

# Colloquium on the Implementation of UNESCO's 2005 Convention in Digital Age

“유네스코 문화적 표현의 다양성 보호와 증진에 관한 협약  
제2차 국내 전문가회의”

-인공지능시대에 변화하는 창작의 개념과 문화적 표현의 다양성-

**유네스코문화적표현의다양성보호와증진에관한협약 - 제2차 국내 킬로쿼엄(자료집)**

펴낸 때 | 2020년 6월 26일

펴낸 곳 | 유네스코한국위원회 문화팀

(04536) 서울특별시 중구 명동길 (유네스코길) 26

유네스코한국위원회 간행물등록번호 | CL-2020-DI-1

©유네스코한국위원회

# 목 차

I. 행사개요 .....	1
II. 연사 & 토론자 소개 .....	3
III. 발표 1 .....	7
인공지능시대, 창작 개념의 재검토 이상욱, 한양대학교 철학과 교수	
IV. 발표 2 .....	18
AI와 인간의 협업가능성과 문화적 표현의 다양성 김재인, 경희대학교 비교문화연구소 학술연구교수	



# 행사 개요

## 1 사업 배경

- 「문화적 표현의 다양성 보호 및 증진을 위한 협약」(이하 2005년 협약)은 문화다양성의 중요성을 확인하고, 각국의 문화적 주권을 인정하며, 발전과 창의성의 원천이 되는 문화의 힘을 강조하기 위해 2005년 채택된 협약으로 현재 149개국(EU포함)이 비준함
- 2018문화가치사슬의 급격한 변화, 대형 온라인 플랫폼의 급성장, 인공지능의 발달 등으로 인한 시장집중, 공공 통계 부족, 인공지능에 대한 독점 등의 새로운 과제가 부상하는 디지털 시대에 맞는 문화다양성협약 이행을 위한 전략이 유네스코차원에서 지속적으로 논의됨
- 급변하는 시대적 환경에 대응하고, 2005년 협약을 효과적으로 이행하기위한 국내외차원의 논의가 필요함

## 2 사업 개요

□ **사업명** : 유네스코 문화적 표현의 다양성 보호와 증진에 관한 협약 - 제1차 국내 전문가회의  
(인공지능시대에 변화하는 창작개념과 문화적 표현의 다양성)

### □ 목적 및 주요내용

- 2005년 협약에 대한 디지털 주체들의 이해 제고
- 문화적 표현의 다양성 관점에서 인공지능 시대의 기회와 위기 논의
- 인공지능시대 2005년 협약이 가지는 의미와 올바른 이행방안 논의

### □ 행사개요

- 일시: 2020년 6월 26일(금) 14:00 - 17:30
- 장소: 서울 중구 웨스틴조선호텔 2층 오키드룸
- 참가자: 2005년 협약 및 인공지능 관련 국내 전문가 및 관계자 약 30여명
- 주요내용
  - 인공지능 시대에 급격히 변화하는 문화가치사슬
  - 인공지능시대 문화적 표현의 다양성과 증진과 보호를 위한 대응방안 논의
  - 인공지능시대 문화분야의 지속가능하고 평등한 문화생태계 구축

□ 프로그램

시 간		세부내용
13:30-14:00	30'	참가자 등록
14:00-14:10	10'	<전문가 회의 개회식> 개회사   유네스코한국위원회 김광호 사무총장
14:10-15:00	50'	발표 1 : 인공지능시대, 창작 개념의 재검토 - 이상욱, 한양대학교 철학과 교수 (25')  발표 2 : AI와 인간의 협업가능성과 문화적 표현의 다양성 - 김재인, 경희대학교 비교문화연구소 학술연구교수 (25')
15:00-15:30	30'	커피브레이크
15:30-17:10	100'	<b>토론 (100')</b> <b>주제: 인공지능 시대에 변화하는 창작개념과 문화적 표현의 다양성</b> <b>좌장: 한경구 서울대학교 자유전공학부 교수</b> • 김용훈, 신승백 김용훈 예술가 듀오 • 박주용, 카이스트 문화기술대학원 교수 • 박찬욱, 한국문화관광연구원 문화산업연구센터 센터장 • 이두갑 서울대학교 서양사학과 교수 • 전진성 유네스코한국위원회 문화팀장 • 정상조 서울대학교 법학전문대학원 교수  ※ 토론 후 행사개최 전 (2020.6.22.~6.25) 접수된 사전질문에 대한 답변

## 연사 & 토론자 소개

### 좌장



한경구  
서울대학교 자유전공학부 교수

한경구 교수는 서울대 인류학과에서 학사·석사 학위를, 미국 하버드대에서 인류학 박사학위를 취득하였다. 일본 동경대 대학원 문화인류학교실에서도 수학한 바 있으며, 제14차 외무고시 수석 합격 후 외무부에서 근무한 경력도 있다. 강원대 인류학과, 국민대 국제학부를 거쳐 현대 서울대 자유전공학부 교수로 재직중이며, 국민대 국제학부장, 서울대 자유전공학부장을 역임하였다. 한국문화인류학회 부회장 및 편집위원장, 한국사회과학연구협의회 편집위원장, 한국국제이해교육학회 회장, 한국이민학회 회장 등을 역임했고, 유네스코 한국위원회 사회과학분과 및 문화분과 위원, 집행위원을 맡은 바 있다.

### 발표자



발표 1 - 인공지능시대, 창작 개념의 재검토

이상욱  
한양대학교 철학과 교수, 과학기술 윤리·법·정책 센터장

서울대학교 물리학과에서 이학사를 마친 후 동대학원 물리학과에서 응집물질 물리학 연구로 이학석사를 취득했다. 그 후 영국 런던대학(LSE)에서 철학박사 학위를 받았으며, 2002년까지 런던정경대학교(LSE) 철학과에서 객원 조교수로 일하다가 현재 한양대학교 철학과 교수로 재직 중이다. 주로 현대 과학기술이 제기하는 다양한 철학적, 윤리적 쟁점을 과학사 및 기술사, 과학기술학(STS)의 시각과 접목하여 연구한다. 2018년부터 UNESCO 과학기술윤리위원회(COMEST)의 위원으로 활동 중이며, 2018년~ 2019년에 UNESCO가 주도한 인공지능 윤리 예비보고서 작성에 참여했다. 『과학은 이것을 상상력이라고 한다』, 『인공지능의 윤리학』, 『인공지능의 존재론』 등의 다양한 저서와 「인공지능의 도덕적 행위자로서의 가능성: 쉬운 문제와 어려운 문제」, 「감정과 의식: 신경과학과 인공지능 시대의 문학」 등의 논문을 집필했다.

발표 2 - AI와 인간의 협업가능성과 문화적 표현의 다양성

김재인

경희대학교 비교문화연구소 학술연구교수



서울대 미학과를 졸업하고 동 대학원 철학과에서 박사학위를 받았다. 고등과학원 초학제연구단 연구원과 서울대학교 철학사상연구소 연구원을 역임했으며, 서울대, 경희대, 한국외대, 홍익대, 한국예술종합학교, 서울여대 등에서 강의했다. 현재는 경희대 비교문화연구소 학술연구교수로 재직하고 있다. 단독 저서로 『코로나 혁명 : 인류의 다음 단계를 위한 새로운 가치의 모색』(2020), 『생각의 싸움 : 철학의 멋진 장면들』(2019), 『인공지능의 시대, 인간을 다시 묻다』(2017), 『혁명의 거리에서 들뢰즈를 읽자』(2016), 『삼성이 아니라 국가가 뚫렸다』(2015) 등이 있다. 『인공지능의 시대, 인간을 다시 묻다』는 제58회 한국출판문화상 본심, 2017년 경향신문 올해의 책, '경기그레이트북스 100권'에 선정되었다. 발표한 논문으로는 「우리 사회에 '사회'가 있는가? - 서구 '사회' 개념의 국내 수용 과정 분석」, 「매끈한 공간 대 흠 파인 공간: 전쟁기계, 또는 공간을 어떻게 구성할 것인가? - 들뢰즈의 공간의 정치철학」, 「들뢰즈의 '아펙트' 개념의 쟁점들: 스피노자를 넘어」, 「바움 가르텐으로 돌아가자 - 감(感)적 앎의 복권을 위한 한 시도」 등이 있다.

토론자

김용훈

신승백 김용훈 예술가 듀오



시드니대학 시각예술학과에서 학사학위를 취득하고, 카이스트 문화기술대학원에서 석사학위를 받았다. 엔지니어 신승백과 함께 2012년 아티스트 그룹, 신승백 김용훈을 결성하였다. 두 사람은 협업을 통해 기술과 인간에 대해 탐구해 오고 있으며 인공지능을 주제로 다수 작품을 만들었다. 독일, 오스트리아, 프랑스, 영국, 네덜란드, 스위스, 미국, 캐나다, 호주, 일본, 대만, 중국 등 전세계에 작품이 전시되고 있다.

	<p>박주용 KAIST 문화기술대학원 교수</p>
	<p>박찬욱 한국문화관광연구원 문화산업연구센터 센터장</p> <p>박찬욱 박사는 현재 한국문화관광연구원 문화산업연구센터 센터장을 맡고 있다. 송실대학교 융합기술원 전임연구원, 산업통상자원부 통상산업포럼 문화콘텐츠분과위원 등을 역임하였다. 경희대학교에서 국제경영학 박사학위를 취득하였고 「융복합 콘텐츠산업 현황진단 및 대응전략」, 「콘텐츠 R&amp;D 고도화를 위한 기초연구」 「신기술 융복합 콘텐츠 관련 법제도 개선방안 연구」 등 다수의 정책 과제에 연구책임으로 참여하였고, Technology Forecasting &amp; Social Change, Online Information Review 등 우수 학술지에 논문을 발표하였다.</p>
	<p>이두갑 서울대학교 서양사학과 교수</p> <p>이두갑 교수는 서울대학교 인문대학 서양사학과 &amp; 자연대 과학사 및 과학철학협동과정 부교수로 재직하고 있다. 프린스턴 대학교에서 역사학(과학기술사) 박사학위를 받았으며 미국 국립보건원역사연구소에서 박사 후 연구원을, 스탠포드대 및 하버드 대학교 방문연구원을 지냈다. 저서로는 The Recombinant University: Genetic Engineering and the Emergence of Stanford Biotechnology (University of Chicago Press, 2015)가 있다.</p>

	<p>전진성 유네스코한국위원회 문화팀장</p> <p>지난 20년간 유네스코한국위원회에서 문화커뮤니케이션팀장, 국제협력팀장 등 다양한 직을 수행했으며 2011년~2012년에는 프랑스 유네스코 본부 대외협력부문 국가위원회 담당과에서 근무하기도 했다. 현재는 유네스코한국위원회 문화팀장으로 근무하면서 성균관대학교 문화융합대학원 겸임교수로도 활동하고 있다.</p> <p>런던시티대학교에서 문화정책 석사학위를 취득했고, 버벡대학교에서 예술경영 디플로마를 취득했으며 유네스코 거버넌스와 문화정책 분야 등에서 「유네스코 조직 개혁 - 유네스코 전략 연구 변화의 시대, 한국의 유네스코 협력 비전」, 「2016년 글로벌 법제 이슈 - 교육/문화 분야」, 『대유네스코 자발적 기여에 관한 실무 안내서』 등 다양한 글을 집필했다.</p>
	<p>정상조 서울대학교 법학대학원 교수</p> <p>정상조교수는 서울대법학과에서 학사 및 석사학위를 취득한 후 영국 런던정경대(LSE)에서 석사 및 박사학위를 취득하였다. 특허법원 사법행정 자문위원회 위원, 한국과학기술한림원 정회원, 국가지식재산위원회 민간 위원장을 역임하고 있으며, 서울대학교 법학연구소 기술과법센터장으로 재직하기도 했다. 『저성장시대의 도시정책(공저, 2011)』, 『예술경영과 예술행정』(공저, 2017), 『문화기본진흥계획』, 「인터넷과 법률」, 「컴퓨터와 법률」, 「Systematic Government Access to Private-Sector Data in the Republic of Korea, International Data Privacy Law」, 「딥러닝에서의 학습데이터와 공정이용」 등 기술과 법률 분야에서 여러 연구를 진행했다.</p>

