

인공 지능 시대의 정보 윤리학:

플로리디의 ‘새로운’ 윤리학

목 광 수[†]

플로리디는 인공지능 시대의 정보 사회에서 나타나는 인공적 해악과 같은 새로운 윤리적 문제들이 기존의 윤리학을 통해 적절하게 대응될 수 없다는 것을 보이고, 이에 대한 대안으로 자신의 정보 철학에 토대를 둔 정보 윤리학을 제시한다. 플로리디의 정보 윤리학은 기존의 철학과 윤리학에서 포착되기 어려운 윤리적 문제와 존재들을 추상화 방법론을 통해 효과적으로 포착할 뿐만 아니라, 이러한 문제들을 해결하기 위해 형이상학적 엔트로피를 줄이고 정보 환경과 정보 유기체의 변형을 도모해야 한다는 정보 윤리의 원칙을 제시한다. 이러한 플로리디의 정보 윤리학은 기존의 윤리학과 차별되는 피동자와 존재 중심적인 거시 윤리학이다. 본 토론 논문의 목적은 이렇게 제시된 플로리디의 윤리학이 이 시대에 적합한 정보 윤리학인지를 비판적으로 검토하고자 한다. 특히 플로리디의 정보 윤리학이 인공지능 시대의 ‘새로운’ 윤리학이 될 수 있을지를 검토하고, 이러한 목표를 위해 보완할 점을 제안하고자 한다. 본 논문은 플로리디의 정보 윤리학이 제시하는 인식론적 방법과 논의가 사실과 가치를 구분하지 못해 야기되는 자연주의적 오류로 인해 도덕적 차원에서는 정당화될 수 있을지 의심스러울 뿐만 아니라, 상대주의적 비판에도 취약하다고 분석한다. 본 논문은, 이러한 한계를 보완하기 위해 플로리디의 정보 윤리학은 이론의 토대가 되는 도덕성 또는 규범성 개념에 대한 설명을 보충해야 할 뿐만 아니라, 정보 윤리학이 작동하는 추상화 층위를 도덕적으로 정당화할 필요가 있으며, 정보 윤리학이 기존 윤리학을 대체하거나 포괄하기보다는 기존 윤리학과 협력적 관계 설정을 통해 체계화할 필요가 있다고 주장한다.

【주요어】 정보 윤리학, 정보 환경, 정보 유기체, 추상화 방법론, 피동자와 존재 중심적인 거시 윤리학

[†] 서울시립대학교 철학과, epistle7@uos.ac.kr

1. 들어가는 글

옥스퍼드 대학의 철학과 정보 윤리 교수인 플로리디(Luciano Floridi)는 이탈리아 로마 라 사피엔차 대학(Sapienza - Università di Roma)에서 철학사를 공부하고 영국의 워릭 대학(University of Warwick)에서 분석철학 분야에서 석사를 마쳤으며 논리학과 인식론에 대한 논문으로 박사 학위를 받았다. 플로리디는 이러한 철학적 토대 위에서 “정보 철학(philosophy of information)은 우리 시대의 철학이다”라는 믿음 아래, 1990년대부터 정보 철학과 관련된 논문을 발표하여 학계의 주목을 받게 된다.¹⁾ 그는 ‘정보 철학의 원리(*Principia Philosophiae Informationis*)’라는 제목의 기획 아래 정보 철학과 관련된 4권의 저술을 계획하면서 자신이 발표한 논문들을 토대로 2011년에 *The Philosophy of Information*을 출판하였고, 그 후속작으로 2013년에 *The Ethics of Information*을 발간했다.²⁾ 플로리디의 논문과 저서는 서구 학계에 큰 관심을 일으켜 정보 철학과 정보 윤리학에 대한 논쟁을 주도해 나가고 있다.³⁾ 본 토론 논문은 플로리디 기획의 두 번째 저술인

1) Floridi (2013), p. 331.

2) Ibid., p. xiv. 플로리디는 발표된 두 권의 책은 서로 독립적이지만, 동시에 상보적이라고 언급하고 있다(Floridi 2013, p. xiv). *The Philosophy of Information*(2011)은 플로리디 정보 철학 기획의 토대가 되는 저서로서, 정보의 개념적 토대를 정립하고 정보 철학이 논의하는 문제 그리고 이에 접근하는 추상화 방법론(the method of abstraction)을 제시하는 메타 이론적 목표, 다양하고 복잡한 정보 개념과 현상에 대한 논의를 다루는 입문적(introductory) 목표, 정보 개념을 통해 의미론적이고 인식론적이며 논리적인 기존의 철학적 논의를 재해석하고 대안적 방안을 제시하는 건설적 목표에 따라 기술되었다. 이러한 이론적 토대 위에서 두 번째 저서인 *The Ethics of Information*(2013)은 정보와 관련된 윤리적 문제들을 다루고 있다. 플로리디는 아직 출판되지 않은 세 번째 책인 *The Policies of Information*은 기존 저서들에서 논의되지 않았던 정보 철학의 정치적 논의들을 다룰 것이라고 소개하고 있다(Floridi 2013, p. xvi).

3) *Ethics and Information Technology* Vol.10(2-3) (2008), *The Information*

*The Ethics of Information*의 전체 16장을 플로리디의 저술 목표인 메타 이론적 목표(1장~5장), 입문적 목표(6장~11장), 이론 구축적(constructive) 목표(12장~16장)에 따라 소개하고, 이 저서에 담긴 플로리디의 정보 윤리학을 비판적으로 검토하고자 한다. 특히 플로리디의 정보 윤리학이 인공지능 시대의 ‘새로운’ 윤리학이 될 수 있을지를 검토하고, 이러한 목표를 위해 보완할 점을 제안하고자 한다.

