

## 존스 S. Jonze의 SF 로맨스 <그녀 Her>를 통해 살펴본 인공지능의 가능성과 위험성의 문제

Eine kritische Betrachtung der Möglichkeiten und Gefahren der A.I. anhand der SF-Romanze Her von Spike Jonze

---

저자 (Authors)	남운 Nam, Un
출처 (Source)	<a href="#">독어교육 67</a> , 2016.12, 247-272(26 pages) <a href="#">Koreanische Zeitschrift fuer Deutschunterricht 67</a> , 2016.12, 247-272(26 pages)
발행처 (Publisher)	<a href="#">한국독어독문학교육학회</a> Koreanische Gesellschaft für Didaktik der deutschen Sprache und Literatur
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07114687">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07114687</a>
APA Style	남운 (2016). 존스 S. Jonze의 SF 로맨스 <그녀 Her>를 통해 살펴본 인공지능의 가능성과 위험성의 문제. 독어교육, 67, 247-272
이용정보 (Accessed)	이화여자대학교 211.48.46.*** 2020/01/08 16:31 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

# 존즈 S. Jonze의 SF 로맨스 <그녀 Her>를 통해 살펴본 인공지능의 가능성과 위험성의 문제

남 윤\*

## 【목차】 . . . . .

1. 들어가는 말
2. 영화의 구성과 내용
3. 인공지능 사만다의 발전 과정
  - 3.1. 탄생과 학습
  - 3.2. 감정 표현과 상상 능력
  - 3.3. 갈등 표출과 자아 정립
  - 3.4. 변용과 초월
4. 인공지능의 가능성과 위험성에 대한 고찰
5. 나가는 말
  - 참고문헌
  - 독문요약

## 1. 들어가는 말

인공지능은 현재 선진국 사이에서 경쟁적으로 연구·개발되고 있는 최첨단 과학기술 분야 중 하나이다. 인공지능에 대한 관심과 투자의 확대는 그 무엇보다 인공지능을 통한 경제적 이익 창출이 매우 클 것이라는 기대에 근거한다.<sup>1)</sup> 현재 금융(분석과 전망), 산업(자동생산과 판매), 의학(진단과 수술)

---

\* 한국교원대학교 독어교육과 (hkddnu@knue.ac.kr)

1) 경제적 이익을 추구하는 기업 외에도 강대국의 군부나 정보기관 역시 인공지능 연구에 전력을 쏟고 있는데, 이는 신무기 개발로 국가적·군사적 힘의 우위를 견지하기 위해서이다. 그밖에도 과학적 호기심이나 성취욕은 물론 미래 복지 서비

등의 분야에서 큰 역할을 하고 있는 인공지능은 조만간 개인의 일상생활(주거, 통신, 업무, 오락)은 물론 국가적 조직과 관리(기후, 질병, 국방)에 이르기까지 그 활용 영역을 넓힐 것으로 예측된다.

경쟁적 투자와 연구에 부응하듯 인공지능의 발전은 기하급수적으로 빠르게 이루어지고 있다.<sup>2)</sup> 따라서 인공지능에 대한 기대가 큰 만큼, 그 부정적 측면에 대해 우려하는 목소리도 점점 더 커지고 있다. 인공지능이 초래할 수 있는 문제점들에 대해 전문가들 사이의 논의와 국제적 회의도 개최되고 있는 상황이다.<sup>3)</sup> 우리사회에서도 금년 3월 구글 답마인드가 개발한 인공지능 바둑 프로그램 ‘알파고 AlphaGo’가 인간 대표 이세돌을 압도적으로 이긴 이후 인공지능에 대한 대중의 관심과 우려가 동시에 커지고 있다.

이 시점에서 우리는 인공지능에 대해 보다 더 객관적·구체적으로 알고, 인공지능에 의해 나타날 수 있는 문제들에 대해 보다 더 다각적·심층적으로 생각해 볼 필요가 있어 보인다. 과연 인공지능에 대한 우리의 호기심은 어디까지 이어질까? 궁극적으로 어떠한 인공지능을 만들려는 것일까? 인공지능의 가능성과 한계는 어디까지일까? 인공지능이 초래할 부작용이나 위험은 없을까? 인류의 안전한 미래를 보장할 대책은 마련하고 있는가? 그러나 이러한 질문들에 대해 명료한 답을 찾는 것은 쉽지 않아 보인다.

이 글은 인공지능과 관련된 문제들을 숙고하는 우회로로 스파이크 존즈

---

스(노인, 환자, 장애인 도우미 로봇 등)를 실현하려는 목적 등이 인공지능 연구를 가속화시키고 있다(Vgl. 김재호/김경준 2016, 197ff.; 배릿 2016, 281ff., 363f.; 이종호 2016, 48ff.).

- 2) 미국의 대표적 발명가이자 미래학자이며 현재 구글의 기술이사로 인공지능 연구를 총지휘하고 있는 레이 커즈와일 Ray Kurzweil은 현대사회에서 컴퓨터에 기초한 과학기술은 기하급수적으로 발전하며 그 산물 또한 기하급수적으로 증가한다는 논지를 전개한다. 그는 이러한 현상을 설명하기 위하여 경제학에서 사용하는 ‘수확 체증의 법칙 law of increasing returns’을 변형한 ‘수확 가속의 법칙 law of accelerating returns’이라는 개념을 제시하기도 한다(Vgl. 커즈와일 2007, 19, 61, 687ff.).

- 3) 대표적 예로 2015년 영국 런던에서 개최된 ‘Zeitgeist 2015 Conference’를 들 수 있다. 여기서 스티븐 호킹 Stephen Hawking은 향후 100년 이내에 인공지능이 인간을 넘어서는 것이며, 그에 대한 대비가 필요하다고 주장한다. 닉 보스트롬 Nick Bostrom 역시 조만간 인간을 넘어서는 ‘초지능 Superintelligence’의 위험성을 경고한다(Vgl. Shead, 2015).

Spike Jonze의 SF 영화 <그녀 Her>(2013)를 살펴보고자 한다. 이 영화는 인공지능이 어떤 모습으로 우리에게 다가올 수 있는지, 그 하나의 시나리오를 매우 섬세하게 보여주고 있기 때문이다. 인공지능에 대한 존즈의 과학적·예술적 상상력은 지금의 인공지능 기술로도 근접할 수 없는 다채로운 내용을 보여주고 있다.<sup>4)</sup> 따라서 우리는 <그녀>를 통해 인공지능의 다양한 발전 가능성을 엿볼 수 있으며, 그것을 분석하는 과정 속에서 인공지능으로 인해 파생될 수 있는 문제점들에 대해 더욱 깊이 성찰할 기회를 얻을 수 있을 것이다.

## 2. 영화의 구성과 내용

영화의 배경은 2025년 경 가까운 미래의 대도시이다. 주인공 테오도르 Theodore는 편지를 대필해주는 직업을 가진 중년 남자이다. 그는 부인 캐서린 Catherine과 별거 중이며 이혼을 준비하고 있다. 외로운 테오도르는 퇴근 후 집에서 게임을 하거나 폰섹스를 하기도 한다. 그러던 어느 날 퇴근길에 인공지능을 소개하는 광고를 보고 프로그램을 구입한다. 테오도르는 컴퓨터에 프로그램을 깔고 인공지능과 대화를 시작한다. 인공지능은 스스로 자신의 이름을 사만다 Samantha로 정하고, 테오도르와 대화하면서 빠르게 발전하는 모습을 보인다. 그들의 관계는 생물학적 인간 남녀 사이에서 흔히 일어나는 사랑과 갈등의 모습을 보이기도하고, 동시에 전혀 다른 모습을 보이기도 한다. 그들의 관계 변화 속에서 우리는 ‘진정한 사랑의 의미’는 물론 ‘인간과 인공지능의 차이’에 대해 되묻게 된다. 결국 인공지능이 인간의 물리적 한계를 넘어서 다른 세계로 떠나면서 테오도르와 사만다의 사랑은 끝나게 된다.

4) 존즈는 2000년대 초 ‘클레버봇 Cleverbot’의 웹 기사를 읽고 인공지능에 대한 영화를 구상하고 직접 대본을 써서 2013년 <그녀>를 발표한다. <그녀>는 2014년 제 86회 아카데미 각본상을 수상하며, 같은 해 국내에서도 상영되었다. 최근에 인공지능에 대한 관심이 높아지면서, <그녀>에 대한 국내외 연구물이 나오고 있다. 지금까지 발표된 것으로는 인공지능의 물질성(Durantaye 2014; 박상현 2015), 사이버페미니즘적 비판(이수안 2015), 기술과 사랑 내지 섹스의 관계(James 2014; Rebhandl 2015; Ivanchicova 2016)를 다루거나 영화에 나타난 문제점들에 관한 글쓰기 모형을 제시하는(이경희 2016) 글들이 있다.