## 발표1

인공지능시대, 창작 개념의 재검토

이상욱 한양대학교 철학과 교수

과학기술 윤리·법·정책 센터 센터장

## 1. 머리말

인공지능(Artificial Intelligence)은 1950년대 잘 만들어진 기계가 인간의 지능을 흉내 낼 수 있다는 튜링의 믿음에서 탄생한 말이다.<sup>1)</sup> 즉 초기 인공지능 제작자들의 원래 의도는 SF 영화의 이미지에 등장하는, 감정이나 자유의지처럼 내적 심리상태에서도 인간과 구별되지 않는 지능적 기계가 아니라 근대 기계제작자 보강송의 오리와 마찬가지로 ‘오리처럼 먹고, 오리처럼 움직이고, 오리처럼 껍딱거리고, 오리처럼 배설하지만 진짜 오리와는 분명히 다른 기계 오리’에 해당되는 기계지능을 만드는 것이었다.<sup>2)</sup>

인공지능이 번역도 하고, 신문 기사도 쓰고, 증시 분석도 하고, 소설도 쓰는 시대가 온 현 시점에서 창작 개념을 재검토하려는 것이 이 글의 목적이다. 이 과정에서 가장 먼저 분명하게 해야 할 점은 최근 대중매체가 인공지능의 미래에 대해 쏟아내는 수많은 자극적 전망을 잠시 접어두고 실제 설계되고 만들어지고 사용되고 있는 인공지능이 무엇인지를 분명하게 이해하는 것이다.<sup>3)</sup> 2절에서 이루어질 이 작업을 통해 우리는 현재까지 (그리고 가까운 미래에 등장할) 인공지능은 그 산출물에 있어서는 평균적인 인간을 부끄럽게 할 정도의 놀라운 성취를 보여주고 있지만 존재론적으로는 여전히 보강송의 기계오리의 연장선상에 있다는 점을 알게 될 것이다. 결정적으로 보강송의 오리가 그 오리를 만들고 문제가 생기면 수리하는 기계제작자 보강송을 필요로 했듯이 현대의 인공지능도 그 인공지능을 만들고 문제가 생기면 보완해야 할 엔지니어가 필수적으로 필요하다는 점이 드러날 것이다.

3절에서는 이런 배경 하에 현재까지 이루어진 인공지능의 ‘창작 행위’에 대해 살펴본다. 이 과정에서 대중매체에서 이들 ‘창작’이 다루어질 때 부각되지 않은, 그럼에도 불구하고 인공지능의 ‘창작’에서 결정적으로 중요한 부분이 지적될 것이다. 그리고 인공지능의 ‘창작’에 대한 논의에서 그다지 주목되지 않는 이 부분이 인공지능에게 창작 개념을 적용하는 것이 불가능하다는 결론을 지지하는 것은 아니지만, 적어도 인공지능의 ‘창작’은 인간의 ‘창작’과 분명히 다른 특징이 있다는 점을 주목하게 해 준다.

이 지점에서 우리는 선택의 기로에 도달하게 된다. 인공지능의 ‘창작’을 수용하는 방식으로 현재의 창작 개념을 재규정할 것인지 아니면 인간의 창작만이 진정한 창작이라는 입장을 고수할 것인지가 그것이다. 필자는 어느 한 입장을 옹호하기 보다는 예술 창작의 역사에서 이런 논쟁은 드물지 않게 발생해 왔으며 그 논쟁을 진행하고 결론을 내리는 과정 자체가 예술 활동의 역사적 전개 과정이었음을 4절에서 지적하겠다.

1) 현재 인공지능 기술에 대한 간략한 소개는 Shanahan 2015, Kaplan 2016a, Mitchell 2019 참조.

2) 보강송의 오리에 대한 자세한 설명과 그것이 인공지능의 역사에서 갖는 의미에 대한 간략한 분석은 Gerrish 2018 참조.

3) 이런 필자의 문제의식과 일치하는 방식으로 인공지능의 여러 쟁점을 21세기 현대 사회의 맥락에서 조명한 책으로는 Ford 2016, Fry 2019 참조.

## 2. 인공지능의 특징 세 가지

이번 절에서는 인공지능 시대의 창작 개념을 재검토하는 과정에서 주목해야 할 인공지능의 특징을 세 가지로 제시한다. 그것은 ‘자각 없는 수행(performance without awareness)’, ‘이해하기 어려운 실패(unintelligible failure), ‘계산과 실재의 간극(gap between computation and reality)’이다. 각각의 특징이 무엇이고 그 특징이 인공지능 시대의 창작 개념을 재검토하는 과정에서 왜 중요한지를 살펴보자.<sup>4)</sup>

첫째 현재까지 개발된, 그리고 당분간 가까운 미래에 등장할 인공지능은 인간이 판단하기에 ‘탁월한’ 결과물을 인간에게는 매우 낮은 방식, 즉 그 결과물에 대한 의식적 경험이 없는 방식으로 생산한다. 이는 이세돌 9단을 가볍게 이긴 알파고가 자신이 바둑을 두었다는 사실 자체를 의식하지 못하는 것에서 잘 드러난다. 이를 ‘자각 없는 수행’이라고 하자.

중요한 점은 이 ‘자각 없는 수행’이 인공지능 기술의 한계가 아니라는 사실이다. 즉, 인공지능의 기술적 목표가 인간처럼 ‘의식적 경험’도 함께 하면서 인간이 판단하기에 ‘탁월한’ 결과물을 산출하는 것이 아니라, 단지 인간이 판단하기에 ‘탁월한’ 지적 결과에 해당하는 것을 가장 효율적으로 산출하는 것이라는 점을 기억할 필요가 있다. 우리에게 익숙한 할리우드 SF 영화의 설정과 달리, 절대 다수의 인공지능 연구자들은 인간 지능의 ‘모든’ 특징을 인간과 정확히 동일한 방식으로 구현하려는 데 관심이 없다. 인간 두뇌 전체를 모사하거나 흉내(simulation) 내려는 시도가 있기는 하지만 이 경우도 정말 인간 의식 경험 자체에 관심이 있어서라기보다는 이런 전체 두뇌 구현(Whole Brain Emulation)을 통해서 현재 인공지능이 ‘자각 없는 수행’에서 겪는 문제점이나 비효율성을 극복하기 위함이다.

기계지능의 가능성에 낙관적이었던 튜링조차도 인간지능과 기계지능은 그 결과물, 즉 외부적으로 확인되는 수행에 있어 동등할 수 있음을 주장했을 뿐 인간지능과 기계지능이 그 작동 방식이나 의식이나 자각처럼 세부적 사항에 있어서 모두 동일할 것이라고 기대하지는 않았다. 오히려 기계지능이 구현되는 물질적 기반이 인간지능이 구현되는 물질적 기반과 다르기에 그로부터 나타나는 지능은 당연히 인간에게는 낮은, 다른 것이 될 것이라고 예상했다.<sup>5)</sup>

미래에 등장할, 기술적으로 훨씬 진보한 인공지능 역시 이 ‘자각 없는 수행’의 특징을 유지할 가능성이 높다. 일단 현재 이루어지고 있는 대부분의 인공지능 연구는 인간 지능이나 수행의 특정 영역, 예를 들면 의학 영상이미지 판독이나 법률문서 요약처럼 맥락과 전문성이 제한된 영역을 대상으로 하는 특수 인공지능 연구이다. 특수 인공지능은 그 기술적 목적이 인간 지능의 특정 기능을 구현하는 것이기에 의식적 경험과 같은 인간 지능의 고유한 특징에 해당되는 회로 자체를 설계 단계에서부터 만들지 않는다.

4) 이 내용에 대한 보다 자세한 논의는 이상욱 2009, 2018a, 2018b, 2019 참조.

5) 이는 인간 진화의 역사에서 의식이 등장하고 변화해 온 과정이나 물질적 기반에서 창발적 속성으로 마음이 등장하는 방식에 대한 표준적인 철학적 설명에서도 수용되는 관점이다. 관련 논의는 Dennett 1997, Floridi 2016 참조.

물론 현재 소수의 인공지능 연구자가 인간지능처럼 다양한 지적 영역에서 결과물을 산출할 수 있는 일반 인공지능을 연구하고 있기는 하다. 하지만 이들 연구자의 연구 내용에서도 ‘의식적 경험’이 이론적 분석이나 설계 내용에 등장하는 경우는 없다. 그 이유는 단순하다. 우리는 현재 우리 두뇌에서조차 ‘의식적 경험’이 정확히 어떤 물질적 인과 작용을 통해 등장하는 지에 대해 아주 제한적인 이해만을 갖고 있다. 그러므로 인공지능에게 ‘의식적 경험’을 만들어 장착하고 싶어도 어떻게 해야 할 지에 대해 출발점부터 전혀 모르는 상태이다.<sup>6)</sup>

결국 ‘의식적 경험’이 미래 인공지능에 등장하리라 기대하는 사람들은 인공지능이 ‘충분히 발전하면’ 인간지능이 진화의 역사에서 그러했듯 ‘의식적 경험’이 자연스럽게 등장할 것이라고 ‘기대’하는 것뿐이다. 하지만 이는 적어도 철학적 관점에서 볼 때는 그저 막연한 기대 이상의 근거를 갖기 어렵다. 인류 진화 과정에서 ‘의식’의 등장이 필연적이었다는 생물학적, 철학적 근거를 찾기 어렵다.

어쩌면 우리에게 매우 소중한, 아마도 우주적 관점에서는 매우 가치 있는 ‘의식’의 등장은 그저 여러 인과 관계가 겹쳐져서 일어난 우주적 ‘우연’일 수도 있다. 설사 지구상의 생명체의 진화 역사에서는 이런 인과 관계의 중첩이 ‘의식’의 등장을 어느 정도 자연스러운 것으로 만들었을 가능성을 배제할 수는 없지만, 우주적 관점에서 볼 때 전혀 다른 물질적 기반을 갖는 기계지능이 ‘자연스럽게’ 이런 의식적 경험을 하리라고 기대하는 것은 다소 유치한 인간중심적 사고라고 해야 할 것 같다.

다음으로 소개할 ‘이해하기 어려운 실패’는 그 자체가 정말 이해하기 쉽지 않다. 인공지능, 특히 인간이 보기에 깜짝 놀랄만한 ‘수행’을 보여주는 인공지능도 완벽하지는 않기에 가끔씩은 주어진 목표 달성에 실패하곤 한다. 단지 인간보다 실패 확률이 낮을 뿐이다. 그런데 문제는 인공지능의 실패는 인간의 실패와 질적으로 다른 것처럼 보인다는 점이다. 이 문제는 최근 유행하고 있는 기계학습의 여러 형태, 예를 들면 딥러닝이나 강화학습 알고리즘을 활용한 인공 신경망 인공지능에서 두드러지게 나타난다.

예를 들어 설명해 보자. 인공신경망이 가장 인상적인 ‘수행’을 보이는 영역이 이미지 판독이다. 가장 최신의 인공 신경망은 (범주 설정의 문제 때문에 다소 논쟁적이기는 하지만) 일반인보다 더 높은 정확도로 사물을 구별해 낸다. 잘 훈련된 의료 인공지능은 이제 영상의학 전문가 수준에 도달한 경우도 있다.

문제는 인공지능이 이미지 판독에서 실패할 때 인간지능에 익숙한 사람이 납득하기 어려운 방식으로 실패한다는 사실이다. 사람이라면 고양이의 여러 종류를 혼동해서 삼 고양이를 페르시아 고양이와 헛갈릴 수 있다. 하지만 일반적인 상식이 있는 사람이라면 말과 고양이를 헛갈리지는 않을 것이다. 그런데 전체적인 이미지 판독 정확도에서는 평균적 사람을 앞서는 인공지능은 고양이 이미지를

---

6) 이 지적은 인간 지능을 수행적 수준에서 넘어서는 초지능(superintelligence)이 등장하여 인류의 생존을 위협할 가능성에 대한 소위 ‘실존적 위험(existential risk)’에 대해 어떤 입장을 취하는 지와 무관하게 성립한다. 초지능이 인간과 유사한 의식적 경험을 할 것이라는 막연한 직관에 대해서 초지능의 가능성을 주장하는 논자들도 분명한 근거를 제시하지 못하고 있다. 관련 논의는 Bostrom 2016, Brockman 2019, Kurzweil 2006 참조. ‘실존적 위험’에 대한 비판적 검토는 이상욱 2020 참조.