2. 정보 윤리의 이론적 토대 (1장~5장)

플로리디는 *The Ethics of Information*의 첫 장을 “정보 혁명”(information revolution)으로 도야 된 정보의 시대에 대한 서술로부터 시작한다.⁴⁾ 플로리디는 과학이 우리의 인식을 변화시키는 두 가지 근본적 방법을 세상과 관련된 외재적 방법과 우리 자신에 대한 내재적 방법으로 구분하면서, 역사에서 나타난 세 번의 과학적 혁명은 이 두 방법 모두를 통해 우리의 외부 세계에 대한 이해를 바꾸면서 우리에게 영향을 끼쳐 우리 자신이 누구인지에 대한 이해를 조금씩 변화시켜왔다고 분석한다.⁵⁾ 코페르니쿠스(Nicolaus Copernicus)가 일으킨 첫 번째 혁명은 인간이 우주의 중심이 아님을 깨닫게 했고, 다윈(Charles Darwin)의 두 번째 혁명은 인간이 생명의 중심이 아니라는 생각의 전환을 초래했으며, 프로이트(Sigmund Freud)가 야기한 세 번째 혁명은 인간이 순수한 합리성 중심의 존재가 아니라는 수정을 하게 했다. 플

Society Vol.25(3) (2009), *Etica & Politica / Ethics & Politics* Vol.XIII (2011), *Minds and Machines* Vol.24(1) (2014) 등의 학술지들은 특집호나 심포지엄 형태로 플로리디의 정보 철학과 정보 윤리학을 집중적으로 검토하고 있다. 플로리디의 논의를 연구한 논문집(Allo (ed.) 2011; Demir (ed.) 2012)과 단행본(Durante 2017) 등이 출판되어, 그의 철학과 윤리학을 심도 깊게 분석하고 세부 주제들에 대해 발전적 논의를 전개하고 있다.

4) Floridi (2013), p. 1.

5) Ibid., pp. 13-14.

로리디는 1950년대 튜링(Alan Turing) 등의 학자들이 주도한 컴퓨터 과학, 그리고 ‘정보와 소통 기술(Information and Communication Technologies, 이하 ICTs)’이 네 번째 혁명인 정보 혁명을 일으켰다고 주장한다.⁶⁾ 정보 혁명은 자연적 실재와 인공적 실재(artificial realities)에 대해 전례가 없을 정도로 강력한 영향력을 인식론적이고 공학적인 차원에서 행사했으며, 이로 인해 우리 자신이 누구인지, 우리는 이러한 세계에 대해 어떻게 관계를 맺어야 하는지에 대한 새로운 시각을 제시하는 방식으로 우리 자신을 변화시키고 있다고 플로리디는 분석한다. 플로리디에 따르면, 이러한 정보 혁명으로 인해 우리는 우리 자신이 뉴턴(Isaac Newton)적 관점에서의 독자적이며 고유한 실체가 아니라 정보적으로 구현된 유기체이며, 우리는 정보 환경에서 상호 연결되고 정보 환경에 밀착된 존재라는 생각을 점진적으로 받아들이고 있다. 이러한 현실 인식에 대한 변화는 우리가 현실에서 다루는 가치들의 변화를 이끌게 되어 정보 혁명 이후의 윤리학을 요구한다. 이런 의미에서 플로리디의 정보 윤리는 ‘정보 유기체(information organisms, 즉 inforgs)’와 ‘정보 환경(information environments, 즉 infosphere)’에 대한 윤리이다.⁷⁾ ICTs에 대한 의존도가 증가하고 있다는 점에서, 정보 윤리는 정보 환경인 현실(reality)에 대한 정보적 해석을 요구하며, 정보 유기체는 정보 환경에서 살아가는 인간뿐만 아니라 인공 행위주체(artificial agents)도 포함해야 한다.

정보 윤리학이라는 명칭은 다양한 학자들에 의해 1980년대부터 계

⁶⁾ Ibid., p. 14.

⁷⁾ 플로리디는 자신의 정보 철학과 정보 윤리학을 전개하면서 많은 신조어를 만들고 있다. 예를 들어, 여기서 언급된 inforgs, infosphere 뿐만 아니라, infraethics, distributed morality, onlife, distributed responsibility, e-nvironmental ethics, ecopoiesis 등의 신조어가 정보 윤리학에 나타난다. 이러한 신조어는 기존 개념에 담겨 있을 수 있는 전통적 가치나 의미를 퇴색시키고 자신이 의도하는 방식으로 논의를 전개하기 위한 포석으로 보인다. 따라서 플로리디의 철학과 윤리학을 심도 있게 연구하기 위해서는, 각 신조어에 담긴 의미를 규명할 필요가 있으며 이러한 규명 아래 각 단어의 적절한 번역어를 제시할 필요가 있다.

속해서 사용되고 논의가 발전되어 왔는데, 플로리디는 2장에서 자신의 정보 윤리학이 기존의 정보 윤리학과 어떤 의미에서 차별화되는지를 중점적으로 설명한다. 플로리디는 1980년대에 처음으로 나타난 정보 윤리학은 자원(resource)으로서의 정보를 다루는 것에 불과했고 이후 발전 과정에서의 논의들은 정보의 산물(products), 정보의 환경 등을 개별적으로 다루는 분절된 논의였지만, 자신의 정보 윤리학은 자원, 산물, 정보 환경 등을 포괄적이고 체계적으로 다루는 거시 윤리(macroethics)라고 주장한다. 이런 의미에서 플로리디의 정보 윤리학은 단순히 어떤 윤리적 문제를 해결하기 위한 이론만이 아니라, 문제 해결 절차를 제시할 수 있는 개념적 토대 또한 제공하려는 기획이다.⁸⁾ 기존의 정보 윤리학과 차별화되는 플로리디의 정보 윤리학의 체계를 제시하기 위한 방법론적 토대는 2011년 저서인 *The Philosophy of Information*의 3장에서도 중요하게 논의된 ‘추상화(층위의) 방법론(the method of (levels of) abstraction)’이다. 이런 점에서 컴퓨터 과학에서 차용한 추상화 방법론은 정보 윤리뿐만 아니라, 플로리디의 정보 철학 전체 기획의 방법론임을 알 수 있다. 추상화 방법론은 특정한 목적 아래 설정된 층위(level)에서 상황과 대상을 인식하는 방법이다. 예를 들어, 중고 자동차에 대해서는 다양한 정보들이 있을 수 있지만 각각의 목적에 따라 설정된 추상화 층위는 자신의 목적에 부합하는 정보만을 파악하고 나머지는 추상화한다.⁹⁾ 구매의 목적을 갖고 추상화 층위를 설정했을 때는 안전성과 차량의 사고 여부 기록 등의 정보를 포착할 것이며, 경제적 목적을 갖고 설정된 추상화 층위에서는 차량의 시장 가치, 유지비용 등의 정보만 파악하게 된다. 플로리디에 따르면, 이러한 추상화 방법론은 기존의 인식 층위에서 포착하지 못하던 정보 영역의 윤리적 문제를 적절하게 설정된 추상화 층위에서 인식하게 한다. 추상화 방법론은 다양한 목적 설정에 따라 다양한 추상화 층위를 형성한다는 특징으로 인해 상대주의라는 비판에 직면할 수 있다. 왜냐하면, 이러한 다원주의적 성격이 다양한 인식적 관점을 허용한다는 의미에서