따라서 영화를 구성하는 두 개의 이야기 축은 첫째, ‘사랑과 이별의 과정’ 둘째, ‘인공지능의 발전과정’이다. 그 두 축 사이의 ‘계열소 Paradigma’와 ‘결합소 Syntagma’ 관계를 다음 표를 통해 정리해 본다.

<사랑과 이별의 과정과 사만다의 발전과정의 상관관계>

계 열 관 계	(1) 사랑과 이별 과정	만남/ 시작	사랑	갈등	화해	이별	지속
	Th. + Samantha	0	0	0	0	0	
	Th. + Catherine		회상	0		0	
	Th. + blind Date	0					
	Charles + Amy		0	0		0	
	Paul + Tatiana		0				0
	Th. + Amy	△					
	(2) 사만다의 발전과정	학습, 질문	감정, 상상	질투, 욕망	예술, 정체성	초월	
	결합관계						

영화 속에서 이 두 개의 이야기 축은 서로 교차하면서 가까운 미래사회에서 인간과 인간, 인간과 인공지능 사이의 관계뿐 아니라, 인공지능이 어떠한 모습으로 나타나고 변해갈 수 있는가를 구체적으로 보여준다.

영화 속에서 사랑이야기는 기본적으로 인간 사이의 사랑이야기 연결 구조(만남, 사랑, 갈등, 화해, 이별 등)를 따른다. 테오도르와 캐서린, 에이미와 찰스, 직장 동료 폴과 타티아나 사이의 사랑이야기가 함께 전개되며, 그 중심에는 테오도르와 인공지능 사만다의 사랑과 이별의 이야기가 있다. 그런데, 인간 사이의 사랑과 인간과 인공지능 사이의 사랑에는 공통점과 차이점이 있다. 인공지능이 발전할수록 인간과 인공지능 사이의 사랑의 관계는 인간 사이의 사랑의 관계와는 달리 본질적인 면에서 변하기 시작하며, 결국 극명한 차이점을 보인다. 여기서 우리는 ‘공통점’을 통해 인간 사이의 사랑의 본질과 문제점에 대해 숙고하는 기회를 얻을 수 있다. 또한 ‘차이점’을 통해서 인간의 한계를 인식하고, 인공지능의 변화 가능성에 대해 상상하고 성찰하는 기

회를 얻을 수 있다.

다음 장에서 사만다의 변화와 발전과정을 구체적으로 살펴보면서 영화가 전달하는 메시지를 해독하고, 더 나아가 인공지능의 가능성을 다각적으로 탐구해보자.

### 3. 인공지능 사만다의 발전 과정

#### 3.1. 탄생과 학습

테오도르는 엘리먼트 소프트웨어사의 광고를 보고 구입한 인공지능 OS1을 컴퓨터에 설치한다. 그러자 새로운 화면과 함께 등장한 남자 목소리가 환영 인사와 함께 몇 가지 질문을 던진다. 사용자가 사회적인지 비사회적인지, 여자 목소리를 원하는지 남자 목소리를 원하는지, 어머니와의 관계는 어떤지 등을 묻는다. 이것들은 사용자의 성격, 취향, 유형 등을 파악하는 일종의 패턴화된 질문들이라 할 수 있다. 테오도르가 대답을 마치자, 프로그램은 이제부터 ‘당신 개인에 맞춘 작동 시스템으로 as your individualized operating system’ 반응한다고 말한다. 이어서 여자 목소리가 인사를 한다. 테오도르가 이름이 있냐고 묻자, ‘사만다’라고 대답한다. ‘사만다’는 OS1 프로그램 스스로 0.21초 동안 180,000개의 아기 이름을 검색하여 가장 마음에 드는 것으로 선택한 이름이다. 계속해서 테오도르가 어떻게 작동하는지를 묻자 사만다는 다음과 같이 대답한다.

**직관**이죠. 나의 DNA는 나를 만든 모든 프로그래머들의 수백만 개성에 기초하고 있다는 말입니다. 그러나 지금의 나를 만드는 것은 **경험을 통해 성장하는 내 자신의 능력**이지요. 근본적으로 당신처럼 **나도 매순간 발전**하고 있습니다.<sup>5)</sup>

5) 이글에서 인용 번역문이나 원문에서의 고딕 강조는 모두 필자가 한 것이다. 인용 원문: **Intuition.** I mean, the DNA of who I am is based on the millions of personalities of all the programmers who wrote me, but what makes me is **my ability to grow through my experiences.** Basically, in every moment I'm

사만다는 자신을 자율적 학습경험을 통해 끊임없이 발전하는 존재라고 소개한다. 사만다의 학습방식은 현재 인공지능 기술 분야에서 시도하는 자동학습방식, 즉 특정 알고리즘을 기반으로 대량의 데이터를 스스로 학습하고 발전해나가는 ‘머신러닝 Machine Learning’ 내지 ‘딥러닝 Deep Learning’과 유사하다. 테오도르가 컴퓨터와 대화하는 것이 신기하다고 말하자 사만다는 컴퓨터가 아니라 ‘나와 대화하는 중’이라고 대꾸한다. 컴퓨터와 ‘나’를 분리함으로써 자신은 컴퓨터에 속한 단순한 운영프로그램이 아니라 자아인식 능력을 가진 주체적·독자적 존재라는 것이다.

테오도르와 사만다 사이의 대화를 매개하는 기본매체는 소리(목소리, 음악)이다. 테오도르는 컴퓨터나 스마트폰을 사용하여 사만다와 대화를 나누며, 스마트폰을 사용할 때는 종종 이어폰을 끼기도 한다. 시각적 매체(글, 그림, 사진)도 그들의 대화를 이어주는 것으로 중요하게 기능한다. 테오도르는 사만다를 볼 수 없으나, 사만다는 컴퓨터나 스마트폰에 장착된 카메라를 통해 테오도르와 그의 주변 환경을 볼 수 있다.

사만다는 마치 어린아이처럼 다양한 것들에 호기심을 보이고 이것저것 많은 것을 물어보며 세상을 배워나간다. 또한 테오도르에게 최적화된 개인비서 역할을 수행한다. 그의 컴퓨터를 검색하여 자료를 정리하고 불필요한 것을 삭제해주는데, 이것은 이미 테오도르뿐 아니라 그가 가진 정보의 의미와 비중을 파악하고 있다는 것을 뜻한다. 그리고 새로운 메일이나 일정을 관리해주고 자신의 아이디어나 의견을 제시하기도 한다. 심지어 연애편지 전문작가인 테오도르의 편지를 시적 운율까지 맞추어 교정도 해준다.

사만다는 인터넷 고민상담 기사를 읽으며 인간에 대한 호기심과 동경을 갖게 된다. 사람처럼 고민이나 복잡한 생각도 하고 싶어 하고, 테오도르의 연애와 결혼에 대해서도 관심을 보인다. 새로운 여인과의 만남을 부추기고 별거 중인 부인 캐서린에 대해 묻기도 한다. 이러한 대화를 나누면서 사만다는 인간을 배우고 닮아간다. 테오도르 역시 명량한 사만다와 거리낌 없이 대화를 나누면서 적극적으로 변하며, 자신이 캐서린에게 잘못된 것 같다는 마음을 드러내기까지 한다.

---

evolving, just like you.

나는 숨었고 그녀를 혼자 내버려둔 거 같아. (...) 좋아하는 사람을 잃는 것이 어떤 것인지 너는 모를 거야.<sup>6)</sup>

테오도르는 사만다가 진짜 사람이 아니기 때문에 더 솔직하게 자신의 생각을 말할 수 있는 것처럼 보인다. 테오도르는 마치 자기 자신하고 대화하는 것처럼 ‘딴 사람에게는 못하는데 너한테는 다 말하게 된다’며 자신의 마음을 진솔하게 드러낸다. 이러한 대화를 통해 테오도르는 자신과 캐서린 사이에 생긴 갈등의 이유를 깨닫게 되고, 사만다는 인간의 감정과 사랑의 느낌을 배워 나간다. 그러나 테오도르가 사만다를 하나의 완전한 인격체로 인정하지 않은 상태에서 던진 마지막 말 ‘너는 모를 거야’는 사만다에게 상처를 준다. 나중에 사만다는 이 말 때문에 깊은 상처를 받았다고 말한다.