사람이 알아차리기 어려울 정도로 조금만 바뀌어도 말로 인식하거나, 거기에서 더 나아가 옛날 송출이 끝난 방송 화면처럼 잡음 이미지를 고양이로 인식하기도 한다. 이는 딥러닝 등 인공지능망이 활용하는 기계학습 알고리즘이 인간이 학습하는 방식과 다르기 때문에 발생하는 근본적인 문제이기에 현재 인공지능 기술이 근본적인 혁신을 이루지 않는 한 쉽게 고쳐질 수 있는 ‘실수’가 아니다. 이런 이유 때문에 최근에 사람은 알아차리기 어려운 변화를 중요한 이미지에 가해서 인공지능 활용 상황에서 혼란을 야기하는 악의적 공격(adversarial attack)이 인공지능 보안 분야의 주요한 쟁점으로 부상하고 있는 것이다.

고양이를 말로 오인하는 것 정도야 가벼운 농담거리일 수 있지만 의학 진단에서 이런 ‘실수’를 한다면 정말 큰 문제일 것이다. 전문의 역시 실수는 하지만 적어도 사람이 ‘납득할 수 있는’ 실수를 하는 경우가 대부분이고 그렇지 않은 경우에는 주의 부족으로 처벌받을 수도 있다. 하지만 전문의라면 현재 의료 인공지능이 보여주는 ‘이해하기 어려운 실패’는 하지 않을 것이다. 핵심은 인공지능이 이런 ‘이해하기 어려운 실패’를 언제 할지에 대해 미리 알기가 거의 불가능하다는 사실이다. 이 사실은 인공지능의 인상적인 수행에 대해 그저 낙관적으로만 평가할 수 없는 불안감의 근거가 될 수 있다.

마지막으로 살펴볼 ‘계산과 실재의 간극’을 이해하기 위해서는 대중들의 인공지능에 대한 이해에서 인공지능과 로봇이 자주 혼용되거나 심한 경우에는 동일시된다는 점에서 시작할 필요가 있다. 이른 컴퓨터공학에서 로봇은 온라인상을 돌아다니며 데이터를 수집하거나 프로그램의 오류를 잡아내는 등의 기능을 수행하는 프로그램을 지칭하기도 하기에 더욱 강화되는 경향이 있다. 하지만 통상적인 의미에서 로봇은 물리적 실체를 갖고 공간을 이동할 수 있는 수단을 갖춘 기계를 의미한다. 당연히 지능적 행동을 할 수 있는 로봇은 인공지능을 활용해야 하기에 우리가 이 글에서 관심을 갖는 로봇은 인공지능을 탑재한 인공지능 로봇이다. 하지만 이 인공지능 로봇과 지적인 결과물을 산출하는 인공지능 자체를 구별하는 일은 이후의 창작에 대한 논의에서 결정적으로 중요하다.

사람들이 잘 인식하지 못하지만 우리가 일상적으로 사용하는 휴대전화에도 인공지능이 여럿 작동하고 있다. 사진을 찍을 때 빛의 양을 조절하고 초점을 맞춰주는 기능이나 검색을 할 때 자동으로 검색어를 완성해 주는 기능 모두 인공지능이 담당한다. 이것이 인공지능이라는 점을 사람들이 잘 느끼지 못하는 이유는 사람들은 인공지능의 ‘지적인 행동’을 그에 대응되는 인간의 지적인 행동을 떠올리며 이해하는 경향이 있고 그러다보니 자연스럽게 SF 영화에 등장하는 안드로이드형 로봇이 ‘전형적인’ 인공지능이라고 생각하는 경향이 있는 것이다.

하지만 구태여 유비할 대상을 찾자면 인공지능의 가장 가까운 유비는 안드로이드형 로봇이 아니라 휴대폰에 내장된 계산기이다. 물론 사칙연산보다는 훨씬 복잡한 일을 순식간에 해치우는 인공지능이 수두룩하지만 기본적으로 인공지능이 하는 일은 매우 복잡한 계산을 하는 일이다. 좀 더 정확하게 이야기하자면 인공지능은 입력 값을 받아 다양한 방식의 계산을 통해 출력 값을 산출함으로써 인공지능 제작자가 의도했던 목적을 효율적으로 달성하는 시스템이다. 특히 최근 인공지능은 인공지능 설계자의 목적을 달성하는 ‘과정’에서 일일이 사람의 결정을 물어보지 않고 ‘자율적으로’

결정하는 모습을 보여주기도 한다. 이런 이유 때문에 세계에서 가장 영향력 있는 전기전자공학자 단체인 IEEE에서는 인공지능이라는 용어 대신에 자율지능시스템(Autonomous Intelligent System)이라는 용어를 사용한다.

이런 논의가 왜 중요하냐면, 인공지능 자체는 (인공지능을 탑재한 로봇이 아니라) ‘몸’이 없다는 점을 늘 기억해야 하기 때문이다. 인공지능은 기본적으로 계산을 하는 장치이기에 그 계산 결과가 물리적 세계에 인과적으로 영향을 주기 위해서는 추가적인 장치가 덧붙여져야 한다. 예를 들어 알파고는 자신이 계산한 수를 대신 바둑판 위에 놓아주는 사람을 필요로 한다. 자율주행차에 탑재된 인공지능은 자신이 자율적으로 결정한 운행 방식을 실현해 줄 물질적 대상, 즉 자동차가 필요하다.

문제는 인공지능의 결정을 물리적으로 구현하는 과정이 탁월한 수행을 보이는 인공지능을 만드는 과정보다 훨씬 더 힘들다는 사실이다. 물리적 세상에는 중력이나 마찰력, 그리고 해석되어야 할 공적 규범처럼 다양한 제약조건이 존재하는 데 이를 모두 ‘적절하게’ 준수하면서 인공지능의 계산 결과를 행동으로 구현하는 일은 대단히 어렵기 때문이다. 이는 인공지능의 탁월한 수행이 온라인 게임처럼 물리적 제약조건이 없는 사이버 세상에서 더욱 두드러진 근본적인 이유이다.

이 문제와 관련된 것이 바로 유명한 모라벡의 역설이다. 고도의 수학적 추론처럼 인간에게 어려워 보이는 지극히 지적인 문제를 인공지능이 해결하는 것은 상대적으로 쉬운 반면, 장애물이 여기저기 놓여있는 방을 통과해서 문 밖으로 나가는 것처럼 어린아이도 별 생각 없이 할 수 있는 쉬운 일을 인공지능을 탑재한 로봇으로 구현하는 일은 매우 어렵다는 것이 그 내용이다. 수학적 추론은 물질적 세계와 무관하게 처리될 수 있지만 장애물에 걸려 넘어지지 않고 공간을 통과하는 일은 물질적 제한조건을 실시간으로 끊임없이 고려해 가면서 인과 작용을 조정해야 하기 때문이다.

그럼 이제 다음 절에서 이런 인공지능의 세 특징이 인공지능의 ‘창작’에 어떻게 관련되는지를 소설 쓰기, 그림 그리기, 작곡하기의 구체적 사례를 들어 살펴보자.

### 3. 인공지능 ‘창작’ 현실에 대한 비판적 검토

인공지능 ‘창작’에 대해 가장 먼저 시도가 된 분야는 원리적으로는 가장 어려운 글쓰기였다. 글쓰기가 그림그리거나 작곡하기보다 원리적으로 어려운 이유는 글의 속성상 의미론적 특징이 창작자에게서 두드러진다는 데 있다. 화가나 작곡가도 물론 자신의 그림이나 음악을 통해 의미를 전달하려고 노력할 수도 있고 어떤 경우에는 글만큼이나 상당히 고정된 의미가 형성되기도 하지만 많은 경우에는 감상자에 의해 의미가 ‘해석’되거나 재구성되는 경우가 많다. 그에 비해 글쓰기는 시인이든 소설가든 의미론적 대상인 개념이나 표현을 일차 재료로 사용하기에, 현 단계에서 의미론적 속성을 ‘이해하고’ 있다고 말하기 어려운 (물론 이 경우도 ‘이해하다’를 철저히 행태론적으로 정의하면 현재 인공지능도 의미를 이해한다고 볼 수 있지만) 현 단계 인공지능이 다루기 까다로운 작업임은 분명하다.

그럼에도 불구하고 인공지능이 ‘저술’해서 출판했다고 알려진 책이 여러 권 존재하고 그 중에서는 상업적으로 판매된 작품도 꽤 있다. 이런 일이 어떻게 가능했을까? 인공지능은 기본적으로 학습 데이터에 존재하는 패턴을 찾아내어 그것을 다시 새로운 방식으로 구현해 낼 수 있는 시스템이다. 그러므로 특정 장르의 글을 수없이 많이 기계학습 시켜서 그 분야 글의 ‘스타일’을 학습한 인공지능이 인간이 판단하기에 상당히 그럴듯한 글을 결과물로 내놓을 수 있다는 사실 자체는 분명 신기할 수는 있어도 원리적으로 놀라운 일은 아니다.

하지만 앞서 설명한 인공지능이 세 특징이 인공지능의 글쓰기에도 고스란히 나타나서 인공지능의 ‘창작’ 능력에 제한을 가한다. 우선 인공지능은 당연히 인간처럼 글을 쓰거나 키보드로 글을 쓸 수는 없다. 실은 글자를 한자씩 시간적으로 써 넣지도 않는다. 그보다는 기존에 학습한 문장이나 표현을 적당히 변형해서 설계자가 미리 설정한 제한 조건을 만족하는 방식으로 ‘생산’해 낸다. 그리고 이 과정에 대한 의미론적 이해가 없는 것은 당연하고, 거기에 더해 우리로서는 요령부득의 이상한 문장을 산출해 내는 경우도 꽤 있다. 시 창작의 경우에는 언어적 파격이 어느 정도 허용 되기에 이런 문장들이 ‘기막힌’ 언어적 유희로 해석될 수도 있겠지만 학술서의 경우에는 용납될 수 없다.

이런 이유로 인공지능이 저술했다는 책들은 모두 그 산출 과정에서 수많은 인간과의 협력이 필수적이었다. 일단 인공지능이 생산해 낸 문장들 중에서 어색하거나 아예 뜻이 통하지 않은 문장은 인간이 일일이 제거하거나 적절한 문장으로 고쳐 써야 한다. 이에 더해서 인공지능의 기계학습 과정에서는 포함되지 않은 여러 글쓰기의 측면, 예를 들어 한 장의 내용을 잘 요약하는 제목 정하기 등은 아예 인간이 도맡아야 하는 경우가 대부분이었다. 요약하자면 인공지능의 글쓰기는 실은 인공지능+인간 연합팀의 글쓰기였던 것이다. 이 특징이 가까운 미래에 극복될 가능성은 여러 기술적 이유로 높지 않아 보인다. 게다가 당연하게도 인공지능은 적당한 출판사를 고르거나 편집자와 책의 구성과 편집에 대해 교섭할 수 없다. 책의 인세도 인공지능이 갖지 못하는데 이는 이 과정을 생각할 때 어찌면 당연하다고 생각할 수 있다.

그림그리기는 어떨까? 최근 네덜란드에서는 국가적 보물로 여겨지는 렘브란트의 작품을 기계학습해서 렘브란트가 그렸을 법한 그림(전문가들이 판단하기에도 그럴 정도의 수준)을 그려내는 인공지능이 등장해서 화제이다. ‘넥스트 렘브란트’로 불리는 이 인공지능이 ‘창작한’ 유사-렘브란트 그림은 유튜브를 비롯한 공개 매체에서 쉽게 접할 수 있는데 그림의 완성도나 물감의 두께 조절까지 깜짝 놀랄 수준을 보여준다.

그런데 여기서 잠시 상황을 꼼꼼하게 따져보자. 앞서 지적했듯이 인공지능은 몸이 없다. 당연히 팔도 없다. 그러므로 ‘넥스트 렘브란트’가 붓을 들고 이 멋진 유사-렘브란트 그림을 그렸을 리는 없다. 관련 내용을 살펴보면 ‘넥스트 렘브란트’가 그림의 3차원적 구성을 기계학습 등의 알고리즘으로 만들어내면 이 ‘정보’를 3차원 프린터로 찍어낸 것이라는 점을 알 수 있다.

‘넥스트 렘브란트’가 그림그리기 하는 과정에서 인간과의 협력이 필수적인 이유는 이것만이 아니다.