8) Ibid., pp. 313-314.

9) Ibid., pp. 31-32.

자의적으로 보일 수 있기 때문이다. 그러나 플로리디는, 각각의 층위들이 목적 지향적이며 서로 의존하는 관계적 성격을 갖고 있다는 점에서 상호 비교 가능하며 평가할 수 있기 때문에, 다양한 층위들이 독립적 이어서 무슨 일이든 허용될 수 있다는 의미의 상대주의는 아니라고 반박한다.¹⁰⁾ 이런 의미에서 플로리디는 자신의 추상화 방법론은 ‘상대주의를 인정하지 않는 다원주의(pluralism without endorsing relativism)’라고 강변한다.

플로리디는 4장에서 이러한 방법론을 토대로, 자신의 정보 윤리학이 ‘확장된 환경에 대한 윤리(e-nvironmental ethics)’임을 강조한다. 윤리학의 역사가 보여주는 것처럼, 윤리는 백인 중심적인 윤리 체계에서 성과 인종을 넘어 인간 종(species) 전반으로의 확장, 그리고 이러한 인간 중심적인 윤리에 대해 도전하는 동물 윤리와 생태 윤리로 인해 그 영역이 점차 확장되고 있지만, 여전히 기존의 윤리학은 새로운 정보 유기체를 도덕의 영역으로 포섭하지 못한다. 이러한 제한성으로 인해 기존의 윤리학은 충분히 보편적이지도 않으며 공평하지도 않다는 비판에 직면하게 한다. 기존의 윤리학에서도 이러한 제한성이 행위주체(agents) 중심의 논의로 인해 야기되었다는 반성으로 인해, 생명의료 윤리나 롤즈(John Rawls)의 논의는 예외적으로 행위객체인 피동자(patients)의 입장(positions)을 고려하여 공정성을 확보하고자 한다. 그러나 이러한 예외에서조차 기존의 윤리학은 ICTs가 초래한 새로운 피동자는 배제한다는 점에서, 이러한 피동자를 도덕 영역으로 포섭할 수 있는 논의가 요구된다. 플로리디는 이러한 ICTs에서 나타나는 행위주체와 피동자를 포괄하는 정보 유기체(inforgs)와 정보 환경(infosphere)에 대한 윤리인 정보 윤리학, 즉 ‘피동자와 존재 중심적인 거시 윤리학(patient-oriented, ontocentric macroethics)’을 제시한다.¹¹⁾ 구체적으로 플로리디가 제시하는 정보 윤리의 4개 윤리 원칙은 다음과 같다.¹²⁾ 0)

10) Ibid., p. 33.

11) Ibid., p. 102.

12) Ibid., p. 71. 플로리디는 1990년대부터 엔트로피 개념을 사용해 오고 있어서 단어를 바꿀 수 없다는 현실을 인정하면서도, 자신이 엔트로피 단어를

형이상학적 엔트로피는 정보 환경에서 야기되지 말아야 한다(null law). 1) 엔트로피는 정보 환경에서 방지되어야 한다. 2) 엔트로피는 정보 환경에서 제거되어야 한다. 3) 전체 정보 환경뿐만 아니라 정보 존재자의 번영이 그들의 웰빙(well-being)이 보존되고 함양되고 풍성하게 되는 방식으로 고취되어야 한다. 이러한 원칙은 모든 정보 유기체들이 동등한 존엄성(dignity)을 가져야 한다는 플로리디의 주장에 토대를 두고 있다. 플로리디에게, 형이상학적 엔트로피를 증가시키는 것은 정보주체를 부패하게 하고 훼손하는 것으로서, 정보주체의 웰빙(well-being)과 번영(flourishing)을 훼손하는 것이기 때문이다.

정보 윤리학의 이론적 토대를 마무리하는 5장에서 플로리디는 자신의 정보 윤리학이 기존의 컴퓨터 윤리(computer ethics)와 어떻게 차별화되고 어떻게 자신의 정보 윤리가 컴퓨터 윤리를 포섭할 수 있는지를 설명한다. 플로리디는 정보 혁명으로 인해 발생한 정책적이고 개념적인 공백 상태에 대응했던 컴퓨터 윤리가 비일관적이고 부적절하기 때문에, 컴퓨터 윤리의 토대에 대해 역사적으로 제시된 5가지 접근법을 검토하면서, 컴퓨터 윤리가 정보 윤리를 통해 혁신되어야 함을 주장한다.

사용했던 것을 후회하는 듯하다(Floridi 2013, p. 66). 왜냐하면, 독자들이 이 개념을 읽을 때마다 열역학과 기존 정보 철학에서 사용되던 엔트로피 개념으로 읽는 것이 너무나 당연해서, 자신의 논의에 대한 오해와 착오를 초래하기 쉽기 때문이다. 플로리디는 자신의 개념은 열역학과 기존 정보 철학과 관련성은 있지만 다른 개념으로 이해해야 하며, 이런 이유로 인해 ‘형이상학적 엔트로피(metaphysical entropy)’로 명명한다. 기존의 개념이 에너지나 정보를 무질서의 양을 지시하는 양적 개념으로 사용해 왔다면, 플로리디의 형이상학적 엔트로피는 정보의 의미론적이며 존재론적 본성과 관련된 질적 개념이다. 따라서 형이상학적 엔트로피가 감소한다는 것은 정보의 내용적인 측면에서 정보의 양이 늘어나며 점진적으로 의미 있고 풍성해지는 것을 의미한다. 반대로, 정보 존재들이 사용하지 못하게 되고 궁극적으로 사라져 정보의 양이 감소하는 것이 형이상학적 엔트로피가 증가하는 것이다.