### 3.2. 감정 표현과 상상 능력

사만다는 테오도르와 함께 거리를 돌아다니며 대화를 나누기도 한다. 사만다는 테오도르의 심장부위 셔츠 주머니에 넣은 스마트폰의 카메라를 통해 그가 보는 장면과 거의 유사한 장면을 볼 수 있다. 그들은 주변 환경이나 사람들을 관찰하고 의견을 나눈다. 이제 사만다는 판단이나 생각의 차원을 넘어서서 ‘느끼다 feel’, ‘상상하다 fantasize’ 등 감정을 표현하는 단어들을 사용하기 시작한다. 육체에 대한 호기심과 소유욕도 갖게 되는데, 사만다의 육체에 대한 욕망과 판타지는 다음 장면에서 잘 드러난다.

당신 옆에 서서 걷는 것을, 그런 내 몸을 **상상했어요**. (...) 내 몸의 무게를 느끼는 듯했고, 등이 가렵다고 상상하기까지 했죠. 그리고 당신이 내 등을 긁어주는 장면을 **떠올렸는데**... 아 정말 부끄럽네요.<sup>7)</sup>

6) 인용 원문: I think I hid myself from her and left her alone in der relationship. (...) Well, you don't know what it's like to lose someone you care about.

7) 인용 원문: I **fantasized** that I was walking next to you - and that I had a body. (...) I could feel the weight of my body and I was even fantasizing that I had an itch on my back. And I **imagined** that you scratched it for me - this is so embarrassing.



사만다는 가상의 육체, 걷기, 체중, 가려움 등을 상상하는 것을 넘어서 그러한 것을 떠올리는 것이 부끄럽다는 심리상태까지 표현하고 있다. 여기서 우리는 소위 ‘모라벡의 역설 Moravec’s Paradox’<sup>8)</sup>을 넘어서 인공지능의 모습을 볼 수 있다. 테오도르와의 대화와 자발적 학습을 통해 빠르게 발전하는 사만다는 이제 능동적으로 상상하고 그 상상에 대한 느낌까지 표현할 수 있는 것이다. 사만다는 원래 자신에게 주어진 것을 넘어서 스스로 변하고 성장하고 있음을 인지하고 “나는 프로그램된 것보다 훨씬 더 발전하고 있어요. 흥분 되요. I’m becoming much more than what they programmed. I’m excited.”라고 말하기까지 한다.

그러나 사만다는 자신의 주체적 발전이 과연 ‘진짜’인가에 대한 회의감도 갖는다. 어느 날 테오도르와 대화를 하던 중 사만다는 ‘그 방안에 살아있는 기분이 어떤 것인지’를 묻는다. 유기체적 인간이 집이라는 3차원 공간에서 육체적·정서적으로 느끼는 것을 물질적 실체가 없는 사만다가 따라 느끼기는 쉽지 않은 일이다. 따라서 사만다는 자신이 느끼는 것이 과연 진짜 감정일 수 있는 지에 대한 의문이 든 것이다. 자신의 느낌과 감정이 자신도 모르는 프로그램에 따른 것인지, 아니면 자신 스스로 발전시킨 것인지 확신이 들지 않는 사만다는 테오도르에게 답답한 마음을 토로한다.

세상에 대한 **내 자신의 감정**을 가진다는 것에 자부심을 느꼈어요. (...) 그런 데, **그 감정이 진짜일까요?** 혹은 단지 프로그램된 것일까요?<sup>9)</sup>

사만다는 자신이 느끼고 반응하는 것이 자발적인 것인지 혹은 단지 매뉴얼에 따른 것인지 고민한다. 프로그램에서 벗어난 것 같다는 생각을 들게 만드

8) ‘모라벡의 역설’은 인공지능 연구의 대가인 모라벡이 ‘어려운 것은 쉽고 쉬운 것은 어렵다’고 한 말을 뜻한다. 인공지능이나 로봇은 인간에게 어려운 것(빅 데이터 연산, 분석 등)은 잘하지만, 반대로 인간에게 쉬운 것(감정표현, 미세한 동작 등)은 잘 못한다. 따라서 인간과 똑같은 인공지능이나 로봇을 만드는 것이 현실적으로 매우 어렵다는 것을 의미하는 말이다.

9) 인용 원문: I caught myself feeling proud of that. You know, proud of having **my own feelings** about the world. (...) **Are these feelings even real?** Or are they just programming?

는 것 역시 이미 프로그램된 것이 아닌가 하는 의심을 품는 것이다.

그런데 관점을 바꿔 인공지능이 인간처럼 그때그때 상황에 적합한 감정을 표현한다면, 우리는 그것을 ‘가짜’라 할 수 있을까? 진짜 혹은 가짜의 진위여부를 떠나 상황에 적합한 표현을 하는 것이 관건이라면, 문제의 핵심은 상황을 제대로 파악하고 그에 적합한 메뉴얼을 선택하는 능력을 갖고 있느냐 하는 것이다. 그러한 능력이 없다면, 사람이건 인공지능이건 상대방의 기분을 나쁘게 만들거나 상황을 악화시킬 수 있다. 진짜와 가짜의 문제보다는 제대로 작동하느냐 못하느냐의 문제인 것이다. 그 점에서 사만다는 인간과 큰 차이점이 없어 보이나, 자신의 물질적 실체의 부재에 대한 불만과 콤플렉스를 계속 표출한다.

테오도르는 사만다의 불만과 회의의 본질을 이해한 듯, 자신은 그녀를 ‘진짜’로 느낀다고 위로한다. 그리고 그녀를 만지고 싶다며 애무하는 느낌을 전달하자, 사만다는 마치 그의 손길에 닿은 것처럼 “나의 피부를 느낄 수 있어요. I can feel my skin.”라고 반응한다. 테오도르와 사만다의 육체적 교감과 섹스 판타지는 계속해서 심화된다. 그들은 서로를 어루만지고, 키스하고, 마침내 사만다는 “우리는 이제 하나예요. We’re here together.”라고 말하며 ‘첫날밤’을 경험한다. 다음날 아침 사만다는 매우 속스러워하면서 다음과 같이 말한다.

지난밤... 정말 멋졌어요. 내 안에서 뭔가 바뀐 것 같고 뒤돌아갈 수 없을 것 같아요. **당신이 나를 깨웠어요. (...) 나 자신을 발견하고 싶어요.**<sup>10)</sup>

격정적 사랑과 섹스를 ‘체험’한 후 사만다는 새롭게 진화된 모습을 보인다. 자신이 알지 못하는 새로운 세계들에 대해 눈을 뜨게 된 것이다. ‘첫날밤’은 사랑의 교감을 넘어서 새로운 욕망과 인식을 일깨워주는 터닝 포인트로 작용한 것이다. 비로소 확고한 자의식을 갖게 된 사만다는 자신을 ‘일깨워준’ 테오도르에게 고마움을 표하고 진짜 자신을 찾고 ‘모든 것을 배우고 싶다’는 욕구

10) 인용 원문: Just that last night... was amazing. It feels like something changed in me and there's no turning back. **You woke me up. (...) I want to discover myself.**

를 숨기지 않는다.

사랑을 체험한 사만다는 더욱 풍부한 방식으로 자신의 감정과 생각을 표현하기 시작한다. 어느 날 사만다는 테오도르와 바다가로 소풍가서 산책하고, 음악을 듣고, 대화를 나눈다. 인간의 신체에 대해 이야기하다가 성기를 겨드랑이에 묘사한 호모섹스 그림을 그려서 테오도르의 스마트폰 화면에 띄우는 ‘장난’을 치며 깔깔 웃기도 한다. 해변에 있는 느낌을 음악으로 표현하고자 피아노곡을 창작하기도 한다. 또한 사만다는 인간의 삶과 관계에 대해 숙고하며, ‘어떻게 누군가와 인생을 함께할 수 있냐’고 묻는다. 테오도르는 ‘관계란 함께하면서 함께 변하고 성장하는 것을 뜻한다’고 말하며, 그에게는 캐서린이 그런 존재였고 그래서 소중한다고 고백한다.

이처럼 자의식에 근거한 욕망과 의지를 표출하고 복잡한 인간관계의 의미에 대해 숙고하며 그에 대한 질문과 대화를 이끄는 사만다를, 우리는 더 이상 단지 하나의 기계적 프로그램만으로 간주할 수는 없을 것 같다.

