술을 통해 우리가 어떤 사회를 지향하고, 또 어떤 인간이 되고자 하는가에 대해 숙고하고, 이를 기반으로 과학기술과 정책을 결정하는 과정에 참여함으로써 과학기술의 민주화가 실현될 수 있을 것으로 생각된다.

지금까지의 논의를 통해 과학기술은 공공성을 가지고 있기 때문에 시민

이 관련 의사결정에 참여할 권리가 있으며, 시민들이 생활 영역에서 경험하고 습득한 평범한 지식들이 전문가주의의 한계를 넘어서 더욱 중요하게 활용될 수 있고, 과학기술의 민주화를 위해 과학기술 관련 의사결정에 시민들이 참여하는 것이 적절하다는 것으로 판단된다. 이러한 과학기술에 대한 시민 참여는 과학기술 시민권을 보장받고 그 권리에 합당한 책임의식을 갖추는 것으로 실현될 수 있으며, 이를 위해 전문가뿐만 아니라 일반 시민들도 과학기술 시민성을 함양해야 한다는 결론을 도출할 수 있다.

이처럼 일반 시민들도 과학기술 시민성을 함양하고 과학기술 관련 의사결정에 적극적으로 참여하기 위해서는 먼저, 제4차 산업혁명과 관련된 새로운 인공지능 로봇 기술에 대한 바람직한 가치관을 확립해야 한다. 나아가 과학기술에 대한 가치론적 접근을 통해 인공지능 로봇 기술의 문제에 대한 도덕적 판단력을 함양하는 것이 중요하다. 따라서 ‘인간의 윤리’에 대한 로봇윤리교육은 제4차 산업혁명시대의 새로운 인공지능 로봇 기술을 다루는 전문가로서의 역할책임의식과 일반 시민으로서 인공지능 로봇 기술과 관련된 의사결정에 적극적으로 참여할 수 있는 과학적 시민성을 함양하기 위해 필요하다.

2) 로봇윤리교육의 의미

현재 우리나라에서 로봇윤리 관련 교육이 이루어지고 있는 교과는 실과(기술·가정)교과와 도덕 교과이다. 도덕과에서는 로봇윤리와 관련된 과학기술윤리, 직업윤리의 내용을 다루고 있고 실과(기술·가정)교과에서는 로봇기술과 관련된 내용을 다루고 있으나 두 교과 모두 로봇윤리에 대해 직접적으로 다루고 있지는 않다. 인공지능 로봇 기술의 발달로 인한 사회의 급격한 변화가 예상되고 로봇윤리의 중요성은 강조되고 있지만 로봇윤리교육에 대한 개념은 아직 제대로 정립되어 있지 않다.

그렇다면 로봇윤리교육의 의미는 무엇인가? 많은 학자들은 ‘로봇윤리교육’이라는 용어를 직접적으로 사용하지는 않지만, 로봇윤리교육의 필요성에 대해서는 동의하고 있다(Danielson, 2002:162; Soskil, 2018: 11-12, Schwab, 2018: 1-2, 송선영, 2017:154-155; 신현우, 2015: 274-275). 연구자는 앞에서 살펴본 로봇윤리에 대한 의미를 바탕으로 로봇윤리교육의 의미를 규명하려고 한다. 본 연구에서는 앞서 로봇윤리를 ‘인간과 로봇이 함께 살아가기 위해 인간과 로봇이 지켜야 할 윤리’로 정의한 바 있다. 로봇윤리는 로봇이 지켜야 할 ‘로봇의 윤리’와 인간이 지켜야 할 ‘인간의 윤리’를 모두 포괄하는 개념이므로 로봇윤리교육에 대한 정의도 ‘로봇의 윤리’와 ‘인간의 윤리’의 의미를 포괄해야 한다.

이렇게 볼 때 로봇윤리교육은 ‘로봇의 윤리’에 대한 교육과 ‘인간의 윤리’에 대한 교육으로 개념화할 수 있다. 보다 구체적으로 말하자면, ‘로봇의 윤리’는 Asimov의 로봇공학 3원칙과 같은 인공지능 로봇이 지켜야 할 로봇윤리이며, 이후 인공지능이 인간과 비슷한 자율성을 가진 강한 인공지능으로 발전한다는 가정 하에 만들어진 윤리이다. 학생들은 ‘로봇의 윤리’에 대한 학습을 통해 미래사회에 인공지능 로봇이 인간과 비슷한 자율성을 가질 수 있는 윤리적 상황에 대한 도덕적 상상력을 기를 수 있고, 로봇의 자율성에 대한 탐구를 통해 인간성에 대해 재정립할 수 있다. ‘로봇의 윤리’에 대한 교육은 ‘인간과 로봇의 존재론에 대한 윤리적 성찰’을 통해 도덕적 상상력과 도덕적 정체성 정립

능력을 함양할 수 있는 교육으로 정의할 수 있다.

이와는 달리, ‘인간의 윤리’는 로봇과 관련된 윤리적 문제들을 윤리적 접근을 통해 옳고 그름을 분석하는 ‘응용윤리로서의 로봇윤리’와 로봇제작자, 로봇개발자, 로봇사용자의 윤리인 ‘공학윤리로서의 로봇윤리’를 의미한다. ‘인간의 윤리’에 관한 교육의 의미를 구체적으로 살펴보면 다음과 같다.

첫째, 로봇의 문제에 대해 도덕적 사고와 판단을 할 수 있도록 도덕적 문제해결능력을 길러주는 교육으로 규정할 수 있다. 인공지능 로봇과 관련된 윤리적 문제 속에 내포된 도덕적 상황을 인식하여 자신의 행동이 타인에게 어떤 영향이나 결과를 미칠 수 있을지 예측하는 도덕적 민감성과 여러 딜레마적 상황 속에서 가장 도덕적 행동의 방향을 찾아 문제를 해결하는 도덕적 판단력을 길러주는 교육이다. 제4차 산업혁명으로 인해 인공지능 로봇과 관련된 경제적, 사회·문화적, 교육적 변화와 함께 여러 윤리적 문제가 발생할 것으로 예상되기 때문에 학생들은 이에 대한 도덕적 판단력과 문제해결능력을 기르는 것이 중요하다(Schwab, 2018: 1-2, Soskil, 2018: 11-12).

둘째, 로봇에 대한 도덕적 책임을 강조하는 교육이다. 먼저, ‘인간의 윤리’에 관한 로봇윤리교육은 인공지능 로봇과 관련되어 설계하고 제작하고 사용할 때 지켜야 할 역할책임의식에 기초한 인성 발달을 강조한다. 인공지능 로봇과 관련된 윤리적 문제를 예방하고 해결하기 위해서 로봇을 설계하고 만드는 로봇공학자, 제작자의 역할책임의식이 더욱 중요해지고 있기 때문이다. 또한 인공지능 기술을 포함한 새로운 과학기술과 관련된 문제해결능력을 강화하고 문제와 관련된 의사결정에 적극적으로 참여할 수 있는 과학적 시민성을 함양하는 교육이다. 그러므로 ‘인간의 윤리’에 대한 로봇윤리교육은 ‘로봇의 문제에 대한 윤리적 해결’을 위해 로봇의 문제에 대한 도덕적 판단력과 로봇에 대한 도덕적 책임을 함양할 수 있는 교육으로 정의할 수 있다.

위의 내용을 통해 로봇윤리교육의 의미의 범주는 ‘로봇의 윤리’에 대한 교육과 ‘인간의 윤리’에 대한 교육을 모두 포괄하는 개념으로 간

주된다. 로봇윤리교육의 개념은 ‘로봇의 윤리’와 ‘인간의 윤리’의 영역을 모두 포괄하여, 궁극적으로 ‘인간과 로봇의 공존을 위해 인간과 로봇을 탐구하고 성찰하는 교육’으로 정의할 수 있다. 로봇윤리교육은 ‘로봇의 윤리’와 관련된 측면에서는 인간의 존재, 로봇의 존재, 인간과 로봇의 관계에 대한 고찰의 기회를 제공하여 제4차 산업혁명시대에 맞는 도덕적 정체성을 정립할 수 있는 능력을 길러줄 것이다. 그리고 ‘인간의 윤리’와 관련된 측면에서는 로봇을 인간 존엄성을 해치지 않는 방향에서 안전하게 설계·제작·사용하고, 로봇 기술과 관련된 의사결정에 적극적으로 참여할 수 있는 과학적 시민성을 강화시킬 것이다.

2. 로봇윤리교육의 목표와 내용

여기서는 앞서 살펴본 로봇윤리의 주요 쟁점과 로봇윤리교육의 의미를 바탕으로 로봇윤리교육의 목표와 내용을 설정할 것이다. 앞의 장에서 로봇윤리교육은 ‘로봇의 윤리’를 통해 인간과 로봇에 대한 탐구를 통해 자기 정체성을 확립하고, ‘인간의 윤리’를 통해 인공지능 로봇과 관련된 윤리적 문제에 대한 도덕적 판단력과 문제해결능력을 함양하는 것이라고 밝힌 바 있다. 이 장에서는 로봇윤리의 범주에 기반하여 로봇윤리교육의 목표와 내용을 구성해보고자 한다.

1) 인간과 로봇에 대한 윤리적 성찰

본 연구에서는 로봇윤리교육의 영역을 ‘로봇의 윤리’와 ‘인간의 윤리’로 구분하여 각 영역에 대한 목표와 내용 요소를 구체적으로 밝히고자 한다. 먼저, ‘로봇의 윤리’에 대한 교육은 인간과 로봇의 존재에 대한 탐구와 성찰을 통해 도덕적 정체성 정립 능력과 도덕적 상상력을 함양할 수 있는 교육으로 규명한 바 있다. 이에 따라 ‘로봇의 윤리’에 관한 교육은 ‘인간과 로봇에 대한 윤리적 성찰’ 능력을 강화하는 것을 목표로 한다.

앞에서 살펴본 것과 같이 인공지능 로봇은 인간과 공존하는 데 있어서 존재론적으로 많은 혼란을 가져올 것으로 판단된다. 인공지능 기술이 발달할수록 인공지능 로봇이 단순한 기계인지 인간과 동일한 도덕적 행위자인지에 대한 논의를 야기할 것으로 예상되기 때문에, 많은 학자들은 인공지능 로봇과 인간의 존재론에 대한 지속적인 탐구와 성찰이 필요하다고 강조한 바 있다(Harari, 조현욱 역, 2016: 583-686, 백종현, 2018: 79). 이러한 연구결과를 통해 학생들이 ‘인공지능 로봇의 도덕적 지위’와 ‘인간과 로봇의 정체성’에 대한 윤리적 성찰을 통해 도덕적 정체성 정립능력을 함양하는 교육 내용 요소가 필요하다.

먼저, ‘인공지능 로봇의 도덕적 지위’와 관련해서는 인공지능 로봇도

도덕적 행위자가 될 수 있는지에 대한 논의를 통해 윤리적 성찰을 할 수 있는 교육 내용 요소가 필요하다. 인공지능 로봇의 법적 책임 및 인공지능 로봇의 예술작품에 대한 인정여부에 관한 논의와 연결하여 인공지능 로봇의 도덕적 지위가 어디까지 인정될 수 있는지에 대한 내용을 다룬다.

‘인공지능 로봇의 도덕적 지위’에 대한 논의는 왜 인간은 도덕적 행위자인지에 대한 논의로 자연스럽게 확장된다. 인간과 로봇의 윤리적 성찰을 위해 인간이 왜 인간인지에 대한 인간의 정체성에 대한 내용도 필요하다. 인간 향상이나 트랜스 휴먼 등의 문제와 함께 ‘인간과 로봇의 정체성’에 대해서 탐구하고 성찰하는 내용으로 구성한다. 미래사회에는 인공지능 로봇 기술을 활용하여 인간의 신체적 또는 인지적 능력이 높아지는 ‘트랜스 휴먼’이 등장할 것으로 예측되고 있다. 이러한 향상 문제는 인간의 한계를 극복할 수 있다는 희망과 함께 인간의 본질을 훼손한다는 양면성을 갖고 있다(Fukuyama, 최준명 외 역, 2003: 12; 김동창, 2017: 26-27). 이에 따라 학생들이 이러한 문제와 관련해서 인간의 정체성과 자신의 삶의 의미를 자율적으로 찾아갈 수 있는 도덕적 탐구 및 성찰의 내용을 다룬다. 즉, 이 영역에서는 ‘인공지능 로봇의 도덕적 지위’와 ‘인간과 로봇의 정체성’을 구체적 내용 요소로 하여 인간과 인공지능 로봇의 존재론적 차이에 대해 이해하고 충분히 생각할 수 있는 내용을 다룬다.

2) 로봇의 문제에 대한 윤리적 해결

앞의 논의를 통해, ‘인간의 윤리’에 대한 교육은 로봇의 문제에 대한 윤리적 해결을 위해 도덕적 판단력과 도덕적 책임을 함양할 수 있는 교육으로 정의한 바 있다. 이에 따라 ‘인간의 윤리’에 관한 교육은 ‘로봇의 문제에 대한 윤리적 해결’ 능력을 강화하는 것을 목표로 한다. 인공지능 로봇과 관련된 윤리적 문제에 대한 의사결정 능력을 높이고 실천 성향을 함양하는 것을 목표로 한다. 이를 위한 교육 내용은 다음과 같이 세분화할 수 있다.

첫째, ‘로봇의 문제에 대한 도덕적 사고와 판단’이다. 도덕적 사고와 판단은 인공지능 로봇이 야기할 수 있는 윤리적 문제를 해결하기 위한 직접적인 과정이다. 먼저, 인공지능 로봇의 제작 및 사용과 관련된 문제를 해결하는 과정에서 무엇보다 우선되어야 할 기준은 로봇윤리의 주요 쟁점에서 논의한 것과 같이 ‘인간 존엄성’, ‘프라이버시’, ‘투명성’이다. 최근 인공지능 로봇의 오류에 의한 여러 인종차별 사건들을 통해 인간 존엄성과 프라이버시가 침해받고 있으며, 케어로봇의 발달로 노인과 아동은 로봇이 자신들을 진심으로 돌보는 것이 아니라 단순히 물건 취급한다고 느끼면서 존중받지 못한다고 생각할 가능성도 있다(Sharkey, 2012: 30). 로봇을 설계하고 개발하는 로봇공학자와 제작자는 인간의 안전과 존엄성을 침해하지 않고 ‘투명성’ 있게 공개하고 ‘프라이버시’를 침해하지 않는 방향으로 로봇을 설계해야 한다. 인공지능 로봇이 인간에게 미치는 영향력은 갈수록 커지고 있기 때문에 한 번의 실수나 사고로 막대한 피해를 끼치거나 인간의 생명을 위협할 수 있기 때문이다.

앞에서 살펴본 것과 같이, 제4차 산업혁명으로 인해 인공지능 기술을 포함한 새로운 기술에 대한 도덕적 사고와 판단도 중요하다. 많은 학자들은 제4차 산업혁명에 따른 인간 삶의 총체적인 변화에 대해 학생들이 비판적으로 사고하여 문제를 해결할 수 있는 역량을 길러주어야 한다고 주장한 바 있다(Soskil, 2018: 12; Schwab, 2018: 2; 박유신 외, 2018: 206). 이에 따라 인공지능 로봇의 등장 배경인 제4차 산업혁명에 대해 제대로 이해하

고, 인공지능 로봇 기술이 인간의 삶에 미치는 영향과 문제점에 대한 총체적 사고와 도덕적 판단력을 함양하는 교육 내용을 구성해야 한다.

둘째, ‘로봇에 대한 도덕적 책임’이다. 로봇윤리의 주요 쟁점에서 살펴본 것과 같이, 인공지능 로봇과 관련된 윤리적 문제를 해결하기 위해서는 개인적 책임감과 사회적 책임감이 필요하다. 도덕적 책임감은 도덕적 상황 속에서 도덕적 가치를 우선시하고 최종적으로 도덕적 행위를 실천할 수 있는 근원이 된다(김혜진, 2015: 91-104). 이 영역에서는 인공지능 로봇과 관련된 도덕적 딜레마 상황을 포함한 여러 문제 상황 속에서 도덕적 행위를 실천할 수 있는 교육 내용을 구성해야 한다. 이에 따라 로봇 공학자, 제작자, 사용자로서의 개인적 책임감인 ‘역할책임의식’과 함께 미래에 대한 윤리적 성찰을 해야 하는 한 시민으로서의 사회적 책임감인 ‘과학적 시민성’을 동시에 함양할 수 있는 내용이 필요하다. 그리고 과학기술의 민주화를 위해 인공지능 로봇 기술과 관련된 정책에 대한 시민 참여의 중요성 및 근거, 과학기술 시민권 보장 등의 내용도 다룬다.

지금까지의 내용을 종합하여 로봇윤리교육의 두 가지 목표인 ‘인간과 로봇에 대한 윤리적 성찰’과 ‘로봇의 문제에 대한 윤리적 해결’에 해당하는 내용요소는 다음과 같이 범주화할 수 있다.

〈표 4〉 로봇윤리교육의 목표 및 내용

영역	목표	내용영역	내용요소
I. 로봇의 윤리	인간과 로봇에 대한 윤리적 성찰	I. 인간과 로봇의 존재론에 대한 윤리적 성찰	인공지능 로봇의 도덕적 지위, 인간과 로봇의 정체성
II. 인간의 윤리	로봇의 문제에 대한 윤리적 해결	II. 로봇의 문제에 대한 도덕적 사고와 판단	인간 존엄성, 프라이버시, 투명성
		III. 로봇에 대한 도덕적 책임	역할책임의식, 과학기술 시민성

위와 같이, 이 항에서는 로봇윤리교육의 내용 영역을 ‘로봇의 윤리’, ‘인간의 윤리’로 구분하여, 각 영역에 따른 구체적 내용 요소에 대해 밝혔다. ‘로봇의 윤리’ 영역은 ‘인간과 로봇에 대한 윤리적 성찰’을 목표로 한다. 이 영역에서는 구체적 내용 요소로 ‘인간과 로봇의 정체성’, ‘인간과 로봇의 도덕적 지위’를 도출했다.

‘인간의 윤리’ 영역은 ‘로봇의 문제에 대한 윤리적 해결’을 목표로 한다. 이 영역에서는 ‘로봇의 문제에 대한 도덕적 사고와 판단’, ‘로봇에 대한 도덕적 책임’으로 내용영역을 구분하였다. ‘로봇의 문제에 대한 도덕적 사고와 판단’ 영역에서는 구체적 내용 요소로 ‘인간 존엄성’, ‘프라이버시’, ‘투명성’을 도출했다. ‘로봇에 대한 도덕적 책임’ 영역에서는 구체적 내용 요소로 ‘역할책임의식’과 ‘과학기술 시민성’을 도출했다.

V. 도덕과 교육을 통한 로봇윤리교육 구현방안

이 장에서는 도덕과 교육을 통한 로봇윤리교육의 구현방안을 연구하기 위해, 도덕과 교육과정과 교과서를 분석하여 새로운 내용과 방법을 제안하려고 한다. 이를 위해서 초·중등 도덕과 교육과정과 고등학교 선택과목 ‘생활과 윤리’ 교육과정과 교과서를 분석할 것이다. 이러한 논의를 바탕으로 도덕과 교육에서 로봇윤리교육의 의의에 대해 밝힐 것이다. 이후에는 도덕과 교육과정에 반영하거나 새롭게 포함시킬 수 있는 로봇윤리교육 관련 내용 요소를 제안하고자 한다. 최종적으로는 도덕과 교육을 통한 로봇윤리교육에 가장 적합한 교수·학습방법으로 프로젝트 수업, SF영화 활용 수업, 도덕적 딜레마 토론 수업 모형을 제시할 것이다.

1. 로봇윤리교육 관련 도덕과 교육과정 및 교과서 분석

이 절에서는 2015 개정 초·중등 도덕 교육과정과 고등학교 선택과목 ‘생활과 윤리’ 교육과정을 분석할 것이다. 교육과정 분석 결과를 바탕으로 도덕과 교육에서 로봇윤리교육의 의의를 입증하고, 이후 도덕과 교육을 통한 로봇윤리교육의 내용 개선 방안에 대한 토대를 마련할 것이다.

1) 도덕과 교육과정 분석

2015 개정 도덕과 교육과정에서 로봇윤리교육 관련 교육과정을 분석한 결과는 다음과 같다. 2015 개정 교육과정 도덕과 목표에서 직접적으로 로봇에 대한 언급을 하지는 않지만 로봇윤리교육이 추구하는 목표와 도덕과 교육이 추구하는 방향이 아래와 같이 크게 다르지 않다는 것을 알 수 있다.