3. 정보 윤리의 기본적 내용(6장~11장)

*The Ethics of Information*의 두 번째 저술 목표인 정보 윤리 입문을 위해 플로리디는 6장에서 정보 환경이 갖는 본질적(intrinsic) 가치를 규명하고자 한다. 기존 윤리학, 예를 들어 칸트(Immanuel Kant)의 의무론적 윤리학에서는 인간만이 수단적 가치가 아닌 본질적 가치를 가지며, 이로 인해 존엄성을 갖는다고 주장하는데, 플로리디는 이러한 입장이 자유로운 행위자와 살아 있는 유기체라는 두 개의 구분되는 본성을 혼동한 결과라고 비판한다.¹³⁾ 피동자 중심의 윤리학인 정보 윤리에서는 더 높은 추상화 층위로의 전환을 통해 존재 중심의 본질적 가치와 존엄성 논의를 제시한다. 이러한 플로리디의 관점에서 볼 때, 모든 존재자는 정보 존재이고, 일부 정보 존재는 행위주체(agent)이며, 일부 정보 행위주체는 인공 정보 행위주체가 되고, 일부 인공 정보 행위주체는 도덕적 인공 행위주체이다.¹⁴⁾ 플로리디는 7장에서 도덕적 인공 행위주체에 대해 상세히 설명하고 있다. 플로리디는 적절한 추상화 방법론을 통해 정보 환경을 인식하게 되면, 행위주체에 대한 적절한 조건을 제시할 수 있는데, 정보 행위주체는 기존의 행위주체 개념보다 약화한 의미 또는 희석된 의미의 상호작용성(interactivity), 자율성(autonomy), 적응성(adaptability)으로 제시할 수 있다.¹⁵⁾ 예를 들어 플로리디의 자율성 개념은 행위주체가 상호작용에 대한 직접적인 대응 없이 자신의 상태를 바꿀 수 있다는 것을 의미한다. 이러한 개념들을 토대로 제시되는 도덕성을 플로리디는 ‘마음 없는 도덕성(mindless morality)’이라고 명명한다.¹⁶⁾ 이러한 마음 없는 도덕성 논의는 추후 플로리디의 정보 윤리학 논의인 ‘해명책임(accountability)’, 그리고 ‘과실 없는 책임(faultless responsibility)’ 논의와 연결된다.¹⁷⁾

¹³⁾ Ibid., p. 116.

¹⁴⁾ Ibid., p. 110.

¹⁵⁾ Ibid., p. 140.

¹⁶⁾ Ibid., p. 148.

¹⁷⁾ Floridi (2016), p. 1. 플로리디는 이 논문에서 *The Ethics of Information*

플로리디의 정보 윤리는 추상화 방법론을 통해 새로운 인식 관점으로 전환하는 구성주의적 입장을 취하고 있는데, 이러한 구성주의적 성격은 8장에서 잘 나타난다. 구성주의를 중요시하는 플로리디에게 성숙한 도덕적 인간을 구성하려는 덕 윤리(virtue ethics)는 매력적이지만, 덕 윤리가 행위주체 중심적이며 인간 중심적 논의라는 점에서 정보 사회에 부적절하다고 비판하다. 플로리디는 이와 대비하여 자신의 정보 윤리가 갖는 구성주의적 특성을 ‘호모 포이에티구스(homo poieticus)’ 논의를 통해 강조한다. 플로리디에 따르면, 인간은 정보 환경에서 개념화되는 실재(reality)를 보호하고 번영하게 하려고 세계를 형성하는 존재인 데미우르고스(demiurge)의 역할을 맡게 되기 때문이다.¹⁸⁾ 데미우르고스가 그랬던 것처럼, 인간은 정보 환경에서 더 많은 힘을 갖게 되면 될수록, 피동자와 존재 중심적 정보 환경이라는 실재를 더 잘 돌보고 보호하고 번성하게 할 더 많은 의무와 책임을 갖게 된다. 인간이 정보 환경에 갖는 이러한 구성주의적 특성으로 인해, 듀란트(Massimo Durante)는 “플로리디는 마지막 휴머니스트이다”라는 주장을 인간 중심적 입장을 포기하였다는 의미보다는 정보 시대에 휴머니즘을 지켜낼 수 있는 유일한 방법이 탈-인간 중심적 입장을 인정하는 것이라는 믿음을 가진 철학자라는 의미로 해석할 수 있음을 강조한다.¹⁹⁾

9장과 10장에서는, 전통적인 기존 윤리학 체계에서 포착하지 못하는 해악(evil)과, 선한 의지를 갖추어도 불구하고 비극을 초래하게 되는 상황들이 정보 윤리에서 어떻게 이해될 수 있는지를 보여주고 있다. 전통적인 도덕적 해악과 자연적 해악 개념이 포착하지 못하는 정보 시대의 새로운 해악이 ‘인공적 해악(artificial evil)’이다. 예를 들어, 수술 도중에 모니터가 재부팅되어 수술을 진행하지 못해 환자가 치명적인 상황에 부딪치는 것과 같은 해악이 인공적 해악에 해당한다.²⁰⁾ 이러한

(2013)에서 충분히 제시하지 못한 책임과 해명책임 논의를 상세히 설명하고 있다.

¹⁸⁾ Floridi (2013), p. 168.

¹⁹⁾ Durante (2017), p. xvi.

²⁰⁾ Floridi (2013), p. 187.

