

2. 최악의 실수와 최고의 축복 사이

“인류 역사상 가장 큰 사건이 될 것” …… “금융 시장보다 더 혁명할 것이고, 인간 과학자들보다 더 나은 것을 발명할 것이고, 인간 지도자들보다 더 정치적 수완이 뛰어난 것이며, 인간이 이해할 수조차 없는 무기도 개발할 수 있을 것” …… “역사상 최악의 실수가 될 수 있다”

물리학자 스티브 호킹^{Stephen Hawking}과 그의 생각을 공유하는 학자들이 급격히 발전하는 AI의 위협에 대해 위와 같이 말했다(*The Independent*, 2014년 5월). 테슬라의 최고경영자 엘론 머스크^{Elon Musk} 또한 “인공지능 연구는 악마를 소환하는 것과 다름없다”는 다소 자극적인 발언을 했다(100주년 심포지엄, 미국 MIT 항공우주공학과 주최, 2014년 10월). 빌 게이츠^{Bill Gates}도 엘론 머스크의 발언에 동조한다고 밝히며 AI의 출현을 경계하는 모습을 보였다. 그는 “기계가 인간을 위해 많은 일을 해줄 수 있지만 ‘초지능^{super intelligent}’은 그렇지 않다”는 말을 남겼다(‘Ask Me Anything’ 행사, 미국 IT 커뮤니티 레드딴주최, 2015년 1월 28일). 이외에도 많은 사회 지도층 인사들이 AI의 위험성을 경고하고 있다.

하지만 이런 부정적 발언에도 불구하고 AI가 가져올 경제적 가치와 미래사회에 대한 장밋빛 전망은 우리 주위에 만연해 있다. 그리고 핵이 그랬듯이 긍정과 부정의 교차 속에서도, 부정이나 우려와 상관없이 AI 개발은 계속될 것이다. 2019년 전 세계의 AI 시장규모가 313억 달러(약 36조7000억 원)에 이를 것이라는 전망(IDC; www.idc.com, 2016년 3월 17일)이나 우리나라 정부가 알파고 이슈로 AI에 5년간 1조원 투자하겠다고 한 것(미래창조과학부, 2016년 3월 17일) 등은 그만큼 이 분야를 밝게 보고 있고, 지속해서 발전시키고자 하는 강한 의지를 밝힌 것으로 보인다.

너무 일상적이어서 체감하기 힘들지만 이미 우리 생활 곳곳에 AI 기술은 적지 않게 사용되고 있다. 센서 작용으로 반응하는 시스템은 기초적인 AI로 볼 수 있는데, 자동문과 주차장의 차량 인식 개폐 장치나 하이패스, 무인 발권기 등이 그 예이다. 다소 차원 높은 AI로는 애플의 ‘시리^{Siri}’ 같은 음성 인식 프로그램이나, 기계 학습 알고리즘으로 이용자 성향에 맞는 것을 추천해주는 ‘아마존닷컴^{amazon.com}(책)’, ‘넷플릭스^{netfix.com}(영화)’, ‘매치닷컴^{match.com}(이성)’ 등이 있다. 또한 구글의 온라인 번역 시스템이나, 무인자동차, 스마트 공장도 고차원의 AI가 결합

되었다고 볼 수 있다.

AI에 대한 연구는 이미 초기 컴퓨터 시대, 즉 디지털 시대 초기부터 본격적인 논의가 있었다. 1950년 앨런 튜링^{A. M. Turing}이 철학 저널 *Mind*의 「컴퓨팅 기계와 지능 *Computing Machinery and Intelligence*」(1950)이라는 논문에서 ‘기계가 생각할 수 있는가?’라는 도발적 질문을 던짐으로써 지능을 갖는 기계에 대한 사유에 자극을 주었다. 그는 ‘튜링 테스트^{Turing test}’라는 이미테이션 게임(흉내 게임)을 제안하여 지능을 가진 개체를 판단하는 기초적 기준을 제시하기도 했다.

1950년대 초반의 많은 연구자는 인공지능시대가 바로 눈앞에 있다고 생각했다. 그들은 10년 안에 컴퓨터가 세계 체스 챔피언이 되고, 언어 간의 완벽한 번역이 가능하며, 컴퓨터 시스템이 제공하는 드라이 마티니로 고된 하루를 달래게 될 것이라고 주장했다. 하지만 우리가 알다시피, 1960년대에 이런 일들은 일어나지 않았다. 컴퓨터(Deep Blue 2)가 인간 체스 챔피언을 이기고 세계 챔피언이 된 때는 1997년이었고, 수준 높은 번역은 아직도 넘어야 할 산이 많다.