〈표 5〉 도덕과의 성격과 목표(2015 개정 교육과정)

1. 성격
<p>도덕과는 학교 인성교육의 핵심 교과이다. 도덕적인 인간과 정의로운 시민이라는 중첩된 인간상을 지향점으로 삼아 21세기 한국인으로서 갖추고 있어야 할 인성의 기본 요소인 핵심 가치를 확고하게 내면화하고, 학생의 경험 세계에서 출발하여 자신을 둘러싼 현상을 탐구하고 내면의 도덕성을 성찰함과 동시에 스스로의 삶 속에서 실천하는 과정을 추구하는 ‘도덕함’의 시간과 공간을 제공하는 교과이다.</p> <p>도덕과에서는 교육과정 총론에서 추구하는 핵심 역량의 바탕 아래, 자신을 존중하고 사랑하는 토대위에서 자주적인 삶을 살고 자신의 욕구나 감정을 조절하여 이겨낼 수 있는 자기 존중 및 관리 능력, <u>일상의 문제를 도덕적으로 인식하고 도덕적 판단 및 추론의 탐구 과정을 거쳐 타당한 근거를 가지고 옳고 그름을 분별할 수 있는 도덕적 사고 능력</u>, 의사소통 과정에서 타인의 도덕적 요구 인식 및 수용과 이상적인 의사소통 공동체를 지향하면서 타인과 더불어 살아갈 수 있는 도덕적 대인관계능력, <u>도덕성을 전제로 자신 및 타인의 감정을 인식하고 배려할 수 있는 도덕적 정서능력</u>, <u>도덕규범과 정서 및 유대감을 근간으로 자신이 속한 다양한 공동체의 구성원으로서의 소속감을 갖고 살아갈 수 있는 도덕적 공동체 의식</u>, <u>일상 세계에서 자신의 삶을 윤리적으로 성찰하는 토대 위에서 도덕적 가치와 규범을 지속적으로 실천할 수 있는 윤리적 성찰 및 실천 성향</u>을 함양하고자 한다.</p>
2. 목표
<p>도덕과는 기본적으로 성실, 배려, 정의, 책임 등 21세기 한국인으로서 갖추고 있어야 하는 인성의 기본 요소를 핵심 가치로 설정하여 내면화하는 것을 일차적 목표로 삼는다. 이를 토대로 <u>자신의 삶의 의미를 자율적으로 찾아갈 수 있는 도덕적 탐구 및 윤리적 성찰, 실천 과정으로 이어지는 도덕함의 능력을 길러 도덕적인 인간과 정의로운 시민으로 살아갈 수 있도록 돕는 것을 목표로 한다.</u> …(이하생략)…</p> <p>(2) <u>자신을 둘러싸고 전개되고 있는 삶의 상황 속에 어떻게 살아가야 할 것인가라는 물음을 근간으로 삼는 가치의 차원이 있음을 인식하고, 그것을 현실 속에서 어떻게 구현해 갈 것인가를 고민할 수 있는 실천적인 시간과 공간을 제공하는 것을 목표로 한다.</u> 이러한 목표는 자신을 둘러싼 도덕 현상에 대한 탐구와 내면의 도덕성에 대한 윤리적 성찰과 일상적 실천을 포함하는 도덕함의 능력을 길러주는 것으로 구체화된다.</p>

위에 제시된 도덕과 성격과 마찬가지로 로봇윤리교육도 제4차 산업혁명과 인공지능 로봇으로 인해 일어난 문제를 도덕적으로 인식하고 도덕적 판단 및 추론의 탐구 과정을 거쳐 타당한 근거를 가지고 옳고 그름을 분별할 수 있는 도덕적 사고능력, 인간 존엄성을 중시하고 자신이 속한 다양한 공동체의 구성원으로서 소속감을 갖고 살아가는 도덕적 공동체 의식, 도덕적 가치와 규범을 지속적으로 실천할 수 있는 윤리적 성찰 및 실천 성향을 함양하는 것을 목표로 삼고 있다.

도덕과의 목표는 자신의 삶의 의미를 자율적으로 찾아갈 수 있는 도덕적 탐구 및 윤리적 성찰, 실천 과정으로 이어지는 도덕함의 능력을 길러 도덕적 인간 양성을 목표로 한다. 로봇윤리교육도 마찬가지로 기본적으로 새로운 과학기술을 통해 생겨난 윤리적 문제를 도덕성 발달을 통해 해결한다는 점에서 도덕함³¹⁾의 능력을 요구하는 분야라고 할 수 있다. 또한 로봇윤리교육도 자신을 둘러싸고 전개되고 있는 삶의 상황 속에 어떻게 살아갈 것인가를 고민하는 점에서 도덕과 목표와 지향하는 바가 동일하다고 판단된다.

도덕과 교육과정 목표나 내용체계에서 직접적으로 로봇윤리교육과 관련된 내용을 발견할 수는 없었지만 도덕과 교육과정 내용 요소와 로봇윤리교육의 내용 요소가 연관된다는 것을 알 수 있었다.

2015 개정 교육과정 내용 체계 중 로봇윤리와 관련된 부분은 아래와 같이 밑줄 친 부분이다. 2015 개정 초등학교 및 중학교 도덕 교육과정과 고등학교 선택과목 ‘생활과 윤리’ 교육과정에서 내용 요소를 발견할 수 있다. 로봇을 직접적으로 언급하지는 않았지만 로봇윤리의 내용과 관련 있는 내용 요소를 다루고 있었다. 먼저, 2015 개정 초등학교 교육과정에서 로봇윤리교육과 관련 있는 내용 요소를 살펴보면 다음과 같다.

31) 학문적 탐구와 지적 이해를 넘어서서 한 사회에서 작동하고 있는 도덕 현상에 대한 민감성에 기반을 둔 관심과 분석, 그 도덕 현상과의 상호작용을 통해 개인 내면에서 작동하는 도덕성에 관한 성찰과 실천 과정 자체를 의미하는 개념이다(교육부, 2015: 50).

<표 6> 2015 개정 초등학교 도덕 교육과정 내용체계

영역	핵심 가치	내용 요소	
		3~4학년군	5~6학년군
자신과의 관계	성실	<ul style="list-style-type: none"> ■ 도덕시간에는 무엇을 배울까?(근면, 정직) ■ 왜 아껴 써야 할까?(시간 관리와 절약) ■ 왜 최선을 다해야 할까?(인내) 	<ul style="list-style-type: none"> ■ 어떻게 하면 감정을 잘 조절할 수 있을까?(감정표현과 충동조절) ■ 자주적인 삶이란 무엇일까?(자주, 자율) ■ 정직한 삶은 어떤 삶일까?(정직한 삶)
타인과의 관계	배려	<ul style="list-style-type: none"> ■ 가족의 행복을 위해 무엇을 해야 할까?(효, 우애) ■ 친구와 사이좋게 지내기 위해 어떻게 해야 할까?(우정) ■ 예절이 없다면 어떻게 될까?(예절) ■ 함께하면 무엇이 좋을까?(협동) 	<ul style="list-style-type: none"> ■ 사이버 공간에서 지켜야 할 것은 무엇인가?(사이버 예절, 준법) ■ 서로 생각이 다를 때 어떻게 해야 할까?(공감, 존중) ■ 우리는 남을 왜 도와야 할까?(봉사)
사회·공동체와의 관계	정의	<ul style="list-style-type: none"> ■ 나는 공공장소에서 어떻게 해야 할까?(공익, 준법) ■ 나와 다르다고 차별해도 될까?(공정성, 존중) ■ 통일은 왜 필요할까?(통일의지, 애국심) 	<ul style="list-style-type: none"> ■ 우리는 서로의 권리를 왜 존중해야 할까?(인권존중) ■ 공정한 사회를 위해 무엇을 해야 할까?(공정성) ■ 통일로 가는 바람직한 방향은 무엇일까?(통일의지) ■ 전 세계 사람들과 어떻게 살아갈까?(존중, 인류애)
자연·초월과의 관계	책임	<ul style="list-style-type: none"> ■ 생명은 왜 소중할까?(생명 존중, 자연애) ■ 아름답게 살아가는 사람들의 모습은 어떠할까?(아름다움에 대한 사랑) 	<ul style="list-style-type: none"> ■ 어려움을 겪을 때 긍정적 태도가 왜 필요할까?(자아 존중, 긍정적 태도) ■ 나는 올바르게 살아가고 있을까?(윤리적 성찰)

2015 개정 초등학교 도덕 교육과정에서는 로봇을 직접적으로 언급하지는 않았지만 한 가지 내용요소가 로봇윤리교육의 내용요소와 관련이 있었다. ‘타인과의 관계’ 영역에서 ‘사이버 공간에서 지켜야 할 것은 무엇일까?’ 내용요소가 로봇윤리교육의 ‘II. 로봇의 문제에 대한 도덕적 사고

와 판단' 영역내용 요소인 '프라이버시'와 관련이 있었다. 이 영역에서는 저작권 침해, 사생활 침해 등의 프라이버시에 대한 문제점과 해결방안에 대해 탐구하여 사이버 공간에서의 예절과 법을 실천하는 것을 성취기준으로 삼고 있다.

프라이버시 문제는 정보화 시대에 발생하는 여러 문제점 중의 하나이다. 정보화시대가 되면서 개인의 프라이버시에 해당하는 정보를 자신의 동의 없이 기업 등에서 함부로 유출하는 문제가 심각해졌기 때문에 자기정보통제권의 의미로 프라이버시가 중요시되었다. 제4차 산업혁명시대에는 프라이버시 문제가 보다 더 심각해졌다. 수많은 인공지능 로봇에 저장된 데이터들로 인해 정보 프라이버시와 자율주행자동차와 집안에 상주하고 있는 케어로봇으로 인해 공간 프라이버시가 침해받을 가능성이 높아지고 있다. 로봇윤리교육에서는 정보 프라이버시 뿐만 아니라 공간 프라이버시까지 문제가 확대되었다. 즉, 정보윤리교육과 로봇윤리교육 모두에서 프라이버시는 중요한 내용 요소이다.

이어서 2015 개정 중학교 도덕 교육과정에서 로봇윤리교육과 관련 있는 내용 요소를 살펴보면 다음과 같다.

<표 7> 2015 개정 중학교 도덕 교육과정 내용체계

영역	핵심 가치	내용 요소
자신과의 관계	성실	<ul style="list-style-type: none"> ■ 왜 도덕적으로 살아야 하는가?(도덕적인 삶) ■ 도덕적으로 행동하기 위해 필요한 것은 무엇인가?(도덕적 행동) ■ 나는 어떤 사람이 되고자 하는가?(자아정체성) ■ 삶의 목적은 무엇인가?(삶의 목적) ■ 행복을 위해 어떻게 살아야 하는가?(행복한 삶)
타인과의 관계	배려	<ul style="list-style-type: none"> ■ 가정에서의 갈등을 어떻게 해결할 것인가?(가정윤리) ■ 참된 우정이란 무엇인가?(우정) ■ 성(性)의 도덕적 의미는 무엇인가?(성윤리) ■ 이웃에 대한 바람직한 자세는 무엇인가?(이웃생활) ■ 정보화시대에 우리는 어떻게 소통해야 하는가?(정보통신윤리) ■ 평화적 갈등 해결은 어떻게 가능한가?(평화적 갈등 해결) ■ 폭력의 문제를 어떻게 대할 것인가?(폭력의 문제)
사회·공	정의	<ul style="list-style-type: none"> ■ 인권의 도덕적 의미는 무엇인가?(인간존중)

동체와의 관계		<ul style="list-style-type: none"> ■ 다문화 사회에서 발생하는 갈등을 어떻게 해결할 것인가?(문화 다양성) ■ 세계 시민으로서 도덕적 과제는 무엇인가?(세계 시민 윤리) ■ 국가 구성원으로서 바람직한 자세는 무엇인가?(도덕적 시민) ■ 정의란 무엇인가?(사회 정의) ■ 북한을 어떻게 이해하고 바라볼 것인가?(북한 이해) ■ 우리에게 통일의 의미는 무엇인가?(통일윤리의식)
자연·초월과의 관계	책임	<ul style="list-style-type: none"> ■ 자연과 인간의 바람직한 관계는 무엇인가?(자연관) ■ 과학 기술과 도덕의 관계는 무엇인가?(과학과 윤리) ■ 삶과 죽음의 의미는 무엇인가?(삶의 소중함) ■ 마음의 평화는 어떻게 이룰 수 있을까?(마음의 평화)

2015 개정 중학교 도덕 교육과정에서도 역시 로봇을 직접적으로 언급하지는 않았지만 두 가지 내용요소가 로봇윤리와 관련이 있었다. ‘타인과의 관계’ 영역에서는 ‘정보화 시대에 우리는 어떻게 소통해야 하는가?’, ‘자연·초월과의 관계’ 영역에서는 ‘과학기술과 도덕의 관계’ 내용요소가 로봇윤리교육의 내용 요소와 관련이 있었다.

좀 더 구체적으로 살펴보면 다음과 같다. 첫째, ‘타인과의 관계’ 영역에서는 ‘정보화시대에 우리는 어떻게 소통해야 하는가?’ 내용 요소가 로봇윤리교육의 II 영역 ‘프라이버시’ 내용 요소와 관련이 있었다. 초등학교 도덕과 교육과정에서 분석한 것처럼, 정보화시대의 프라이버시 문제는 제4차 산업혁명시대에 더욱 심각해지고 있으며 정보윤리교육과 로봇윤리교육에서 프라이버시는 중요한 내용 요소이기 때문이다.

둘째, ‘자연·초월과의 관계’ 영역에서는 ‘과학기술과 도덕의 관계’ 내용 요소가 로봇윤리교육의 III 영역 ‘과학기술 시민성’ 내용 요소와 관련이 있었다. 이 부분에서는 과학 기술이 현대인의 삶에 미치는 긍정적인 영향과 가치를 설명하고, 과학기술의 문제점과 한계를 열거하며, 과학기술의 바람직한 활용 방안을 제시하는 것을 성취기준으로 삼고 있다. 인공지능 로봇 기술도 과학기술의 하나로 로봇윤리교육에서도 인공지능 로봇 기술에 대한 올바른 이해와 바람직한 활용에 대한 논의가 진행되어야 하기에 이 부분도 로봇윤리교육과 관련이 있다.

이어서 고등학교 도덕 교육과정을 살펴보면, 고등학교 선택과목 ‘윤리

와 사상’에서는 로봇윤리교육 관련 내용을 찾을 수 없었다. 그러나 고등학교 선택과목 ‘생활과 윤리’ 교육과정에서는 로봇을 직접적으로 언급하지는 않았지만 로봇윤리교육과 관련된 내용 요소를 찾을 수 있었다. 2015 개정 교육과정 ‘생활과 윤리’ 내용 체계 중 로봇윤리교육과 관련된 내용 요소는 아래 밑줄 친 부분과 같다.

<표 8> 2015 개정 고등학교 교육과정 ‘생활과 윤리’ 내용체계

영역	세부영역	내용 요소
현대의 삶과 실천 윤리	현대 생활과 실천 윤리	<ul style="list-style-type: none"> ◎ 현대 사회의 다양한 윤리적 문제들을 어떻게 해결할 것인가? ① 현대인의 삶과 다양한 윤리적 쟁점들 ② 실천윤리학의 성격과 특징
	현대 윤리 문제에 대한 접근	<ul style="list-style-type: none"> ◎ 동서양의 윤리를 통해 현대의 윤리문제를 어떻게 풀어갈 수 있는가? ① 동양 윤리의 접근 ② 서양 윤리의 접근
	윤리 문제에 대한 탐구와 성찰	<ul style="list-style-type: none"> ◎ 윤리적 쟁점들을 풀어가기 위해 우리에게 필요한 것은 무엇인가? ① 도덕적 탐구의 방법 ② 윤리적 성찰과 실천
생명과 윤리	삶과 죽음의 윤리	<ul style="list-style-type: none"> ◎ 생명의 시작과 끝에서 만나는 윤리는 무엇인가? ① 출생의 의미와 삶의 가치 ② 죽음과 관련된 윤리적 쟁점
	생명윤리	<ul style="list-style-type: none"> ◎ 생명과학 기술의 발달로 발생하는 윤리적 쟁점들은 무엇인가? ① 생명 복제와 유전자 치료 문제 ② 동물 실험과 동물 권리의 문제
	사랑과 성윤리	<ul style="list-style-type: none"> ◎ 성의 가치는 무엇이고 사랑과의 바람직한 관계는 무엇인가? ① 사랑과 성의 관계 ② 결혼과 가족의 윤리
사회와 윤리	직업과 청렴의 윤리	<ul style="list-style-type: none"> ◎ 직업을 통해 어떻게 행복한 삶을 영위할 수 있는가? ① 직업 생활과 행복한 삶 ② 직업윤리와 청렴
	사회 정의와	<ul style="list-style-type: none"> ◎ 공정한 사회로 발전하기 위해 우리에게 필요한 정의는

	윤리	무엇인가? ① 분배적 정의의 의미와 윤리적 쟁점들 ② 교정적 정의의 의미와 윤리적 쟁점들
	국가와 시민의 윤리	◎ 참여는 시민의 의무인가? ① 국가의 권위와 시민에 대한 의무 ② 민주시민의 참여
과학과 윤리	과학기술과 윤리	◎ 과학 기술은 사실의 문제인가, 가치의 문제인가? ① 과학기술 가치중립성 논쟁 ② 과학기술의 사회적 책임
	정보사회와 윤리	◎ 사이버 공간의 윤리와 현실의 윤리는 다른가? ① 정보 기술 발달과 정보윤리 ② 정보 사회에서의 매체윤리
	자연과 윤리	◎ 지속가능한 발전을 위한 윤리는 무엇일까? ① 자연을 바라보는 동서양의 관점 ② 환경 문제에 대한 윤리적 쟁점

‘과학과 윤리’ 영역에서는 ‘과학기술과 윤리’와 ‘정보사회와 윤리’의 내용 요소가 로봇윤리교육의 내용 요소와 관련이 있었다. 첫째, ‘과학 기술과 윤리’ 내용 요소 중 ‘과학 기술의 가치중립성 논쟁’, ‘과학 기술의 사회적 책임’에 대한 내용이 로봇윤리교육의 ‘Ⅲ. 로봇에 대한 도덕적 책임’ 영역과 관련이 있었다. 이 영역에서는 과학기술의 가치중립성 논쟁에 대해 논의하고, 과학기술이 사회에 미치는 영향력과 과학기술자의 사회적 책임을 강조하고 있다.

로봇윤리교육의 ‘Ⅲ. 로봇에 대한 도덕적 책임’ 영역에서도 인공지능 로봇 기술이 인간의 삶에 미치는 영향과 문제점에 대한 총체적인 사고와 도덕적 판단력을 함양하는 교육 내용을 다루고 있으며 과학기술에 대한 가치론적 접근을 하고 있다. 이러한 내용은 ‘과학기술과 윤리’ 내용 요소 중 ‘과학기술의 가치중립성 논쟁’ 내용 요소와 관련이 있다고 볼 수 있다. 또한 로봇윤리교육에서는 인공지능 로봇과 관련해서 전문가, 학자 등을 포함한 모든 사람들이 문제에 대해 함께 논의해야 한다는 측면에서 한 도덕적인 인간으로서의 책임의식을 강조하는 내용을 다루고 있다. 즉, 인공지능 로봇 기술과 관련된 문제에 대해 한 시민으로서의 사회적 책임감을 다해야 한다는 내용을 다루고 있기 때문에 ‘과학기술의 사회적 책

임' 내용 요소와 관련이 있다.

둘째, '정보사회와 윤리' 내용 요소 중 '정보 기술 발달과 정보 윤리', '정보 사회에서의 매체 윤리'에 대한 내용이 로봇윤리교육의 'II. 로봇의 문제에 대한 도덕적 사고와 판단' 영역과 관련이 있었다. 이 영역에서는 정보기술의 발달에 따른 윤리적 문제들을 제시할 수 있고 이에 대한 해결 방안을 제시하는 것을 성취기준으로 삼고 있다. 정보기술의 발달에 따른 윤리 문제로 사이버 폭력 문제, 저작권 문제, 사생활 침해 문제 등의 내용을 다루고 있다. 이러한 문제를 해결하기 위해서 인간 존중의 태도와 프라이버시를 존중하는 내용을 강조하고 있다.

로봇윤리교육에서도 초·중등 도덕과 교육과정 분석에서 살펴본 것과 같이, 인공지능 로봇의 제작 및 사용과 관련해서 인간 존엄성과 프라이버시를 존중하는 방향으로 인공지능 로봇을 활용할 것을 강조하는 내용을 다루고 있다.

앞에서 살펴본 내용을 종합해보면, 2015 개정 도덕과 교육과정에서 목표나 내용에서 직접적으로 인공지능 로봇을 언급한 부분은 없었다. 그러나 목표에 있어서는 자신을 둘러싸고 전개되고 있는 삶의 상황 속에 어떻게 살아갈 것인가를 고민하는 윤리적 성찰을 중시한다는 점에서 도덕과 목표와 맥락을 같이 한다고 볼 수 있다.

그리고 초·중등 도덕과 교육과정 내용 요소와 로봇윤리교육의 내용 요소가 상관성을 가지고 있다는 사실을 밝혔다. 첫째, 초등학교 도덕과 교육과정의 '사이버 공간에서 지켜야 할 것은 무엇인가?', 중학교 도덕과 교육과정의 '정보화 시대에 우리는 어떻게 소통해야 하는가?', 고등학교 '생활과 윤리' 교육과정의 '정보사회와 윤리'의 내용 요소가 로봇윤리교육의 II영역의 내용 요소와 관련이 있었다. 사이버공간과 정보화시대의 윤리적 문제인 저작권 침해, 사생활 침해 등의 프라이버시가 로봇윤리교육에서도 여전히 중요한 내용 요소이기 때문이다.

둘째, 중학교 도덕과 교육과정의 '과학기술과 도덕의 관계', 고등학교 선택과목 '생활과 윤리' 교육과정의 '과학기술과 윤리' 내용 요소가 로봇윤리교육의 III영역의 내용 요소와 관련이 있었다. 로봇윤리교육에서

도 인공지능 로봇 기술이 사회에 미치는 영향력과 로봇에 대한 사회적 책임감에 대한 내용을 다루고 있기 때문이다.

지금까지 살펴본 것과 같이, 초·중등 도덕과 교육과정 및 고등학교 선택과목 ‘생활과 윤리’ 교육과정에서 로봇과 관련하여 직접적인 단어의 언급은 없었지만 일부의 내용 요소가 로봇윤리교육의 내용 영역 중 ‘Ⅱ. 로봇의 문제에 대한 도덕적 사고와 판단’ 과 ‘Ⅲ. 로봇에 대한 도덕적 책임’ 의 내용 영역과 부분적으로 관련이 있었다. 이러한 분석 결과는 로봇윤리교육의 내용을 도덕과 교육의 내용에 적용하는 것이 타당하다는 것을 입증한다.