‘넥스트 렘브란트’가 렘브란트 그림을 기계학습 한다고 할 때 그 일을 정확히 어떻게 수행할까? 핵심은 이 질문에 ‘넥스트 렘브란트’는 대답할 수 없다는 사실이다. 대답은커녕 그 질문을 (통상적 의미에서) 이해하지도 못한다. 그러므로 ‘넥스트 렘브란트’가 렘브란트 화풍에 내재된 패턴을 성공적으로 ‘학습’하기 위해서는 이 과정 전체를 기획하고 매 단계마다 실행하는 인간 엔지니어의 존재가 필수적이다. 렘브란트 그림을 방 안 가득 늘어놓고 휴머노이드 로봇에게 ‘자 이제 렘브란트 식으로 그려봐!’라는 식의 설정은 오직 SF 영화에서만 가능할 뿐이고 실제로는 렘브란트 식으로 그린다는 것이 무엇인지를 인간 엔지니어가 고민하고 그 화풍을 어떻게 프로그램적으로 구현할지를 인간 엔지니어가 고민하고 그 고민의 내용을 알고리즘으로 ‘넥스트 렘브란트’에게 집어넣고, 학습 과정에서 등장하는 수많은 문제를 해결하는 것은 모두 인간의 몫이다. ‘넥스트 렘브란트’는 이런 인간의 협력을 기반으로 다양하고 복잡한 계산을 하여 결과 값을 산출할 뿐이다. 그러므로 그림 그리는 인공지능의 경우에도 인공지능이 그림을 그리는 것이 아니라 인공지능-인간 연합팀이 그림을 그리는 것이다.

마지막으로 작곡을 살펴보자. 작곡하는 인공지능은 여럿 있지만 그 중에서도 가장 유명한 것 중 하나가 EMI(Experiments in Musical Intelligence)이다. 이 인공지능은 음악학자 데이비드 코프가 오랜 연구를 통해 완성했는데 그 과정에서 바흐, 비발디, 모차르트 등 위대한 작곡가들의 작품을 어떻게 코딩하여 인공지능에게 학습 데이터로 제공할 것인지, 기존 작품과 충분히 달라서 새로운 작품이라는 느낌을 주되 여전히 각 작곡가가 작곡했을 법한 작품의 테두리 안에 남아 있다는 평가를 받기 위해서는 무작위 변이를 어떤 방식으로 얼마만큼 주어야 하는지 등에 대해 수많은 시행착오를 통해 연구했다고 한다.

EMI의 창시자 코프조차 EMI가 자신이 작곡을 하고 있다는 점을 의식하지 못하는 고도의 복잡한 컴퓨터 프로그램이라는 점을 거리낌 없이 인정한다. 그런데도 놀라운 점은 이 컴퓨터 프로그램이 작곡한 곡을 감상한 수많은 사람들이 어떤 경우에는 ‘영혼의 울림’을 느낄 정도로 감탄했다는 사실이다. 그러므로 EMI는 앞서 지적했던 ‘자각 없는 수행’을 가장 인상적으로 보여주는 예에 해당된다. 실제로 EMI의 능력을 의심한 현대 작곡자와의 공개 경연에서 청중들은 압도적으로 EMI의 손을 들어 주었다. 놀라운 점은 청중들이 EMI가 바흐를 흉내 내 작곡한 바흐 곡이 바흐가 정말로 작곡한 바흐 곡보다 더 바흐스럽다고 판단했다는 사실이다. 적어도 결과물에 있어서 인공지능의 수행능력은 인간의 ‘고유한 영역’이 어딘가에는 분명히 남아 있을 것이라는 낭만적 생각을 비웃는 것처럼 보인다.

하지만 앞서 글쓰기와 그림그리기 사례와 마찬가지로 EMI는 몸이 없다. 그러므로 자신이 ‘작곡한’ 곡을 피아노로 연주하거나 성악으로 공연할 수 없다. 이는 모두 다른 사람의 도움을 받아야 한다. 하지만 이 부분은 인간 작곡가도 부분적으로 같은 처지에 있다고 볼 수 있으므로 결정적인 문제가 되지는 않는 것 같다. 게다가 EMI가 작곡한 곡들에는 약간의 ‘진부함’은 느껴져도 ‘이해할 수 없는 실패’에 해당되는 이상한 음은 느껴지지 않는다.

하지만 이런 생각은 EMI의 작동과 결과물 산출이 코프라는 위대한 음악학자-컴퓨터 프로그래머의 세밀한 조정과 감시 하에 이루어진다는 점을 간과한 것이다. 예를 들어 EMI에게 바흐의 음악 스타일을 학습시키는 데는 부지런한 코프의 노력을 통해서도 7년이 걸렸지만 일단 기계학습이 완료된 EMI는 하루 사이에 수천 곡의 ‘바흐 스타일’ 곡을 작곡해 낸다. 이 중에는 코프가 대중에게 내놓기에는 부족한 곡들이 다수 포함되어 있었을 것이다. 결국 EMI의 작곡 능력이 인상적이 이유는 코프라는 철저하게 헌신적인 인간 협력자가 인공지능의 작곡 능력의 한계를 탐색하기 위한 목적으로 끊임없는 실험적 시도를 통해 가장 인상적인 작품을 산출하려고 노력했고 그 결과물을 인간 연주자들이 탁월한 기량으로 공연했기 때문이다. 정도의 차이는 있지만 이 지점에도 여전히 인공지능 혼자 작곡을 하는 것이 아니라 인공지능-인간 연합팀이 작곡을 하는 것이다.

#### 4. 맺음말: 창작 개념을 재규정해야 할까?

인공지능 시대의 예술에 대해 자주 등장하는 질문이 있다. “인공지능이 ‘진정으로’ 창작한다고 할 수 있을까?”가 그것이다. 필자의 견해로는 이 질문 자체가 잘못 제기된 질문이다. 인공지능은 앞서 설명한 세 특징 때문에 인간이 지적인 수행과는 다른 방식으로 작동한다. 그 결과물은 분명 인간이 판단하기에 지적이지만 인간에게 ‘익숙한’ 방식으로 그 결과물을 만들어 내지도 않고, 실수하는 방식도 낯설고, 결정적으로 몸이 없기에 물리적 세계에 개입하기 위해서는 항상 인간이나 다른 인과적 대상과 결합해야 한다.

인공지능이 진정으로 창작할 수 있는지 여부는 우리가 창작으로 무엇을 의미하는지에 따라 달라진다. 우리가 앞서 지적한 인공지능이 인간지능과 다른 특징을 강조하고 인간이 창작하는 방식으로 인공지능이 창작할 수 있기 전까지는 인공지능은 진정한 의미에서 창작을 한다고 볼 수 없다고 판단한다면, 그 결정은 창작을 그렇게 인간중심적으로 정의한다는 전제 하에서 타당한 결정이다. 좀 더 구체적으로 말해서 창작 행위는 반드시 그 창작 행위를 기획하고 그 의미를 이해하면 그 과정을 몸을 통해 직접 참여하는 과정을 거쳐야 한다고 생각이 설득력이 있다면 가까운 미래에 등장할 인공지능을 포함해서 인공지능은 결코 창작할 수 없다. 오직 인간이 워드프로세서로 글을 쓰거나, 태블릿으로 그림을 그리거나, 신시사이저를 사용해 작곡을 하는 것처럼 인간의 창작 활동에 유용한 도구로 활용될 뿐이다.

하지만 현재 수준의 인공지능에 한정하더라도 인공지능의 최종 산출물은 그것에 창의성이나 창작 개념을 부여하지 않기 어려울 정도로 감탄스러운 것이 사실이다. 물론 현재까지 인공지능이 ‘창작’한 결과물은 인간과의 협력이 절대적으로 중요했고 이 점은 당분간 바뀌지 않겠지만, EMI가 작곡한 음악을 듣고 감동하는 상황과 인간 작곡가의 음악을 듣고 감동하는 상황에 질적인 차이가 있다고 주장하기는, 특히 설사 질적인 차이가 있다고 해도 그것이 ‘음악을 이해 못하는’ 인공지능이 작곡했기 때문이라고 주장하는 것은 약간 억지스러워 보인다. 이미 우리는 보수적인 고전음악 작곡가가 보기에는 제대로 음악을 이해 못하는 수많은 대중음악가들이 작곡한 음악에 영혼의 울림을 경험하고 있기 때문이다.

결국 보다 거시적인 수준에서 말하자면 인공지능이 진정한 의미에서 창작을 할 수 있는지 여부는 예술의 역사가 그러했듯이 새로운 기술적 장치들이 창작 행위에 도입되면서 각 시대마다 그 기술적 장치를 어떤 사회적 논의를 통해 규정해 왔는지와 마찬가지로 시작에서 논의될 필요가 있다. 이런 구성적 논의의 결과 ‘창작’ 개념은 온전히 인간에게만 귀속시킬 수 있는 (의식적 경험을 하고 인간과 유사한 방식으로 욕구와 가치론적 지향점을 추구하는 인공지능의 등장 전까지는) 개념이라는 결론에 도달할 수 있다. 혹은 그런 ‘고전적’ 창작 개념을 그대로 둔 채, 수용자의 측면에서 예술을 감사하는 사람들에게 끼치는 영향에만 집중하여 창작물을 규정하는 새로운 개념적 재정의가 시도될 수도 있다. 우리에게 필요한 것은 이 두 가능성을 포함하여 ‘창작’ 개념을 포함한 수많은 인간적 개념에 대한 재검토를 진지하게 시도하는 것이다. 이 재검토가 충분히 이루어지기 전에 “인공지능은 진정으로 창작할 수 있는가?”라는 물음에 답하려는 시도는 그저 자신의 직관적 선호를 내세우는 것 이상이 되기 어려울 것이다.

## 5. 참고문헌

신상규 외, 『포스트휴먼이 몰려온다』, 아카넷, 2020.

이상욱, 「인공지능의 한계와 일반화된 지능의 가능성: 포스트휴머니즘적 맥락」, 『과학철학』 11(1)(2009), 49-70쪽

이상욱, 「감정과 의식: 신경과학과 인공지능 시대의 문학」, 『한국언어문화』 65(2018), 5-28쪽

이상욱, 「인간, 낮은 인공지능과 마주하다」, 이종원 엮음, 『인공지능의 존재론』, 한울, 2018, 285-316쪽

이상욱, 「인공지능의 도덕적 행위자로서의 가능성: 쉬운 문제와 어려운 문제」, 『철학연구』 125(2019), 125-159쪽

이상욱, 「인공지능과 실존적 위험 - 비판적 검토」, 『인간연구』 40(2020), 107-136쪽

이종원 엮음, 『인공지능의 존재론』, 한울, 2018.

이종원 엮음, 『인공지능의 윤리학』, 한울, 2019.

Agrawal, A. et al. *Prediction Machines: The Simple Economics of Artificial Intelligence*, Cambridge, MA: Harvard Business School Press, 2018

Bostrom, N. *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press, 2016

Brockman, J. (ed.) *Possible Minds: 25 Ways of Looking at AI*, New York: Penguin Press, 2019.

Brynjolfsson, E. and McAfee, A. *Machine, Platform, Crowd: Harnessing Our Digital Future*, New York: W.W. Norton, 2018.

Damasio, A. *The Strange Order of Things: Life, Feeling, and the Making of Cultures*, New York: Vintage Books, 2019

Credo P. C. *Invisible Women: Exposing Data Bias in a World Designed For Men*, London: Chatto & Windus, 2019

Dehaene, S. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*, New York: Penguin Books, 2014

Dennett, D. *Kinds of Minds: Toward an Understanding of Consciousness*, New York: Basic Books, 1997

Floridi, L. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford: Oxford University Press, 2016.

Ford, M. *Rise of the Robots: Technology and the Threat of a Jobless Future*, New York: Basic Books, 2016

Fry, H. *Hello World: How to Be Human in the Age of Machine*, London: Transworld Publisher Co., 2019

Gerrish S. *How Smart Machines Think*, Cambridge, MA: The MIT Press, 2018.

Hume, D. *A Treatise on Human Nature*, second edition with text revised and notes by P. H. Nidditch, Oxford: Oxford University Press, 1978[1739-1740].

Kaplan, J. *Artificial Intelligence: What Everyone Needs to Know*, Oxford: Oxford University Press, 2016a

Kaplan, J. *Humans Need Not Apply*, Ithaca, NJ: Yale University Press, 2016b

Kurzweil, R. *The Singularity is Near: When Humans Transcend Biology*, New York: Penguin books, 2006

Mitchell, M. *Artificial Intelligence: A Guide for Thinking Humans*, New York: Farrar, Straus and Grioux, 2019.

Norvig, P. and Russell, S. *Artificial Intelligence: A Modern Approach*, 3<sup>rd</sup> Edition, New York: Pearson Education, 2016.

Pinker, S. *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*, New York: Penguin Books, 2019.

Pinker, S. 'Tech Prophecy and the Underappreciated Causal Power of Ideas', in Brockman 2019.