그렇다면 더디던 AI 연구가 근시기에 급물살을 타게 된 것은 왜일까? 바로 인터넷의 발달, 빅데이터^{big data} 축적, 향상된 알고리즘의 결합으로 이전과는 상황이 달라졌기 때문이다. 다시 말해, 현재와 같은 모습의 AI는 네트워크(인터넷)가 연결되면서 형성된 빅데이터와, 그 방대한 데이터를 분석하여 통계적 상관성을 알아내서 스스로 프로그램 개선하는 기계 학습 알고리즘이 향상됐기 때문이다.

현재의 AI 수준을 보여준 알파고도 빅데이터와 클라우드 컴퓨팅, 기계 학습 알고리즘 등 최첨단 ICT^{Information Communication Technology}(정보기술과 통신기술의 결합) 기술이 총동원된 시스템으로 초당 10만 개에 달하는 수를 고려한다고 한다. 알파고를 개발한 구글 딥마인드의 창업자이자 최고경영자인 데미스 하사비스^{Demis Hassabis}는 알파고에 3,000만 건의 프로바둑기사 대국 기보를 입력시켰고, 이것을 바탕으로 쉬지 않고 바둑을 두며 배우도록 기계 학습 알고리즘을 설계했다. 이로써 딥마인드의 연구총괄 데이비드 실버^{David Silver}가 말했듯이 “1,000년에 해당하는 시간만큼 바둑을 학습”했다. 그렇다면 이세돈 9단이 다섯 번 중 한 번 승리했던 것도 대단한 성과라 볼 수 있을 듯하다.

알파고가 1,000년의 시간 동안 바둑을 두며 시스템을

최적화시켰듯이, AI는 인간이 남긴 기록을 파헤치며 스스로 학습하고 지속해서 개선할 것이다. 이 같은 AI는 우리의 예상보다 빨리 현실화 될 것 같다. 제임스 배럿 James Barrat이 AI 연구자 200명을 대상으로 AGI(Artificial General Intelligence)(인공범용성지능)의 실현 시기에 대해 비공식적인 조사를 해본 결과 42%가 2030년, 25%가 2050년, 20%가 2100년으로 나왔다. 절대 이루어지지 않을 것이라는 답변은 겨우 2%에 머물렀다. 또한 상당수의 과학자는 의견란에 더 빠른 선택지인 2020년도 있어야 한다고 적었다.¹⁾

그렇다면 이 이후에는 무엇이 있을까? AGI에 도달하면 AI가 자신의 설계를 개선하고, 소프트웨어를 교체하여 진화된 새로운 설계를 제작할 것이며, 이 순환을 거듭하면서 최적화하는 진화 프로그래밍 기법을 적용할 것이라고 의견을 모은다. 이렇게 되면 인간으로서는 이해할 수 없을 정도의 연산 속도와 정보 접속 능력을 갖게 될 것이다. 최적화를 거듭하는 과정에서 어떤 인간보다도 수백만 배 더 현명한 기계의 모습을 갖게 될 것이고, 우리의 상상을 초월한 모습으로 우리 앞에 나타날 것이다. 실로 미래학자 레이 커즈와일 Ray Kurzweil의 말대로 “역사의 틀을 해집어놓을” ‘특이점(singularity)²⁾이 올 지도 모르는 일이다.

이 상황에 이르면 ‘포스트휴먼’을 꿈꾸며 AI를 긍정하는 커즈와일의 바람처럼 생각하는 기계가 인류에게 무조건적인 도움을 줄지는 의문이다. 특히 인간의 모든 것(빅데이터)을 알 수 있는 AI가 인간을 긍정적으로만 판단할까? 과연 예외 상황에서 우리는 이 AI 시스템을 통제할 수 있는 방법은 있을까?

3. 언어 잃은 이아모스

“숫자들과 더불어 모든 것들이 진행된다. 즉, 변조 modulation, 변형 transformation, 동기화 synchronization가 가능해지고 지연 delay, 저장, 전위 transposition 뿐만 아니라 혼합 scrambling, 스캐닝 scanning, 매핑 mapping도 가능해진다. 디지털 토대 위에서 진행 되는 전체적인 미디어 연결 media link은 미디어 개념 그 자체를 지우게 될 것이다.”³⁾

앨런 튜링은 제2차 세계대전 중 기계 장치를 제어하는 수리 논리적인 추상적 구조를 세움으로써 디지털 컴퓨팅 머신의 싹을 틔웠다. 이것은 기계의 논리적 형식을 물질

적 토대에서 떼어낸 것으로 ‘알고리즘’의 시초가 되었다. 이후 인간이 사용하는 ‘정현파(Sine Wave)’를 ‘이진 신호(Binary Signal)’로 변경하는 방식이 상용화되면서 이진 신호로 변환하는 ‘디지털화(Digitalization)’가 전 영역에 걸쳐 괴물처럼 모든 것을 흡수하기 시작했다.