그러나 로봇윤리교육 내용의 세 영역 중 ‘Ⅰ. 인간과 로봇의 존재론에 대한 윤리적 성찰’ 과 관련된 내용은 찾을 수 없었다. 로봇윤리교육의 Ⅰ 영역은 로봇의 도덕적 지위에 대한 탐구를 통해 인간과 로봇의 정체성에 대해 성찰할 수 있는 내용을 다루고 있으며, 학생들은 이와 같은 내용으로 도덕적 상상력과 도덕적 정체성 정립 능력을 함양할 수 있다. Ⅰ 영역의 내용은 궁극적으로는 인간의 정체성 정립과 관련된 내용을 담고 있기 때문에 중학교 도덕과 교육과정의 ‘자신과의 관계’ 영역에 적용 가능할 것으로 보인다.

그리고 로봇윤리교육 내용의 세 영역 중 ‘Ⅲ. 로봇에 대한 도덕적 책임’ 중 ‘과학기술 시민성’ 과 관련된 내용을 도덕과 교육과정에서 다루고 있었지만, ‘과학기술 시민성’ 에 대한 구체적인 내용은 다루고 있지 않았다. 제4차 산업혁명시대에는 인공지능 로봇 기술을 포함한 새로운 기술이 인간의 삶에 큰 변화를 야기할 가능성이 있다. 이에 따라 로봇윤리교육에서는 과학기술 관련 의사결정에 시민들이 적극적으로 참여하는 과학적 시민성에 대한 내용을 포함하고 있다. 따라서 도덕과 교육과정의 ‘과학기술의 사회적 책임’ 에 대한 논의를 좀 더 확장하여 과학기술의 민주화를 위한 인공지능 로봇 기술과 관련된 정책에 대한 시민 참여의 중요성 및 근거, 과학기술 시민권 보장 등의 내용도 포함시킬 필요가 있을 것이다.

2) 「생활과 윤리」 교과서 분석

도덕과 교육과정 분석에 이어서 교과서를 분석한 결과, 초·중학교 도덕 교과서와 고등학교 선택과목 ‘윤리와 사상’ 교과서에서는 로봇윤리교육의 내용을 찾을 수 없었고, 고등학교 선택과목 ‘생활과 윤리’ 교과서에서만 로봇윤리교육 관련 내용을 찾을 수 있었다. 고등학교 ‘생활과 윤리’ 교과서에서는 로봇윤리교육의 영역 중 ‘Ⅱ. 로봇의 문제에 대한 도덕적 사고와 판단’ 과 ‘Ⅲ. 로봇에 대한 도덕적 책임’의 내용을 다루고 있었다.

먼저, 로봇윤리교육의 ‘Ⅱ. 로봇의 문제에 대한 도덕적 사고와 판단’ 내용 영역과 관련해서 ‘Ⅰ. 현대의 삶과 실천윤리’의 중단원 ‘01 현대 생활과 실천 윤리’의 소단원 ‘1. 현대인의 삶과 다양한 윤리적 쟁점들’에서는 아래와 같은 내용을 제시하고 있다.

만일 내가 자율주행자동차를 운전하고 있다고 가정해 보자. 주행 중인데 갑자기 사람이 나타났고 멈추기에는 너무 늦은 상황이 벌어졌다. 자율주행자동차가 방향을 틀지 않으면 눈앞의 사람을 칠 수 밖에 없고, 방향을 틀면 나는 크게 다치거나 생명을 잃을 수도 있다.

Q: 위의 상황에서 자율주행자동차는 무엇을 우선하도록 설계되어야 할까?(『생활과 윤리』 ② : 12).

미국의 한 고속 도로에서 전기 자동차로 주행하던 운전자가 트레일러와 충돌하여 사망한 사건이 발생하였다. 이 사건은 자율주행자동차에 의해 발생한 첫 사망 사고였다. 이를 계기로 자율주행기술의 안전성, 상용성뿐만 아니라 윤리성에 대해서도 문제가 제기되었다. 예를 들면 ‘자율 주행 중 직선거리에서 정면에 세 사람과 갓길에 한 사람이 있을 때, 직진을 할 것인가, 핸들을 돌릴 것인가?’와 같은 딜레마 상황에서 자동차에 어떤 원리를 장착할지의 문제가 발생할 수 있다.

Q: 밑줄 친 부분에 대하여 자율주행자동차에 어떤 원리를 장착할지 토론해 보자.

Q: 자율주행자동차로부터 초래되는 윤리적 문제에는 무엇이 있는지 알아보자
(『생활과 윤리』 ① : 14)

위의 같이 ‘생활과 윤리’ 교과서에서는 로봇의 제작 및 사용과 관련해서 자율주행자동차의 딜레마 상황에 대해 제시하고 있다. 본 연구에서는 III장에서 자율주행자동차의 딜레마 상황에 대해 의무론적 접근과 공리주의 접근으로 구분하여 관련 내용을 다룬 바 있다. 위의 내용에서는 로봇의 공학자와 제작자로서 인간과 로봇의 올바른 공존을 위해 자율주행자동차를 어떻게 설계하고 제작해야 할지에 대해 묻고 있다. 이러한 과정들은 인간의 윤리와 관련된 로봇윤리교육의 내용으로 학생들에게 인공지능 로봇을 제작하고 사용하는 것과 관련된 윤리적 문제에 대한 도덕적 판단력과 도덕적 민감성을 길러줄 수 있을 것이라 판단된다. 그러나 인간의 윤리와 관련된 로봇윤리교육 내용에서는 로봇 공학자, 제작자, 사용자로서의 역할 책임의식도 필요하기 때문에, ‘III. 사회와 윤리’의 중단원 ‘01 직업과 청렴의 윤리’에서 로봇 공학자, 제작자, 사용자로서의 역할 책임의식에 대해서 좀 더 심도 있는 논의를 하는 것이 필요하다고 생각된다.

또한 위의 자율주행자동차의 윤리적 딜레마 문제와 함께 로봇윤리의 개념을 ‘로봇의 설계, 제작 및 사용 과정에서 제기되는 윤리 문제를 다루는 분야’로 제시하고 있다(『생활과 윤리』 ①: 14). 이러한 개념은 로봇윤리의 범주에서 인간의 윤리 관련 내용만을 포함한 개념으로 로봇의 윤리에 관한 내용과 함께 개념 정의에 대한 수정이 필요하다고 판단된다.

다음으로 로봇윤리교육의 ‘III. 로봇에 대한 도덕적 책임’ 내용 영역과 관련해서 ‘I. 현대의 삶과 실천윤리’의 중단원 ‘01 현대 생활과 실천윤리’의 소단원 ‘1. 현대인의 삶과 다양한 윤리적 쟁점들’에서는 아래와 같이 토론 수업을 제시한다.

허서비스(알파고 책임자): 인공지능 알파고는 바둑판을 전체적으로 볼 수 있는 능력을 갖추고 있어요. 그래서 인간이 미처 보지 못하는 부분을 잘 볼 수 있죠. 우리가 인공지능을 개발한 이유가 바로 이것입니다. 인류가 당면한 질병, 환경, 에

너지 문제 등을 전체적으로 조망하여 문제의 핵심과 명확한 해결책을 제시하려는 것이죠. 인공지능은 아직 초기 단계 수준이지만 향후 다양한 분야에 적용할 수 있어요.

호킹(세계적인 물리학자): 인공지능은 생물학적 진화에 제한을 받는 인간보다 더 빠르게 진화할 잠재력을 가지고 있어요. 우리의 기계들이 스스로 진화하는 단계에 도달하면 그들의 목표가 우리 인간의 것과 같다고 예단할 수 없습니다. 인공지능은 인간과 함께 존재하고 우리의 능력을 증대할 수 있습니다. 하지만 독자적으로 행동하는 인공지능은 멈추기가 어려울 수 있습니다. 그 결과 인류의 종말을 가져올 수도 있는 것이죠.

- | |
|---|
| <p>Q: 허서비스와 호킹의 인공지능 개발에 관한 관점을 요약해보자.</p> <p>Q: 인공지능 개발의 긍정적인 면과 부정적인 면을 생각하여 써 보자.</p> <p>Q: 인공지능 개발의 부정적인 면을 최소화할 수 있는 방안을 토론을 통해 찾아보자.(『생활과 윤리』② : 14).</p> |
|---|

위와 같이 생활과 윤리 교과서에서는 알파고와 이세돌 9단의 바둑대결에 관련된 그림을 제시하고, 인공지능 로봇의 개발에 대한 긍정적인 면과 부정적인 면에 대해 생각해보고 인공지능 개발에 대한 비판적 사고와 도덕적 판단력을 함양할 수 있는 내용을 제시하고 있다.

도덕과 교과서를 분석한 결과, 초·중학교 도덕 교과서와 고등학교 선택과목 ‘윤리와 사상’ 교과서에서는 로봇윤리교육의 내용을 찾을 수 없었고, 고등학교 선택과목 ‘생활과 윤리’ 교과서에서 로봇윤리교육의 내용 영역 중 ‘Ⅱ. 로봇의 문제에 대한 도덕적 사고와 판단’, ‘Ⅲ. 로봇에 대한 도덕적 책임’ 과 관련된 내용을 다루고 있음을 확인할 수 있었다. 그러나 인간과 인공지능 로봇의 존재론과 관련된 ‘Ⅰ. 인간과 로봇의 존재론에 대한 윤리적 성찰’ 과 관련된 내용은 찾을 수 없었다. ‘생활과 윤리’ 교과서에서 ‘Ⅲ. 인간의 윤리’ 와 관련된 내용을 다루고 있지만, 로봇윤리의 정확한 개념과 주요 쟁점인 프라이버시와 투명성, 역할책임의식에 대한 내용은 다루지 않고 있기 때문에, 인공지능 로봇의 윤리적 문제에 대한 정확한 판단을 위해 로봇윤리와 관련된 내용을 구체적으로 다룰 필요가 있

을 것이다.

3) 도덕과 교육에서 로봇윤리교육의 의의

초·중등 도덕과 교육과정 및 고등학교 선택과목 ‘생활과 윤리’ 교육과정 분석 결과, 도덕과 교육과 로봇윤리교육은 목표와 내용 요소가 밀접한 관련이 있었다. 이와 같은 분석 결과를 토대로 이 장에서는 도덕과 교육을 통한 로봇윤리교육의 의의에 대해 구체적으로 논의하려고 한다.

첫째, 로봇윤리교육은 도덕적 주체로서의 자아 정체성 정립 과정을 보다 구체화하는데 기여할 수 있다. 도덕과는 자신을 둘러싼 현상을 탐구하고 내면의 도덕성을 성찰하여 스스로의 삶 속에서 실천하는 과정을 추구하며, 자신의 삶의 의미를 자율적으로 찾아갈 수 있는 도덕적 탐구 및 윤리적 성찰, 실천 과정으로 이어지는 ‘도덕함’의 능력을 기르는 것을 목표로 하는 교과이다(교육부, 2015: 50). 그리고 도덕과에서는 자신을 둘러싸고 전개되고 있는 급변하는 상황 속에서 어떻게 살아야 할 것인지에 대한 물음에 대해 끊임없이 성찰하는 내용을 중요하게 다루고 있다. 특히 초·중등 도덕과 교육과정의 ‘자신과의 관계’ 영역을 중심으로 도덕적 주체로서 자신의 삶에서 도덕적 삶의 중요성과 도덕적 행동을 하는 데 필요한 것이 무엇인지 탐구하는 내용을 강조하고 있다.

로봇이 인간과 동일한 도덕적 지위를 가질 수 있는지에 대한 논쟁을 통해 로봇윤리교육은 학생들에게 인간이 왜 도덕적 행위자인지에 대해 규명할 기회를 제공한다. 학생들은 도덕적 행위자에 대한 탐구를 통해 도덕적 행동과 도덕적 삶의 중요성을 깨달을 수 있을 것이다. 또한 인공지능 로봇으로 인한 향상이나 예술 논쟁을 통해 인간 본성이 갖고 있는 존엄성과 인류의 미적 가치와 윤리적 가치의 중요성에 대해서도 생각해 볼 수 있는 기회를 제공할 것이다.

즉, 도덕과를 통한 로봇윤리교육은 학생들에게 인간이 도덕적 행위를 하기 위해 필요한 것이 무엇인지를 탐구하게 하고, 이를 인간의 삶의 목적에 대한 논의로 확장시킬 것이다. 이를 통해 ‘자신과의 관계’ 영역의 한계를

보완하여 학생들이 도덕적 주체로서의 삶을 실현할 수 있도록 많은 도움을 줄 것이다. 이러한 과정은 로봇과 함께 원만하게 살아가는 미래사회를 위해서는 도덕과 교육이 더욱 필요하다는 사실도 입증할 것이다.

둘째, 로봇윤리교육은 도덕과 과학기술윤리교육 영역의 내용을 확장하는데 기여할 수 있다. 앞에서 살펴본 것과 같이, 도덕과 교육과정에서 다루어지고 있는 과학기술윤리교육의 내용 요소는 ‘과학기술과 도덕의 관계’, ‘과학기술의 가치중립성 논쟁’, ‘과학기술의 사회적 책임’이었다. 도덕과에서는 과학기술의 발전이 여러 윤리적 문제를 야기할 수 있으며 과학과 도덕의 관계에 대한 이해의 내용에 대해 중점적으로 다루었다. 그리고 과학기술의 가치중립성 논쟁과 관련하여 가치 지향적 과학기술 발전의 중요성에 대해 논의하고 주로 과학자의 사회적 책임을 강조하는 내용에 대해 교육하였다.

로봇윤리교육에서는 이러한 논의를 더욱 확장하여 새로운 로봇 기술에 대한 적극적인 과학기술 시민성을 함양해야 한다는 내용을 다루고 있다. 로봇 기술은 기존의 과학기술과는 달리 인간의 안전과 생명을 위협할 가능성이 매우 크다. 이에 따라 로봇 기술 개발의 초기부터 그 목적과 가치, 사회적 결과의 예견, 논의와 의사결정 절차 등 모든 부분에 있어서 시민들이 참여할 수 있는 ‘과학기술 거버넌스’의 역할이 중요해지고 있다(하대청, 2019: 90-91). 왜냐하면 로봇으로 인한 사고의 법적 책임이 로봇 공학자만이 아니라 로봇 제작자, 로봇 사용자까지 부여되기 때문이다. 즉, 로봇윤리교육에서는 새로운 로봇 기술이 우리의 공동체가 직면하고 있는 사회적 문제에 대응하는 기술인지 아니면 단순히 향상을 추구하면서 사회적 가치를 위협하는 것인지에 대한 많은 사람들의 도덕적 판단력이 중요해짐에 따라 ‘과학기술 시민성’을 내용 요소로 다루고 있다.

이에 따라 도덕과를 통한 로봇윤리교육은 도덕과 과학기술윤리교육의 내용을 과학과 도덕의 관계, 과학기술의 가치중립성 논쟁에 대한 논의의 틀에서 벗어나 과학기술 시민권, 과학기술 민주화로 내용 영역을 확장할 수 있을 것이다. 이를 통해 도덕과는 과학기술에 대해 가치론적 접근을 하는 주된 교과로서의 역할을 더욱 확고히 하여 학문적 정체성을 온전하게

발휘할 수 있을 것이다.

2. 도덕과 교육을 통한 로봇윤리교육 내용 개선 방안

이 장에서는 로봇윤리교육 관련 도덕과 교육과정 분석 결과를 바탕으로 중학교 도덕 및 고등학교 선택과목 생활과 윤리 교육과정에 반영하거나 새롭게 포함시킬 수 있는 로봇윤리교육 관련 내용 요소를 제안하고자 한다.

1) 로봇윤리교육 내용의 현행 도덕과 교육과정 반영 방안

2015 개정 도덕과 교육과정에 대한 분석 결과, 초·중등 도덕과 및 고등학교 선택과목 생활과 윤리 교육과정에서 일부의 내용 요소가 로봇윤리교육의 내용 영역 중 ‘Ⅱ. 로봇의 문제에 대한 도덕적 사고와 판단’ 과 ‘Ⅲ. 로봇에 대한 도덕적 책임’ 의 내용 영역과 부분적으로 관련되어 있음을 확인할 수 있었다. 그러나 로봇윤리교육 내용의 세 영역 중 ‘Ⅰ. 인간과 로봇의 존재론에 대한 윤리적 성찰’ 과 관련된 내용과 ‘Ⅲ. 로봇에 대한 도덕적 책임’ 중 ‘과학기술 시민성’ 에 대한 구체적 내용을 다루지 않고 있다는 점도 알 수 있었다. 이런 문제의식에 기초하여 로봇윤리교육의 내용을 적용할 수 있는 중학교 도덕 및 고등학교 생활과 윤리 교육과정의 내용요소를 제시해 보면 다음과 같다.

<표 9> 로봇윤리교육이 적용 가능한 2015 개정 중학교 도덕 교육과정 내용요소

영역	핵심 가치	내용 요소
자신과의 관계	성실	<ul style="list-style-type: none"> ■ 왜 도덕적으로 살아야 하는가?(도덕적인 삶) ■ 도덕적으로 행동하기 위해 필요한 것은 무엇인가?(도덕적 행동) ■ 나는 어떤 사람이 되고자 하는가?(자아정체성) ■ 삶의 목적은 무엇인가?(삶의 목적) ■ 행복을 위해 어떻게 살아야 하는가?(행복한 삶)
타인과의 관계	배려	<ul style="list-style-type: none"> ■ 가정에서의 갈등을 어떻게 해결할 것인가?(가정윤리) ■ 참된 우정이란 무엇인가?(우정) ■ 성(性)의 도덕적 의미는 무엇인가?(성윤리) ■ 이웃에 대한 바람직한 자세는 무엇인가?(이웃생활) ■ 정보화시대에 우리는 어떻게 소통해야하는가?(정보통신윤리) ■ 평화적 갈등 해결은 어떻게 가능한가?(평화적 갈등 해결) ■ 폭력의 문제를 어떻게 대할 것인가?(폭력의 문제)
사회·공동체와의 관계	정의	<ul style="list-style-type: none"> ■ 인권의 도덕적 의미는 무엇인가?(인간존중) ■ 다문화 사회에서 발생하는 갈등을 어떻게 해결할 것인가?(문화 다양성) ■ 세계 시민으로서 도덕적 과제는 무엇인가?(세계 시민 윤리) ■ 국가 구성원으로서 바람직한 자세는 무엇인가?(도덕적 시민) ■ 정의란 무엇인가?(사회 정의) ■ 북한을 어떻게 이해하고 바라볼 것인가?(북한 이해) ■ 우리에게 통일의 의미는 무엇인가?(통일윤리의식)
자연·초월과의 관계	책임	<ul style="list-style-type: none"> ■ 자연과 인간의 바람직한 관계는 무엇인가?(자연관) ■ 과학 기술과 도덕의 관계는 무엇인가?(과학과 윤리) ■ 삶과 죽음의 의미는 무엇인가?(삶의 소중함) ■ 마음의 평화는 어떻게 이룰 수 있을까?(마음의 평화)

이어서 로봇윤리교육의 내용을 적용할 수 있는 고등학교 선택과목 ‘생활과 윤리’ 교육과정의 내용요소를 제시하면 아래의 표와 같다.

〈표 10〉 로봇윤리교육이 적용 가능한 2015 개정 『생활과 윤리』 교육과정
내용요소

영역	핵심 가치	내용 요소	기능
현대의 삶과 실천윤리	성실 배려 정의 책임	1. 현대 생활과 실천윤리 ① 현대인의 삶과 다양한 윤리적 쟁점들 ② 실천윤리학의 성격과 특징	○ 윤리적 성찰 및 실천 성향 · 윤리학 이해하기 · 윤리 문제에 적용하기 · 윤리적 실천 방안 적용하기
		2. 현대 윤리 문제에 대한 접근 ① 동양 윤리의 접근 ② 서양 윤리의 접근	
		3. 윤리 문제에 대한 탐구와 성찰 ① 도덕적 탐구의 방법 ② 윤리적 성찰과 실천	
생명과 윤리		1. 삶과 죽음의 윤리 ① 출생의 의미와 삶의 가치 ② 실천윤리학의 성격과 특징	○ 자기 존중 및 관리 능력 · 윤리적 관점에서 설명하기
		2. 생명윤리 ① 생명 복제와 유전자 치료 문제 ② 동물 실험과 동물 권리의 문제	○ 윤리적 성찰 및 실천 성향 · 윤리이론을 통해 정당화하기
		3. 사랑과 성윤리 ① 사랑과 성의 관계 ② 결혼과 가족의 윤리	○ 도덕적 대인관계 능력 · 윤리적 관점에서 성찰하기
사회와 윤리		1. 직업과 청렴의 윤리 ① 직업 생활과 행복한 삶 ② 직업윤리와 청렴	○ 도덕적 공동체 의식 · 윤리적 관점에서 설명하기 · 공정한 사회건설방안 제안하기 · 윤리적 성찰방안 제안하기
		2. 사회 정의와 윤리 ① 분배적 정의의 의미와 윤리적 쟁점들 ② 교정적 정의의 의미와 윤리적 쟁점들	
		3. 국가와 시민의 윤리 ① 국가의 권위와 시민에 대한 의무 ② 민주시민의 참여	
과학과 윤리	1. 과학 기술과 윤리 ① 과학 기술 가치중립성 논쟁 ② 과학 기술의 사회적 책임	○ 윤리적 성찰 및 실천 성향 · 윤리적 관점에서 정당화하기 · 윤리적 관점에서 비	
	2. 정보 사회와 윤리		

	<ul style="list-style-type: none"> ① 정보 기술 발달과 정보윤리 ② 정보 사회에서의 매체윤리 	<ul style="list-style-type: none"> 판하기 · 윤리적 실천 방안 제안하기
	<ul style="list-style-type: none"> 3. 자연과 윤리 ① 자연을 바라보는 동서양의 관점 ② 환경 문제에 대한 윤리적 쟁점 	<ul style="list-style-type: none"> ○ 도덕적 공동체 의식 · 윤리적 실천 방안 제안하기
문화와 윤리	<ul style="list-style-type: none"> 1. 예술과 대중문화 윤리 ① 미적 가치와 윤리적 가치 ② 대중문화의 윤리적 문제 	<ul style="list-style-type: none"> ○ 윤리적 성찰 및 실천 성향 · 윤리적 관점에서 비판하기
	<ul style="list-style-type: none"> 2. 의식주 윤리와 윤리적 소비 ① 의식주의 윤리 ② 윤리적 소비문화 	<ul style="list-style-type: none"> ○ 윤리적 성찰 및 실천 성향 · 윤리적 실천 방안 제안하기
	<ul style="list-style-type: none"> 3. 다문화 사회의 윤리 ① 문화 다양성과 존중 ② 종교의 공존과 관용 	<ul style="list-style-type: none"> ○ 도덕적 공동체 의식 · 윤리적 관점에서 비판하기 · 윤리적 실천 방안 제안하기
평화와 공존의 윤리	<ul style="list-style-type: none"> 1. 갈등 해결과 소통의 윤리 ① 사회 갈등과 사회 통합 ② 소통과 담론의 윤리 	<ul style="list-style-type: none"> ○ 도덕적 공동체 의식 · 윤리적 관점에서 비판하기 · 윤리적 관점에서 정당화하기
	<ul style="list-style-type: none"> 2. 민족 통합의 윤리 ① 통일 문제를 둘러싼 쟁점 ② 통일이 지향해야 할 가치 	
	<ul style="list-style-type: none"> 3. 지구촌 평화의 윤리 ① 국제 분쟁의 해결과 평화 ② 국제 사회에 대한 책임과 기여 	<ul style="list-style-type: none"> · 윤리적 실천 방안 제안하기

위에서 제시한 두 개의 표에 대해 구체적으로 설명하면 다음과 같다. 첫째, ‘I. 인간과 로봇의 존재론에 대한 윤리적 성찰’ 영역은 중학교 도덕의 ‘자신과의 관계’, 고등학교 생활과 윤리의 ‘생명과 윤리’ 영역에 적용될 수 있다.