Shanahan, M. *The Technological Singularity*, Cambridge, MA: The MIT Press, 2015.

Tallinn, J. 'Dissident Messages', in Brockman 2019

Tegmark, M. *Life 3.0: Being Human in the Age of Artificial Intelligence*, New York: Penguin Books, 2016

Tegmark, M. 'Let's Aspire to More Than Making Ourselves Obsolete', in Brockman 2019

## 발표2

AI와 인간의 협업가능성과 문화적 표현의 다양성  
김재인 경희대학교 비교문화연구소 학술연구교수

## AI와 인간의 협업가능성과 문화적 표현의 다양성

김재인 (철학, 경희대학교 비교문화연구소 학술연구교수)

오늘의 주제를 논의하기 위해서는, 몇 가지 기초 사실을 점검해야 한다고 생각한다. 그 전에 아젠다를 발제 내용에 맞게 재구성해 보겠다. 1) 인공지능과 작품 창작의 관계, 2) 문화적 표현의 다양성을 확보하기 위한 방안, 3) 지속가능하고 평등한 문화 생태계 구축 방안.

이 주제를 다루려면 몇 가지 답변이 선행되어야 한다. 이를 물음의 형태로 구성하면 다음과 같다. 이들 물음에 대한 답변이 정확하지 않거나, 답을 얼버무린다면, 후행 논의는 근거가 없거나 무의미한 것이 될 것이다.

<먼저 답해야 할 기본적인 물음>

1. 창조성/창의성(creativity)이란 무엇인가?
  - 1.1 예술과 창조성의 관계는 무엇인가?
  - 1.2 현대 예술과 낭만주의적 예술의 차이는 무엇인가?
  - 1.3 상식적이며 대중적인 수준에서, 예술 창작이란 무엇인가?
2. 인공지능의 작동 원리는 무엇이며, 무엇을 할 수 있고, 무엇을 할 수 없는가?
  - 2.1 인공지능의 바탕인 소프트웨어란 무엇인가?
  - 2.2 소프트웨어의 바탕인 디지털 컴퓨터란 무엇이며, 무엇을 할 수 있고, 무엇을 할 수 없는가?
  - 2.3 디지털 컴퓨터의 바탕인 컴퓨터 하드웨어란 무엇이며, 어떻게 만들어진다?
  - 2.4 컴퓨터 주변장치의 의미는 무엇인가?
3. 인공지능의 예술 창작을 둘러싼 쟁점은 새로운 것인가?
  - 3.1 디지털 세계와 피지컬 세계는 어떤 차이가 있는가?
  - 3.2 디지털, 인터넷, 플랫폼은 쟁점과 어떤 관계가 있는가?
  - 3.3 '창작 - 생산 - 유통 - 접근'이라는 항목 구별은 오늘날 적절한가?
  - 3.4 '~의 흥수' 속에서 홍보라는 행위의 본질은 무엇인가?

### 1. 창조성/창의성(creativity)의 규정

(물음) 1. 창조성/창의성(creativity)이란 무엇인가?

나는 인공지능에 대한 책을 쓰면서 '창조성'의 본질과 '창의성을 기르는 교육'에 대한 얘기로 마무리했다.<sup>7)</sup> 창조성이란 생각이건 물건이건, 행위건 제도건 간에 새로운 무언가를 최초로 만들어내는 데서 성립한다. 물론 새롭다고 다 창조적이라는 평가를 받는 것은 아니다. 사회가 그것을 가치 있는 것으로 받아들여야만 '그것이 창조적이었구나' 하고 회고적으로 확인되기 때문이다. 창조적인 것은

7) 김재인, 『인공지능의 시대, 인간을 다시 묻다』, 동아아시아, 2017, 360~366쪽. 본고에서 사용한 '창조성'과 '창의성'은 영어로는 creativity이지만, 사물이나 사태에 적용할 때는 주로 '창조성'이라는 용어를 쓰고 사람에게 적용할 때는 '창의성'이라는 용어를 쓴다는 점에서 표현을 달리했다.

견뎌 배기고 살아남아 승인된 것이다. 사전에 그것이 창조적인 것인지 아닌지 알 수 없다는 뜻이다. 창조성은 일단 긍정적인 결과라고 사회적으로 인정받은 후, 그것을 만들었던 행위에 내려지는 평가다. 창조성은 개인 차원의 사안이 아니다.

그런데 창조적 결과는 반드시 실험을 거쳐야 생겨날 수 있다. 시간순으로 보면 실험이 원인이 되어 창조적 결과가 나오지만, 논리적으로 보면 창조적 결과가 원인이 되어 실험 시점의 창조성이 승인되고 봉인된다. 실험은 결과와 무관하게 독자적이고 독립적이다. 실험은 결과에 의해 평가되겠지만, 실험은 결과를 모르는 채 무작정 행해질 수밖에 없기 때문이다. 성공할 거라는 보장이 없더라도, 결과를 예측할 수 없는 실험 속에서만, 창조적 결과가 잉태될 수 있다.

창조성의 본질을 이해하지 못하면, ‘창조성이 중요하다’라는 하나 마나 한 말만 되풀이할 뿐 ‘어떻게 해야 창의성을 계발할 수 있는지’에 대해서는 한마디도 할 수 없다. 나아가 창조성의 본질인 실험을 가로막는 일까지도 서슴지 않고 하게 된다. 실험은 미리 결과를 알 수 없다는 이유로 위험하다고 여겨지기 때문이다. 그러나 실험을 막으면서 동시에 창조적 결과를 기대하는 건 자기 모순이다. 창조성은 실험의 위험성을 이겨내는 개인적이고 사회적인 용기와 힘, 그리고 그것이 실천될 수 있는 자유를 전제로 한다. 실패해도 괜찮아야 한다. 개인도 사회도 크게 상처 입지 않을 수 있는 여건이 되어야 실험은 허용된다. 자유와 실험과 창조성의 관계를 강조하는 것도 이런 까닭이다.<sup>8)</sup>

## 2. 인공지능의 작동 원리

(물음) 2. 인공지능의 작동 원리는 무엇이며, 무엇을 할 수 있고, 무엇을 할 수 없는가?;  
2.1 인공지능의 바탕인 소프트웨어란 무엇인가?; 2.2 소프트웨어의 바탕인 디지털 컴퓨터란 무엇이며, 무엇을 할 수 있고, 무엇을 할 수 없는가?; 2.3 디지털 컴퓨터의 바탕인 컴퓨터 하드웨어란 무엇이며, 어떻게 만들어진다?; 2.4 컴퓨터 주변장치의 의미는 무엇인가?; 3.1 디지털 세계와 피지컬 세계는 어떤 차이가 있는가?

인공지능은 컴퓨터 프로그램, 즉 소프트웨어의 일종이며, 다른 모든 소프트웨어와 마찬가지로 컴퓨터 하드웨어에서 돌아간다. 컴퓨터 하드웨어의 핵심은 실리콘 반도체, 즉 NAND 게이트를 물리적으로 구현한 트랜지스터이다. 따라서 인공지능 기술의 발전은 두 가지에 의존하는데, 하나는 트랜지스터의 발전이고 또 하나는 소프트웨어 논리의 발전이다. 둘 중 하나에 ‘비약적’ 발전이 있으면, 현재까지 가능한 인공지능보다 ‘비약적으로’ 뛰어난 인공지능이 나올 여지가 있다.

한편 인공지능이 소프트웨어인 한에서 그것은 물리 세계에 영향을 줄 수 없다. 소프트웨어는 어디까지나 논리 회로에 불과하며, 시공간을 차지하지 않는 순수 추상이다. 소프트웨어로서의 컴퓨터는 “단지 램과 중앙처리장치”가 전부이다.<sup>9)</sup> 소프트웨어는 입출력 주변장치를 통해 비로소

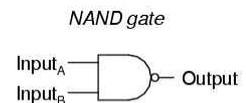
8) 이상 세 문단은 곧 출판될 김재인, 『뉴노멀의 철학 : 인류의 다음 단계를 위한 새로운 가치의 모색』, 동아시아, 2020의 6장 1절의 내용이다.

9) J. 클라크 스코트, 『그래서 컴퓨터는 어떻게 동작하나요?』, 지유록 옮김, 인사이트, 2019, 221쪽. (원서 : John Clark Scott, *But How Do It Know? - The Basic Principles of Computers for Everyone*, 2009.) 이 책은 폰노이만 아키텍처를 이용한 컴퓨터의 작동 원리를 일반인도 알 수 있게 잘 설명으며, 컴퓨터공학의 기본적인 작동 원리를 알고 싶은 모든 이에게 권한다. 특히 강인공지능(Artificial General Intelligence, AGI)과 특이점(singularity)을 주장하는 이라면 꼭 반박해야 할 핵심을 담고 있다. 부디 특이점을 공상하지 말고, 그 전에 새로운 수학/논리학의 가능성부터 제안하면 좋겠다. 아울러 다음 두 책도 참고하는 것이 좋다. 찰스 펫졸드, 『CODE : 하드웨어와 소프트웨어에 숨어 있는 언어』, 김현규 옮김,

물리 세계와 연결되고 연동된다. 키보드, 마우스, 마이크, 카메라, 모니터, 스피커, 프린터 등이 대표적인 주변장치이다. 인공지능 예술작품은 한편으로 소프트웨어를 만드는 문제이지만, 궁극적으로는 주변장치를 통해 물질적·감각적으로 구현되지 않으면 존재하지 않는 것과 다름없다. “고양이 코드가 고양이다워지는 것은 어디까지나 컴퓨터가 아닌 주변 장치에 달려 있다.”<sup>10)</sup>

### 폰노이만 아키텍처 : 디지털 컴퓨터와 소프트웨어

현존하는 모든 컴퓨터는 헝가리 출신 미국 수학자 존 폰노이만(John von Neumann, 1903~1957)이 1945년에 처음으로 설계한 디지털 컴퓨터의 논리구조에 따라 작동하며, 이를 ‘폰노이만 아키텍처(Von Neumann architecture)’라고 부른다. 70년이 훨씬 넘는 시간이 지났지만, 오늘날 가장 복잡한 슈퍼컴퓨터라 할지라도 폰노이만 아키텍처를 벗어나 작동하는 것은 없다. 인공지능은 컴퓨터 프로그램(소프트웨어)이므로, 마찬가지로 폰노이만 아키텍처 디지털 컴퓨터에서 작동한다. 컴퓨터 하드웨어가 아무리 발전했다 하더라도, 이는 단지 연산 속도의 발전에 불과하며, 소프트웨어가 할 수 있는 일과 없는 일을 아는 문제와 관련해서는 폰노이만 아키텍처를 분석하고 이해하는 것으로 충분하다는 뜻이다. 폰노이만 아키텍처는 그림에 보는 논리 연산을 수행하는 ‘낸드 게이트(NAND gate)’ 단 하나의 여러 조합만으로 작동하며, 그 다음엔 오로지 순수 수학·논리학의 영역이기 때문이다. 적어도 현재의 모든 컴퓨터는 그러하다. “컴퓨터는 설계된 대로 정확히 작업을 수행하는 기계일 뿐이며, 그 이상은 아니다.”<sup>11)</sup> 컴퓨터는 아무것도 모른다. 사람이 시키는 일을 할 뿐이다.



A	B	Output
0	0	1
0	1	1
1	0	1
1	1	0

“어떤 식으로 들여다보면 뇌와 마음은 적어도 우리가 배운 컴퓨터가 작동하는 원리와 전혀 다른 방식으로 돌아가는 것 같다. 물론 ‘우리가 배운 컴퓨터’라는 말은 미래에 다른 방식으로 설계된 컴퓨터가 등장할 수 있다는 가능성을 내포한다. 하지만 현대의 컴퓨터는 모두 ‘프로그램을 내장한 디지털 컴퓨터’로, 모든 컴퓨터는 이 책에서 소개한 것과 동일한 원리를 따른다. 지금까지 생각하는 컴퓨터에 관해 많은 부정적 ‘증거’를 제시했지만 그것이 인조인간은 결코 만들어질 수 없다는 ‘증명’이 될 순 없다. 언젠가는 생각하는 컴퓨터를 만들 수 있을지도 모른다. 다만 이 책에서 소개한 현재 컴퓨터가 동작하는 원리로는 불충분하다는 뜻이다. 완전히 다른 원리에 기반을 두고 동작하는 전혀 다른 장치를 설계하고 만들 수 있다면 가능할 수도 있다. 하지만 누군가 그런 장치를 실제로 발명할 때까지는 가능하나 불가능하냐를 놓고 따져봐야 아무 의미가 없다.”<sup>12)</sup>

인사이트, 2015 (원서 : Charles Petzold, *CODE : The Hidden Language of Computer Hardware and Software*, Pearson Education, Inc., 2009.) 박준석, 『세상을 만드는 글자, 코딩』, 동아시야, 2018.