해악은 정보 윤리가 제시하는 인공적 해악 개념을 통해 포착될 수 있는데, 인공적 해악은 피동자의 웰빙을 부패하게 하거나 훼손하는 것이라고 규정되기 때문이다. 플로리디는 그리스 신화인 카산드라의 곤경이 보여주는 것처럼 선의를 가진 행위주체들의 비극을 ICTs를 잘 활용하는 정보 시대에는 정보 윤리를 통해 막아낼 수 있다고 주장한다.

*The Ethics of Information*의 두 번째 목표인 정보 윤리의 입문 목표를 마무리하는 11장에서 플로리디는 자아들(selves)의 정보적 본성을 다룬다. 플로리디는, 자아의 통합성에 대한 플라톤의 마차 비유 문제를 기존의 철학, 즉 로크(John Locke)적 방식이나 서사적 접근이 효과적으로 설명하지 못함을 지적하고, 정보의 과정을 통해 제시된 모델을 통해 자아 통합성을 설명할 수 있음을 보여준다. 이러한 설명은, 정보 혁명이 인간 자신에 대한 이해를 수정해 정보적 유기체로 인식하게 되었다는 1장에서의 주장과 일관성을 유지한다.

4. 정보 윤리의 적용(12장~16장)

*The Ethics of Information*의 세 번째 저술 목표인 이론 구축을 위해 플로리디는 정보 윤리가 적용될 수 있는 정보 프라이버시(12장), 분산된 도덕(13장), 정보 경영 윤리(14장), 전지구적 정보 윤리(15장)를 검토하고, 마지막 장에서는 정보 윤리에 제기된 비판과 질문에 대응하고 있다. 플로리디는, 정보 프라이버시 논의가 기존 윤리학에서 중요한 역할을 하지 못하지만, 정보 주체의 웰빙과 통합성을 증시하는 정보 윤리에서는 중요하다고 역설한다.²¹⁾ 왜냐하면, ICTs로 인해 개인의 정보 프라이버시가 침해될 우려가 높다는 기존의 해석과 달리, 새로운 ICTs의 정보 시대에는 오히려 정보 프라이버시를 보호할 수 있다는 정보 윤리를 제시하고 있기 때문이다.

플로리디는 정보 사회에서 복잡성, 장기적 영향력, 상호작용이 기하

21) Ibid., p. 258.

급수적으로 증가하고 있다고 분석한다.²²⁾ 이러한 특징으로 인해, 행위 주체들(agents) 사이의 관계와 행위가 중첩되고 복잡해진 상황에서 발생하는 윤리적 문제는 특정 행위주체로부터라기보다는 행위주체들의 네트워크로부터 야기되기도 한다. 기존 윤리학은 이러한 현상에서 야기되는 윤리적 문제들에 대응하기에 어려움을 겪고 있는데, 플로리디는 이러한 현상을 ‘분산된 도덕 행위(distributed moral actions)’라는 개념으로 접근하여 어려움을 극복하고자 한다.²³⁾ 플로리디는 분산된 도덕 개념을 가지고, 결과주의나 의무론과 같은 기존의 윤리학이 포착하지 못하는 정보 시대의 윤리적 문제, 즉 도덕적으로 무시 가능한(morally negligible) 또는 중립적 행위가 초래하는 해악과 좋음을 설명하며 이를 통해 해악을 막고 좋음을 증진하는 방안을 제시한다. 분산된 도덕 행위는 과학기술의 발전으로 인해 더욱 가속화된 전지구화 사회에서 그 자체로는 도덕적으로 무시해도 될 정도의 각 개인의 작은 행위들이 결합하여 전지구적 윤리 문제에 효과적으로 대응할 수 있게 한다.²⁴⁾ 왜냐하면, 예를 들어 혼자서 했을 때는 별 영향이 없을 것 같은 개인의 작은 기부가 ICTs의 도움으로 가능해진 전지구적 네트워크를 통해 결합하였을 때는 엄청난 영향력을 행사하기 때문이다.²⁵⁾

플로리디는 분산된 도덕성 개념에 적합한 사례로 14장에서 경영에 대한 정보적 분석과 이해를 제시한다. 플로리디의 정보 윤리에 따르면, 경영 사업체와 조직은 도덕적 행위주체이며 복잡한 관계 속에서 도덕적 해악과 좋음을 초래할 수 있다. 따라서 플로리디는 정보 경영 윤리적 관점에서 경영 행위주체는 4장에서 제시된 정보 윤리의 4개 윤리

22) Floridi (2016), p. 11.

23) Floridi (2013), p. 262.

24) 목광수는 기존의 책임 논의가 전지구화 시대의 책임 논의로 한계를 갖고 있으며, 플로리디의 분산된 책임 논의를 전지구적 맥락으로 재해석한다면 전지구화 시대에 적합한 책임 논의가 될 수 있다고 주장한다(목광수 2017, pp. 33-59).

25) Floridi (2013), pp. 268-269. 플로리디는 각 개인의 물건 구매를 통해 수익금의 최대 50%가 AIDS 환자를 돕는데 기부되는 홈페이지(red.org)를 이러한 방식의 하나로 소개하고 있다.

원칙인 번영과 웰빙을 도모하고 엔트로피를 감소할 것을 요구한다. 플로리디는 15장에서 정보 환경이 세계화된 환경 속에서 자신의 도덕적 행위주체와 분산된 도덕 개념이 효과적으로 작동할 수 있는지를 검토하고 있다.