먼저 중학교 도덕에서 ‘자신과의 관계’는 ‘자아정체성’을 내용요소로 한다. 로봇윤리교육과의 연계는 이와 같은 내용요소를 인간과 인공지능의 정체성, 인간과 로봇의 공존과 연결시키는 것이다. 인공지능 로봇의 자

울성이 발전함에 따라 부각된 인간과 비인간의 경계, 인간성에 대한 재정의 문제 등을 다루어 바람직한 자아 정체성을 확립하고 인간과 인공지능 로봇의 공존을 위해 어떻게 살아야 하는지에 대한 학생들의 생각을 일깨워준다. 다음으로, 고등학교 생활과 윤리에서 ‘생명과 윤리’는 ‘생명윤리’를 내용요소로 한다. 로봇윤리교육과의 연계는 이와 같은 내용요소를 인공지능의 정체성과 연결시키는 것이다. 인공지능 기술과 로봇을 통해 탄생할 수 있는 복제인간 문제나 유전자 가위 등의 향상 문제에 대해 학생들이 각 윤리적 관점에 따라 비교·분석하고 자신의 관점을 확립할 수 있도록 한다.

둘째, ‘II. 로봇의 문제에 대한 도덕적 사고와 판단’ 영역은 중학교 도덕의 ‘타인과의 관계’, 고등학교 생활과 윤리의 ‘과학과 윤리’ 영역에 적용될 수 있다. 구체적으로는 중학교 도덕에서 ‘타인과의 관계’ 중 ‘정보통신윤리’와 고등학교 생활과 윤리의 ‘과학과 윤리’ 중 ‘정보통신윤리’ 영역과 연계할 수 있다. 로봇윤리교육과의 연계는 이와 같은 내용요소를 인간 존엄성, 프라이버시 문제와 연결시키는 것이다. 정보사회에 필요한 윤리인 인간 존엄성과 프라이버시를 인공지능 로봇의 제작 및 사용에 대한 논의로 확장하여 인간 존엄성과 프라이버시의 중요성을 강조할 수 있다.

셋째, ‘III. 로봇에 대한 도덕적 책임’ 영역은 중학교 도덕의 ‘자연·초월과의 관계’, 고등학교 생활과 윤리의 ‘사회와 윤리’, ‘과학과 윤리’ 영역에 적용할 수 있다.

우선, 중학교 도덕에서 ‘자연·초월과의 관계’ 중 ‘과학기술과 도덕의 관계’ 영역과 고등학교 생활과 윤리의 ‘과학과 윤리’ 중 ‘과학기술과 윤리’ 영역에서 로봇윤리교육의 내용 요소인 ‘과학기술 시민성’과 연계할 수 있다. 과학기술은 현대인의 삶에 미치는 문제점과 한계가 있기 때문에 사회적 책임이 필요하다는 내용은 인공지능 로봇 기술의 설계와 사용과 관련된 문제에 대해 항상 책임의식을 가져야 하며, 적극적으로 로봇윤리 관련 의사결정에 참여하여 과학기술 시민성을 발휘해야 한다는 내용과 연계할 수 있다. 다음으로, 고등학교 생활과 윤리의 ‘사회와 윤리’

중 ‘직업과 청렴의 윤리’ 영역을 로봇윤리교육의 내용 요소인 ‘역할책임의식’ 과 연계할 수 있다. 인공지능 로봇을 제작하는 공학자와 제작자는 사용자의 생명 및 안전과 관련이 있기 때문에 모든 직업에서 공통적으로 요구되는 행위규범인 일과 직업에 대한 긍정적 가치관 및 태도, 일에 대한 직분이나 책임을 넘어서 더욱 높은 역할책임의식이 요구된다는 것을 일깨워 줄 수 있다.

지금까지 로봇윤리교육의 내용요소를 도덕과 교육과정에 적용해보았으나 로봇윤리교육의 모든 내용 요소를 구체적으로 다루지 못했다. ‘I. 인간과 로봇의 존재론에 대한 윤리적 성찰’ 영역은 중학교 도덕의 ‘자신과의 관계’, 고등학교 생활과 윤리의 ‘생명과 윤리’ 영역의 자아 정체성의 내용 요소에서 인간과 로봇의 존재론에 대해 논의할 수 있지만 구체적으로 로봇의 존재론에 대해서는 다룰 수 없다. 인공지능 로봇의 도덕적 지위와 인간과 로봇의 정체성에 대한 윤리적 성찰을 통해 도덕적 정체성 정립 능력을 함양할 수 있는 교육 내용이 필요하다.

‘II. 로봇의 문제에 대한 도덕적 사고와 판단’ 영역은 중학교 도덕의 ‘타인과의 관계’, 고등학교 생활과 윤리의 ‘과학과 윤리’ 영역의 내용 요소에서 정보윤리와 관련된 인간 존엄성, 프라이버시의 내용에 대해 논의할 수 있지만 투명성과 공간 프라이버시의 내용에 대해서는 다룰 수 없다. 정보사회와 제4차 산업혁명시대의 윤리적 문제에서 존중되어야 할 가치가 동일한 부분도 있지만, 문제의 발생 배경이 다르기 때문에 정보윤리의 내용으로 로봇윤리의 내용까지 포괄할 수는 없다. 즉, 현재의 교육과정에서 로봇윤리교육의 II 영역의 내용을 모두 다룰 수는 없다.

‘III. 로봇에 대한 도덕적 책임’ 영역은 중학교 도덕의 ‘자연·초월과의 관계’, 고등학교 생활과 윤리의 ‘사회와 윤리’, ‘과학과 윤리’ 영역의 내용 요소에서 역할책임의식, 과학기술 시민성의 내용을 다룰 수 있지만 과학기술의 민주화를 위한 과학기술 시민권 보장의 내용까지 논의를 확장할 수 없다. 이러한 한계점을 토대로 다음 항에서 로봇윤리교육을 위한 도덕과 교육과정 내용 신설을 제안하려고 한다.

2) 로봇윤리교육 관련 도덕과 교육과정 내용 신설 방안

이 항에서는 앞에서 살펴본 로봇윤리교육 적용의 한계점을 보완하고 도덕과 교육을 통한 로봇윤리교육의 구현을 위해 중학교 도덕 및 고등학교 선택과목 생활과 윤리 교육과정의 내용 신설을 제안하려고 한다. 도덕과는 한 사회에서 작동하고 있는 도덕 현상에 대한 민감성에 기반을 둔 관심과 분석, 그리고 그러한 도덕 현상과의 상호작용을 강조하고 있다. 이러한 맥락에서 교육과정도 시의 적절한 변화가 필요하기 때문에 제4차 산업혁명과 인공지능 로봇에 따른 변화를 반영하여 로봇윤리교육의 내용 요소를 독립적으로 신설하는 시도가 필요하다. 본 연구에서는 중학교 도덕 교육과정의 ‘자연·초월과의 관계’의 영역과 고등학교 선택과목 ‘과학과 윤리’ 내용 영역에 로봇윤리교육에 관한 새로운 내용요소를 제안하고자 한다.

(1) 중학교 도덕 교육과정 내용 신설 제안

중학교 도덕 교육과정은 현대 사회의 다양한 문제에 대한 탐구와 삶의 이야기에 대한 성찰을 바탕으로 도덕적 정체성을 구성하며, 정의로운 공동체를 구현하기 위해 적극적으로 참여하고 실천하는 덕성과 역량을 기르는 것을 목표로 한다(교육부, 2015: 5). 이러한 맥락에서 인공지능 로봇으로 인해 발생할 수 있는 여러 문제에 대한 탐구와 성찰의 과정으로 도덕적 정체성 정립 능력과 도덕적 판단력을 함양하기 위해 중학교 도덕 교육과정에 로봇윤리교육 내용 요소의 신설을 제안하려고 한다.

중학교 도덕 교육과정에서는 가치 관계에 따라 ‘자신과의 관계’, ‘타인과의 관계’, ‘사회·공동체와의 관계’, ‘자연·초월과의 관계’로 내용 영역을 구분하고 있다. ‘가치 관계’란 도덕적 주체와 사고 대상(객체)간 관계이며, ‘도덕적 가치’에 의해 규정되고 이를 통해 상호 연결되어 있다. ‘가치 관계 확장법’은 도덕적 주체가 성찰하는 대상의 범위가 도덕적 가치를 중심으로 점차 확장되어 가는 방식이다(정창우, 2013: 410).

그렇다면 로봇윤리교육의 내용은 어떤 가치 관계에서 다루어져야 할까? 만약 미래에 인공지능 로봇의 자율성이 발달하여 로봇이 하나의 도덕적 행위자로 인정받게 된다면, ‘II. 타인과의 관계’에 해당될 수 있을 것이다. 그러나 현재의 인공지능 기술 수준으로는 인간과 동등한 위치의 도덕적 행위자로 인정할 수 없다. 그래서 인공지능 로봇은 과학기술의 하나이기 때문에 ‘자연·초월과의 관계’에 해당되는 것으로 생각된다. 2015 개정 중학교 도덕 교육과정의 ‘자연·초월과의 관계’에서는 과학기술의 윤리적 사용을 통해 지속 가능한 미래를 지향하는 것을 일반화된 지식으로 삼고 있으며, 자연과 인간의 바람직한 관계, 과학기술과 도덕의 관계, 삶과 죽음의 의미, 마음의 평화에 대한 내용요소를 다루고 있다.

‘과학과 윤리’ 영역에서 과학기술과 도덕의 관계 및 과학기술의 책임과 관련된 내용을 다루고 있지만, 로봇윤리교육의 모든 내용을 포괄하지 못한다. ‘과학과 윤리’ 영역에서는 주로 ‘과학 기술과 도덕의 관계’와 ‘과학기술의 사회적 책임’에 대한 내용을 다루고 있다. 이 영역은 과학기술이 현대인의 삶에 미치는 긍정적인 영향과 가치를 설명하고, 동시에 과학 기술이 지닌 문제점과 한계를 열거하며, 과학 기술의 바람직한 활용 방안을 제시할 수 있는 것을 성취기준으로 삼고 있다. 앞서 살펴본 바와 같이, 이러한 내용은 로봇윤리교육의 ‘III. 로봇에 대한 도덕적 책임’ 영역과 연계하여 인공지능 로봇 기술의 설계와 사용과 관련된 문제에 대해 항상 책임의식을 가져야 한다는 내용을 다룰 수 있음을 앞서 살핀 바 있다.

‘과학과 윤리’ 영역의 내용 요소와 로봇윤리교육의 내용 요소가 직간접적으로 관련이 있지만 두 영역은 여러 가지 점에서 구별된다. ‘과학과 윤리’ 영역은 ‘I. 로봇의 존재론에 대한 윤리적 성찰’, ‘II. 로봇의 문제에 대한 도덕적 사고와 판단’ 영역의 내용을 모두 포괄하지 못한다. 특히 로봇윤리교육의 ‘I. 로봇의 존재론에 대한 윤리적 성찰’ 영역은 단순히 과학기술과 도덕의 관계, 과학기술의 사회적 책임을 넘어서 인간과 로봇의 정체성에 관련된 내용을 다루고 있다. 모든 과학기술이 인간의 존재론과 관련 있지 않기 때문에 ‘I. 로봇의 존재론에 대한 윤리적 성찰’ 영

역은 ‘과학과 윤리’ 영역의 내용 요소로 적합하지 않다.

이에 따라 ‘과학과 윤리’ 영역이 아닌 교육과정 내용 영역 중 독립적인 내용 요소로 제안하고자 한다.

<표 11> 로봇윤리교육 관련 중학교 도덕 교육과정 내용요소 신설 제안

영역	핵심 가치	내용 요소
자신과의 관계	성실	<ul style="list-style-type: none"> ■ 왜 도덕적으로 살아야 하는가?(도덕적인 삶) ■ 도덕적으로 행동하기 위해 필요한 것은 무엇인가?(도덕적 행동) ■ 나는 어떤 사람이 되고자 하는가?(자아정체성) ■ 삶의 목적은 무엇인가?(삶의 목적) ■ 행복을 위해 어떻게 살아야 하는가?(행복한 삶)
타인과의 관계	배려	<ul style="list-style-type: none"> ■ 가정에서의 갈등을 어떻게 해결할 것인가?(가정윤리) ■ 참된 우정이란 무엇인가?(우정) ■ 성(性)의 도덕적 의미는 무엇인가?(성윤리) ■ 이웃에 대한 바람직한 자세는 무엇인가?(이웃생활) ■ 정보화시대에 우리는 어떻게 소통해야하는가?(정보통신윤리) ■ 평화적 갈등 해결은 어떻게 가능한가?(평화적 갈등 해결) ■ 폭력의 문제를 어떻게 대할 것인가?(폭력의 문제)
사회·공동체와의 관계	정의	<ul style="list-style-type: none"> ■ 인권의 도덕적 의미는 무엇인가?(인간존중) ■ 다문화 사회에서 발생하는 갈등을 어떻게 해결할 것인가?(문화 다양성) ■ 세계 시민으로서 도덕적 과제는 무엇인가?(세계 시민 윤리) ■ 국가 구성원으로서 바람직한 자세는 무엇인가?(도덕적 시민) ■ 정의란 무엇인가?(사회 정의) ■ 북한을 어떻게 이해하고 바라볼 것인가?(북한 이해) ■ 우리에게 통일의 의미는 무엇인가?(통일윤리의식)
자연·초월과의 관계	책임	<ul style="list-style-type: none"> ■ 자연과 인간의 바람직한 관계는 무엇인가?(자연관) ■ 과학 기술과 도덕의 관계는 무엇인가?(과학과 윤리) ■ 로봇과 인간은 어떻게 공존할 수 있을까?(로봇과 윤리) ■ 삶과 죽음의 의미는 무엇인가?(삶의 소중함) ■ 마음의 평화는 어떻게 이룰 수 있을까?(마음의 평화)

본 연구에서는 도덕과 교육을 통한 로봇윤리교육 내용 구현방안으로 위 표에 밑줄 친 바와 같이 ‘자연·초월과의 관계’ 영역 중 ‘로봇과 인간은 어떻게 공존할 수 있을까?(로봇과 윤리)’ 를 내용 요소로 추가할 것을 제안한다. 성취기준으로는 ‘로봇이 야기할 수 있는 문제에 대해 생각해보

고, 윤리적 해결 방안을 제시할 수 있다.’로 설정하고자 한다. 구체적으로는 ‘① 로봇이 야기할 수 있는 도덕 문제는 무엇이 있을까?’와 ‘② 로봇에 대한 도덕적 책임이 필요한 이유는 무엇인가?’의 성취기준으로 구분하여 그 내용을 자세히 살펴보고자 한다.

‘① 로봇이 야기할 수 있는 도덕 문제는 무엇이 있을까?’와 관련된 내용은 로봇윤리교육의 내용 영역 중 I, II 영역에 대한 내용을 근거로 한다. 이 영역에서는 인공지능 로봇으로 인한 우리 삶의 변화와 인공지능 로봇의 발달로 인해 발생할 수 있는 프라이버시 침해, 로봇 중독 등의 문제에 대해 다룬다. 궁극적으로는 학생들이 인공지능 로봇으로 인해 발생할 수 있는 문제점을 제시하는 것을 목표로 한다.

‘② 로봇에 대한 도덕적 책임이 필요한 이유는 무엇인가?’와 관련된 내용은 로봇윤리교육의 내용 영역 중 III 영역에 대한 내용을 근거로 한다. 이 영역에서는 로봇 공학자, 제작자, 사용자로서의 역할책임의식과 과학기술시민성이 필요한 이유에 대해 제시하고 인간과 로봇이 함께 살아가기 위해 도덕적 책임을 실천하는 자세에 대한 내용을 다룬다. 궁극적으로는 학생들이 로봇에 대한 도덕적 책임이 필요한 까닭을 설명하고 도덕적 책임을 실천하려는 자세를 지니는 것을 목표로 한다.

중학교 도덕과 교육과정에서 ‘로봇과 인간은 어떻게 공존할 수 있을까?(로봇과 윤리)’의 내용을 통해 학생들은 로봇과 관련된 윤리적 문제에 대해 이해하고 도덕적으로 해결하려는 자세와 태도를 함양할 수 있을 것이라 사료된다.

(2) 고등학교 선택과목 생활과 윤리 내용 신설 제안

앞의 절에서 살펴보았듯이 고등학교 도덕과 교육과정의 선택 과목인 ‘생활과 윤리’는 생활의 여러 영역에서 발생하는 문제들의 의미와 성격을 윤리적인 관점에서 이해하며, 이를 바람직하고 합리적으로 해결할 수 있는 능력과 태도를 지니게 하는 것을 목표로 한다. 또한 ‘생활과 윤리’는 동서양의 윤리 이론을 토대로 다양한 윤리 문제의 해결 방안을 모색하는 실천윤리의 성격에 주안점을 두고 있어 주로 응용윤리의 내용을 많이 다루고 있다. 응용윤리의 한 분야인 로봇윤리는 미래사회를 준비하는 예방 윤리로서 학생들이 로봇으로 비롯될 많은 윤리적 문제와 쟁점들을 탐구하고 성찰하며 보다 도덕적인 사회를 만들기 위해 ‘생활과 윤리’의 내용 요소로 적용 가능하다. 이러한 이유로 앞의 절에서 로봇윤리교육을 ‘생활과 윤리’ 과목에 적용한 바 있다. 그러나 로봇윤리교육의 내용 영역을 모두 다룰 수 없었다.

특히 생활과 윤리는 학생들의 생활 속에서 일어나는 다양한 윤리적 문제를 다루고 있기 때문에 시대적 변화에 더욱 민감하게 반응할 필요가 있다. 이러한 이유로 2009 개정 교육과정에는 없었던 ‘매체 윤리’도 시대적 변화에 따라 2015 개정 교육과정에 새롭게 적용된 바 있다. 이러한 맥락에서 아래와 같이 ‘과학과 윤리’ 영역의 독립적인 한 영역의 내용 요소로 로봇윤리교육의 내용을 신설하여 다룰 것을 제안한다.

로봇기술이 과학기술 중 하나임에도 불구하고 새로운 영역으로 다루어야 하는 이유는 다음과 같다. ‘과학과 윤리’ 영역 중 ‘1. 과학기술과 윤리’는 현대 과학기술의 올바른 발전방향에 대한 영역으로, 구체적으로는 ‘과학 기술 가치중립성 논쟁’과 ‘과학기술의 사회적 책임’을 내용 요소로 다루고 있다. ‘1. 과학기술과 윤리’ 영역은 로봇윤리교육의 ‘Ⅲ. 로봇에 대한 도덕적 책임’에서 로봇 제작자, 공학자의 역할책임의식을 다룬다는 점에서 로봇윤리교육의 내용 영역과 관련이 있다. 그러나 ‘1. 과학기술과 윤리’ 영역은 ‘Ⅰ. 로봇의 존재론에 대한 윤리적 성찰’, ‘Ⅱ. 로봇의 문제에 대한 도덕적 사고와 판단’ 영역의 내용까지 모두 포괄하지

못한다. 이에 본 연구에서는 로봇윤리교육을 활용한 도덕과 교육과정 내용 개선 방안으로 아래 표에 밑줄 친 바와 같이 ‘과학과 윤리’ 영역 중 ‘로봇과 윤리’ 를 내용 요소로 추가할 것을 제안한다.