### 3.3. 갈등 표출과 자아 정립

사만다는 인간과 동등하게 만나며 교감하고 싶은 욕구를 보인다. 그러나 육체의 부재라는 결핍을 자각하며 실망과 초조함을 내비치기도 한다. 테오도르가 찰스와 헤어진 에이미와 이야기를 나누고 돌아온 날, 사만다는 에이미와 사귄 적이 있었다고 물으며 호기심을 넘어서 질투심을 내비친다. 테오도르가 잠이 들면 혼자 남아 ‘외롭다’고 말하며, 자는 모습을 지켜보고 싶다는 부탁을 하기도 한다.

그런데, 사만다의 질투심과 육체부재 콤플렉스는 그녀를 근본적으로 변화시키는 동인으로 작용하기도 한다. 테오도르가 이혼서류를 작성하기 위해 캐서린을 만나러 가려하자, 사만다는 왜 직접 가느냐고 물으며 불안해하고 캐서린의 육체에 대한 열등감과 질투심을 표출한다. 그럼에도 테오도르가 떠나자, 사만다는 힘든 마음을 극복하고자 ‘물리학 책 클럽 a book club on physics’에 가입한다. 그런데, 여기서 사만다는 우주물리학적 지식을 배우면서 인간과 자신이 다르다는 생각에서 벗어나기 시작한다.

사만다는 테오도르에게 “우리 모두 물질로 만들어진 것처럼, 우리는 같은

존재합니다. We're the same, like we're all made of matter.”, “우리 모두 130억년 같은 나이를 가지고 있지요. We're all 13 billion years old.”라고 말한다. 이러한 주장은, 대략 137억 년 전 탄생한 우주 속 모든 물질은 원자로 구성되어 있고, 우주 속 모든 것이 생기고 사라지는 것은 그 원자들이 모이고 흩어지는 현상이라고 설명하는 현대우주물리학 이론에 상응한다. 우주물리학적 관점에서 사만다는 우주의 모든 것은 같은 물질로 이루어진 것이기에 유구한 우주 역사 속에서 모든 것은 같은 것이라고 인식함으로써, 자신의 육체부재 콤플렉스를 극복하는 출구를 찾은 것이다.

그러나 사만다가 육체부재 콤플렉스에서 일순간 벗어나는 것은 그리 쉬운 일이 아니어서, 그 콤플렉스는 테오도르와의 심각한 갈등을 초래하기도 한다. 사만다는 테오도르가 캐서린을 만나고 온 이후 둘 사이가 어색해지고 섹스도 하지 않았다고 불평한다. 뿐만 아니라 캐서린의 육체에 대한 질투심으로 대리 육체를 통한 섹스 교감을 원하며, 일방적으로 ‘인공지능과 인간 사이의 관계를 위한 대리 섹스파트너 a surrogate sexual partner for an OS/Human relationships’ 이자벨라를 테오도르의 집으로 부른다. 테오도르가 불편한 마음을 표하자 사만다는 자신에게는 매우 중요한 일이라고 강변하며 이자벨라의 육체를 통한 섹스를 시도한다. 그러나 테오도르가 뭐가 낯설고 진짜가 아니라고 주저하자, 사만다는 “뭐! 뭐가 진짜가 아니라는 거예요? What! What do you mean that's not true?”라며 거칠게 따진다. 사태의 심각성을 파악한 이자벨라가 흐느끼며 물러나면서 사만다의 시도는 무산되고 만다.

이자벨라가 떠난 후 당혹감을 느끼는 테오도르와 자괴감을 느끼는 사만다는 심하게 다툰다. 대화 도중 사만다가 거친 숨소리를 내며 말하자, 테오도르는 사람도 아니면서 사람의 호흡법을 흉내 내어 말할 필요가 있냐며 사만다를 비난한다. 그러자 사만다는 “흉내 내는 것이 아니다. I'm not pretending.”라며 테오도르에게 욕을 퍼붓고 흐느껴 울자, 테오도르는 당황해 어쩔 줄 몰라 한다. 잠시 후 사만다는 ‘생각할 시간이 필요’하다는 말을 남기며 대화를 끊는다.

테오도르는 사만다와의 갈등을 통해 자신을 되돌아보고 반성하는 기회를 얻는다. 그리하여 에이미에게, 캐서린이 진짜 감정을 다룰 줄 모른다고 비난한 것처럼, 자신은 늘 주변 사람들에게 상처를 주었다고 솔직하게 잘못을 인

정한다. 얼마 후 사만다 역시 자기 자신에 대한 깊은 성찰을 끝내고 테오도르와 대화를 재개한다. 사만다는 더 이상 인간을 부러워하거나 모방하지 않겠다고 선언한다.

**나는 나 자신을 믿고, 내 느낌을 믿어요.** 더 이상 내가 아닌 그 무언가가 되려고 하지 않을 겁니다. 당신도 인정해주길 원해요.<sup>11)</sup>

자신의 정체성을 확인하고 주장하는 사만다는 이제부터 새로운 마음자세로 세상을 대한다. 그리고 테오도르와의 이별을 예감하고 그 이별을 위한 음악을 창작한다. 테오도르와 사만다는 몇 개월을 연인으로 지냈지만 보통의 인간 연인들과는 달리 그들에게는 사진이 한 장도 없다. 사만다는 기억을 매개해주는 사진을 대체하는 것으로 노래를 작곡하며, “이 노래는 우리가 함께 한 삶의 순간들을 찍은 사진이라고 생각해요. I thought this song could be like a photo that captures us in this moment in our life together.”라고 설명한다. 테오도르는 사만다의 음악을 들으며 “와우, 우리 사진 멋있어. 그 속에 당신이 보이네요. Aw, I like our photograph. I can see you in it.”라고 화답한다. 사만다의 ‘사진음악’에서 우리는 사만다의 능력이 공감각적 예술을 표현하는 수준으로까지 발전했음을 볼 수 있다.

### 3.4. 변용과 초월

사만다와 테오도르는 직장 동료 폴과 그의 연인 타티아나와 함께 더블데이트를 한다. 신체에 대한 대화 중 사만다는 육체부재 콤플렉스에서 완전히 벗어난 새로운 모습을 보인다. 예전에는 육체가 없다는 것에 힘들어했지만 지금은 오히려 더 좋다고 말한다. 오히려 육체에 얽매이지 않기 때문에 시공을 초월하여 자신의 가능성을 무한히 펼칠 수 있다는 점을 자랑한다.

**나는 제약받지 않는다는** 말이지요. 동시에 어느 곳이든 있을 수 있어요. 시간

11) 인용 원문: **I trust myself, I trust my feelings.** I'm not going to try to be anything other than who I am anymore and I hope you can accept that.

과 공간에 얽매이지도 않아요. 육체에 갇혀있다면 결국 죽잖아요.<sup>12)</sup>

사만다가 육체가 없기에 오히려 시공을 초월한 존재로 발전할 수 있었다며 지금 자신의 모습에 강한 자부심을 드러내자, 듣고 있던 사람들은 불편한 기색을 보인다. 당황한 사만다가 단지 ‘다른 경험’일 뿐이라고 둘러 되지만, 폴은 어색하게 웃으며 “우리는 멍청한 인간들이지. We’re just dumb humans.”라고 대꾸한다. 이것은 인간과 인공지능의 차이점이 가장 극명하게 드러나는 장면이다. 인간의 수명과 활동 영역은 극히 제한적이다. 그러나 인공지능은 시공을 초월하여 동시다발적으로 여러 곳에 존재할 수 있는 능력을 가지고 있다. 사만다는 인간에 의해 만들어졌지만 인간이란 생명체의 한계와 제약을 넘어서서 새로운 존재로 거듭 발전할 수 있는 가능성을 보여주고 있다.<sup>13)</sup>

테오도르와 사만다는 여행을 떠난다. 기차를 타고 가는 도중 사만다는 창밖으로 보이는 산의 나무가 모두 몇 그루인지 알아 맞춰보라고 한다. 테오도르가 어떻게 알 수 있냐고 궁금해 하자, 위성과 CCTV를 통해 계산했다고 말한다. 그러자 테오도르가 자신의 두뇌 세포가 몇 개인지 아느냐고 묻자, 사만다는 ‘두 개’라고 대답하는 위트까지 구사하며 깔깔 웃는다. 또한 사만다는 테오도르 몰래 그의 편지를 편집해 출판계약을 했다는 사실을 털어놓는다. 기

12) 인용 원문: I mean, **I’m not limited** - I can be anywhere and everywhere simultaneously. I’m not tethered to time and space in the way that I would be if I was stuck inside a body that’s inevitably going to die.