10) J. 클라크 스코트, 같은 책, 315쪽.

11) 같은 책, 257쪽.

12) 같은 책, 325쪽.

## 하드웨어의 문제 : 실리콘 대신 양자 컴퓨터?

컴퓨터의 비약적인 발전으로 인간처럼 작동할 수 있는 인공지능이 나올 수 있을 거라는 기대도 많다. 그렇게 되면 인공지능의 예술 창작은 자동으로 가능하게 될 것이다. 하지만 컴퓨터 하드웨어의 발전을 기대하는 것은 난망하다. 트랜지스터의 수가 18개월마다 2배가 된다는 저 유명한 무어의 법칙이 한계에 이른 건 꽤 된 일이다. 인텔은 2016년 실리콘 소재를 이용한 프로세서 축소를 축소했다.<sup>13)</sup>

또한 한 전자공학자는 현재의 반도체 기술 수준에 대해 다음과 같이 증언한 바 있다.<sup>14)</sup> 이세돌은 여러가지 다 하는데, 알파고는 왜 바둑밖에 못 둘까? 알파고의 초기 버전(AlphaGo Fan)은 10만 개 정도 뉴런을 흉내 낸 것으로 추정되는데, 인간 이세돌은 약 1000억 개의 뉴런이 있다. 신경망의 규모가 100만 배 차이 나기 때문에, 인간은 알파고가 하지 못하는 큰 틀의 패턴 인식 가능하다고 추정할 수 있다. 그러면 알파고를 100만 대 만들어 연결하면 가능하지 않을까 하고 양을 늘려 보는 것을 고려해 볼 수도 있다. 그런데 이세돌은 저 규모를 처리하는 데 20W 전력을 사용했고, 알파고는 200kW 사용했다. 결론적으로 효율을 100억 배 올려야 가동할 수 있는데, 현실에서는 반도체 에너지 효율이 발전하지 못하고 정체 상태이다. 아마 노벨상 받을 만한 발견·발명이 10개 연속으로 나오면 가능할 수도 있겠다고 공학자들은 생각하고 있다.

현재의 실리콘 반도체가 한계에 이르렀다면, 양자컴퓨터가 등장하면 어떨까? 하지만 양자컴퓨터 또한 쉬운 문제가 아니다. 양자컴퓨터에 대해 국내 최고 전문가의 한 사람인 조동현 고려대 물리학과 교수(원자물리학)는 이렇게 말했다. 양자컴퓨터가 구동되었다는 보도가 잇따르던 시점이다.

**“양자컴퓨터 개발은 될 일이 아니다. 내가 죽을 때까지 안 될 것이고 앞으로도 영원히 구현되지 않는다고 생각한다. 이런 일(양자정보학 연구)을 해본 물리학자라면 그런 걸 다 알고 있다고 생각한다. 양자컴퓨터를 만든다는 건 그냥 불가능(simply impossible)하다. 1만년 걸릴 계산, 500년 걸릴 계산이라는 말을 하는데 중요한 건 그게 아니다. 계산이 의미 있으려면 하고 싶은 계산을 할 수 있어야 하고, 계산 결과가 맞아야 한다. 그런데 그렇지 못하다. 구글의 머신은 양자컴퓨터가 아니라 난수 발생기(random number)이다. 난수는 맞고 틀리냐가 없다. 얼마나 임의적(random)이나 하는 문제이다. 구글이 시커모어라는 프로세서를 갖고 난수 만들기 가능하다는 건 보였다. 그런데 그래서 어쨌다는 것인가(So What?). 나는 그걸 묻고 싶다. 구글의 양자컴퓨터는 쉬운 계산도 못 한다. 구글은 우리가 하고 싶은 계산이 아니라, 자신이 할 수 있는 계산을 한 거다. 나는 양자컴퓨터가 가능하다고 진짜 그들이 믿는 것인지, 아니면 회의적이면서도 이 연구를 하는 것인지 모르겠다.”<sup>15)</sup>**

이 말을 내 식으로 톺아보면, 양자컴퓨터는 과학기술의 영역이 아니라 주식시장의 영역이라고 요약할 수 있다. 따라서 양자컴퓨터의 실현을 전제로 인공지능의 발전을 운위하는 것은 논리적으로 신중하지 못할 뿐 아니라, 긴박한 문제에 집중할 에너지를 분산한다는 윤리적 문제도 있다.

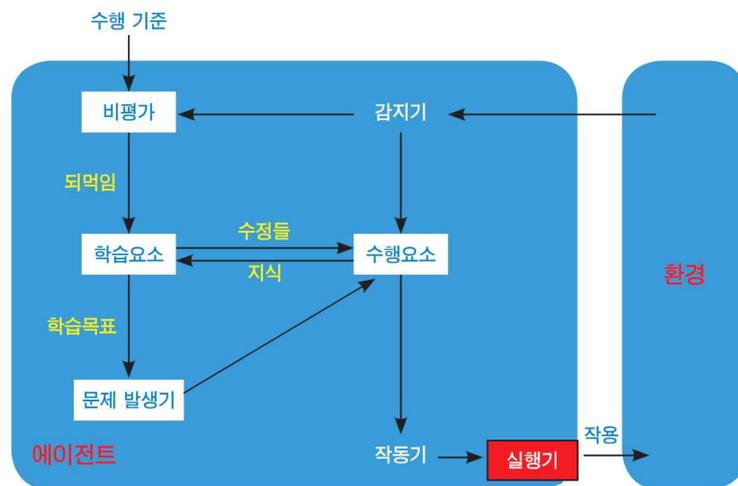
13) 무어의 법칙의 축소와 관련해서는 장가브리엘 가나시아, 『특이점의 신화 : 인공지능을 두려워해야 하는가』, 이두영 옮김, 글항아리, 3장 참조. (원서 : Jean-Gabriel Ganascia, *Le mythe de la Singularité : Faut-il craindre l'intelligence artificielle?*, Seuil, 2017.)

14) 감동근, 고등과학원 ‘인공지능을 둘러싼 몇몇 철학적 쟁점’ 논평, 2017년 11월 29일. 홍성욱 편, 『인공지능 : 과학역사철학』, 고등과학원 초학제연구단 자료집, 2018.

15) 최준석, “[스페셜 리포트] 양자컴퓨터 회의론 조동현 고려대 물리학과 교수”, 주간조선 2590호, 2020.01.06.

## 소프트웨어의 문제 : 인공지능의 논리

인공지능이 무엇이며, 무엇을 할 수 있고 또 무엇을 할 수 없는지에 대해, 나는 『인공지능의 시대, 인간을 다시 묻다』(2017)에서 자세히 논했다. 특히 인공지능의 가장 일반적인 구조는 아래 그림과 같이 제시할 수 있다. “요점은 이렇습니다. 전반적으로는 에이전트의 작용에 도움이 될 학습이 이루어진다는 거고, 그 학습이 잘 이루어지기 위해 **에이전트 바깥에 있는** 수행 기준을 근거로 비평가가 에이전트를 평가하고, 정해진 틀대로만 행할 경우 수행은 개선될 여지가 전혀 없을 테니 조금씩 변칙적인 탐험도 해보게끔 무작위로 문제도 발생시킨다는 거예요.”<sup>16)</sup>



일반적 에이전트의 구조 (변형)

여기서 ‘수행 기준’이 에이전트 바깥에, 즉 인공지능 바깥에서 주어진다는 점은 매우 중요하다. 이는 인공지능이 동물적 수준의 자율성도 갖고 있지 못하다는 뜻이기 때문이다. 둘 사이의 차이를 러셀과 노빅은 최고의 인공지능 교과서에서 이렇게 말한다.

“무엇이 좋고 무엇이 나쁜지에 대한 되먹임이 없으면, 에이전트는 다음 수를 결정하기 위한 근거가 없을 것이다. (...) 에이전트에 대한 우리의 구상은 보상을 입력 지각의 일부로 간주한다. 하지만 지각의 일부가 단지 또 다른 감각 입력이 아니라 하나의 보상임을 에이전트가 알아채기 위해서는 [판별 기준이] ‘내장(hardware)되어’ 있어야만 한다. 동물들은 고통과 배고픔을 부정적 보상으로 알아채고 쾌락과 음식 섭취를 긍정적 보상으로 알아채도록 내장되어 있는 것처럼 보인다.”<sup>17)</sup>

나는 이를 바탕으로 페드로 도밍고스를 인용하며 인공지능의 한계와 본질을 언급했다.

“초인공지능이 지각 능력을 갖게 되고 로봇 군대로 인류를 진압한다는 <터미네이터>의 시나리오는 우리가 이 책에서 만나게 될 그런 종류의 학습 알고리즘들로는 실현될 가능성이 없다. 컴퓨터가

16) 김재인, 『인공지능의 시대, 인간을 다시 묻다』, 75~76쪽.

17) Stuart J. Russell & Peter Norvig, *Artificial Intelligent: A Modern Approach, Third Edition*, Prentice Hall, 2010, p. 830. cf. p. 56.

학습할 수 있다는 것이 컴퓨터가 마법적으로 자신의 고유한 의지를 획득한다는 뜻은 아니니까. 학습자는 우리가 정해진 목표를 성취하는 법을 배운다. 그들은 목표를 바꾸는 데까지 가지 않는다. (...) 안심하시라. 마스터 알고리즘을 장착한 인공지능이 세상을 지배할 가능성은 제로이다. 그 이유는 간단하다. 인간과 달리 컴퓨터는 자신의 고유한 의지가 없다. 컴퓨터는 진화가 아니라 공학의 산물이다. 무한히 강력한 컴퓨터라도 여전히 우리 의지의 확장일 뿐, 두려워할 건 아무것도 없다. (...) 인공지능 시스템은 하는 일을 변경할 수도 있고, 심지어 놀라운 계획들을 내놓을 수도 있지만, 우리가 정해진 목표를 위해서만 그렇게 한다.”<sup>18)</sup>

요약하면, 인공지능은 사람이 시키는 일을 아주 잘 해내는 아주 유용한 도구이다. 인공지능에게 동물적인 수준의 자율성마저 기대해서는 안 된다.

### 3. 인공지능과 예술작품 창작

(물음) 1.1 예술과 창조성의 관계; 1.2 현대예술과 낭만주의 예술의 차이; 1.3 상식적이며 대중적인 수준에서, 예술 창작의 의미

나는 이 주제들에 대해 「인공지능은 예술작품을 창작할 수 있을까?」<sup>19)</sup>라는 글에서 자세히 논했다. 요점은 ‘인공지능 예술작품’의 문제를 그 누구도 ‘순수 공학’의 문제로 생각하지 않는다는 점에 있다. 말하자면 1) 우리가 예술을 공학이 아니라 미학의 문제로 본다는 점, 그리고 2) 미학이 18세기에 탄생한 가장 중요한 근대적 가치의 구현인 ‘예술(fine arts system)’을 위해 창조한 학문이라는 점<sup>20)</sup>, 바로 이런 점에서 3) 인간의 예술작품 창작을 창조의 ‘으뜸 본보기’로 승격시킨 낭만주의 미학은 현대인에게 아직도 가장 호소력이 있다.

#### 3.1 사례1. 구글의 ‘딥드림’과 마이크로소프트의 ‘넥스트 렘브란트’의 예

“저 두 인공지능은 예술작품을 창작한 것일까? 아니면 둘 다 예술가가 활용할 수 있는 유용한 도구일 뿐, 그 이상의 창조성을 발휘하지는 못한 것일까? 이런 물음 앞에서 우리는 비교적 쉽게 답할 수 있다. 결론부터 말하면, 결과물이 아무리 높은 미적 성취를 이룩했다 해도, 그건 ‘예술작품 창작’이라고 보기 어렵다. 나는 오늘날 낭만주의적 예술관, 즉 천재와 창조성을 강조하는 관점이 아무리 퇴조했다 하더라도, 우리가 저 결과물을 예술작품으로 여길 수 없는 데는 중요한 근거가 있다고 본다. 저 결과물들은 작품성 자체는 뛰어날지 몰라도 최소한 고흐나 렘브란트를 흉내 낸 것에 지나지 않는다. 우리는 이런 것들을 아류, 짝퉁, 심지어 표절이라고까지 부른다. 독창성이 없는 작품을 놓고, 비록 그 자체로서 아무리 뛰어나다는 평가를 받을 수 있다 하더라도, ‘예술작품을 창작했다’고 평가하기 어렵다. 거리의 화가의 작품이나 이른바 달력 그림을 즐기는 것은 충분히 가능하지만, 그것이 ‘창작된 예술작품’이라서 즐기는 건 아니다.”<sup>21)</sup>

18) Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, 2015. pp. 45, 283.