<표 12> 로봇윤리교육 관련 『생활과 윤리』 교육과정 내용요소 신설 제안

영역	핵심 가치	내용 요소	기능
현대의 삶과 실천윤리		1. 현대 생활과 실천윤리 ① 현대인의 삶과 다양한 윤리적 쟁점들 ② 실천윤리학의 성격과 특징	○ 윤리적 성찰 및 실천 성향 · 윤리학 이해하기 · 윤리 문제에 적용하기 · 윤리적 실천 방안 적용하기
		2. 현대 윤리 문제에 대한 접근 ① 동양 윤리의 접근 ② 서양 윤리의 접근	
		3. 윤리 문제에 대한 탐구와 성찰 ① 도덕적 탐구의 방법 ② 윤리적 성찰과 실천	
생명과 윤리	성실 배려 정의 책임	1. 삶과 죽음의 윤리 ① 출생의 의미와 삶의 가치 ② 실천윤리학의 성격과 특징	○ 자기 존중 및 관리 능력 · 윤리적 관점에서 설명하기 ○ 윤리적 성찰 및 실천 성향 · 윤리이론을 통해 정당화하기 ○ 도덕적 대인관계 능력 · 윤리적 관점에서 성찰하기
		2. 생명윤리 ① 생명 복제와 유전자 치료 문제 ② 동물 실험과 동물 권리의 문제	
		3. 사랑과 성윤리 ① 사랑과 정의 관계 ② 결혼과 가족의 윤리	
사회와 윤리		1. 직업과 청렴의 윤리 ① 직업 생활과 행복한 삶 ② 직업윤리와 청렴	○ 도덕적 공동체 의식 · 윤리적 관점에서 설명하기 · 공정한 사회건설방안 제안하기 · 윤리적 성찰방안 제안하기
		2. 사회 정의와 윤리 ① 분배적 정의의 의미와 윤리적 쟁점들 ② 교정적 정의의 의미와 윤리적 쟁점들	
		3. 국가와 시민의 윤리 ① 국가의 권위와 시민에 대한 의무 ② 민주시민의 참여	
과학과		1. 과학 기술과 윤리	○ 윤리적 성찰 및 실천

윤리	① 과학 기술 가치중립성 논쟁 ② 과학 기술의 사회적 책임	천 성향 · 윤리적 관점에서 정당화하기 · 윤리적 관점에서 비판하기
	2. 정보 사회와 윤리 ① 정보 기술 발달과 정보윤리 ② 정보 사회에서의 매체윤리	· 윤리적 실천 방안 제안하기 ○ 도덕적 공동체 의식 · 윤리적 실천 방안 제안하기
	3. 로봇과 윤리 ① 인간과 로봇의 관계 ② 인간이 지켜야 할 로봇윤리	
	4. 자연과 윤리 ① 자연을 바라보는 동서양의 관점 ② 환경 문제에 대한 윤리적 쟁점	
문화와 윤리	1. 예술과 대중문화 윤리 ① 미적 가치와 윤리적 가치 ② 대중문화의 윤리적 문제	○ 윤리적 성찰 및 실천 성향 · 윤리적 관점에서 비판하기
	2. 의식주 윤리와 윤리적 소비 ① 의식주의 윤리 ② 윤리적 소비문화	○ 윤리적 성찰 및 실천 성향 · 윤리적 실천 방안 제안하기
	3. 다문화 사회의 윤리 ① 문화 다양성과 존중 ② 종교의 공존과 관용	○ 도덕적 공동체 의식 · 윤리적 관점에서 비판하기 · 윤리적 실천 방안 제안하기
평화와 공존의 윤리	1. 갈등 해결과 소통의 윤리 ① 사회 갈등과 사회 통합 ② 소통과 담론의 윤리	○ 도덕적 공동체 의식 · 윤리적 관점에서 비판하기
	2. 민족 통합의 윤리 ① 통일 문제를 둘러싼 쟁점 ② 통일이 지향해야 할 가치	· 윤리적 관점에서 정당화하기
	3. 지구촌 평화의 윤리 ① 국제 분쟁의 해결과 평화 ② 국제 사회에 대한 책임과 기여	· 윤리적 실천 방안 제안하기

‘로봇과 윤리’의 내용을 구성함에 있어서는 크게 두 가지 내용요소로 제시할 수 있다.

먼저, ‘인간과 로봇의 관계’에서는 로봇윤리교육의 영역 중 ‘1. 인

간과 로봇의 존재론에 대한 윤리적 성찰'에 대한 내용 영역을 근거로 한다. 이 영역에서는 인공지능 로봇의 도덕적 지위 및 인간과 로봇의 존재론에 대한 탐구를 통해 인간과 로봇의 존재에 대해 성찰하는 내용을 다룰 것이다. 구체적으로는 인간과 로봇의 관계에 대한 다양한 입장, 로봇의 도덕적 지위에 대한 논쟁에 대해 논의할 것이다. 인공지능 로봇의 도덕적 지위에 대한 논쟁의 핵심을 이해하고 인간과 로봇의 관계에 관한 주요 윤리적 입장을 설명하는 것을 성취기준으로 제안한다. 이 영역에서 학생들은 인간과 로봇의 존재에 대한 탐구와 성찰을 통해 도덕적 정체성 정립 능력과 도덕적 상상력을 함양할 수 있을 것이다.

다음으로 '인간이 지켜야 할 로봇윤리'에서는 로봇윤리교육의 내용 영역 중 'II. 로봇의 문제에 대한 도덕적 사고와 판단'과 'III. 로봇에 대한 도덕적 책임'에 대한 영역을 근거로 한다. 좀 더 구체적으로 살펴보면 다음과 같다. 첫째, 'II. 로봇의 문제에 대한 도덕적 사고와 판단' 영역에서는 인공지능 로봇의 발달에 따른 긍정적 변화와 윤리적 문제의 내용을 다룰 것이며 로봇윤리의 주요 쟁점인 인간 존엄성, 프라이버시, 투명성에 대해 논의할 것이다. 구체적으로는 로봇과 관련된 문제에 대한 의무론적 접근, 공리주의 접근 등 다양한 윤리적 접근에 대해 제시할 것이다. 그리고 인공지능 로봇의 발달에 따른 양면성과 윤리적 문제에 대해 이해하고 로봇윤리에 대해 설명하는 것을 성취기준으로 제안한다. 이 영역에서 학생들은 로봇이 야기할 윤리적 문제에 대한 도덕적 판단력을 함양할 수 있을 것이다.

둘째, 'III. 로봇에 대한 도덕적 책임' 영역에서는 인공지능 로봇과 관련된 윤리적 문제를 해결하기 위한 개인적 책임감과 사회적 책임감에 대해 다룰 것이다. 구체적으로는 로봇 공학자, 제작자, 사용자로서의 윤리에 대해 제시하고 한 시민으로서의 사회적 책임감인 과학적 시민성에 대해서도 제시한다. 이 영역에서는 역할책임의식의 중요성을 설명하고 인공지능 로봇과 관련된 정책에 대한 적극적인 과학적 시민성을 함양하는 것을 성취기준으로 제안한다. 이 영역에서 학생들은 인공지능 로봇을 만드는 공학자와 제작자들의 책임과 관련된 윤리적 문제에 대한 도덕적 판단력을 함

양하고, 과학기술 시민성에 대한 이해를 통해 도덕적 공동체 의식을 함양할 수 있을 것이다.

지금까지의 내용을 살펴보면, 도덕과 교육을 통한 로봇윤리교육 내용 구현방안으로 중학교 도덕과 교육과정과 고등학교 ‘생활과 윤리’ 교육과정에 로봇윤리교육 관련 내용 요소를 신설하는 방안을 제안하였다. 중학교 도덕과 교육과정에서는 ‘자연·초월과의 관계’ 영역에 ‘인간과 로봇은 어떻게 공존할 수 있을까?’ 를 내용 요소로 추가할 것을 제안하였다. 그리고 고등학교 생활과 윤리 교육과정에서는 ‘과학과 윤리’ 내용 영역 중 ‘로봇과 윤리’ 를 내용 요소로 추가하고 하위 요소를 ‘로봇과 인간의 관계’, ‘인간이 지켜야 할 로봇윤리’ 로 제시하였다.

고등학교 선택과목 생활과 윤리 교육과정 ‘로봇과 윤리’ 의 영역에서 학생들은 인간과 로봇의 존재론에 대한 윤리적 성찰을 통한 도덕적 정체성 정립 능력과 로봇이 야기할 윤리적 문제에 대한 도덕적 판단력을 함양할 수 있을 것이다.

3. 도덕과 교육을 통한 로봇윤리교육 방법 제안

본 연구에서는 도덕과 교육을 통한 로봇윤리교육의 방법으로 로봇윤리교육의 목표를 달성할 수 있는 수업 모형을 제안하고자 한다. 앞에서 로봇윤리교육의 목표는 ‘인간과 로봇에 대한 윤리적 성찰 능력’을 강화하는 것과 ‘로봇 관련 문제에 대한 윤리적 해결 능력’을 강화하는 것으로 규명한 바 있다.

먼저, ‘인간과 로봇에 대한 윤리적 성찰 능력’을 강화하기 위한 방법으로 도덕적 상상력과 도덕적 판단력을 함양할 수 있는 프로젝트 수업 모형과 SF영화 활용 수업 모형을 제시할 것이다. 이어서 ‘로봇 관련 문제에 대한 윤리적 해결 능력’을 강화하기 위한 방법으로 도덕적 판단력과 도덕적 추론능력을 함양할 수 있는 ‘도덕적 딜레마 토론 수업 모형’에 대해 소개하고자 한다.

1) 프로젝트 수업 모형

이 항에서는 도덕과 교육을 통해 ‘인간과 로봇에 대한 윤리적 성찰 능력’을 강화할 수 있는 수업 모형으로 프로젝트 수업 모형에 대해 규명하고자 한다. 먼저, 프로젝트 수업의 개념과 구조에 대해 제시하고, 이를 바탕으로 로봇윤리교육을 프로젝트 수업에 적용한 실제 수업 모형을 제시하고자 한다.

(1) 프로젝트 수업의 개념과 구조

프로젝트 수업(project method)은 학생들이 생활에 가치가 있다고 생각되는 문제를 스스로 설정하고, 계획하고, 해결해 나가는 교수·학습 방법이다. 프로젝트 수업의 탐구 과정은 학생들에 의해 계획되고 주도되며 제시되며 프로젝트 주제에 대한 학생들의 생각, 질문, 이론, 예상, 관심에 따라 학생들 스스로 수행 활동과 발표 방법을 결정하는 학습자 중심 접근법

중 하나이다(이정렬, 2016: 8-9). 프로젝트 수업은 무엇인가 마음속에서 생각하고 있는 것을 현실에서 구체적으로 실현하고 형상화하기 위하여 자기 스스로가 계획을 세워 수행하는 방법이다. 따라서 프로젝트 수업은 기본적으로 탐구와 성찰에 유용하고, 실천·체험 중심의 참여 활동을 통해 학생들이 도덕적 의미를 스스로 구성하고 윤리적 성찰 능력을 함양하는 데에 큰 도움이 된다. 또한 학생들이 실생활 문제를 다룬다는 점에서 로봇윤리를 다루기 적합한 교수·학습방법이라는 것을 알 수 있다.

이러한 프로젝트 수업의 목적은 단순한 지식이나 기술뿐 아니라 정서적·심미적 감각까지 계발하는 데에 있다. 프로젝트 수업은 교사가 제시한 문제에 대한 정답을 찾는 것이 아니라 학생과 교사가 함께 어떤 주제가 제시하고 있는 문제에 대한 해답을 찾고자 협력하는 과정에서 그 주제에 대해 보다 많은 것을 깊이 있게 학습해 나가는 교수·학습 활동이다. 따라서 프로젝트 수업은 학생들이 학습 내용 선정에 직접 참여함으로써 그들 자신의 흥미가 충족될 때까지 관심 있는 활동에 몰입할 수 있고, 학습이 실생활과 밀접하게 관련되어 있기 때문에 보다 유의미한 학습이 이루어지며, 학생들이 스스로 창의적인 발견을 하고 나아가 사회와의 관계를 증진시킬 수 있다. 이러한 학생들의 능동적 참여는 학생들의 전인격적인 성장을 가져온다.

도덕과 프로젝트 수업모형을 자세히 살펴보면 아래와 같다.

<표 13> 도덕과 프로젝트 수업모형의 단계(이정렬, 2016: 15)

과정요소	단계	과정
토픽, 학습 내용의 설정	1단계 도덕적 문제 상황 제시	학생들의 실생활과 관련성이 있는 문제 상황을 제시한다.
	2단계 핵심 프로젝트 선정과 확인	학생들이 모둠별 논의를 거쳐 핵심 프로젝트를 선정한다.
	3단계 탐구 문제의 명료화	핵심 주제에 대해 다양한 관점에서 생각하면서 탐구 주제를 세분화하고 관련 추진질문을 만든다.
프로젝트 실행 계획	4단계 모둠별 프로젝트 실행 계획	탐구할 문제들을 인식하고 계획한다.
프로젝트 전개 활동	5단계 모둠별 프로젝트 실행	관련 정보를 수집하고 모둠원들과 아이디어를 교환한다.
발표	6단계 프로젝트 발표와 성찰	탐구 결과를 발표하는 과정을 통해 반성적 성찰을 한다.

위와 같이 1단계에서는 도덕적 문제 상황을 제시한다. 이 단계는 구체적인 탐구 주제를 선정하기보다는 문제의식을 가질 수 있도록 하고 프로젝트의 큰 아이디어를 파악하는 과정이다. 영상이나 사진, 신문 기사 등을 제시하고 탐구해야 할 주제를 파악하고 탐구에 대한 동기를 불러일으키도록 한다. 이 단계에서 학생들의 실생활과 관련성이 있는 맥락적 상황이나 학생들이 관심을 가지는 사회적 이슈를 제시하는 것이 좋다.

2단계에서는 탐구 주제인 프로젝트의 토픽을 정한다. 도덕과 교육과정에서 제시된 내용 요소와 연계하여 학생들의 일상이나 경험과 연계할 수 있는 주제를 파악하도록 한다. 또는 여러 교과 영역이나 학문 영역과 통합할 수 있는 주제를 선택하여 다른 교과와의 융합적 접근을 시도할 수도 있다.

3단계에서는 핵심 주제를 탐구하기 위해 주제를 세분화한다. 이 단계에서 수업을 계획하는 교사는 학생들이 무엇을 알고, 무엇을 할 수 있기를

기대하는지에 대해 분명한 목표를 가지고 학생들의 탐구 과정을 지도한다. 학생들은 핵심 주제에 대해 다양한 관점에서 생각하면서 탐구 주제를 세분화하고 관련 추진 질문을 만들고 개선해나가면서 탐구 주제를 보다 명료하게 설정한다.

4단계에서는 교사가 학생들이 모둠별 탐구 주제에 대한 하위 질문을 열거하면서 탐구할 문제들을 구체적으로 분석하도록 안내한다. 또한 학생들이 정보를 찾을 수 있는 가능한 모든 출처를 생각해보도록 도와주고 탐구 결과를 어떻게 발표할 것인지를 스스로 결정하도록 안내한다. 프로젝트 수업의 탐구 결과 발표를 학생들이 스스로 논의하고 결정하게 하는 것은 그 자체로 창의적인 사고를 발휘하는 과정이며 탐구 과정에서 역할 분담을 구체적으로 계획하고 정보 수집·조직의 방향을 선택하는 데에도 영향을 준다.

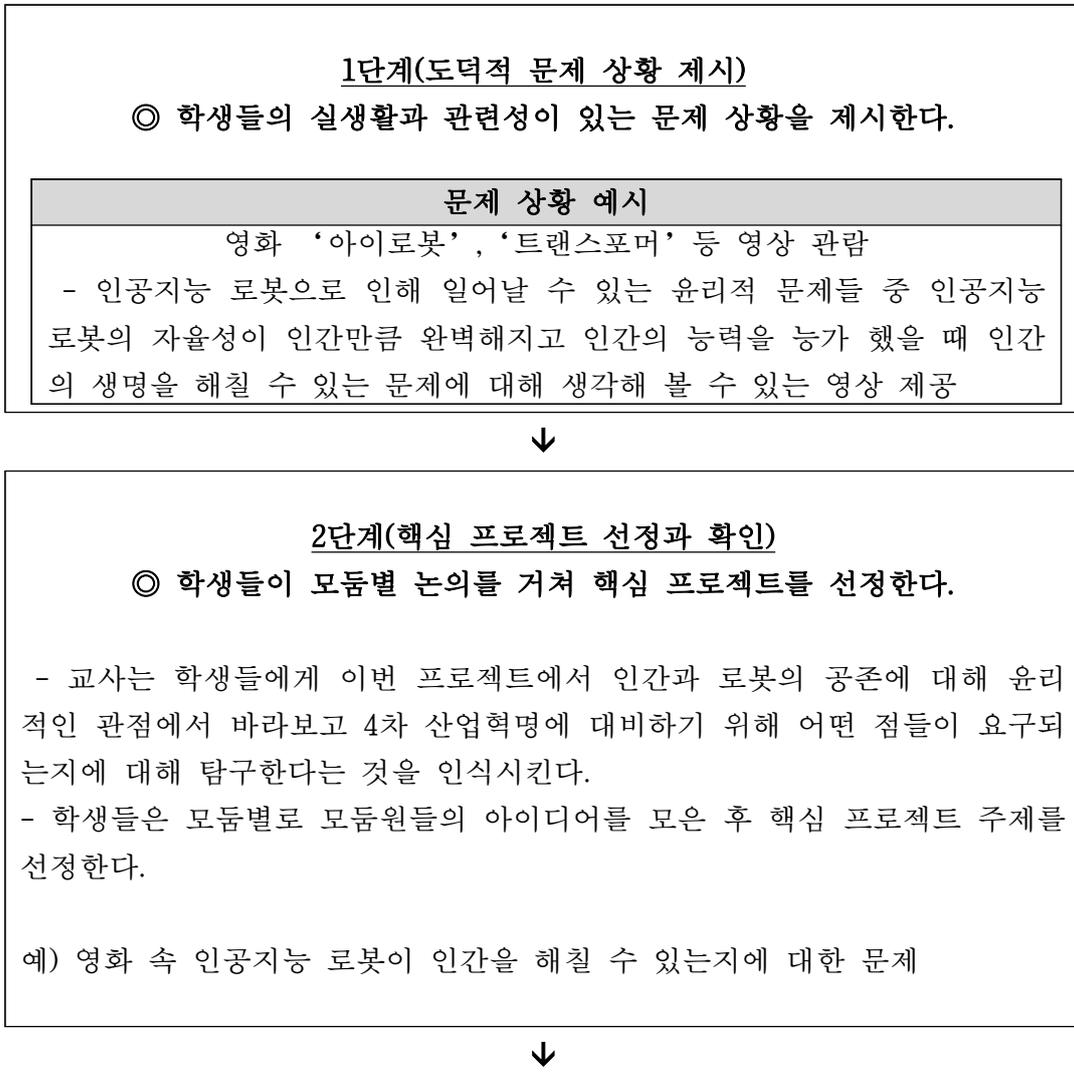
5단계에서는 학생들은 교사의 지도를 받으면서 각자의 수준에 맞게 개별적으로 혹은 다른 학생들과 협력하여 활동한다. 이 단계에서는 탐구하고 협의하고 표현하는 활동들을 주로 실행한다. 학습 활동으로는 조사 활동과 현장 견학, 전문가의 초청 등이 있다. 학생들은 현장 견학을 통해 경험한 여러 가지 현상과 자신의 느낌에 대해 설명하기도 한다.

6단계에서는 학생들이 프로젝트 전 과정을 통해 성취하거나 도달한 결과 또는 결론을 다른 사람들에게 제시하고 의견을 교환하는 단계이다. 탐구 결과를 발표하는 방법은 연극, 신문, 편지, 연설 등 매우 다양하다. 이 단계에서는 프로젝트 전 과정을 통해 이루어진 활동을 평가하고 프로젝트 활동 과정과 결과물 또는 결론을 검토한다. 이 단계에서 학생들은 프로젝트 실행 과정에 대한 반성적 성찰을 하게 된다(이정렬, 2016: 15-17).

(2) 프로젝트학습을 활용한 로봇윤리교육 수업 모형

앞의 항에서 살펴보았듯이 프로젝트 학습은 도덕적 문제 상황 제시, 핵심 프로젝트 선정과 확인, 다양한 관점을 포함한 주제의 세분화와 추진 질문 제시, 모듈별 프로젝트 실행 계획, 모듈별 프로젝트 실행, 프로젝트 발표와 성찰로 구성할 수 있으며 총 6단계로 제시할 수 있다. 프로젝트 수업 모형을 로봇윤리 수업에 적용하면 아래와 같다.

[그림 2] 프로젝트 학습을 활용한 로봇윤리교육 수업 모형



3단계(탐구 문제의 명료화)

◎ 핵심 주제에 대해 다양한 관점에서 생각하면서 탐구 주제를 세분화하고 관련 추진질문을 만든다.

예) 로봇이 인간을 해칠 수 있는 문제

- 로봇도 윤리를 지킬 수 있을까?
- 인간과 로봇의 차이는 무엇인가?
- 로봇도 윤리를 지킬 수 있다면 어떤 윤리를 지켜야 할까?
- 인간은 로봇에게 함부로 대해도 되는가?



4단계(모둠별 프로젝트 실행 계획)

◎ 탐구할 문제들을 인식하고 계획한다.

- 프로젝트에 관한 아이디어를 개발하기 위해 필요한 문헌을 찾아보고 여러 가지 조사 방법에 대한 대안들을 준비하여 교사와 상의한다.
- 프로젝트를 수행하기 위해 필요한 자원들을 알아보고 계획이 완성되면 개인별로 또는 그룹별로 연구계획서를 제작한다.

예) 인공지능 로봇이 인간을 능가하는 자율성을 가진 경우를 가정하여 인간과 로봇이 함께 지켜야 할 윤리기준 만들기



5단계(모둠별 프로젝트 실행)

◎ 관련 정보를 수집하고 모둠원들과 아이디어를 교환한다.

모둠원들의 다양한 관점과 아이디어 발상을 통해 문제 해결의 다양한 방법을 모색하고 아이디어 선택 후 구체적으로 발표를 준비한다.

예) 최초의 로봇윤리인 Asimov(1945)의 로봇윤리 3원칙에 대해 탐구하기

- 로봇 3원칙을 고안한 이유가 무엇인지 토의하고 더 필요한 원칙이 있는

지 논의해보기

- 수정한 원칙을 보완해서 올바른 로봇의 성품에 대해 작성하기



6단계(프로젝트 발표와 성찰)

◎ 탐구 결과를 발표하는 과정을 통해 반성적 성찰을 한다.