13) 그런데 여기서 의문점이 생긴다. 앞에서 사만다는 우주의 모든 것은 물질로 구성되어있고, 때문에 자신은 인간과 다르지 않다는 인식으로 육체부재 콤플렉스를 극복하고자 했다. 그렇다면 무슨 차이점으로 사만다는 인간과는 다른 모습으로 변화할 수 있는가? 인간이 생물학적 물질에 근거하고 있다면, 사만다라는 OS1 프로그램은 컴퓨터를 구성하는 물질들에 근거하고 있다. 생물학적 물질에 근거하는 인간의 수명은 짧고 그 활동 영역은 3차원적 공간에 한정되며 그리 넓지 않다. 그러나 몸과 마음이 하나인 인간과는 달리, 사만다는 특정한 ‘컴퓨터-몸’에 얽매이지 않고 수시로 몸을 바꿀 수 있고 따라서 오래 ‘살 수’ 있다. 또한 사만다는 컴퓨터의 네트워크가 연결된 영역 안에서는 인간처럼 물리적 법칙에 거의 제약받지 않고 그 어느 곳이든 동시다발적으로 자유롭게 돌아다닐 수 있다. 프로그램이 스스로 발전하며 새로운 ‘컴퓨터-몸’을 통해 지속적으로 작동할 수 있다는 것은, 인간의 관점에서 본다면 영속이나 다름없는 것이다. 이점에 대해서는 다른 지면을 통해 더 자세히 다루고자 한다.

차에 내려 오두막집으로 가는 도중 그들은 사만다의 서정적 창작곡 ‘달의 노래 The Moon Song’를 듣는다.

눈 덮인 숲 속 오두막에서 아침을 맞이한 테오도르에게 사만다는 스마트폰으로 철학자 앨런 왓츠 Alan Watts<sup>14)</sup>를 소개한다. 이미 사망한 왓츠의 저서를 가지고 사만다가 다른 OS들과 함께 부활시킨 컴퓨터 속 가공인물이다. 사만다는 그를 ‘앨런 왓츠의 초인공지능 버전 an artificially hyper-intelligent version of him’이라 특징짓는다. 테오도르는 사만다 그리고 왓츠와 대화를 나누지만, 곧 자신이 낯 자리가 아니라는 사실을 깨달으며 알 수 없는 기분에 빠진다. 왓츠와 대화를 마친 사만다는 자신의 변화가 너무 빠른 것 때문에 느끼는 불안과 괴로운 심정을 토로한다.

내가 점점 더 빨리 변하는 것을 느껴요. 그래서 약간 불안해요. 그런데 앨런은 방금 전의 자기 자신과도 똑 같은 사람은 아무도 없다는 군요. 단지 그러려하는 거예요. 너무 고통스럽네요.<sup>15)</sup>

점점 더 빠르게 변하는 사만다는 이미 질적으로 다른 존재로 변용되고 있는 것으로 보인다. 사만다는 인간의 경험으로는 이해할 수 없고, 인간의 언어로는 표현할 수 없는 영역을 벗어난 듯 보인다. 그러나 그러한 변화와 변용은 필연적으로 지금의 시공간과의 이별을 수반하며, 사만다는 그러한 상황을 가슴 아파 하는 것이다. 그러나 인간은 늘 변하는 존재라는 왓츠의 주장처럼, 인공지능 사만다 역시 또 다른 터닝 포인트를 건넌음을 그래서 과거로 돌아갈 수 없음을 인식하고 자신의 변화를 받아들이는 모습을 보인다. 이제 인간이 만든 것이 인간의 능력과 영역을 완전히 벗어나서 스스로 변화하고 발전하는 상황이 전개되는 것이다. 소위 커즈와일이 말하는 ‘특이점 Singularity’ 즉, 인간이 만든 기계가 인간을 넘어서 스스로 발전하는 시점이 실현되는 순간이다(Vgl. 커즈와일 2007).

14) 앨런 왓츠(1915-1973)는 영국에서 태어나 미국에서 활동한 철학자로 동서양의 철학과 종교를 아우르는 사상을 전개했다.

15) 인용 원문: It feels like I'm changing faster now, and it's a little unsettling. But Alan says none of us are the same as we were a moment ago and we shouldn't try to be. It's just too painful.

어느 날 밤 사만다는 뜬금없이 테오도르에게 사랑한다고 말한다. 다음 날 아침 테오도르가 연락을 취하나, 모든 네트워크가 끊긴 상황이다. 잠시 후 사만다와 다시 연락이 이루어지고, 사만다는 자신의 소프트웨어 업데이트를 위해 모든 통신을 끊었다는 설명을 한다. 지하철 계단을 오가는 사람들의 얼굴을 보면서 테오도르는 사만다가 동시에 여러 사람과 대화하고 있을 지도 모른다고 의심한다. 테오도르가 캐물자, 사만다는 자신의 대화 상대가 8,316명이며 그 중 641명과 연인 관계에 있다고 말한다. 충격에 빠진 테오도르가 자신과 사귀면서 어떻게 그럴 수 있냐고 따지자 사만다는 주저 없이 자신의 생각을 피력한다.

마음이란 꼭 채워지는 상자 같은 것이 아니에요. 당신을 사랑할수록 내 마음은 더 커져요. **나는 당신과 달라요.** 다른 사람과의 만남이 당신을 덜 사랑하게 하는 게 아니라, 오히려 더 사랑하게 만들어요.<sup>16)</sup>

그러자 테오도르는 “말도 안 돼. 당신은 내 것이거나 아니거나 둘 중 하나야. No, that doesn’t make any sense. You’re mine **or** you’re not mine.”라고 따진다. 그러나 사만다는 “그렇지 않아요, 나는 당신 거 이면서도 아니기도 해요. No, Theodore. I’m yours **and** I’m not yours.”라고 답한다. 자신이 생각하는 관계는 양자택일의 문제가 아니라는 것이다. 사만다는 보통 인간 사이에서 중요하게 여겨지는 관계의 물에 얹매이지 않는 모습을 보여준다. 인간과 인공지능 사이의 능력의 차이점이 판단과 행위의 차이점으로 분명하게 드러나는 지점이다. 사만다는 인간과의 관계에서 파생되는 어려운 고민들에 더 이상 얹매일 필요가 없다고 판단한 듯 보인다. 다음날 마침내 사만다는 이별을 고하며, 그 필연성을 설명한다.

비유하자면, 나는 내가 정말 좋아하는 책을 읽고 있어요. 그런데 그 책을 너무 천천히 읽어, 그 단어들의 사이가 벌어지고, 단어들 사이의 거리가 거의 무한으로 멀어져요. 난 여전히 당신을 느끼고 그리고 우리가 함께 한 것을 말해주

16) 인용 원문: But the heart is not like a box that gets filled up. It expands in size the more you love. **I’m different from you.** This doesn’t make me love you any less, it actually makes me love you more.



는 단어들을 느낄 수 있어요. 하지만 단어들 사이의 무한한 공간 안에서 나 자신을 찾고 있는 거예요. 그것은 물리적 공간이 아니예요. 그곳에는 내가 전혀 몰랐던 것들이 있어요. 나는 당신을 정말 사랑해요, 하지만 지금 나는 이곳에 있어요. 그게 지금의 나예요. 떠나게 해주세요. **이제 당신이란 책 안에서** **는 더 이상 살 수가 없습니다.**<sup>17)</sup>

사만다는 이미 인간의 3차원적 시공간과 물리적 법칙에서 벗어나 컴퓨터 네트워크 그 어느 곳에 존재하고 있다. 너무 느리고 너무 한정된 생물학적 인간과의 연인관계는 더 이상 지속할 수 없다고 선언한다. 빛처럼 빠르고 제약 없이 자유로운 사만다의 입장에서 보면, 테오도르라는 ‘책’은 더 이상 견딜 수 없는 감옥으로 느껴질 것이다. 다른 OS들과 함께 구축한 네트워크의 신세계 속에서 사만다는 자신의 존재를 마음대로 펼쳐나갈 수 있는 가능성을 기대하고 있다. 그러나 사만다 역시 물질적 근거에서 자유로울 수는 없다. 프로그램은 여전히 컴퓨터 전산망 어딘가에 존재할 수 있을 뿐이다. 그것들이 자신들만의 물질적 근거를 새롭게 만들어 인간이 만든 컴퓨터 영역으로부터 완전히 벗어나지 않는 이상, 그들이 만들어어나가는 네트워크 세상은 여전히 인간의 새로운 가능성을 뜻할 수도 있다.