19) 김제인, 「인공지능은 예술작품을 창작할 수 있을까?」, 유현주 엮음, 『인공지능시대의 예술』, 도서출판b, 2019.

20) 김제인, 「바움가르텐으로 돌아가자: 감(感)적 앎의 복권을 위한 한 시도」, 『근대철학』 제9집, 2017 참조.

### 3.2 사례2. 렛거스 대학 팀의 ‘AICAN’

생산물이 예술작품일 수는 있어도 예술작품을 창작한 것은 아니다:

“십분 양보해서, 인공지능이 만든 작품, 즉 그 결과물에 대해서는 백 프로 인정하고 받아들일 수 있다. 미적 가치를 담고 있다는 점이 예술작품의 중요한 의미인데, 인공지능이 만든 작품을 보고 감동이 느껴지기도 한다. 그런 면에서는 인공지능이 만든 작품도 예술일 수 있다. 결과물만 놓고 보면 인공지능이 창의적인 활동을 한다고 볼 수도 있다. 문제는 인공지능 스스로가 그게 새로운지 모른다는 점이다. 그게 창의적이라고 평가해주는 건 결국 인간일 수밖에 없다. 오직 인간만이 인공지능이 만든 작품을 보고 ‘와 이거 새롭다!’라고 이야기한다.”<sup>22)</sup>

생산은 무작위적이며, 평가는 인간만 한다:

“인공지능은 작품을 일관되게 내놓지도 못한다. (...) 진짜 작가는 같은 수준의 그림을 계속해서 만들어낼 수 있지만, 인공지능은 무작위로 작품을 내놓는다. (...) 인공지능은 미적 가치를 평가하지 못한다. (...) 인공지능에게 작품을 무작위가 아닌 스스로 내린 평가 순서대로 내놓으라고 할 수 있을까? 자기 작품 중에 제일 좋은 것 10개를 순서대로 꼽아보라고 할 수 있을까? 적어도 AICAN의 작업에서는 그것이 불가능하다. AICAN은 (...) 작품을 무작위로 생산하는 일 이상은 하지 못하기 때문이다. (...) 나아가 인공지능은 자기 작품은 물론 다른 작품도 평가하지 못한다. (...) 원리상 인공지능은 평가 기준을 자기 바깥에 둘 수밖에 없기 때문이다. 그 기준은 인간이 준 것이다. 나는 이 점을 논증한 바 있다.”<sup>23)</sup>

3.3 소결론. 인공지능은 ‘창작’의 주체가 아니며, 예술가가 사용하는 유용한 도구이다:

“창작이라는 말을 제한된 의미로 쓴다면, 즉 넓은 의미에서 자연의 창작도 창작이겠지만 이 제껏 인간이 행했던 창작에 국한해서 고찰한다면, 인공지능은 ‘창작’의 주체가 되기 어렵다. 예술 ‘창작’이라는 말이 성립하려면 자신이 내놓는 작품에 대한 평가가 필수적이기 때문이다. (...) ‘인공지능이 예술작품을 창작한다’는 말이 어떤 의미인지 계속해서 물어야 한다. 인공지능이 만든 작품이 작품으로서의 가치가 없는 건 아니지만, 예술가가 하는 작업 과정을 거쳐서 나온 결과물과는 전혀 다르기 때문이다. 동시에 인공지능의 등장으로 ‘평가자’ 역할을 하는 인간의 안목이 굉장히 중요해지는 시대가 왔다. (...) 인공지능은 작가가 사용하는 도구들 하나에 지나지 않으므로, 단순한 작업 도구 이상으로 어떻게 나아갈 것인지는 작가의 몫으로 남을 것이다. 작가가 인공지능을 사용할 건지 아니면 다른 도구를 사용할 건지, 작가가 인공지능을 이용해 어떤 작업을 할 수 있을지는, 결국 작가가 무엇을 하려고 하느냐에 따라서 달라질 문제이다. 인간과 인공지능의 대립이 아니라 인공지능을 잘 다루는 인간과 그렇지 못한 자들의 차이가

21) 같은 책, 65~66쪽.

22) 같은 책, 77쪽.

23) 같은 책, 77~79쪽.

더 중요하다. 아이디어가 풍부한 작가가 뛰어난 작품을 만들 때 인공지능은 아주 좋은 도구로 작용할 수 있을 것이다.”<sup>24)</sup>

#### 4. 문화적 표현의 다양성 문제

– 디지털, 인터넷, 인공지능, 플랫폼은 창작과 표현의 다양성을 가속화했고, 차별적 취향을 가진 소비자의 편익을 증대시켰다

(물음) 3.2 디지털, 인터넷, 플랫폼은 쟁점과 어떤 관계가 있는가?; 3.3 ‘창작 - 생산 - 유통 - 접근’이라는 항목 구별은 오늘날 적절한가?; 3.4 ‘~의 흥수’ 속에서 홍보라는 행위의 본질은 무엇인가?

아젠다와 관련해서, 인공지능이 문화적 표현의 다양성을 증진시킨다는 점은 충분히 논증한 것 같다. 인공지능은 새로운 표현을 가능케 하는 유용한 도구이기 때문이다. 필요한 지원 방안이 있다면, 창작자에게 좋은 개인용 및 클라우드 컴퓨터를 보급하고, 소프트웨어 교육을 시키고, 개인이 구비할 수 없는 ‘인공지능 창작센터’ 같은 공간을 주고, 협업을 위한 미디어를 열어주는 일일 것이다. 나머지는 창작자가 알아서 하게 두면 된다.

이제 이상의 내용을 바탕으로 남은 물음들에 답해 보자. 디지털과 인터넷이 창작, 유통, 접근에 모두 긍정적이라는 점은 반박이 어렵다. 나는 이 논점에 대해서는 ‘디지털 르네상스’라는 연구의 결과로 논의를 대신하겠다.

“인터넷의 무한한 서고 공간에 의해 발생하는 전통적 롱테일의 편익은 상당히 크다. 이것은 특히나 다른 사람들과 차별적인 선호를 가진 소비자들에게 매우 유용하다. 그러나 아무도 알지 못하는 불확실성 효과는 훨씬 더 크다. 아마, 디지털화의 가장 중요한 편익(디지털 르네상스의 추진력)은 많은 창작자에게 시장에서의 기회를 제공하고, 많은 좋고 새로운 상품이 성공을 거두는 것이다.

디지털화가 해온 새로운 작업과는 별개의 다른 의미에서 디지털화의 편익을 볼 수도 있다. 공간 혹은 시간에 의해 한때 소비가 제한받았지만, 이제는 그런 제약이 없다. (...)

디지털화는 여러 가지 방법으로 소비를 해방시켰다. 음악의 소비자들은 물리적 상품을 소유하는 것로부터 자유로워졌고, 한 번에 12곡을 사야 하거나 다루기 힘든 장비 옆에 있어야 할 필요도 사라졌다. 넷플릭스와 스포티파이 같은 뷔페식 서비스를 이용하는 소비자들은 추가적인 영화와 노래를 이용하는 데에 추가적인 돈을 지불하지 않아도 되었다. 전부는 아니지만, 그런 접근 방식이 보고, 듣고, 읽기에 최적의 방법은 아니지만 말이다. 요약하자면, 우리는 디지털 르네상스에 살고 있고, 새로운 상품과 편익으로부터 얻는 소비자 편익은 엄청나게 크다.”<sup>25)</sup>

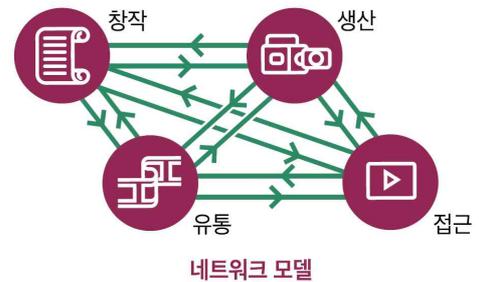
24) 같은 책, 84~87쪽.

25) 조엘 월드포겔, 『디지털 르네상스 : 데이터와 경제학이 보여주는 대중문화의 미래』, 임정수 옮김, 한울아카데미, 2020, pp. 210~211. (원서: Joel Waldfogel, *Digital Renaissance : What Data and Economics Tell Us about the Future of Popular Culture*, Princeton UP, 2018.) 강조는 모두 인용자.

디지털과 인터넷에 대해서는 이견이 많지 않지만, 플랫폼에 대해서는 논란이 많다. 특히 유튜브, 넷플릭스, 페이스북 등 글로벌 플랫폼의 독점적 지위가 문제시된다. 이 점은 <유네스코 문화적 표현의 다양성 보호와 증진에 관한 협약 제1차 국내 전문가회의 - 디지털 플랫폼과의 협력을 통한 2005년 협약의 이행><sup>26)</sup>에서도 집중적으로 토의되었다.

나는 플랫폼 자본주의에 대해서는 매우 비판적이지만, 2005년의 유네스코 「문화적 표현의 다양성 보호와 증진 협약」과 관련해서 미디어 플랫폼의 부정성에 대한 비판의 논거는 대단히 빈약하며, 오히려 미디어 플랫폼의 긍정성을 살리는 논의로 가야 한다는 입장이다.

이 점을 논하기에 앞서 유네스코에서 발간한 <문화 다양성협약 글로벌 리포트 : 문화정책의 (재)구성>(2017)에 언급된 ‘창작 - 생산 - 유통 - 접근’이라는 도식<sup>27)</sup>이 매우 부정확하며 시대착오적이라는 점을 지적하고 싶다. 이 도식은 “디지털 기술이 문화가치사슬을 변화시켰다”는 점을 설명하기 위해 도입되었는데, 지금 시점에서 이 네 항을 분리하는 것 자체가 적절하지 않다. 왜냐하면 네 항의 구분은 ‘피지컬’에 어울리는 것이지만, 현재의 ‘디지털 네트워크’에서는 무엇보다 동시성이 중요하다. 창작과 생산은 매체가 분리되었을 때 가능한 구별이며(가령 시나리오와 영화), 유통과 접근은 물리적 간격이 존재할 때 성립하는 개념이다(거대 배급사의 매개). 그러나 오늘날 디지털 네트워크 안에서 모든 것은 ‘미디어’이며, 매체와 매개는 별도로 존재하지 않는다. 그래서 소셜미디어<sup>28)</sup>이다. 최소한 유튜브와 페이스북은 소셜미디어라는 관점에서 분석해야 한다. 따라서 위의 도식은 개념 설계부터 다시해야 한다. 재설계에서 관건은 컴퓨터 하드웨어, 소프트웨어, 통신망, 입력장치(키보드, 마우스, 카메라, 마이크, 센서 등), 출력장치(모니터, 스피커, 3D프린터, 로봇, VR-AR장비 등)의 위상과 역할을 어떻게 설정하고, 이들 간의 관계를 어떻게 연결하느냐에 있을 것이다. 나는 들뢰즈와 과타리가 『안티 오이디푸스』에서 사용한 ‘생산-분배-소비’라는 도식이 여기에도 들어맞을 것으로 짐작한다.<sup>29)</sup>



이제 플랫폼을 살펴보자. 디지털 네트워크에서 플랫폼이 등장한 것은 오래 되지 않았으며, 그로 인해 아직까지 그 성격이 명확히 이해되지 못하고 있다. 어느 정도 이견이 없는 점은, 1) 그 안에서 수십 억명에 육박하는 사람들이 교류할 정도로 ‘집중적·독점적’이고, 2) 콘텐츠의 생산(업로드)-분배(검색, 추천)-소비(읽기, 보기, 듣기)가 동시에 일어나고 있으며, 3) 어떤 콘텐츠가 전 지구적으로 ‘대박’을 칠지는 아무도 예측하지 못한다는 정도이리라. 이런 상황에서 몇 가지 비판이 제기된다.

먼저 이동연의 논의부터 검토하겠다. 논지는 이렇게 요약될 수 있다. “플랫폼의 지배력이 너무

26) 유네스코한국위원회 11층 대회의장, 2020년 5월 22일(금) 14:00 - 18:00. 발제는 이동연, 「디지털 미디어 플랫폼의 부상과 문화다양성」과 이광석, 「디지털 환경의 미래와 문화적 표현의 다양성」이었다.