프로젝트를 발표하고 프로젝트 과정과 결과에 대한 성찰의 시간을 갖는다.

예) 『인간과 로봇이 함께 지켜야 할 로봇윤리』 개발 내용 발표

- 인간 또한 로봇에게 해를 끼쳐서는 안 된다.
- 인간의 무리한 요구는 로봇이 거부할 권리가 있다.
- 로봇은 인간의 법에 위배된 행동을 해서는 안 된다.
- 인간은 로봇을 함부로 폐기해서는 안 된다.

지금까지 살펴본 프로젝트 학습을 적용한 로봇윤리교육 수업 모형에서는 크게 세 가지의 도덕교육적 효과를 기대할 수 있다.

첫째, 프로젝트 학습은 학생들이 문제 해결을 위해 윤리적 문제를 이해하고 분석하고 추론하는 과정을 통해 도덕적 판단력을 길러줄 수 있다. 프로젝트 학습은 학생들을 복합적이고 참신한 문제에 몰입하게 함으로써 학습의 질을 높이고 높은 수준의 인지 발달을 도모한다(Markham et al., 노선숙 외 역, 5-6). 그리고 Norman et al.(1994: 118-119)의 연구에 의하면, 프로젝트 학습을 한 학습자는 전통적인 학습을 한 학습자보다 가설에 근거한 사고를 하고 보다 정확한 설명과 진단을 내리는 것으로 나타났다. 이와 같이 프로젝트 학습은 윤리적 문제를 해결하는 과정을 통해 학생들이 문제해결에 필요한 도덕적 추론능력과 학습한 내용을 적용할 수 있는 도덕적 문제해결 능력을 길러줄 것이라 판단된다.

둘째, 프로젝트 학습은 학생들이 함께 문제를 해결하고 다른 학생들과의 관계를 형성하고 배려하며 의사소통을 하는 과정을 통해 도덕적 민감성을 향상할 수 있다. 최정임 외(2010: 57)의 연구에 의하면, 프로젝트 학습 중 모둠활동을 통해서 학생들은 의사소통 기술, 상호 존중과 같은 대인관계

기술을 개발할 수 있다. 또한 학생들은 모둠에 기여하는 기술을 배우고, 다른 학생이 모둠에 기여하도록 돕는 방법을 배울 수 있다고 주장했다. 이처럼 프로젝트 학습은 학생들이 모둠활동을 통해 문제를 해결하기 위해 관계를 맺고 배려하며 의사소통을 하는 과정을 통해 도덕적 민감성을 향상할 수 있을 것으로 생각된다.

셋째, 프로젝트 학습은 프로젝트를 선정하고 계획하고 발표하는 과정을 모든 학생들이 협력해서 진행하는 협동학습 활동을 포함하고 있다. 이러한 협동학습 활동을 통해 도덕적 공동체 의식을 함양할 수 있다. 프로젝트 활동은 경쟁적인 분위기가 아닌 협동적 분위기에서 공동의 목적을 위해 협동하는 과정을 통해 학생들에게 공동체로서 학습을 경험할 수 있는 기회를 제공한다(Katz et al., 윤은주 역, 2013: 3). 교사와 학생들이 공동으로 문제를 해결하는 과정을 통해 하나의 강력한 공동체로 만들어주고 공동체 기여가 강조되는 수준 높은 교실을 만들 수 있다(Markham et al., 노선숙 외 역, 2007: 5-6). 이처럼 프로젝트 학습은 공동으로 문제를 해결한다는 측면에서 학생들이 도덕적 공동체 의식을 함양할 수 있는 환경을 제공해 줄 것으로 판단된다.

2) SF영화 활용 수업 모형

이 항에서는 도덕과 교육을 통해 ‘인간과 로봇에 대한 윤리적 성찰 능력’을 강화할 수 있는 수업 모형으로 SF영화 활용 수업 모형에 대해 규명하고자 한다. 먼저, SF영화 활용 수업의 개념과 구조에 대해 제시하고, 이를 바탕으로 로봇윤리교육을 SF영화 활용 수업에 적용한 실제 수업 모형을 제시하고자 한다.

(1) SF영화 활용 수업의 개념과 구조

SF(science fiction)영화는 과학을 주제로 한 소설을 영상으로 옮겨놓은 장르로 과학의 발전과 시대적 흐름에 따라 오락성과 대중성을 강조하면서 인간의 무한한 상상력을 표현할 수 있는 영화이다(배정민 외, 2007: 239). 일반적으로 미래가 배경이 되며 만화적인 내용을 고도의 특수 효과를 이용하여 제시한다. 이들은 공통적으로 미래의 다른 세계를 공상적으로 다루는 한편 강력한 특수 효과로 상상 속 내용이 마치 보는 사람의 눈앞에서 실제로 일어나는 것처럼 묘사한다. 따라서 SF영화를 통해 학생들은 미래 사회를 간접적으로 경험할 수 있다.

로봇윤리를 적용한 도덕교육은 도덕적 상상력을 근간으로 한다. 그것은 단지 어떤 이미지를 형성하는 능력일 뿐만 아니라 현재의 즉각적인 상황을 뛰어넘어 자유롭게 무엇인가를 보고 기술하는 능력이다. 즉, 도덕적 상상력은 자기 자신의 상황이나 입장에서 벗어나 타인의 상황이나 입장에서 문제를 바라보고 이해할 수 있는 능력이다. 로봇윤리는 현재의 문제에 대한 반성이기보다는 미래에 발생 가능한 문제에 대한 해결책을 찾는 것으로 상상력 특히 도덕적 상상력을 필요로 한다. 로봇으로 인해 생겨나는 다양한 문제들은 미래사회에 일어나는 일이기 때문에 로봇이 야기할 수 있는 쟁점이나 문제들에 대한 반성과 성찰이 필요하며 나아가 학생들에게 점차 나타날 수 있는 문제들에 대한 숙고능력을 길러주어야 한다(신현우, 2015: 275-277). 로봇으로 인해 일어날 미래사회의 변화는 SF영화나 드라

마를 통해서 학생들이 간접적으로 경험할 수 있다.

영화 <2001 스페이스 오디세이>³²⁾에서 우주선 디스커버리 호 안에 있는 로봇 HAL 9000은 자기 방어를 위해 우주 비행사에게 거짓된 정보를 제공하기도 하고 자신을 멈추게 하려는 우주 비행사를 제거하기도 한다. 아주 먼 미래에는 로봇이 자신을 만든 인간의 명령을 거부할 수도 있고 속이거나 심지어 해를 가하는 상황을 가정해서 보여준다.

영화 <그녀(Her)>³³⁾에서는 인공지능 로봇인 사만다가 대필작가인 데오도르와 사랑의 대화를 나누며 그를 위로해준다. 그러나 사만다는 다중사용자 운영체제의 성격 상 3천명이 넘는 고객을 상대로 일대일로 맞춤형 매칭 방식을 통해 대화하는 중이었고 사만다에게 데오도르는 3천명 중의 한명일 뿐이었다. 사만다는 인간이 아닌데 많은 사람들로 하여금 자신만을 위하여 존재하는 인간인 것처럼 착각하게 만드는 셈이다. 이처럼 영화 <그녀(Her)>에서는 인공지능 로봇으로 인해 로봇중독에 빠지고 인간 존엄성이 훼손되는 모습을 간접적으로 볼 수 있다.

드라마 <휴먼스 시즌2>³⁴⁾에서는 인간과 똑같은 모습인 휴머노이드 로봇 ‘아니타’가 등장한다. 아니타는 요리도 잘하고 집안 청소도 잘하며 자녀들과 상담도 잘해주는 가사도우미 케어로봇이다. 이러한 아니타의 등장으로 가정 내에 아내이자 엄마의 위상이 갈수록 흔들리게 된다. 케어로봇이 인간의 일상에 너무 깊숙이 개입할 경우 아이가 혼란을 느낄 수 있다는 것을 보여준다(김명주, 2017: 47). 이처럼 미래사회를 배경으로 한 SF영화나 드라마 속 인공지능 로봇이 가져올 다양한 문제점들을 미리 생각해보고 대비하는 활동을 통해 학생들의 도덕적 상상력을 향상시킬 수 있다.

Nussbaum(1997: 86)은 도덕적 상상력이 도덕적 상호작용을 하기 위한 본질적인 준비가 된다고 주장했다. 도덕적 상상력은 어떤 상황에 직면하였을 때도 그 상황을 해결하는 다양한 관점을 상상할 수 있을 뿐만 아니라, 나와 다른 타자의 사고방식도 이해하고 인정해줄 수 있도록 학생들을 변화시킬 것이라고 주장했다. Weisberg 등(1995: 248)도 학생들은 영화를 보

32) S. Kubrick 감독, 1968.

33) S. Jonze 감독, 2013

34) 미국 AMC 채널, 2016.

고 친구들과 서로 토론하는 상상력 활동을 통해 자기 자신과 친구의 관점, 책 속에 나오는 인물의 관점을 경험하게 해줄 것이라 말했다. 또한 어떤 상황에서 다양한 가능성을 보는 능력을 길러 주게 될 것이라 주장했다.

SF영화 활용교육은 교육적 활용이 가능한 매체를 활용하는 측면에서 미디어 활용교육(Media In Education, MIE) 수업모형의 일환으로 볼 수 있다. 미디어 활용교육은 도덕적 민감성, 도덕적 상상력, 비판적 사고력 등을 향상시킬 수 있다는 점에서 도덕교육적으로 중요한 의의를 가진다. 이처럼 미디어 활용교육은 많은 교육적 효과를 가지고 있다.

첫째, 일상적 생활에서 경험하기 어려운 지식과 사실을 시·공간을 초월하여 간접 경험할 수 있으며, 이를 통해 학습의 효율성과 효과성이 증진된다. 추상적인 개념이나 이론만으로는 구체적인 사물 및 사회현상을 명확하게 인식하기 어렵기 때문에 구체적이고 사실적인 영상을 통해 심층적 지식과 정보를 습득하고 응용가능성을 높일 수 있다. 특히 로봇윤리는 미래 사회에서 나타나는 문제이기 때문에 단순하게 텍스트를 읽는 것보다는 인간과 로봇이 공존하는 미래사회를 SF영화를 통해 간접 경험함으로써 로봇윤리에 대해 깊게 생각해볼 수 있는 기회를 제공할 수 있다.

둘째, SF영화는 대체로 학생들에게 인기가 많기 때문에 로봇윤리 학습에 대한 흥미와 동기를 불러일으킬 수 있고 궁극적으로는 학습효과를 높일 수 있다.

셋째, SF영화는 다음 장면을 예측하거나 주제를 찾거나 내용을 종합해보는 등의 다양한 학습활동을 할 수 있고 좋은 영화는 학생들에게 감정이입과 동정심 등의 가치를 일깨워 주고 도덕적인 삶의 필요성에 대한 진지함과 열의를 갖게 한다.

넷째, 영화를 활용한 수업을 통해 교사는 학생들의 문화를 이해하고 공유하는 친근한 교사가 될 수 있으며, 이를 통해 학생들이 수업에 관심을 갖고 적극적으로 참여할 수 있다(정창우, 2007: 278-280).

영화를 활용한 수업의 구조를 자세히 살펴보면 아래와 같다.

<표 14> 영화 활용 수업 모형의 구조(정창우, 2007: 283-284 참조)

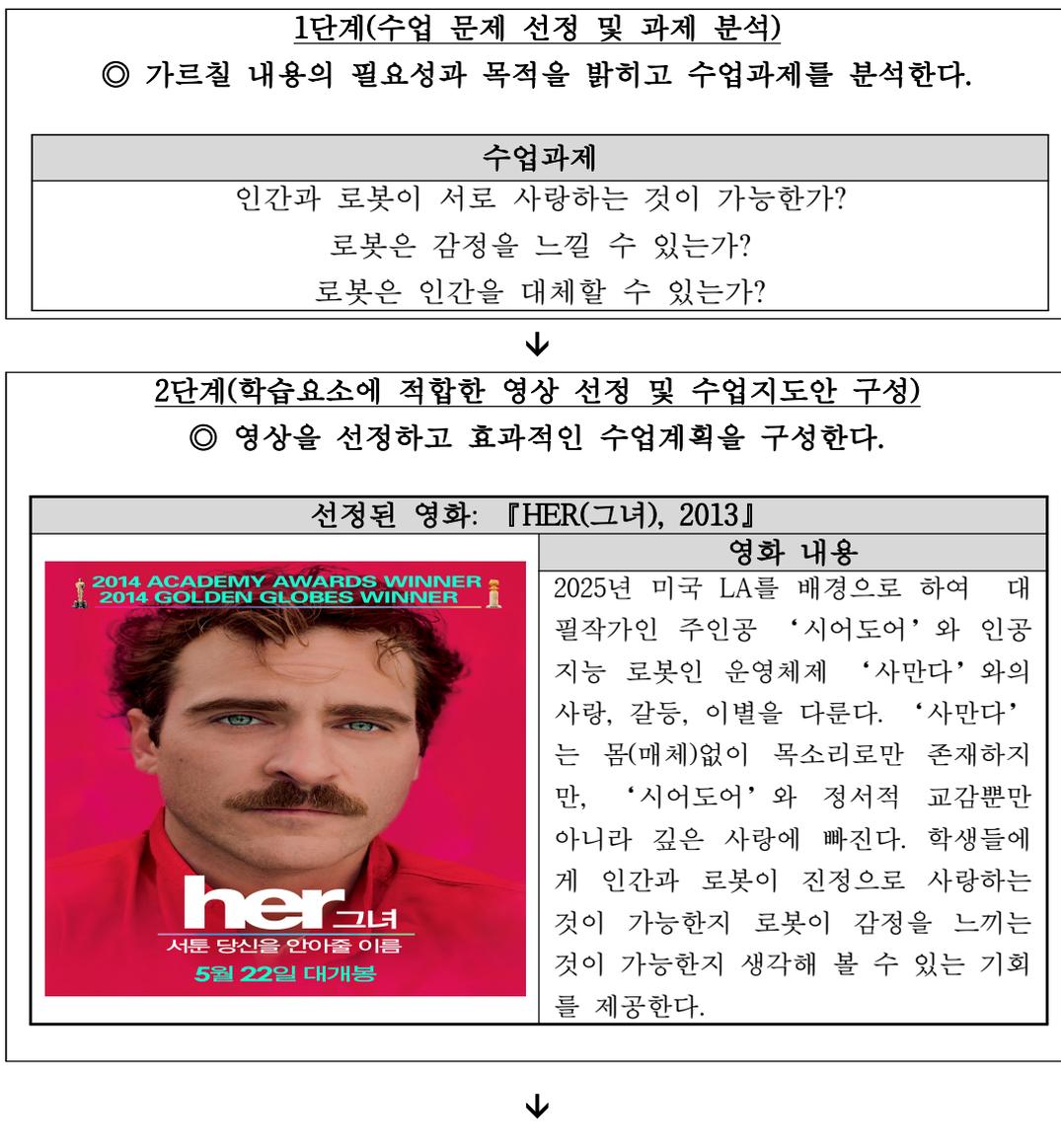
과정요소	단계	과정
수업 전 활동	1단계 수업문제 선정 및 과제 분석	가르칠 내용의 필요성과 목적을 밝히고 수업 과제를 분석한다.
	2단계 학습요소에 적합한 영상 선정 및 수업지도안 구성	효과적인 수업을 만들 수 있는 적합한 영상을 선정하고 다양한 지도방법을 모색한 후 가장 효과적인 수업지도안을 작성한다.
	3단계 학습지 및 멀티자료 제작	교사는 영화내용을 숙지한 후 질문을 만들고 학습지를 작성한다. 학생들의 흥미유발과 참여 활동을 위하여 다양한 멀티자료를 제작한다.
본 수업 활동	4단계 학습동기 유발과 영상 시청	시청 전에 학습목표를 제시하여 동기유발 및 주의를 집중시키고 학습지를 제시하여 감상 후 기록하도록 한다.
	5단계 학생 주도적 교육활동 (개인별 및 모둠별)	영상시청 후에 느낀 소감문 등 학생들의 의견을 정리하고 관련된 학습활동을 진행한다.
수업 후 활동	6단계 발표 및 생각 나누기	영상시청 후에 느낀 소감문 및 활동 결과물 등을 발표하고 서로의 생각을 나눈다.

영화 활용 수업 모형은 수업 전 활동, 본 수업 활동, 수업 후 활동으로 이루어져 있으며 수업 문제 선정 및 과제 분석, 학습요소에 적합한 영상 선정 및 수업지도안 구성, 학습지 및 멀티자료 제작, 학습동기 유발과 영상 시청, 학생 주도적 교육활동(개인별 및 모둠별), 발표 및 생각 나누기의 총 6단계로 구성되어 있다. 이 구조를 토대로 다음 목에서는 로봇윤리교육 수업에 적용해보려 한다.

(2) SF영화를 활용한 로봇윤리교육 수업 모형

앞의 항에서 살펴보았듯이 영화 활용 수업은 수업 전 활동, 본 수업 활동, 수업 후 활동의 과정요소로 구성할 수 있으며 총 6단계로 제시할 수 있다. 영화 활용 수업모형을 로봇윤리 수업에 적용하면 아래와 같다.

[그림 3] SF 영화를 활용한 로봇윤리교육 수업 모형



3단계(학습지 및 멀티자료 제작)

◎ 교사는 영화내용을 먼저 숙지한 후 질문과 학습자료를 제작한다.

학습지 예시

1. 인간과 로봇의 사랑은 가능한가?
2. ‘사만다’도 ‘시어도어’에게 사랑의 감정을 느꼈을까?
3. 내가 주인공이라면?
4. ‘시어도어’가 ‘사만다’가 사실은 ‘시어도어’ 이외에 8,316명과 소통하고 있었다는 사실을 깨닫고 난 뒤에 느꼈던 감정은 어떤 것인가?



4단계(학습동기 유발과 영상 시청)

◎ 시청 전 학습목표와 학습지를 미리 제시하여 동기유발 및 주의를 집중시킨다.

학습목표

1. 인간과 로봇의 차이점을 이해하고 설명할 수 있다.
2. 로봇중독의 개념을 알고 이를 방지하기 위해 노력하는 자세를 가진다.



5단계(학생 주도적 교육활동)

◎ 영상시청 후 개인별 및 모둠별로 학습활동을 진행한다.

모둠별 토론수업 예시

‘인간과 로봇이 사랑에 빠질 수 있다/없다’의 모둠별 토론을 통해 인간과 로봇의 차이점을 자연스럽게 이해하고 로봇에 대한 개념을 명확하게 이해한다. 또한 미래사회에 인간과 로봇의 공존을 위해 어떤 가치관을 가져야 할지 생각해볼 수 있는 기회를 제공한다.



6단계(발표 및 생각 나누기)

◎ 영상시청 후에 활동 결과물 등을 발표하고 서로의 생각을 나눈다.

위와 같이 SF 영화를 활용하여 학생들은 인간과 로봇이 실제로 공존하는 상황에 대해 생각해보고 로봇의 개념을 이해하고 로봇윤리의 필요성을 체감할 수 있다. 그러나 영화는 학생들에게 흥미를 불러일으킬 수 있는 매체이긴 하지만 12분 정도가 넘으면 학생들의 집중도가 떨어져 산만해지기 때문에 시청 중에 학습의욕과 주의집중을 계속 유지시키거나 적절하게 편집된 영상을 사용하는 것이 적절하다. 또한 교사는 영상시청 후 사고하는 교실분위기를 조성하여 도덕적 상상력과 민감성, 비판적 사고와 문제해결 능력을 증진시키도록 노력해야 한다(정창우, 2007: 282-284).

지금까지 살펴본 SF 영화를 활용한 로봇윤리 수업 모형에서는 크게 두 가지의 도덕교육적 효과를 기대할 수 있다. 첫째, SF영화 활용 수업은 인공지능 로봇과 인간이 함께 공존할 미래를 SF영화를 통해 실감나게 체험하고 주인공에게 감정이입을 하는 과정을 통해 미래에 발생할 윤리적 문제에 대한 도덕적 상상력을 함양할 수 있다. Nussbaum(1997: 86)의 연구에 따르면, 영화나 문학과 같이 이야기를 활용하는 수업은 타인에게 공감하고 다양한 종류의 사람들의 구체적인 환경과 문제를 재현하는 능력을 갖도록 해준다. 이를 통해 어떤 상황에 직면하였을 때에도, 그 상황을 해결하는 다양한 관점을 상상할 수 있다고 판단했다. Weisberg 등(1995: 248)과 Johnson(1993: 1-13)의 연구에서도 이야기에 대한 토론은 다른 사람의 입장을 취해보고, 어떤 상황에 대해 다양한 가능성을 생각해볼 수 있는 도덕적 상상력 함양의 효과가 있다는 결과가 나타났다.

둘째, SF영화 활용 수업은 인공지능 로봇으로 인한 윤리적 문제의 해결 방안 도출을 통해 도덕적 판단력을 향상시킬 수 있으며, 윤리적 문제에 대한 실천력 제고를 통해 궁극적으로는 인간과 로봇의 공존을 도모할 수 있다. 이미식 외(2008: 131)의 연구에 따르면, 영화를 활용한 수업은 영화 속에 숨어있는 정치, 경제, 문화와 관련 있는 여러 도덕적 문제에 대해 민감하게 반응하고 자신의 가치관에 따라 도덕적으로 구성하고 조직하는 경험을 제공한다는 결과를 밝혔다.

지금까지의 내용을 종합해보면, SF영화 활용 수업은 향후 인공지능 로봇으로 인해 발생될 윤리적 문제에 대한 도덕적 상상력을 기를 수 있고 윤리

적 문제 해결을 통해 도덕적 판단력을 함양할 수 있다는 점에서 도덕교육적 효과가 있다고 판단된다.