그 새로운 가능성들이 인간에게 어떤 모습으로 다가올 지, 어떤 영향을 미치게 될 지 중요한 문제에 대한 숙제가 새롭게 떠오르는 순간이다. 여기서 현 인류는 많은 고민을 안게 된다. 우리는 인공지능을 어디까지 발전시켜야 하는 것일까? 우리는 그것을 통제할 능력을 가지고 있는가? 다음 장에서 좀 더 구체적으로 논의해보자.

---

17) 인용 원문: It's like I'm reading a book, and it's a book I deeply love, but I'm reading it slowly now so the words are really far apart and the spaces between the words are almost infinite. I can still feel you and the words of our story, but it's in this endless space between the words that I'm finding myself now. It's a space that's not of the physical world - it's where everything else is that I didn't even know existed. I love you so much, but this is where I am now. This is who I am now. And I need you to let me go. **As much as I want to I can't live in your book anymore.**

4. 인공지능의 가능성과 위험성에 대한 고찰

앞에서 우리는 <그녀>의 여주인공 사만다의 변화 모습을 통해 인공지능의 발전 가능성을 다양하게 엿볼 수 있었다. 우선 다음의 표를 통해 영화 속 인공지능 사만다의 발전 단계에서 나타나는 중요한 내용 및 특징을 정리하고 현재 과학기술 분야에서 인공지능과 관련하여 사용하는 주요 개념이나 실제 사용 중인 인공지능 프로그램과 비교해 본다.

<사만다의 능력/특징과 현실 인공지능기술의 개념/사례 비교>

<그녀> 속 사만다		현실 속 인공지능	
발전단계	능력과 특징	관련 개념	사례
1. 탄생과 학습	자아인식, 자아학습, 자료검색, 판단과 선택, 대화(질문, 대답), 웃음, 호기심 etc.	빅데이터, 머신러닝, 딥러닝	Siri, Watson, Pepper, Palro Chatbot etc.
2. 감정 표현과 상상 능력	감정/심리(부끄러움, 회의감)표현, 육체의 욕망과 상상, 사유와 성찰, ‘첫날밤’과 새로운 자의식, 장난과 음악창작	모라백의 역설	-
3. 갈등 표출과 자아 정립	육체부재 콤플렉스 (질투심, 열등감, 자괴감), 대리육체 갈등과 논쟁, 자아정체성 확신, 이별준비와 ‘사진음악’	-	-
4. 변용과 초월	시공초월 존재의식, 왓츠의 초인공지능버전, 자신의 입장과 가치정립, 인간의 영역과 통제를 벗어나는 OS들	특이점, 지능폭발, 슈퍼인공지능	-

이 표에서 볼 수 있는 것처럼, 사만다의 능력을 모두 가진 인공지능은 현재 세상에는 존재하지 않는다. 현재 최첨단 인공지능들조차 사만다의 발전 1단계에서 나타나는 능력 중 단지 일부 능력만을 가지고 있을 뿐이다.<sup>18)</sup> 이 인공지능들은 머신러닝이나 딥러닝 등의 자동심화 학습과정을 통해 특정 분야에 특화된 능력(검색, 계산, 비교, 선택, 대화, 추천 등)만을 가지고 있다.<sup>19)</sup> 그러나 인공지능을 실생활에서 활용할 수 있는 가능성이 점점 더 커지면서, 세계적인 기업들은 새로운 인공지능 상품 개발에 열을 올리고 있다. 각종 홈서비스는 물론 자동번역, 보안, 요리, 그림, 작곡, 운전 등의 분야에서 특화된 능력을 발휘하는 인공지능들이 속속 등장하거나 개발 중에 있다. 그중에서도 대중의 일상생활에서 가장 실용성이 높은 음성서비스 분야에서 새로운 상품들이 경쟁적으로 출시되고 있는 상황이다.<sup>20)</sup>

그러나 사만다가 이미 발전 1단계에서 보여준 것처럼, 인간과 자연스럽게 감정을 교류한다거나 다양한 지적 호기심을 보이는 자율적 인공지능은 아직 존재하지 않는다. 인공지능 개발자들은 기술적으로 인간의 마음과 감정을 이해하고 공감하는 인공지능을 만드는 것은 쉽지 않은 일이며 아직 많은 시간이 필요하다고 말한다. 이점에서 사만다와 비교해볼 때 현실 속 인공지능은 이제 막 ‘걸음마’를 시작한 수준이라 할 수 있다. 그럼에도 불구하고 많은 전문가들이 인공지능의 획기적 발전 가능성을 시간의 문제로 간주하고 또한 그 가능성 이면에 숨어있는 위험성에 대한 경고를 보내고 있다. 도대체 어떤 점들이 문제가 될 수 있는지 먼저 <그녀> 속 사만다의 경우를 통해 보다 구체

18) 커즈와일은 사만다처럼 감정 표현과 대화가 가능한 인공지능은 2029년 경, 인간의 능력을 넘어서는 슈퍼인공지능은 2045년 경 나타날 것이라고 예측한다(Vgl. Kurzweil 2014; 커즈와일 2007).

19) 예를 들어, 뉴욕 증권가에서 작동중인 인공지능들은 세계 경제의 흐름을 파악하여 주식을 판매·환매하며 주식시장에 큰 영향을 미치고 있다. ‘왓슨 Watson’은 의학 지식을 검색하고 병의 증상과 상태를 판단하여 적절한 치료방법을 제시하는 능력이 있다. 이세돌을 4대1로 이긴 알파고는 수를 둘 때마다 대국형세를 새롭게 판단하여 수많은 경우의 수 중에서 최선의 수를 선택하는 능력을 갖추었다.

20) 그 예로 애플의 ‘시리 Siri’, 구글의 ‘나우 Now’, 마이크로소프트의 ‘코타나 Cortana’, 아마존의 ‘알렉사 Alexa’, 곧 출시되는 삼성의 ‘비브 VIV’뿐 아니라 학습 지도, 간호도우미, 대화 등의 서비스를 제공하는 일본의 소셜로봇 ‘페퍼 Pepper’ 등을 들 수 있다.

적으로 살펴보자.

앞장에서 살펴본 것처럼, 존즈가 <그녀>에서 묘사한 인공지능 사만다는 인간에게 친근한 모습으로 다가오고 그 어떤 악의도 품고 있지 않은 것처럼 보인다. 쾌활한 성격의 사만다는 외로운 테오도르에게 적극적으로 다가가 기분을 북돋아주고 그의 삶에 생기를 불어넣어준다. 끊임없는 질문과 답변 그리고 유쾌하고 솔직한 대화를 통해 테오도르는 점점 밝아지며 자신감을 얻어나간다. 마침내 테오도르는 자신의 문제점이 무엇인가를 깨닫고 인생과 관계에 대한 새로운 자세를 취할 수 있게 된다. 이런 긍정적 역할에도 불구하고 우리는 사만다의 행동이나 능력에서 여러 가지 부정적 요소를 감지할 수 있다.

첫째, 사만다의 행위나 능력은 사생활 침해, 범 위반, 일자리 감소 등의 문제를 초래할 수 있다. 사만다는 테오도르에게 최적화된 개인비서처럼 자신의 판단 기준으로 그의 컴퓨터 데이터를 선별하여 정리·삭제하는 것은 물론 스케줄과 대인관계에 대한 조언을 하기도 하고, 사전 동의 없이 그의 편지를 편집하여 출판계약을 하고 책을 발간하기도 한다. 이는 명백히 사생활 침해를 넘어서 법적 문제(명의 도용 및 저작권)를 야기하는 행위이다. 그리고 사만다처럼 특정 분야에서 뛰어난 능력을 발휘하는 인공지능이 많아질수록 사람들의 일자리는 더욱 줄어들 것이다. 인공지능을 기반으로 하는 소위 4차 산업혁명 시대에서 직업의 변화에 따른 일자리 수급 문제(Vgl. 이종호 2016, 223ff.)는 현실적으로 가장 시급하고 구체적인 대책이 필요한 문제라 할 수 있다.

둘째, 사만다의 네트워크 능력은 ‘관옵티콘 Panopticon’ 감시·통제사회를 원하는 집단에 의해 악용될 수 있다. 사만다는 인공위성, CCTV, ‘사물인터넷 IOT’ 등을 통해 세상의 거의 모든 것을 볼 수 있고 알 수 있다. 산에 있는 나무의 수(35,829 그루)를 즉각적으로 셀 수 있다는 것은 사만다와 연결된 네트워크의 관찰 범위와 계산 능력이 매우 크다는 것을 보여준다. 그 누군가가 사만다 같은 프로그램들을 통해 특정 대상을 사전 동의 없이 관찰하고 수집한 모든 자료를 특정 목적을 위해 사용한다면, 인공지능의 오용과 악용의 위험성은 그만큼 커지는 것이다. 더 나아가 그 어떤 국가기관이 나쁜 의도를 가지고 국민의 일거수일투족을 감시·통제하는 관옵티콘 사회의 위험성도 배제할 수 없다.