27) 옥타비오 콜레스, 「제3장 플랫폼 시대의 문화정책」, 유네스코, 『문화다양성협약 글로벌 리포트 : 문화정책의 (재)구성』, 이구표 옮김, 유네스코한국위원회, 2018, 72쪽.

28) 한국에서만 부당하게 SNS라고 불린다. 하지만 어디까지나 미디어이다. 이를 둘러싼 난맥상에 대해서는, 김재인, 「한국언론은 왜 ‘소셜미디어’가 아니라 ‘SNS’를 주로 쓰나」, 뉴스톱, 2018년 12월 5일 칼럼을 참조.

29) 질 들뢰즈 & 펠릭스 과타리, 『안티 오이디푸스 : 자본주의와 분열증』, 김재인 옮김, 민음사, 1장 참조. (원서 : Gilles Deleuze & Félix Guattari, L’Anti-Oedipe, Capitalisme et schizophrénie, Minuit, 1972/3.)

강해질 경우에는 콘텐츠 창작자들의 권리가 축소되고, 콘텐츠의 플랫폼 쏠림현상이 심화될 우려를 낳고 있다. 이는 문화다양성의 관점에서 충분히 우려되는 지점이고, 많은 논쟁지점들을 낳는다.”<sup>30)</sup> 나는 창작자의 권리가 보장되고, 특히 적절한 보상이 있어야 한다고 생각한다.<sup>31)</sup> 이 문제는 ‘디지털’과 ‘글로벌’이 엮여 작동하는 시대의 난점을 모조리 내포하고 있기 때문에 별도의 논의가 필요하다.<sup>32)</sup> 하지만 이와는 별도로 ‘콘텐츠의 플랫폼 쏠림현상’이 우려된다는 점을 나는 납득하기 어렵다.

플랫폼이 중립적인 놀이터라고 생각한다는 말은 아니다. 그러나 각 주제별로 나눠서 생각해 보면, 플랫폼은 모두에게 이익이다. 콘텐츠 생산자는 가장 많은 기회와 잠재 고객이 있는 곳으로 가는 것이 당연하고, 소비자는 가장 다양한 콘텐츠가 있는 곳으로 가려 할 것이며, 광고주는 가장 많은 사람이 있는 곳을 타겟으로 해야 마땅하다. 플랫폼 기업의 이익이 천문학적이라고 해서, 단지 그 자체만으로 비판받을 이유는 없다. 이동연은 “넷플릭스가 제작 투자를 결정하는 기준은 콘텐츠의 잠재력보다는 자신이 보유한 고객의 시청 습관과 잠재 고객의 수요를 분석한 데이터의 힘”<sup>33)</sup>이라고 말하는데(이는 많은 비판자와 동일한 어조이다), 나는 이 구절에서 ‘콘텐츠의 잠재력’과 ‘데이터의 힘’이 무엇을 지칭하는지 도무지 알 수 없다. 데이터만 잘 분석하면, 아무 콘텐츠라도 상관없다는 말인가? 사실 인공지능이 분석해서 제시하는 넷플릭스의 추천 알고리즘은 기업의 핵심 노하우이다.

말이 나온 김에 넷플릭스와 유튜브에 대한 비판을 좀 더 경청하자. 생산 측면에 초점을 맞추면서 이동연은 말한다. “유튜브 같은 미디어 플랫폼을 제국의 플랫폼으로 비판할 수밖에 없는 것은 그러한 모든 행위들이 플랫폼 안에서 이루어진다는 점 때문이다. (...) 제3세계 저항의 영상들은 유튜브에 업로드되는 순간 문화 콘텐츠의 자산으로 전환되는 것을 피할 수 없다. 일반인이 구루(guru)가 되고, 놀이가 엔터테인먼트가 되며, 저항이 상품 형식으로 전환되는 것은 플랫폼 제국의 숙명이다. 유튜브 안에 존재하는 수백억 개의 동영상은 훈육되지는 않겠지만, 통제되는 것을 막을 수는 없다. (...) 그것은 문화다양성의 실현이기보다는 문화적 미국화의 진화된 형태다. 문화적 미국화는 문화적 다양성을 거대하고 강력한 플랫폼으로 흡수한다.”<sup>34)</sup> 한편 소비 측면에 초점을 맞추면서 이광석은 “이용자들이 알고리즘 분석과 취향 예측에 최적화된 기형의 문화 소비 주체가 되는 일은 시간 문제”라면서, “가령, 맞춤형 콘텐츠를 서비스하는 넷플릭스나 유튜브 등은, 이용자의 문화적 관심을 분할하고 최적화된 맞춤형 콘텐츠를 제공하면서 세분화된 취향을 구분해내지만, 동시에 우리 취향을 한곳에 가두면서 점점 ‘납작하게’ 만들고 정해진 알고리즘 회로 안에서만 각자의 문화 소비를 행하도록 가둘 확률 또한 높”이며, 그 결과 “소비나 이용의 장르적 널뛰기가 서로 이뤄지기가 어렵다”고 지적한다. “빅데이터 기술문화는 이미 존재하는 문화적 선호와 편견을 더 단단히 만드는 반면, 새롭고 이질적인 것들에 대한 대중의 접촉면을 현저히 낮춘다는 점에서 대단히 문화 보수적”이며, “소셜미디어 가입자는 공급자의 설계가 강요하는 닷컴 감정과 정서 패턴의 소모라는 사회적 편견을 강화하는 의식의 소통 회로 방식에 길들여질 수밖에 없다”는 것이다.<sup>35)</sup>

30) 이동연, 같은 글, 7~8쪽.

31) 이 점에서 “글로벌 미디어 플랫폼이 콘텐츠 창작자들의 창작물들을 플랫폼 서비스에 연결하면서 받아가는 비율이 너무 많고, 창작자들과 소비자들이 문제제기 하지 않는 이상, 이 거래의 방식들은 전적으로 플랫폼 서비스 업자들이 주도할 수밖에 없”(이동연, 17쪽)기 때문에, 꼭 풀어야 한다는 주장은 십분 동의한다.

32) 국경 없이 넘나드는 글로벌 정보통신기술(ICT) 기업을 어떻게 규제할 것이냐 하는 문제이다.

33) 이동연, 11쪽.

34) 이동연, 15쪽.

35) 이광석, 같은 글, 23~24쪽.

나는 이런 비판이 프랑크푸르트학파의 대중문화 비판과 얼마나 다른지 구별하기 어렵다. 하지만 영화와 텔레비전 같은 대중매체(mass media)와 페이스북, 넷플릭스, 유튜브 같은 소셜미디어(social media)는 성격이 많이 다르다. 전자가 ‘일대다’의 구도라면 후자는 ‘다대다’의 구도이다. 즉 후자는 생산과 소비의 조합을 바꾸는 게 얼마든지 가능하다. 생산의 측면부터 보자. 이동연은 콘텐츠 크리에이터의 위상이 변했다고 말한다. “제도화된 문화자본을 갖고 있지 않”지만, “그들은 자신이 좋아하는 문화적 취향에 기반 해 동영상을 올리기도 하고, 자신이 관심 있는 여러 분야에 자신의 경험적인 노하우를 유튜브 이용자와 공유하고 싶어 올리기도 한다.”<sup>36)</sup> 이광석도 비슷하게 말한다. “표준화된 자본주의 순환 체제에서 벗어나 문화 소비나 향유의 민주화는 말할 것도 없고, 누구든 문화 제작과 표현의 기회를 가질 수 있는 문화적 수단에 대한 접근의 대중화와 데이터 생산과 소비의 다양화가 이뤄지고 있다. 적어도 기술 여건에서만 보자면, 오늘날 누구든 원하면 데이터 생산과 유통의 주체들로 나설 수 있게 됐다.” 나아가 소비 측면을 주목하면, “빅데이터 알고리즘 예측은 콘텐츠 소비자들 각자 지닌 특정 취향의 독특한 문화 소비의 다양한 결들을 확대한다는 점에서 긍정적”이며, “질적으로 보면, 이는 단순히 이용자의 콘텐츠 제작 참여라는 수준을 넘어서서 그들 스스로 문화 시장 내 지분을 획득하면서, 대중문화의 생산과 제작, 소비에 있어 질적인 영향력을 미치고 때론 주류 대중문화와 색다른 언더그라운드 문화적 다양성의 가치를 확대 생산하는 근거가 되고 있다.”<sup>37)</sup>

이처럼 대부분의 비판자까지도 생산과 분배와 소비에서 일어난 긍정적 변화를 잘 알고 또 지적한다. 하지만 동시에 비판자는 생산과 소비 면에서 한편으로는 ‘플랫폼 형식에 맞추는 생산의 종속’이 일어나고 다른 한편 ‘취향이 납작해지면서 다양한 문화 소비가 축소’된다고 지적한다. 나는 이런 비판이 정확히 무엇을 의미하는지 알 수가 없다. 플랫폼이 없었더라면 생산-분배-소비되지 않았을 극도로 다양한 콘텐츠가 생겨났기 때문에 플랫폼은 문화 표현의 다양성뿐 아니라 문화 향유의 다양성까지도 가능케 했다. 이러한 활동이 플랫폼 기업의 배를 불린다는 것은 별개의 문제이다. 나는 문화 다양성과 관련해서 대부분의 플랫폼 비판자가 ‘다양성의 조건 확보’와 ‘이윤 극대화 및 독점’이라는 서로 별개인 두 문제를 섞어버리는 오류를 범한다고 지적하고 싶다.

문제를 잘못 설정하면 해법도 엉뚱한 법이다. 이동연은 시민사회 관점에서 문화 다양성을 보장하기 위한 법적·제도적 대안을 제안하고 있으나, 제시한 대안은 영화 배급을 모델로 하고 있기 때문에, 현재의 디지털 네트워크에 들어맞지 않는다. 게다가 디지털 네트워크는 국내법으로 규제할 수 있는 범위를 넘어선다. 이광석 역시 ‘플랫폼 협동주의’를 제안하면서, “대안 공정문화 시장을 만들기 위한 플랫폼 커먼즈 조합”이라고 규정한다. 생산자에게는 정당한 이익을, 소비자에게는 다양한 문화 콘텐츠를 제공하자는 취지이다. 그러나 대부분의 협동조합 운동이 실패한 것은 자생을 위한 규모를 확보하기 전에 망해버렸기 때문이다. 앞에서 지적했듯, 내가 생산자라도 기회와 잠재 고객이 많은 곳을 먼저 노릴 것이고, 내가 소비자라도 콘텐츠 품종이 더 많은 곳부터 기웃거리리라.

넷플릭스와 유튜브의 인공지능 추천 알고리즘이 취향을 좁히고 다양성을 줄인다는 문제도 다른 각도에서 이해할 수 있다. 문화 콘텐츠와 관련해서 인간의 가장 중요한 특징은 지루하고 재미 없으면 소비하지 않는다는 점이다. 종수가 제한되어 있을 때는 어쩔 수 없지만, 즐길 거리가 무궁무진할 때는 금세 다른 대안을 찾아간다. 추천 알고리즘이 문제라는 지적은 사태를 너무 알파하게 보기

36) 이동연, 14쪽.

37) 이광석, 21~22쪽.

때문에 생겨난다. 우리는 적절한 추천을 언제나 환영한다. 미술관, 박물관, 음악회는 모두 추천, 즉 큐레이션에 의해 작동한다. 무차별적으로 제공되는 콘텐츠에는 사실 관심을 가질 시간 여유가 없다. 우리는 전문가의 안목으로 잘 고른 콘텐츠를 원한다. 취향의 강요는 폭력이지만, 우리는 취향의 유도를 원한다. 삶과 시간의 가성비 때문이다. 인공지능이 추천하지 못하란 법도 없으며, 때로는 그게 더 잘 맞아떨어진다. 인공지능이 가장 잘 하는 것은 패턴 분석, 분류, 예측, 이 세 가지이다. 넷플릭스와 유튜브는 소수 및 개인 취향까지도 분석해서 추천하는 놀라움 덕에 호응을 얻었으며, 이는 아마존이 도서 추천을 통해 초기의 성과를 올린 것과 같은 이치이다. 홍보 또는 광고가 함께 가는 것도 비슷한 이치이다. 인공지능 추천을 가볍게 보는 건 아직 너무 인간주의에 젖어 있기 때문일 것이다. (끝)