16장은 그동안 자신의 논의에 제기된 질문과 비판에 답변하려는 시도로 마련되었다. 이러한 시도는, *The Ethics of Information*이 장마다 요약과 결론 부분을 두어 독자들의 이해를 돕고자 하는 것과 같은 맥락이다. 자신의 정보 윤리학에 대한 비판과 회의에 대응하기 위해 고안된 16장은, 모두 20가지의 비판에 대해 대응하고 있다. 그러나 플로리디는, 이러한 대응이 제기된 비판 모두가 틀렸다는 강한 대응이 아니라, 부분적 오해에는 적극적으로 해명하지만 논란이 있는 부분에는 계속 토론의 여지가 있다는 것을 보여주는 것임을 밝힌다.²⁶⁾

5. ‘새로운’ 윤리학으로서의 정보 윤리학에 대한 비판적 검토

정보 시대에 나타나는 윤리적 상황을 설명하고 윤리적 문제들에 대응하려는 플로리디의 정보 윤리학은, 새롭게 제기된 윤리적 문제들에 대해 임기응변(ad hoc)적으로 대응하려는 기존 논의와 달리, 인식론적이고 존재론적인 토대 아래 체계적인 윤리 이론을 제시한다는 점에서 기존 윤리와 구분될 뿐만 아니라, 정보 시대의 윤리학으로 주목할 만하다. 이러한 플로리디의 정보 윤리학은 ‘피동자와 존재 중심적인 거시 윤리학(patient-oriented, ontocentric macroethics)’이라는 점에서 기존의 윤리학과 차별화되는 ‘새로운’ 윤리학으로 볼 수 있다. 인간 중심의 윤리로부터 시작된 기존의 윤리학은 동물과 생태에 대한 윤리적 물음으로 인해 동물 윤리학과 생태 윤리학으로 조금씩 그 영역을 확장하고 있다. 플로리디는, 기존의 윤리학에서 가장 확장된 영역으로 제시된 생

²⁶⁾ Ibid., p. 307.

태 윤리학으로부터도 더 나아간 윤리학, 즉 존재 중심의 윤리학을 제시한다. 왜냐하면, 인공지능 시대의 정보는 우리에게 새로운 윤리적 물음을 끊임없이 제시하고 있으며, 이러한 물음에 대응하기 위해서는 존재 중심의 논의가 필요하기 때문이다. 또한, 정보 윤리학은 기존의 윤리학과 달리 피동자 중심의 접근을 옹호한다. 기존의 윤리학이 몇몇 예외를 제외하고는 행위주체 중심적 윤리학인데, 이러한 시도들은 정보 시대에 야기되는 윤리적 문제를 포착하지 못하기 때문에 플로리디는 행위주체가 아닌 피동자 중심의 윤리학을 제시하고자 한다. 더 나아가, 플로리디의 정보 윤리학은 기존의 윤리학을 통해 정보의 시대에 대응하는 미시 윤리학이 아닌 기존의 윤리학을 포괄하는 거시 윤리학이다.

플로리디의 정보 윤리학은 체계적이며 풍성한 논의를 제시하고 있음에도 불구하고, 정보 시대에 적합한 ‘새로운’ 윤리학으로 자리매김하기에는 보완하거나 해명할 부분이 여전히 남아 있는 것으로 보인다. 첫째, 플로리디는 정보 윤리학의 토대가 되는 도덕성(morality) 또는 규범성(normativity)에 대한 설명을 보충할 필요가 있다. 세상을 정보의 관점에서 인식하려는 추상화 층위에서 세상은 정보 세상이 되고 존재자들은 정보 유기체로 인식될 수 있다. 그러나 이렇게 인식된다고 해서, 그러한 존재자들을 모두 도덕적 존재라고 부를 수 있는 것은 아니다.²⁷⁾ 도덕적 지위(moral status)를 갖는 도덕적 존재가 되기 위해서는 어떤 정당화된 논리적 근거가 있어야 하며, 이를 위해서는 도덕성 또는 규범성에 대한 논의가 선행해야 한다. 윤리학의 역사에서 볼 때, 인간 중심적 논의를 넘어 동물의 권리를 주장하기 위해서는 1살 이상이 된 포유류가 미래감 등을 갖는 ‘삶의 주체(subject of a life)’라는 근거가 제시되었고, 동물의 복지는 ‘쾌고 감수 능력(sentience)’을 가진

27) 플로리디가 정보 존재들이 모두 동등한 도덕적 지위를 가진다는 부분에 대해서는 많은 비판들이 제기되어 오고 있다. 왜냐하면, 모든 정보 존재자들과 정보 환경이 본질적 가치로서의 존엄성을 갖는다는 플로리디의 주장이 충분하고 설득력 있는 근거 없이 제시되고 있기 때문이다(Doyle 2010, pp. 106-107; Himma 2004, pp. 149-156; Brey 2008, pp. 111-113).

동물까지로 한정되었다.²⁸⁾ 그런데 플로리디는 어떤 이론적 근거나 논의 없이 정보 존재가 있다는 사실로부터 모든 정보 존재가 평등한 존엄성을 가진 도덕적 존재라고 주장하는데, 이러한 추론은 플로리디가 ‘자연주의적 오류(naturalistic fallacy)’를 범하고 있는 것은 아닌지 의문스럽게 만든다. 플로리디는 자신의 정보 윤리에 대한 비판과 오해에 해명하려는 목적으로 쓰인 16장에서 자신의 논의는 영적(spiritualistic)이며, 존재(Being)와 좋음(Goodness)을 본질적으로 연결 지었던 플라톤이나 스피노자의 논의에서 볼 수 있는 것처럼 사실을 통해 좋음을 설명하는 것이 가능하기에 자신의 논의가 자연주의적 오류가 아니라고 강변한다.²⁹⁾ 그런데 이러한 플로리디의 설명은 불충분해 보인다. 왜냐하면, 플라톤이나 스피노자와 달리, 플로리디는 존재 자체가 가치를 어떻게 담지하고 있는지에 대한 이론적 설명과 논의를 충분히 그리고 설득적으로 제시하고 있지 않기 때문이다. 이와 관련해서, 플로리디가 사용하는 개념들의 도덕적 의미 또한 충분히 설명되어야 한다. 예를 들어, 정보 윤리의 원칙에서 제기된 정보 유기체의 웰빙과 번영이 무엇인가라는 질문에 대해, 플로리디는 정보 유기체가 존재하며 그 존재가 현상유지 상태에 있고, 그 유기체의 수가 증가하는 것을 의미한다고 대답한다.³⁰⁾ 이에 반대되는 해악인 형이상학적 엔트로피 증가는 정보 존재자의 파괴와 오염을 의미한다. 그런데 이러한 개념 설명이 대답 전부라면 이러한 언급은 사실에 대한 기술에 불과한 것으로 보이는데, 이러한 설명이 어떻게 도덕적 의미를 담아낼 수 있을지 의문스럽다. 따라서 플로리디의 정보 윤리학이 ‘새로운’ 윤리학으로 자리매김하기 위해서는 정보 윤리의 적합한 도덕성과 규범성을 제시하면서, 이에 부합하는 윤리적 개념들을 재정립할 필요가 있다.