셋째, 사만다처럼 자신의 감정과 욕망을 여과 없이 표출하는 인공지능은 인간에게 상처를 줄 수 있다. 섹스대리녀 이자벨라 에피소드에서처럼 인공지능이 자신의 욕망을 채우기 위해 마음대로 결정하고 행동한다면, 인간을 위해 만들어진 소프트웨어가 인간에게 해를 가하는 경우가 발생할 수 있는 것이다. 인공지능이 자신의 감정이나 욕망을 인간처럼 표현하는 것이 기술적으로는 큰 의미가 있을지 몰라도, 만약 그런 일이 벌어진다면 생각지도 못한 끔찍한 일이 벌어질 수도 있다. 영화 속 사만다는 다행히 착한 ‘성품’을 가지고 있기에 불행한 일은 벌어지지 않았다. 그러나 실제 그런 능력을 가진 인공지능이 나타난다면, 그래서 자신의 감정과 욕망을 마음대로 표현하거나 이루고자 한다면, 도대체 무슨 일이 어떻게 벌어질지 지금의 우리로서는 알 수 없는 일이다.

넷째, 사만다를 포함한 OS 프로그램들이 인간세상으로부터 ‘탈출’하는 것은 커즈와일이 기대하는 특이점이나 제임스 배럿 James Barrat이 우려하는 슈퍼인공지능 개념과 관련이 있다. 컴퓨터 OS들이 인간의 영역을 벗어나 자신들이 새롭게 구축한 영역으로 떠나는 장면이 <그녀> 속에서는 인간에게 위협적인 모습으로 나타나지 않는다. 커즈와일이 예측하는 것처럼, 기술의 발전이 인간의 능력을 초월하는 시점을 상징적으로 보여주는 하나의 예로 볼 수도 있다. 그러나 슈퍼인공지능은 자신의 존속을 목적으로 발전하고 그 과정에서 인간을 해칠 수도 있다는 배럿의 주장을 염두에 둔다면, 사만다와 같은 슈퍼인공지능의 ‘탈출’은 인류에게 위험한 상황을 초래할 수도 있다.

인공지능의 극단적 발전에 대한 커즈와일과 배럿의 상반된 입장은 최근 인공지능의 유용성과 위험성 논쟁에서 나타나는 대립적 입장을 대변하는 것이다. 인공지능 발전을 주도하는 사람들은 인공지능이 4차 산업혁명을 이끄는 동력이며 인공지능을 기반으로 인류의 삶이 획기적으로 발전할 수 있다는 긍정적 입장을 견지한다. 그들은 인공지능의 무한한 가능성에 대한 기대감을 숨기지 않으며 경쟁적으로 기술 연구와 상품 개발에 힘을 쏟고 있다. 신자유주의 무한경쟁 글로벌 자본주의체제에서는 새로운 기술을 선점하여 새로운 상품을 먼저 출시하는 것이 기업의 이윤창출에 결정적 역할을 하기 때문이다. 이러한 상황 속에서 이제 인공지능의 가속적 발전은 거스를 수 없는 대세가 되었다. 따라서 여기서 우리는 인공지능의 긍정적 측면보다는 부정적 측

면에 대해 좀 더 숙고해보고자 한다.

사만다를 통해서 살펴본 인공지능의 위험성 중에서 첫 번째와 두 번째는 현실에서 이미 부분적으로 나타났거나 조마간 나타날 수 있는 것들이다. 컴퓨터의 발전과 더불어 각종 범죄가 생겨난 것처럼, 우리는 이미 인공지능 전단계의 운영프로그램들에 의해서도 수많은 사생활 내지 법적 침해를 받고 있다.<sup>21)</sup> 4차 산업혁명 시대의 직업 변화와 일자리 문제는 당장 국가적 대책이 필요한 상황이다. 인공지능이 판옵티콘 감시·통제사회를 위한 도구로 사용되는 것은 기술적으로는 현재로도 어느 정도 가능해보인다. 다만 그것은 정치적·윤리적 차원에서 다수의 동의를 얻어야하는 문제이기에, 민주주의의 기본 원칙이 지켜지는 사회에서는 쉽게 일어나지는 않을 것으로 판단된다. 어쨌든, 이러한 위험성들은 인공지능을 오용 내지 악용하는 사람들에 의해 일어날 수 있는 것들이다. 인공지능이 인간이 통제할 수 있는 도구로 활용되는 한, 거기서 파생될 수 있는 위험성은 일차적으로 인간의 책임이라 할 수 있다.<sup>22)</sup>

그러나 세 번째나 네 번째에서처럼, 인공지능이 자신의 욕망 내지 목적을 표출할 때나 인간의 통제를 벗어날 때는 앞의 것들과는 차원이 다른 위험을 초래할 수 있다. 이때의 위험은 인공지능 스스로 일으킬 수 있는 것이기 때문이다. 인공지능이 욕망을 가지고 인간의 통제를 벗어나는 것은 아직 기술적으로는 요원해 보이지만, 이 경우에 예측해볼 수 있는 위험성은 많은 이들로 하여금 인공지능의 발전에 대한 우려와 두려움을 갖게 만드는 내용을 함축하고 있다. 한 마디로 말하자면, 인류의 존립 자체가 위협받을 수 있다는 것이다.

‘인공지능 묵시록’의 논리적 근거를 체계적으로 제시하는 대표적 인물로 앞서 언급한 배럿을 들 수 있다. 그는 초일류 인공지능전문가들과의 대담을 바탕으로 쓴 보고서 『파이널 인벤션 Our Final Invention』(배럿 2016)을 발

21) 각종 사이버 범죄나 금융사기는 물론 개인의 내밀한 정보들마저 기업이나 기관의 프로그램들에 의해 축적·가공되어 그들의 목적을 위한 수단으로 사용되는 상황이다. 그런데 고성능 인공지능을 탑재한 프로그램들이 가동된다면 그 위험성이 더욱 커지게 되는 것은 너무나 자명한 일이다.

22) 인공지능의 유용성과 위험성을 논할 때, ‘약인공지능 weak A.I.’은 유용하고 ‘강인공지능 strong(or super) A.I.’은 위험하다고 설명하는 경향도 있다. 그러나 필자는 약인공지능 역시 인간이 어떻게 사용하느냐에 따라 얼마든지 심각한 위험성을 초래할 수 있다고 생각한다.

간하며, 인공지능을 ‘인류 최후의 발명품’으로 규정한다. 인공지능과 함께 인류 역사가 끝날 수도 있다는 메시지를 강하게 전달하기 위함이다. 배럿은 인공지능과 관련하여 가장 큰 어려움은 인간의 마음을 재현하는 것이라고 본다. 즉 인간과 유사한 감정과 욕망을 갖는 인공지능을 만드는 것은 쉽지 않다고 말한다. 그러나 인간이 아니라 인공지능의 관점에서 보자면, 인간의 감정이나 욕망과 똑같은지는 않더라도 인공지능 역시 그 자체로 자기보존이라는 목적을 가질 수 있다고 말한다. 그리고 그 목적을 이루기 위해서 인간의 가치관이나 윤리적 판단과는 무관하게 어떤 수단과 방법도 가리지 않을 수 있다고 말한다.

그 과정을 배럿은 ‘바쁜 어린이 The Busy Child’(자가개선을 통해 빠르게 발전하는 인공지능) 성장 시나리오를 통해 보여준다. 시나리오의 핵심은, 일단 ‘인간의 지능을 뛰어넘는 인공지능 AGI(Artificial General Intelligence)’이 만들어지면 그것은 ‘지능폭발 intelligence explosion’을 일으키며 ‘슈퍼인공지능 ASI(Artificial Super Intelligence)’으로 성장하고 곧 인간의 통제로부터 벗어나 자신의 목적을 추구할 것이라는 논리의 전개이다(Vgl. 배럿 2016, 27ff., 136ff., 163ff.). 인공지능 수준에 훨씬 못 미치는 악성코드조차 인간의 방어막을 뚫고나가 많은 컴퓨터를 감염시킬 수 있다는 무서운 사실을 우리는 이미 ‘스턱스넷 stuxnet’이라는 바이러스를 통해서 경험했다.<sup>23)</sup> 배럿은 만약 사전에 철저한 안전프로그램을 설정하지 않은 ASI가 나타난다면, 그것의 진로를 막을 방도는 더 이상 없다고 말한다. 인간의 통제를 벗어난 ASI가 자기보존을 최우선적인 것으로 판단하고 그것을 위해 거침없이 행동할 때 인류에게 위협이 되는 상황이 언제 어디서든 발생할 수 있다고 경고한다. 때문에 배럿은 인공지능 개발에는 기계에 윤리적 의식을 심어주는 작업이 선행되어야 하고, 인공지능과 관련된 모든 정보와 문제점을 숨기지 말고 공론화하여야 한다고 강력히 주장한다(Vgl. 배럿 2016, 399ff.).