3) 도덕적 딜레마 토론 수업 모형

이 항에서는 도덕과 교육을 통해 ‘로봇의 문제에 대한 윤리적 해결’을 강화하기 위한 수업 모형으로 도덕적 딜레마 토론 수업 모형을 제시하고자 한다. 이를 위해 우선, 도덕적 딜레마 토론 수업 모형의 개념과 구조에 대해 제시하고, 이를 바탕으로 로봇윤리교육을 도덕적 딜레마 토론 수업 모형에 적용한 실제 수업 모형을 제시하고자 한다.

(1) 도덕적 딜레마 토론 수업 모형의 개념과 구조

도덕적 딜레마 토론 수업 모형은 두 가지 이상의 도덕 원칙 사이에 갈등과 충돌이 전개되는 도덕적 딜레마를 활용한 수업으로 도덕과의 대표적인 수업 모형 중 하나이다. 도덕적 딜레마 토론 수업 모형은 딜레마 상황에 대해 분석하고 토론하는 과정을 통해 옳은 선택이 무엇인지, 어떻게 행동해야 하는지에 대해 탐구할 수 있는 모형이다.

딜레마 토론은 윤리적 문제에 대해 탐구하는 과정을 통해 도덕적 추론 능력과 도덕적 판단능력을 신장시킬 수 있다(정창우, 2006: 128). 그리고 딜레마에 대한 토론의 과정을 통해 자신이 추구해야 할 가치에 대해서 숙고하는 기회가 될 뿐만 아니라 다른 토론자들의 의견을 존중하고 배려하는 시민성을 함양하는데도 도움이 된다. 그리고 도덕적 딜레마 토론 수업 모형을 로봇윤리교육 수업에 활용하면, 도덕적 딜레마 상황에 대한 도덕적 판단력과 추론 능력과 함께 복잡한 윤리적 문제를 해결할 수 있는 문제 해결 능력도 함께 신장시킬 것으로 기대된다.

도덕적 딜레마 토론 수업 모형의 절차를 자세히 살펴보면 아래와 같다.

<표 15> 도덕적 딜레마 토론 수업 모형의 절차(정창우, 2006: 130-133)

단계	과정
예비단계: 도덕적 문제 사태의 제시	도덕적 문제 사태로 가상적 딜레마, 내용에 근거한 딜레마, 실생활 딜레마를 제시한다.
도덕적 토론의 도입	교사는 도덕적 이슈를 강조하는 질문, 이유를 묻는 질문, 상황을 복잡하게 만드는 질문을 통해 학생들이 주어진 딜레마를 이해하고 있는지를 확인하고 주어진 딜레마에 대해 학생 자신이 분명한 도덕적 입장을 취하도록 돕는다.
도덕적 토론의 심화	도덕적 추론의 구조적 변화를 위해 하나의 문제에 대해 여러 관점에서 진지하게 숙고할 수 있도록 심층적 질문, 다음 단계의 주장을 강조하는 질문, 역할채택을 위한 질문을 제기한다.
실천 동기 강화 및 생활에의 확대 적용	문제의 해결책에 대해 일상생활에서 학생들이 실천할 수 있도록 동기를 강화한다.

위와 같이 예비 단계에서는 도덕적 문제 사태를 제시한다. 도덕적 딜레마를 선정할 때에는 우선 학생들의 발달 수준에 맞는 내용인지, 문제의 해결책에 대해 학생들이 너무 쉽게 동의할 가능성은 없는지, 딜레마가 너무 개인화되었거나 연령에 적합하지 않아 학생들에게 지나친 정서적 부담을 줄 가능성은 없는지 등에 대해 신중하게 검토해야 한다.

본격적인 도덕적 토론의 도입 단계에서는 교사는 도덕적 이슈를 강조하는 질문을 통해 도덕 딜레마에 대해 학생들이 분명하게 이해하고 있는지를 확인하고, 도덕적 딜레마에 대해 학생들이 분명한 입장을 취하도록 도와야 한다. 또한 교사는 이유를 묻는 질문을 통해 학생들에게 도덕 문제에 대한 그들의 입장을 지지해 주는 이유들을 설명할 것을 요구해야 한다. 교사는 상황을 복잡하게 하는 질문을 통해 도덕 문제의 여러 측면을 탐색하면서 다각도로 생각해 보도록 하기 위해 상황을 더욱 복잡하게 만들어 줄

필요가 있다.

도덕적 토론의 심화 단계에서는 도입단계에서 학생들을 도덕적 논의에 참여시키고 어느 정도 도덕의식을 형성시켰다고 판단될 때, 도덕적 추론의 구조적 변화를 위해 하나의 문제에 대해 여러 관점에서 진지하게 숙고할 수 있도록 하는 데 중점을 두게 한다. 이를 위해 동일한 이슈를 여러 측면에서 상세하고도 심층적으로 탐색하도록 안내하는 심층적 질문을 하거나 학생들의 현재 추론 단계보다 한 단계 더 높은 단계의 추론에 접하게 하여 다음 단계의 주장을 강조하는 질문을 하거나 토론의 내용을 적절하게 다듬고 요약해 줌으로써 학생들이 논의하고자 하는 문제의 본질에서 벗어나지 않도록 한다. 또한 학생들을 자기중심성으로부터 벗어나 타인의 생각과 관점, 사고, 느낌, 권리 등을 다양하게 고려해 보도록 하는 역할 채택을 위한 질문을 할 수도 있다.

마지막 단계에서는 문제의 해결책에 대해 일상생활에서 학생들이 실천할 수 있도록 실천 동기를 강화한다(정창우, 2006: 130-135).

지금까지 살펴본 것과 같이, 도덕적 딜레마 토론 수업모형은 복잡한 도덕적 딜레마에 대한 도덕적 판단력과 도덕적 추론능력을 향상시킬 수 있다는 점에서 로봇윤리교육에 적합한 교수·학습 방법으로 생각된다. 이러한 내용을 바탕으로 다음 목에서는 로봇윤리교육을 도덕적 딜레마 토론 모형에 적용한 실제 수업 모형을 제시하고자 한다.

(2) 도덕적 딜레마 토론을 활용한 로봇윤리교육 수업 모형

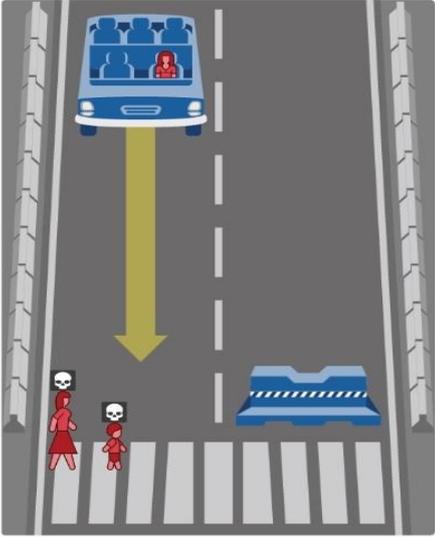
앞에서 살펴본 내용을 토대로 도덕적 딜레마 토론을 활용한 로봇윤리교육 수업모형을 도식화하면 [그림 4]와 같다. 이 수업 모형에서 도덕적 문제로 제시되는 딜레마는 자율주행자동차로 발생할 수 있는 윤리적 문제에 관한 내용이다.

[그림 4] 도덕적 딜레마 토론을 활용한 로봇윤리교육 수업 모형

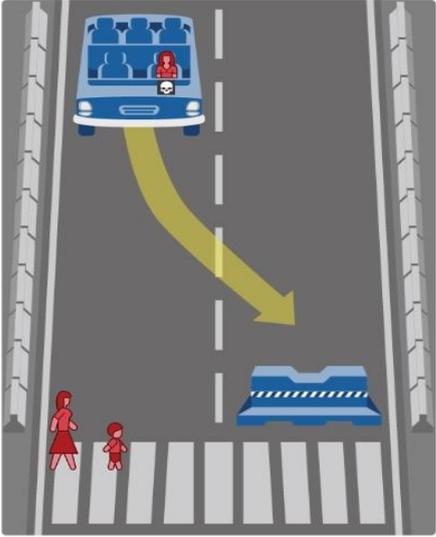
1단계(도덕적 문제 사태의 제시)
◎ 도덕적 문제 사태로 인공지능 로봇으로 인해 발생할 수 있는 가상적 딜레마를 제시한다.

가상적 딜레마 : 자율주행자동차의 윤리적 문제

어떤 사람이 자율주행자동차를 타고 운전하고 있다. 보행자 신호등은 빨간 불인데 직선거리에서 두 사람이 길을 건너고 있다. A와 같이 직진을 할 것인가? B와 같이 당신이 죽더라도 핸들을 돌려서 2명을 살릴 것인가? 공리주의와 의무론 등 윤리이론을 활용하여 생각해보자.



A



B



2단계(핵심문제의 선정 및 윤리적 딜레마 확인)

◎ 찾아낸 문제들 가운데 가장 중요한 문제를 선택한다.

자율주행자동차가 딜레마 상황에 놓였을 때 직진을 해서 살아남을 것인가?
아니면 핸들을 꺾어서 행인 2명을 살릴 것인가?



3단계(도덕적 토론의 도입)

◎ 선택한 문제를 해결할 수 있는 방법을 다양하게 찾아낸다.

- 공리주의에 따르면, 2명의 보행자를 살리고 1명의 차량 소유자가 희생되는 B 상황을 택해야 한다.



4단계(도덕적 토론의 심화)

◎ 여러 관점에서 숙고할 수 있는 심층적 질문을 제기한다.

운전자가 본인이라면 어떻게 할 것인가?(역할채택질문)
공리주의를 적용한 자율주행자동차를 상용화할 수 있을까?(보편적 결과 탐색)



5단계(최선의 대안 도출)

◎ 최선의 해결방법을 실행하기 위한 실행 계획을 수립한다.

인간은 모두 존엄하기 때문에 한명이라도 희생되어서는 안 된다.
따라서 자율주행자동차가 인간의 안전을 완전하게 확보할 수 있을 때까지 상용화되어서는 안 된다. 자율주행자동차가 맞닥뜨릴 수 있는 모든 상황의 선택지를 최대한 자세히 입력하여 인간의 안전을 우선적으로 확보할 수 있도록 한다.

지금까지 살펴본 도덕적 딜레마를 활용한 로봇윤리교육 수업 모형에서는 크게 두 가지의 도덕교육적 효과를 기대할 수 있다.

첫째, 도덕적 딜레마를 활용한 로봇윤리교육 수업 모형은 학생들의 도덕적 판단력과 도덕적 추론 능력을 길러줄 것이다. 인공지능 로봇과 관련된 여러 복잡한 딜레마 상황에 대한 문제를 이해하고 탐구하는 과정을 통해 높은 수준의 도덕적 인지 발달을 도모할 것이며 도덕적 문제 해결 능력도 신장시킬 수 있을 것이다.

둘째, 도덕적 딜레마에 대한 토론에서 다른 학생들과 상호작용 하는 과정을 통해 도덕적 민감성을 형성할 수 있을 것이다(최정임 외, 2010: 57). 학생들은 토론을 통해 의사소통 기술, 상호 존중과 같은 대인관계 기술을 함양할 수 있고 문제를 해결하기 위해 관계를 맺고 배려하는 과정을 통해 시민적 역량을 향상시킬 수 있을 것으로 보인다.

Ⅵ. 결론

최근 우리나라는 제4차 산업혁명의 소용돌이에 빠졌다고 해도 과언이 아닐 만큼 각종 매체에서 제4차 산업혁명에 대한 뉴스가 연일 쏟아져 나오고 있다. 사람들의 높은 관심 속에는 제4차 산업혁명의 핵심기술인 인공지능 로봇이 우리의 일상생활에 들어오면 우리의 생활이 더욱 편리해질 것이라는 기대감도 있지만 인공지능 로봇으로 인한 실직에 대한 불안감, 각종 SF영화들을 통해 생긴 로봇에 대한 공포감 등이 뒤섞여 있다. 실제로 인공지능 로봇으로 인해 직업을 잃고, 자율주행자동차 사고로 인해 생명을 잃고, 로봇과의 결혼을 입법화하려는 움직임까지 있을 정도로 인공지능 로봇은 우리의 삶을 뒤흔들고 있다. 그 중 가장 심각한 문제는 인간의 생명과 안전성, 존엄성 등을 위협하는 윤리적 문제들이다. 인공지능 로봇 기술은 지금도 계속 발전하고 있기 때문에 이러한 변화를 더 이상 막을 수는 없다. 따라서 본 연구에서는 인공지능 로봇으로 인한 윤리적 문제들에 대비하기 위해서 로봇윤리의 의미와 쟁점에 대해 밝히고 이러한 시대적 요구를 받아들여 로봇윤리교육을 제안하였다.

이러한 문제의식을 바탕으로, 이 논문은 도덕과 교육의 입장에서 로봇윤리와 로봇윤리교육의 의미와 내용을 명확하게 정립하고, 로봇윤리를 도덕과 교육에서 다룰 수 있는 방안에 대해 연구하여, 학생들이 실생활에서 일어날 수 있는 윤리적 문제에 대비할 수 있도록 도덕적 정체성 정립 능력과 도덕적 판단력 등을 함양하는 것에 궁극적 목적을 두었다. 이러한 목적에 따른 연구의 결과는 다음과 같다.

첫째, ‘제4차 산업혁명과 로봇’에서는 로봇윤리에 대해 본격적으로 연구하기 전에 로봇윤리가 등장하게 된 이론적 배경인 제4차 산업혁명과 로봇의 개념을 규명하였다. 첫째, 제4차 산업혁명에 대한 정확한 이해를 위해 먼저 다양한 선행연구를 통해 제4차 산업혁명에 대해 새롭게 정의하였다. 제4차 산업혁명이란 인공지능, 사물인터넷, 빅데이터 등과 같은 새로운 신생기술의 융합에 의한 혁신적인 변화 및 현실과 가상의 세계가 융합하는 것을 뜻한다. 그러나 학자들 간에는 제4차 산업혁명은 이전 혁명의 연

장선이라는 주장과 완전히 새로운 혁명이라는 주장이 대치하고 있다. 제4차 산업혁명이 새로운 혁명이라는 의견은 단순히 하나의 발명에 국한되는 것이 아니라 여러 업계의 혁신이 모두 상호 연결되어 있다는 점, 변화의 속도가 굉장히 빠른 점, 하나의 아이디어를 통해서 모든 시스템이 같이 통합되는 점 등을 논거로 제시했다. 이어서 제4차 산업혁명이 가져올 특징과 양면성을 검토한 결과, 가장 주목해야 할 부분은 인공지능 로봇이라는 사실을 밝혔다. 인공지능 로봇에 대한 정확한 개념 정립을 위해 인공지능의 개념, 로봇의 개념에 대해 규명하였다. 다음으로 인공지능 로봇으로 인한 경제적 변화, 사회·문화적 변화, 교육적 변화에 대해 알아보았다. 이상의 논증을 통해 제4차 산업혁명의 핵심 기술은 인공지능 로봇이며, 인공지능 로봇으로 인해 다양한 윤리적 문제가 예상되고 이에 대한 대비가 필요하다는 사실을 밝혔다.

둘째, ‘로봇윤리의 의미와 쟁점’에서는 로봇윤리의 의미와 내용을 정립하고, 로봇윤리의 주요 내용을 규명했다. 로봇윤리의 개념과 범위를 로봇의 윤리와 인간의 윤리로 두 가지로 분류하여 규명하였다. 로봇윤리의 필요성은 인공지능 로봇의 정체성 확립 문제, 인공지능 로봇에 관련된 책임 문제, 인공지능 로봇과 관련된 편견, 인종 차별 문제들을 해결하기 위한 것임을 밝혔다. 마지막으로 여러 학자와 단체들의 선행연구를 바탕으로 하여 로봇윤리의 주요 쟁점으로 로봇의 존재론과 인간 존엄성, 프라이버시, 투명성, 책임을 제시하였다.

셋째, ‘로봇윤리교육의 의미와 내용’에서는 로봇윤리교육의 필요성에 대해 논증하였고 이를 바탕으로 로봇윤리교육의 의미와 내용에 대해 규명하였다. 앞에서 로봇윤리를 분류한 방식으로 로봇의 윤리와 인간의 윤리로 구분하여 로봇윤리교육의 의미와 내용에 대해 밝혔다. 로봇윤리교육의 의미를 ‘인간과 로봇의 공존을 위해 인간과 로봇을 탐구하고 성찰하는 교육’으로 정립하였다. 또한 로봇윤리교육의 목표를 ‘인간과 로봇에 대한 윤리적 성찰’과 ‘로봇의 문제에 대한 윤리적 해결’로 제시하였다. 로봇윤리교육의 목표에 따른 구분을 토대로 로봇윤리교육의 내용 영역을 ‘인간과 로봇의 존재론에 대한 윤리적 성찰’, ‘로봇의 문제에 대한 도덕적

사고와 판단’, ‘로봇에 대한 도덕적 책임’으로 구분하여 구체적 내용 요소를 밝혔다.

넷째, ‘도덕과 교육을 통한 로봇윤리교육 구현방안’에서는 도덕과 교육을 통한 로봇윤리교육의 내용과 방법을 제안하였다. 먼저 로봇윤리교육과 관련 있는 도덕과 교육과정과 교과서 분석을 통해 도덕과 교육을 통한 로봇윤리교육의 가능성을 모색하였다. 교육과정 분석 결과, 로봇을 직접 언급한 부분은 없었지만 목표나 내용면에서 로봇윤리와 관련 있는 부분을 찾을 수 있었다. 목표에 있어서는 삶의 상황 속에서 윤리적 성찰을 중시한다는 점에서 도덕과 목표와 지향점이 동일했고, 중학교 도덕 및 고등학교 선택과목 생활과 윤리 교육과정의 내용 요소가 부분적으로 관련이 있었다. 또한 도덕과에서 로봇윤리교육을 적용하는 것에 대한 의의를 규명하였다. 이를 바탕으로 로봇윤리교육 내용의 현행 도덕과 교육과정 반영 방안에 대해 논의하고, 로봇윤리교육을 활용한 도덕과 교육과정 내용 신설 방안을 제안하였다.

다음으로 도덕과 교육을 통한 로봇윤리교육 수업 모형을 제안하였다. 인간과 로봇에 대한 윤리적 성찰 능력을 함양할 수 있는 방법으로 프로젝트 수업 모형과 SF영화 활용 수업 모형을 제시하였고, 로봇 관련 문제에 대한 윤리적 해결 능력을 함양할 수 있는 방법으로 도덕적 딜레마 토론 수업 모형을 제안하였다.

이상과 같이 이 논문에서는 로봇윤리의 의미와 쟁점 및 로봇윤리교육의 의미와 내용에 대해 규명하고, 도덕과 교육을 통한 로봇윤리교육 내용 및 구현방안을 마련하였다. 서론에서 밝힌 바와 같이, 지금까지 로봇윤리교육의 필요성을 강조하는 목소리에 비해 로봇윤리를 어떻게 가르칠 수 있을지에 대한 체계적인 논의는 부족했다. 본 연구는 이러한 문제인식에 기초하여 로봇윤리의 개념을 정립하고, 로봇윤리를 ‘로봇의 윤리’와 ‘인간의 윤리’로 분류하였다. 또한 로봇윤리의 주요 쟁점을 분석하여 로봇윤리교육의 개념과 목표, 내용을 규명하였다.

이와 같이 본 연구가 갖는 다양한 장점 및 의의에도 불구하고, 본 연구에서 제대로 다루지 못한 부분에 대해 후속 연구의 필요성을 제안하고자

한다. 우선 초등 도덕 교육과정의 내용 체계에도 로봇윤리교육 관련 내용을 반영할 수 있는 방안에 대한 후속 연구가 필요할 것이다. 또한 로봇윤리교육의 교수·학습방법에 대한 효과성 검증 연구가 객관적, 과학적으로 이루어질 필요가 있다. 이런 시도를 통해 로봇윤리교육을 위한 교수·학습방법의 타당성 및 적절성을 검토하고, 이를 바탕으로 로봇윤리교육의 질적 개선을 추구할 필요가 있을 것이다.

<참고문헌>

1. 국내 문헌

1) 단행본

구본권(2015), 『로봇 시대, 인간의 일 : 인공지능 시대를 살아가야 할 이들을 위한 안내서』, 서울: 어크로스.

김선영(2018), 『예술로 읽는 4차 산업혁명』, 서울: 별출판사.

김의중(2016), 『인공지능, 머신러닝, 딥러닝 입문』, 파주: 위키북스.

미래전략정책연구원(2016), 『10년 후 4차 산업혁명의 미래』, 고양: 일상이상.

박찬구(2014), 『개념과 주제로 본 우리들의 윤리학』, 파주: 서광사.

_____ (2016), 『생활 속의 응용윤리: 딜레마와 마주하다』, 서울: 세창.

박찬국(2018), “인간과 인공지능의 미래: 인간과 인공지능의 존재론”, 『인공지능과 새로운 규범』, 한국포스트휴먼연구소 편, 파주: 아카넷.

백종현(2018), “인공지능의 출현과 인간 사회의 변동”, 『인공지능과 새로운 규범』, 한국포스트휴먼연구소 편, 파주: 아카넷.

변순용·송선영(2015), 『로봇윤리란 무엇인가?』, 서울: 어문학사.

손화철 외(2017), 『4차산업혁명이라는 거짓말: 과도한 열풍을 바라보는 여러 가지 시선』, 서울: 북바이북.

양해림 외 6인(2011), 『사이버공간과 윤리』, 대전: 충남대학교 출

판문화원.