둘째, 정보 윤리학의 방법론적 토대인 추상화 방법론에 대한 도덕적 정당화가 보장될 필요가 있다. 왜냐하면, 추상화 방법론이 인식론적 관점을 전환하는 방법으로 사용될 때에는 충분히 그 자체로 정당화가 가

²⁸⁾ Regan (2004), p. 243, 싱어 (2013), p. 53.

²⁹⁾ Floridi (2013), p. 324.

³⁰⁾ Ibid., pp. 70-71, p. 75.

능하지만, 도덕적 논의의 방법론으로서는 도덕적 상대주의라는 비판에 직면할 수 있기 때문이다. 플로리디는 추상화 층위들이 관계적(relational)이기 때문에 자의적인 의미의 상대주의의 비판에 대응할 수 있다고 주장하지만, 여기서 언급하는 상대주의는 그러한 관계 속에서도 제기되는 도덕적 상대주의이다. 상호 관계를 갖는 추상화 층위들을 의미하는 ‘결속된 추상화 층위들(nested gradient levels of abstraction)’조차 세계에 존재하는 모든 존재자에 대한 추상화를 의미하는 것이 아니다.³¹⁾ 따라서 다양한 결속된 추상화 층위들이 존재한다고 할 때, 서로 다른 결속된 추상화 층위는 서로 다른 도덕적 존재를 설정할 뿐만 아니라, 도덕적 좋음과 도덕적 해악인 형이상학적 엔트로피에 대해 서로 다른 규정을 할 것이며, 이로 인해 도덕적 상대주의는 불가피하기 때문이다. 플로리디는 이러한 비판이 전제한다고 추측되는 도덕적 가치의 절대성과 상대성을 구분하는 방식이 ‘잘못된 이분법(false dichotomy)’이라고 대응한다.³²⁾ 그런데 여기서 제기하는 도덕적 상대주의는 이러한 이분법과 무관하게 서로 다른 층위에 대해 서로를 평가할 객관적 기준이 없기 때문에 도덕적 우위를 설명할 수 없다는 의미의 도덕적 상대주의이다.³³⁾ 이러한 비판에 대응하는 한 가지 방법을 제안해 본다면, 정보 윤리 논의에서 플로리디의 추상화 방법론은 특정 추상화 층위를 설정하고 정당화하는 과정을 제시할 필요가 있다. 이러한 논의는, 도덕적으로 정당화된 추상화 층위에서 모두가 동일한 도덕적 개념을 가지고 논의를 진행하여 상대주의라는 비판을 피할 수 있게 해 줄 것이다.

셋째, 플로리디의 정보 윤리학과 기존 윤리학의 관계를 분명하게 제시하고 재정립할 필요가 있다. 플로리디는 기존 윤리학을 대체하거나 포괄하는 거시 윤리학(macroethics)으로 자신의 정보 윤리학을 자리매김하는 것으로 보인다.³⁴⁾ 플로리디는 적절한 추상화 층위 설정을 통해

31) Floridi (2011), pp. 54-58.

32) Ibid., p. 32.

33) 레이첼즈 (2006), p. 63.

34) Floridi (2011), pp. 74-79.

인간 중심적인 기존 윤리학의 한계를 극복하고 정보 유기체를 도덕적 존재로 포착하는 존재 중심적인 논의를 제시하고자 하는데, 이러한 과정에서 기존 윤리학의 도덕 개념들을 약화하거나 희석하는 것으로 보인다. 예를 들어, 플로리디는 책임(responsibility) 대신에 ‘해명책임(accountability)’을 제시한다.³⁵⁾ 그런데, 해명책임이 도덕적 비난이 부여된 전통적 책임 개념이 갖는 도덕적 무게 보다 약화된 개념으로 보이는데, 만약 그렇다면 이러한 해명책임이 인공 도덕적 행위주체와 관련된 행위에 대해서는 적합한 개념일 수 있겠지만 인간의 행위에도 동일하게 적용된다면 책임 개념과 같은 실천적 힘(practical force)을 확보할 수 있을지 의심스럽다.³⁶⁾ 따라서 정보 윤리가 기존 윤리를 대체

35) 현재 플로리디의 ‘accountability’에 대한 번역어는 ‘책무’(이중원 2017, p. 187), ‘책무성’(신상규 2017, p. 287) 등이 있다. 그런데 기존의 윤리학 영역에서 ‘obligation’에 대한 번역어로 ‘책무’ 또는 ‘책무성’이 많이 사용되고 있어서 혼란을 막기 위해서라도 ‘accountability’에 대한 다른 번역어가 필요한 것으로 보인다. 더욱이 책무(責務)라는 단어는 어떤 계약이나 약속 또는 직무로 인한 책임이라는 의미여서 그 자체로 도덕적 비난 등이 포함될 여지가 많은데, 이러한 의미를 담을 수 있는 번역어가 플로리디의 의도와 부합하는 번역어인지 의문스럽다. 플로리디가 자신의 저서(Floridi 2013)에서 충분히 다루지 못한 정보 시대의 책임 논의를 다루는 논문(Floridi 2016)에서 잘 부각한 것처럼, 분산된 책임(distributed responsibility)으로 제시되는 ‘accountability’는 도덕적 비난 등의 의미 대신에 설명하고 해명하는 투명성과 관련되기 때문이다. 따라서 본 글은 이러한 플로리디의 의도를 반영하는 번역어로 ‘해명책임’을 제안한다. 해명책임은 책임의 하위 개념이나 한 종류가 아닌, 책임이 갖는 실천적 힘(practical force)이라는 현실성(feasibility) 특성을 갖는 개념으로 이해할 필요가 있다.