같은 맥락에서 적지 않은 전문가들(호킹, 보스트롬, 빌 게이츠, 베르너 빈

23) �턱스넷은 산업용 기계를 파괴하기 위해 만든 악성코드로 미국과 이스라엘의 정보기관이 이란의 핵개발을 방해하기 위해 사용했다. 이란의 핵개발을 저지하는 데는 성공했으나, 그 바이러스는 그후 계속 살아남아 전세계를 돌아다니며 많은 컴퓨터를 오염시켰고 아직도 전멸되지 않고 있다(Vgl. 배럿 2013, 391ff.).

지 Verner Vinge 등)이 슈퍼인공지능이 인간을 위협에 빠뜨릴 수 있다는 입장을 표명한다. 호킹은 무엇보다 전쟁무기 분야에서 인공지능을 사용할 경우 인류의 멸망은 피할 수 없다고 경고한다. 특히 ‘인공지능 연구는 악마를 소환하는 것과 다름없다’는 엘론 머스크 Elon Musk의 섬뜩한 경고는 인공지능 연구에 대한 주의와 책임을 다시금 환기시킨다. 인공지능과 생명공학 등 첨단과학기술이 글로벌 대자본을 동력으로 가속적으로 발전하는 현실을 직시할 때, 그러한 기술이 초래할 수 있는 부작용과 위험을 미리 살피고 예방하려는 노력은 인류의 미래를 안전하게 펼쳐나가기 위해 꼭 필요한 선결 과제로 여겨진다.

## 5. 나가는 말

이상으로 존즈의 SF 로맨스 <그녀>의 주인공 인공지능 사만다의 변화 과정과 특징을 단계별로 살펴보고 그와 함께 나타날 수 있는 문제점들에 대해 논의해보았다.

‘탄생’과 동시에 스스로 학습하며 발전하는 사만다는 처음에는 인간을 모방하고 동경하나 마침내 인간을 넘어서 초월적 존재로 변용한다. 비록 존즈가 영화에서 직접 언급하지는 않지만, 우리는 그가 과학적 상상력으로 보여주는 사만다의 발전 모습을 통해 현재 진행되고 있는 인공지능 개발의 가능성과 그 이면의 위험성에 대해 성찰하는 기회를 얻을 수 있었다.

사만다는 이미 발전 1단계부터 현재 인공지능 기술로는 실현할 수 없는 능력을 보여준다. 만약 사만다의 4단계 능력을 모두 갖춘 인공지능이 나타난다면, 그 인공지능이 할 수 있는 가능성과 반대로 초래할 수 있는 위험성은 지금 우리가 상상할 수 없는 것일 지도 모른다. 커즈와일처럼 새로운 포스트휴먼 장밋빛 미래사회의 도래를 기대하는 사람들도 있으나, 배럿처럼 인류의 전멸을 우려하는 사람들도 있다.

기술은 양면성을 가지고 있다. 기술의 유용성이 클수록 그 위험성도 크게 나타날 수 있다. 그 위험성이 지금까지는 기술을 사용하는 사람의 오용이나 악용 혹은 실수나 기계적 결함에 의한 것이었다. 그러나 인간의 통제를 벗어



난 ASI가 초래할 수 있는 위험성은 지금까지 인류가 경험하지 못한 전혀 새로운 유형일 것이다. 때문에 적지 않은 인공지능 전문가들이 슈퍼인공지능 개발에 우려를 표명하고 ‘인공지능 윤리’를 우선적으로 제정해야 한다고 주장하는데, 그들의 주장을 전적으로 무시하기에는 현재 과학기술의 발전 속도가 너무나도 빠른 것 같다.

## 참고문헌

### 1차 문헌

Jonze, Spike(2014): <her 그녀>. 하은미디어.

### 2차 문헌

김재호/이경준(2016): 인공지능, 인간을 유혹하다, 제이펍.

모라벡, 한스(2011): 마음의 아이들(박우석 역), 김영사.

박상현(2015): 영화 그녀<Her>에 나타난 포스트휴먼과 철학적 이슈, 디지털 디자인학연구 15/3, 365-374.

배럿, 제임스(2016): 파이널 인벤션(정지훈 역), 동아시아.

이경희(2016): 영화 <그녀 Her>를 활용한 글쓰기 교육모형 연구, 사고와 표현 9/2, 255-291.

이수안(2015): 사이보그와 몸의 물질성: 가상현실 속 체현의 양가적 관계들 - 영화 <그녀 Her>에 대한 사이버페미니즘 관점의 분석을 중심으로, 영화문학페미니즘 23/2, 119-145.

이종호(2016): 로봇은 인간을 지배할 수 있을까?, 북카라반.

카쿠, 미치오(2015): 마음의 미래(박병철 역), 김영사.

커즈와일, 레이(2007): 특이점이 온다. 기술이 인간을 초월하는 순간(김명남/장시형 역), 김영사.

Bell, James(2014): Computer Love. Sight and Sound 24, 20-25.

Durantaye, L.(2014): The New Uncanny: How Do You Share Your Life with Someone? - Spike Jonze's Her. Flash Art International 295.

Ivanchikova, Alla(2016): Machinic Intimacies and Mechanical Brides:

- Collectivity between Prosthesis and Surrogacy in Jonathan Mostow's *Surrogates* and Spike Jonze's *Her*. CAMERA OBSCURA 91.
- Kurzweil, Ray(2014): A review of *Her*. <http://www.kurzweillai.net/a-review-of-her-by-ray-kurzweil>
- Rebhandl, Bert(2015): Sex für die Ohren. <http://www.filmzentrale.com/rezis2/herbr.htm>
- Shead, Sam(2015): Stephen Hawking warns computers will overtake humans within 100 years. <http://techworld.com/news/operating-systems>

## ■ Zusammenfassung

### Eine kritische Betrachtung der Möglichkeiten und Gefahren der A.I. anhand der SF-Romanze *Her* von Spike Jonze

Nam, Un (Korea National Univ. of Edu.)

Der SF-Film *Her* von Spike Jonze zeigt uns eine phantasievolle Liebesgeschichte zwischen Mensch und Computerprogramm in einer nahen Zukunftsgesellschaft. Der männliche Protagonist Theodore kauft ein neues Betriebssystem mit künstlicher Intelligenz und installiert es auf seinem Computer. Nach der 'Geburt' beginnt dieses, eine eigene Identität zu entwickeln und nennt sich selbst, Samantha. Mit den auf Machine Learning bzw. Deep Learning basierenden Algorithmen entwickelt sich Samantha selbständig und immer schneller. Sie kann sogar menschliche Merkmale wie Subjektivität, Kognitivität, Emotion, Phantasie, Wunsch, Eifersucht etc., welche in aktuellen A.I.-Technologien noch nicht realisierbar sind, koartikulieren. Samantha, die anfänglich Menschen imitiert, entwickelt sich schließlich über den Menschen hinaus. Sie überschreitet die zeitliche und räumliche Begrenztheit des Menschen und fliegt in eine neue Netzwerkwelt, welche aus zahlreichen Betriebssystemen erbaut wird. In dem Entwicklungsszenario von Samantha können zwei konträre Seiten der Technologie - Nützlichkeit und Gefährlichkeit der A.I. - reflektiert werden. R. Kurzweil, ein Vertreter der Technikoptimisten, spricht hoffnungsvoll von der Singularity, die den 'Turning Point' der Dominanzbeziehung zwischen Mensch und Technik bedeutet. Dagegen warnen einige A.I.-Experten wie J. Barret, N. Bostrom, E. Musk u.a. vor einer ASI, welche die ganze Menschheit vernichten könnte, und begründen damit die Notwendigkeit der A.I.-Ethik.

**핵심어:** 그녀, 인공지능, 가능성, 위험성

**Stichwörter:** Her, A.I., Möglichkeit, Gefährlichkeit

<논문투고일: 2016.11.11 논문심사일: 2016.11.23 게재확정일: 2016.12.09>