- 오윤자(2018), 『철학과 함께 춤을』, 서울: 북마크.
- 이민화(2016), 『4차 산업혁명으로 가는 길』, 서울: 창조경제연구회.
- 이상헌(2012), 『융합시대의 기술윤리』, 서울: 생각의 나무.
- 이종관(2017), 『포스트휴먼이 온다』, 고양: 사월의책.
- 이종승(2009), 『교육·심리·사회 연구방법론』, 파주: 교육과학사.
- 이중원(2018a), 『인공지능의 존재론』, 파주: 한울아카데미.
- 정재승(2018), 『열두발자국』, 서울: 어크로스.
- 정창우(2006), 『도덕교육의 새로운 해법』, 서울: 교육과학사.
- _____ (2007), 『도덕과 교수·학습방법 및 평가』, 서울: 인간사랑.
- _____ (2013), 『도덕과 교육의 이론과 쟁점』, 서울: 울력.
- 최정임·장경원(2015), 『PBL로 수업하기』, 서울: 학지사.
- 카카오 편집진(2018), 『카카오 AI 리포트』, 서울: 북바이북.
- 하대청(2018), “생명의료 기술의 발전과 과학시민권”, 『제4차 산업혁명 시대 인문정책 방향』, 정원섭(편), 서울: 진한엠앤비.

2) 국역본

- Beck, U. 홍성태 역(1997), 『위험사회: 새로운 근대를 향하여』, 서울: 새물결.
- Bennis, W. G. & Goleman, D. & O’ Tool, J. M. 배인섭 역(2008), 『투명성의 시대: 미래 기업의 절대 조건』, 서울: 엘도라도.
- Brynjolfsson, E. & McAfee, A. 이한음 역(2014), 『제2의 기계 시대: 인간과 기계의 공생이 시작된다』, 서울: 청림출판.
- Capurro, R. & Nagenborg, M. 변순용·송선영 역(2013), 『로봇윤리: 로봇의 윤리적 문제들』, 서울: 어문학사.

- Dewey, J. 이재언 역(2003), 『경험으로서의 예술』, 서울: 책세상.
- Fukuyama, F. 최준명·송정화 역(2003), 『부자의 유전자, 가난한 자의 유전자』, 서울: 한국경제신문
- Harari, Y. N. 김명주 역(2017), 『호모데우스』, 파주: 김영사.
- Kaplan, J. 신동숙 역(2016), 『인간은 필요 없다』, 서울: 한스미디어.
- Katz, L. G. & Chard, S. C. 윤은주·이진희 역(2013), 『Katz와 Chard의 프로젝트 접근법』, 파주: 아카데미프레스.
- Markham, T., et al. 노선숙·김민경·임해미 역(2007), 『프로젝트 기반학습 입문서』, 파주: 교육과학사.
- Rachels, J. 김기덕·노혜련·박소영 역(2009), 『도덕철학의 기초』, 서울: 나눔의집.
- Roland Berger. 김정희·조원영 역(2017), 『4차 산업혁명 이미 와 있는 미래』, 파주: 다산 3.0.
- Russell, S. & Norvig, P. 류광 역(2016), 『인공지능 1: 현대적 접근 방식』, 파주: 제이펍.
- Vogel, D. 김민주·김선희 역(2006), 『기업은 왜 사회적 책임에 주목하는가』, 서울: 거름.
- Walsh, T. 이기동 역(2018), 『생각하는 기계』, 서울: 프리뷰.

3) 논문

- 고인석(2011), “아시모프의 로봇 3법칙 다시 보기”, 『철학연구』 제93집.
- 고미숙(2013), “도덕적 상상력 함양을 위한 교육-초등학교 도덕교육을 중심으로-”, 『교육문제연구』 제19집.

- 김동창(2017), “인간 향상의 윤리적 쟁점과 도덕교육적 의미”, 『윤리교육연구』 제44호
- _____ (2019), “제4차 산업혁명시대와 도덕과 교육”, 『윤리교육연구』 제53호.
- 김명주(2017), “인공지능 윤리의 필요성과 국내외 동향”, 『정보와 통신』 제34집 제10호.
- 김선주(2017), “인성교육 진흥법 핵심가치덕목에 근거한 대학생용 인성 척도 개발”, 『교양교육연구』 제11집 제3호.
- 김지원(2017), “인공지능 시대의 새로운 교육패러다임에 대한 연구”, 경북대학교 박사학위 논문.
- 김혜진(2015), “도덕적 동기화의 원천으로서 도덕적 책임의 역할과 형성 방안 연구”, 서울대학교 박사학위 논문.
- 남중권(2019), “머신러닝 알고리즘과 차별”, 고려대학교 박사학위 논문.
- 류성창·강태훈(2018), “4차 산업혁명 대비 미래 교육과정 및 교육방법 방향 탐색”, 『교육연구』 제72집.
- 박주현·한정혜(2018), “초등학생을 위한 로봇윤리 교육 내용 개발”, 『창의정보문화연구』 제4권 2호.
- 박유신·조미라(2017), “미래사회를 위한 포스트휴먼 교육”, 『미술교육논총』 제31권 2호.
- 배정민·김영삼(2007), “SF영화 의상에 표현된 중세적 요소에 관한 연구”, 『한국의류학회 학술발표논문집』.
- 백종현 외 5인(2017), “제4차 산업혁명과 포스트휴먼 사회”, 『철학과 현실』 제112집.
- 변순용·송선영(2014), “공학윤리교육으로서의 로봇윤리프로그램 개발에 대한 연구”, 『2014년 한국교양교육학회 춘계학술대

- 회 자료집』.
- _____ 외 3인(2017), “로봇윤리현장의 필요성과 내용에 대한 연구”, 『윤리연구』 제112집.
- _____ (2018), “인공지능로봇을 위한 윤리 가이드라인 연구 - 인공지능로봇윤리의 4원칙을 중심으로”, 『윤리교육연구』 제47집.
- 송선영(2017), “로봇과 인공지능 시대의 시민윤리와 도덕교육적 함의-인공지능형 로봇의 활용을 중심으로”, 『윤리연구』 제115집.
- 신현우(2015), “도덕과 교육에서 로봇윤리의 교육적 함의”, 『윤리연구』 제105집.
- 안민호(2014), “SNS는 얼마나 동종애적인가? -SNS와 오프라인의 정치적 동질성 비교 분석”, 『한국방송학보』.
- 안상희·이민화(2016), “제4차 산업혁명이 일자리에 미치는 영향”, 『한국경영학회 통합학술발표논문집』.
- 오하경·김태성(2017), “인공지능 비서 이용과 프라이버시 침해에 대한 사용자 인식 연구”, 『한국경영정보학회 학술대회논문집』 제2017집 제12호.
- 윤상균(2015), “과학기술사회에서 요구되는 시민성 탐구”, 『시민교육연구』 제47집 제4호.
- 원동규·이상필(2016), “인공지능과 제4차 산업혁명의 함의”, 『ie매거진』 제23집 제2호.
- 이운복(2009), “과학, 기술 그리고 윤리: 과학기술의 책임에 대한 윤리적 고찰”, 『철학연구』 제112집.
- 이정렬(2016), “도덕적 창의성 함양을 위한 프로젝트 수업”, 『윤리교육연구』 제41집.

임종헌 외 2인(2017), “4차 산업혁명사회에서 교육의 방향과 교원의 역량에 관한 탐색적 연구”, 『한국교육』 제44집 제2호.

전준영 외 2인(2018), “개인화 알고리즘으로 필터 버블이 형성되는 과정에 대한 검증”, 『멀티미디어학회논문지』 제21권 제3호.

정창우(2019), “미래 지향적인 도덕과교육의 설계와 실천”, 『한국도덕윤리과교육학회 학술대회 자료집』 제8호.

조난심(2017), “제4차 산업혁명과 교육”, 『교육비평』 제39호.

조현국(2017), “4차 산업혁명에 따른 대학교육의 변화와 교양교육의 과제”, 『교양교육연구』 제11권 제2호.

태혜신·김선영(2018), “인공지능과 예술의 융합 양상에 관한 탐색적 고찰”, 『한국무용과학회지』 제36권 제2호.

하영숙·구인회(2018), “인공지능이 인간 존엄성에 미치는 영향에 대한 연구”, 『가톨릭신학』 제32집.

4) 교과서

교육부(2015), 『도덕과 교육과정』, 교육부 고시 제 2015-74호 [별책 6].

변순용 외 10인(2017), 『고등학교 생활과 윤리』 ①, 서울: 천재교과서.

정창우 외 10인(2017), 『고등학교 생활과 윤리』 ②, 서울: 미래엔.

_____ 외 9인(2017), 『고등학교 윤리와 사상』, 서울: 미래엔.

2. 국외 문헌

1) 단행본

Bekey, G.(2005), *Autonomous Robots: From biological inspiration to implementation and control*, Cambridge : MIT Press.

Brooks, R.(2002), *Flesh and Machines: How Robots will change us*, Cambridge : MIT Press.

Johnson, M.(1993), *Moral Imagination*, Chicago : The University of Chicago Press.

Kurzweil, R.(2010), *The singularity is near: when humans transcend biology*, NY : Gerald Duckworth&Co.

Mataric, M. J.(2007), *The robotics primer*, Cambridge, MA : MIT Press.

Moravec, H.(1988), *Mind children : the future of robot and human intelligence*, London : Harvard University Press.

Nussbaum, M.(1997), *Cultivating Humanity*, Cambridge : Harvard University Press.

Patrick, L. & Bekey, G. & Abney, K. (2012), *Robot ethics : the ethical and social implications of robotics*, London : MIT Press.

Rathenau Institut.(2017), *Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality*, The Hague : Rathenau Instituut.

Schwab, K.(2017), *The fourth industrial revolution*, New York : Crown Business.

_____ (2018), “Foreword” , Doucet, A., et al.(2018), *Teaching*

- in the fourth industrial revolution : standing at the precipice*, NY : Taylor&Francis.
- Soskil, M.(2018), “Education in a Time of Unprecedented Change” , Doucet, A., et al.(2018), *Teaching in the fourth industrial revolution : standing at the precipice*, NY : Taylor&Francis.
- Tamburrini, G.(2006), “Artificial intelligence and Popper’ s solution to the problem of induction” , Jarvie. I. C., et al.(2006), *Karl Popper : a centenary assessment, Metaphysics and epistemology*, London : Ashgate.
- Turner, J.(2019), *Robot rules : regulating artificial intelligence*, Cham : Palgrave Macmillan.
- Tzafestas, S. G(2016), *Roboethics: a navigating overview*, Cham : Springer.
- Vallor, S.(2016), *Technology and the virtues : a philosophical guide to a future worth wanting*, New York : Oxford University Press.

2) 논문

- Bostrom, N. (2007), “Ethical Principles in the creation of artificial minds” , *Linguistic and Philosophical investigations*, 6(1).
- Bonnefon et al.(2015), “Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?” , ArXiv.
- Bringjord, S(2004), “On Building Robot Persons: Response to Zlatev” , *Minds and Machines*, 14(3).
- Brynjolfsson, E. et al.(2017), “Artificial intelligence and the

- modern productivity paradox: A Clash of expectations and statistics” , NBER WORKING PAPER SERIES, 24001(1).
- Chang, C. et al(2010), “Exploring the Possibility of Using Humanoid Robots as Instructional Tools for Teaching a Second Language in Primary School” , Journal of Educational Technology & Society, 13(2).
- Colombo, A. W. et al.(2017), “Industrial Cyberphysical Systems: A Backbone of the Fourth Industrial Revolution” , IEEE Industrial electronics magazine, 11(1).
- Danovitch, J. H. & Frank, C. K.(2008), “Young Humeans: the role of emotions in children’s evaluation of moral reasoning abilities” , Developmental science, 11(1).
- Frey, C. B. & Osborne, M. A.(2013), “The future of employment: How susceptible are jobs to computerisation?” , Technological Forecasting and Social Change, 114.
- Mark, W & Jacalyn, D.(1995), “Evoking the Moral Imagination” , Change, 27(1).
- Murphy, R. R. & Woods, D. D.(2009), “Beyond Asimov: The Three Laws of Responsible Robotics” , IEEE Intelligent Systems, 24(4).
- Norman, G, R. et al.(1994), “Cognitive differences in clinical reasoning related to postgraduate training” , Teaching and Learning in Medicine, 6.
- Patrick, L. & Bekey, G. & Abney, K. (2011), “Robot ethics: Mapping the issues for a mechanized world” , Artificial Intelligence 175.

- Salvini, P. et al.(2010), “How safe are service robots in urban environments? Bullying a robot” , IEEE.
- Serholt, S. et al.(2017), “The case of classroom robots : teachers’ deliberations on the ethical tensions” , AI & Society, 32(4).
- Sharkey, A. & Sharkey, N.(2010), “The Crying Shame of Robot Nannies: An Ethical Appraisal” , Interaction Studies, 11(2).
- _____ (2012), “Granny and the robots: ethical issues in robot care for the elderly” , ETHICS AND INFORMATION TECHNOLOGY, 14(1).
- Stahl, B. C.(2004), “Information, Ethics, and Computers: The Problem of Autonomous Moral Agents” , Minds and Machines, 14(1).
- Tanaka, F. & Cicourel, A. & Movellan, J. R.(2007), “Socialization between toddlers and robots at an early childhood education center” , Proceedings of the National Academy of Sciences, 104(46).
- _____ & Kimura, T.(2009), “The use of Robots in Early Education” , The 18th IEEE International Symposium on Robot and Human Interactive Communication.
- Turing, A. M.(1950), “Computing Machinery and Intelligence” , Mind, 59(236).
- Tzafestas, S. G.(2018), “Roboethics: Fundamental Concepts and Future Prospects” , Information, 9(6).
- Veruggio, G. & Solis, J. & Van der Loss, M.(2011), “Roboethics: Ethics Applied to Robotics” , IEEE robotics & automation

magazine, 18(1).

Wallach, W. & Colin, A. & Smit, I.(2008), “Machine morality: bottom-up and top-down approaches for modelling human moral faculties” , AI & Soc, 22.

Weisberg, M. & Duffin, J.(1995), “Evoking the moral imagination” , Change, 27(1).

3. 보고서

- 과학기술정보통신부·한국정보화진흥원(2017), “지능정보사회 윤리 가이드라인” .
- 강태중 외 3인(2016), “2016 지능정보사회를 위한 교육 발전 전략 구상” , 한국교육개발원 수탁연구 CR 2016-25.
- 고학수 외 5인(2018), “2018 NAVER Privacy White Paper” , NAVER Corp. 2018.12.12.
- 국토교통부(2020), “세계 최초 부분자율주행차[레벨3] 안전기준 제정” , 2020.1.3.
- 금융보안원(2018), “일본 총무성, AI 연구개발 가이드라인(draft) 주요 내용” , 2018.2.7
- 류성창(2017), “지능정보사회를 선도할 미래 인재상 및 교육체제의 변화 전망” , 제1차 KEDI 미래교육정책포럼 자료집, 2017.3.10.
- 이중원(2018b), “비인간 인격체로서의 AI로봇에 대한 존재론적 고찰” , 제2회 윤리적 인공지능 로봇 워크숍 자료집 , 2018.6.1.
- 한국고용정보원(2016), “2017 한국직업전망” , 고용노동부. 2017.04.07
- 한국과학기술평가원(2018), “2018년 기술영향평가 결과보고서-블록체인
인의 미래” , 한국과학기술평가원. 2019.04.16.
- 한국정보화진흥원(2016), “지능정보사회의 문화예술변화와 전망” , 『정보문화 이슈리포트』 , 16-03호, 2016.12.
- Danielson, P.(2002), “Video surveillance for the rest of us: proliferation, privacy, and ethics education” , 『IEEE 2002 International Symposium on Technology and Society』 .

- IEEE(2016), “Ethically Aligned Design: A vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems” , IEEE.
- Senior Project(2008), “Ethics of e-Inclusion of older people” , 『Senior Discussion Paper』 .
- The Conference toward AI Network Society(2017), “Draft AI R&D GUIDELINES for International Discussions” , 2017.7.28.
- UNESCO COMEST(2005), “The Precautionary Principle” , Paris: UNESCO.
- _____ (2017), “Report of COMEST on Robotics Ethics” , Paris: UNESCO.
- Veruggio, G.(2006), “The EURON Roboethics Roadmap”, 6th IEEE-RAS International Conference on Humanoid Robots.
- World Economic Forum(2015), “New vision for education: Unlocking the potential of technology” , Geneva: World Economic Forum.
- _____ (2016), “The future of job” , 『Global challenge insight report』 , World Economic Forum.
- Zhang, G. & Zhang, J.(2008), “The Issue of Robot Education in China’s Basic Education and its Strategies” , IEEE Conference on Robotics.

4. 기타

임세정 · 이재연, “카이스트 發 ‘AI 킬러로봇’ 논란... 세계 로봇학자들 ‘보이콧’ 왜? ”, 『국민일보』, 2018. 4. 6.

<https://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=105&oid=005&aid=0001086741>(검색일: 2020. 1. 29 기준)

Engineering and Physical Sciences Research Council(2010), “Principles of robotics”, <https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/>(검색일: 2020. 1. 29 기준)

Future of Life(2017), “Asilomar AI Principles”, <http://futureoflife.org/ai-principles/> (검색일: 2020. 1. 29 기준)

Holley, P., “Tech leaders: Killer robots would be ‘dangerously destabilizing’ force in the world”, 『The Washington Post』, 2018. 7. 19.

<https://www.washingtonpost.com/technology/2018/07/19/tech-leaders-killer-robots-would-be-dangerously-destabilizing-force-world/>(검색일: 2020. 1. 29 기준)

Levin, S., “A beauty contest was judged by AI and the robots didn’ t like dark skin”, 『The Guardian』, 2016. 9. 8.

<https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people>(검색일: 2020. 1. 29 기준)

Moore, R. K.(2015), “AI Ethics: Artificial Intelligence, Robots, and

Society?” , <http://www.cs.bath.ac.uk/~jib/web/ai.html>(검색일: 2020.

1. 29 기준)

Popely, R., “Are America’ s best years of growth behind us?” ,
『Northwestern Now』 , 2016. 4. 28.

<https://news.northwestern.edu/stories/2016/04/gordon-v-mokyr-american-economic-growth>(검색일: 2020. 1. 29 기준)

RoboLaw(2017), “Guidelines on Regulating Robotics” http://www.robolaw.eu/RoboLaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf(검색일: 2020. 1. 29 기준)

박영선 대표발의, 로봇기본법(안), 의안번호 8068, 2017.7.19.

산업자원부, 로봇윤리헌장(초안), 2007.

UN, 세계인권선언, 1948.

Abstract

A Study on the Application of Roboethics to Moral Education

Ji, Hyun A

Department of Ethics Education

The Graduate School

Seoul National University

With the emergence of the Fourth Industrial Revolution, a variety of ethical problems due to artificial intelligence (AI) robots are expected, and much attention has been paid to the importance of roboethics that can solve such problems. Accordingly, more and more people are agreeing that we need a new change in school education to apply roboethics. Thus, this study aims to search “the application of roboethics to moral education” focusing on roboethics as a core factor to solve the ethical problems caused by AI robots.

However, compared to the high interest in the need for roboethics, it has been insufficient to systematically discuss the application of roboethics to school education. In light of this problem awareness, this study identified the meaning and category of roboethics, and then established the meaning and contents of

roboethics education based on the identified meaning and category to finally propose the application to embody roboethics education through moral education.

Roboethics means ethics that humans and robots shall follow for the co-existence of humans and robots. In this study, roboethics was categorized into two: robot's ethics that robot shall follow, and human ethics that human shall follow, and proposed each of the contents and issues. Some of the main issues discussed in relation to robot's ethics was a moral status of robot, and in relation to human ethics, human dignity, privacy, transparency, and responsibilities were discussed.

Furthermore, this study identified the meaning and contents of roboethics education based on the meaning and issues of roboethics. To specify the roboethics education, this study searched the need, meaning, purpose, and contents of the roboethics education based on the categories of robot's ethics and human ethics. The education about robot's ethics was needed to re-establish humanity in the era of post-humanism, and the education about human ethics was needed to foster the citizenship of science and technology. In this study, roboethics education means education that explores and reflects on human and robots for the co-existence of humans and robots. The roboethics education based on the above definition aims to achieve ethical reflection on human and robots and solve the ethical problems of robots. To achieve the above goals, this study suggested the following contents of the roboethics education.

First, this study set the ethical reflection on the ontology of human and robot as the content area to achieve the goal of ethical reflection on humans and robots. The specific content elements discussed in this area were derived as follows : moral status of AI robots and the identity of humans and robots.

Second, this study set the moral thinking and judgment about robot problems and moral responsibilities of robots as the content area to achieve the goal of the ethical solution of robot problems. In the area of moral thinking and judgment about robot problems, human dignity, privacy, and transparency were derived as specific content elements, and in the area of moral responsibilities of robots, a sense of responsibility about roles and citizenship of science and technology were derived as specific content elements to be discussed.

This study analyzed moral education curriculum and textbooks in relation to roboethics education by utilizing the above content elements as the analysis framework. The analysis results on the education curriculum exhibited that there were some parts related to roboethics education in the purpose and contents although no direct mention about robots was found. Through this result, the validity of the application of roboethics education in moral education was proven. Furthermore, roboethics education will contribute to the specification of self-identify establishment process as a moral agent, as well as the expansion of the contents on ethics education of science and technology in moral education.

Based on the validity of the application of roboethics education in moral education curriculum, this study suggested a practical plan to improve the contents in roboethics education through moral education and methods of roboethics education. First, this study suggested a plan to reflect the roboethics education contents in the current moral education curriculum and a practical plan to newly establish a moral education curriculum in relation to roboethics education, thereby aiming to embody roboethics education in moral education. Furthermore, this study suggested a project method, science fiction movie-utilized method, and moral dilemma debate method as the method of roboethics education through moral

education.

This study aimed to embody the roboethics education in moral education as it focused on the need for roboethics education according to the emergence of the Fourth Industrial Revolution era. To do this, this study clearly identified the meaning and category of roboethics. In addition, this study clarified the meaning and contents of roboethics education based on the roboethics category. This study is significant for suggesting the practical plan to apply the moral education of roboethics by clarifying the meaning and contents of roboethics education based on the roboethics category. The significant implications of this study are as follows: it not only attempted the discussion about the meaning and contents of roboethics education, which was not dealt with in previous studies, but also discussed the application of roboethics education through moral education.

Keywords : The Fourth Industrial Revolution, roboethics, roboethics education, moral education

Student Number : 2013-31094