36) Miller (2001), p. 457. 플로리디의 저서에 나타난 개별 표현에만 주목해 본다면, 그가 인간 행위주체에게는 책임을 부여하고 인공 행위주체에게는 해명책임을 나눠서 제시하고 있다고 볼 여지도 있다(Floridi 2013, p. 154). 그러나 플로리디 저서의 전체적인 관점에서 볼 때 이러한 구분이 분명하게 제시되지 않았을 뿐만 아니라, 그가 제시하는 마음 없는 도덕성 개념이나 분산된 책임 개념은 오히려 이러한 구분 없이 해명책임으로 모든 행위주체의 책임을 설명하려는 시도로 보인다.

하거나 포괄하기보다는, 적어도 현 단계에서는 정보 윤리가 기존 윤리와 적절한 협력 관계를 설정할 필요가 있는 것으로 보인다. 플로리디는 정보 유기체들 사이에서의 차이에 주목하지 않는데, 정보 유기체인 인간과 다른 정보 유기체들 사이의 발전 단계를 구분하고 각각의 단계에 부합하는 정보 윤리를 제시하면서 기존 윤리학과의 관계를 재설정한다면 더욱 실천성 높은 논의가 될 수 있을 것이다.

6. 나오는 글

플로리디의 정보 윤리학은 우리가 뿌린 디지털의 씨앗들이 초래한 우리 시대의 윤리적 문제들의 근본적인 물음을 다루고 있다.³⁷⁾ 인공지능 시대인 정보 사회에 부합하는 정보 윤리학을 제시하려는 플로리디의 시도는 그 자체로 가치가 있지만, 이러한 시도가 ‘새로운’ 윤리학으로 자리매김하기 위해서는 보장되어야 할 부분이 적지 않은 것으로 보인다. 이러한 이론적 여백은 *The Ethics of Information*을 읽는 독자들의 자리로 남겨둔 플로리디의 의도하지 않은 배려일 수 있다. 플로리디는 정보 윤리학의 영역이 “매우 흥미로우며 거대한 개념적 공간이 여전히 미개척”된 영역임을 강조하면서, 자신과 함께 연구의 여정에 “합류하자”고 초대하고 있다.³⁸⁾ 인공지능 시대의 정보 사회를 살아가는 철학자와 윤리학자들이 플로리디의 초대에 어떻게 반응하는가에 따라 이 시대와 사회는 달라질 것이다.

³⁷⁾ Floridi (2013), p. xvi.

³⁸⁾ Ibid., p. 331.

참고 문헌

- 레이첼즈, 제임스 (2006), 『도덕철학의 기초』, 노혜련, 김기덕, 박소영 옮김, 나눔의 집.
- 목광수 (2017), 「전지구화 시대에 적합한 책임 논의 모색」, 『철학·사상·문화』 제25호, pp. 33-59.
- 신상규 (2017), 「인공지능은 자율적 도덕행위자일 수 있는가?」, 『철학』 제132집, pp. 265-292.
- 싱어, 피터 (2013), 『실천윤리학』3판, 황경식·김성동 옮김, 연암서가.
- 이중원 (2017), 「빅데이터가 던지는 도전적인 철학적 문제들에 대한 고찰」, 『도시인문학연구』 9권1호, pp. 168-205.
- Allo, P. (ed.) (2011), *Putting Information First: Luciano Floridi and Philosophy of Information*, Wiley-Blackwell.
- Brey, Philip (2008), “Do we have moral duties towards information objects?”, *Ethics and Information Technology* 10, pp. 109-114.
- Demir, H. (ed.) (2012), *Luciano Floridi's Philosophy of Technology: Critical Reflections*, Springer.
- Doyle, Tony (2010), “A Critique of Information Ethics,” *Knowledge, Technology and Policy* 23(1-2), pp. 163-175.
- Durante, Massimo (2017), *Ethics, Law and the Politics of Information: A Guide to the Philosophy of Luciano Floridi*, Springer.
- Floridi, Luciano (2011), *The Philosophy of Information*, Oxford University Press.
- Floridi, Luciano (2013), *The Ethics of Information*, Oxford University Press.
- Floridi, Luciano (2016), “Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions,” *Philosophical Transactions of the Royal Society of*

London A Mathematical, Physical and Engineering Sciences
374(2083), pp. 1-13.

Himma, Kenneth Einar (2004), "There's something about Mary: The moral value of things *qua* information objects," *Ethics and Information Technology* 6, pp. 145-159.

Miller, David (2001), "Distributing Responsibilities," *The Journal of Political Philosophy*, 9(4): pp. 453-371.

Regan, Tom (2004), *The Case for Animal Rights* (2nd Edition), University of California Press.

| | |
|--------|---------------|
| 논문 투고일 | 2017. 11. 12. |
| 심사 완료일 | 2017. 11. 18. |
| 게재 확정일 | 2017. 11. 20. |

A Theory of Information Ethics in an Era of Artificial Intelligence:

Luciano Floridi's Ethics of Information

Kwangsu Mok

Luciano Floridi shows that established theories of ethics, such as utilitarianism and deontology, cannot respond to new issues such as artificial evil, which happens in an era of artificial intelligence, and proposes that his ethics of information can do effectively. Floridi argues that his ethics of information is effective to find out new issues through the method of (levels of) abstraction and proposes ethical principles, in which metaphysical entropy should be prevented or removed and information entities should be flourished. Floridi's ethics of information is a patient-oriented, ontocentric macroethics, which is differentiated from established theories of ethics. The purpose of this discussion paper is to examine critically whether Floridi's ethics of information is morally justifiable. I doubt that Floridi's theory of ethics is morally justifiable in that his theory is not only vulnerable to the critique of moral relativism, but also is hard to escape from the problem of the naturalistic fallacy. In order to respond to these critiques, I recommend that Floridi's theory of information should include a concept of morality which can effectively support his theory from a normative perspective, that his theory should argue which level of abstraction is appropriate to his theory of ethics, and that his theory should try to explicate how his theory can relate to established theories of ethics. I argue that if Floridi's theory of ethics is supplemented through these recommendations, it could be an appropriate ethics of information in an era of artificial intelligence.

Key words: information ethics, infosphere, inforgs, method of abstraction, patient-oriented, ontocentric macroethics