‘On(1)’과 ‘Off(0)’의 이진 신호 체계는 완전히 같은 신호를 만들어 내기 힘든 아날로그 신호 체계와는 달리 완벽한 복제가 가능한 체계로, 소통 관계에서 수신자와 발신자의 내용이 완전히 일치하는 모습을 보인다. 매체학자 프리드리히 키틀러 Friedrich Kittler는 이 이진 신호 체계가 이미지, 소리, 텍스트 등을 모두 숫자(0, 1)로 전환시킨다고 말한다. 그 결과 광학적 매체(이미지)나 청각적 매체(소리), 문자 매체(텍스트)가 지닌 매체의 개별성은 사라진다고 주장한다.⁴⁾ 근대 시기까지 다양한 물리적인 힘, 다시 말해 타점, 주파수, 전기 신호 등을 이용하던 기록 매체는 이제 0, 1의 신호를 사용하는 디지털 매체로 수렴되는 상황이다.

문제는 아날로그 매체보다 디지털 매체가 실행과정을 더욱 감추고 있으며, 기록된 것을 해석하기 더 힘들다는 사실이다. 기계어(Binary Code)라 불리는 디지털 신호가 목적 프로그램(object program)에서 실행되기 위해서는 컴파일(compile)(번역)이라는 과정을 거쳐 된다. 번역을 수행하는 프로그램들이 바로 컴파일러(compiler)와 어셈블러(assembler)인데, 컴파일러는 FORTRAN, COBOL, ALGOL, C와 같은 인간이 이해하기 쉬운 고급 프로그래밍 언어로 작성된 프로그램을 번역하고, 어셈블러는 컴퓨터 언어에 가까운 저급 프로그래밍 언어인 어셈블리어(assembly language)로 작성된 원시 프로그램을 번역한다. 그래서 디지털 매체 안에서는 ‘사용자 언어 → 번역 → 기계어 → 실행’의 과정이 우리가 상상할 수 없을 정도로 무한 반복된다. 이 과정에서 인간의 감각은 디지털 데이터로 전환(혹은 기록)되는 것을 포착할 수 없다. 그리고 그 간극은 프로그램이 사용자 친화적으로 변하면서 점점 더 벌어지고 있다. 마이크로소프트 운영체제만 보더라도 MS-DOS였을 때에는 사용자가 여러 컴퓨터 명령어를 알아야 기능을 구동할 수 있었지만, WINDOWS 체제로 변하면서 마우스 클릭만으로 가능한 직관적 시스템으로 바뀌었다. 리눅스(Linux)나 맥 OS(Mac OS) 등의 운영체제도 직관적으로 변하였다. 현재 우리는 마우스를 요술 지팡이처럼 너무도 쉽게 사용하고 있지만, 클릭하는 순간 얼마나 복잡한 과정이 진행되는지 알 수 없다. (전문가조차도 이론적으

로만 감지할 뿐 감각으로 포착할 수는 없다.) 기계어라는 미지의 세계를 모른다고 해서 컴퓨터 사용에 크게 불편함을 느끼지 못하는 상황은 ‘블랙박스^{black box}’를 강화하는 방향으로 흐르게 한다. 디지털 프로그래밍 언어로 설계된 소프트웨어 알고리즘은 공학적·물리적 기술에 비해 그 구조가 노출되지 않는데, 이것을 ‘블랙박스’ 기술이라 한다. 디지털 미래를 낙관하는 사람들은 작동 방식을 상세히 알 수 없는 ‘블랙박스’라도 그것을 벽돌로 삼아 여러 개를 차곡차곡 쌓아서 예측하지 못할 혁명적인 애플리케이션의 건물을 지을 수 있다고 전망한다. 하지만 현재도 기술 독점으로 발생하는 ‘블랙박스’로 인해 사용자가 거대한 힘을 통제하지 못하고 도리어 그 힘에 통제당하는 상황에서, 기술 혁신에 대한 낙관적 전망을 기쁘게만 받아들일 수 없다. 이미 우리는 불필요하거나 원하지 않는 프로그램을 (개발자의 의도에 따라) 어쩔 수 없이 컴퓨터에서 실행해야하는 상황이다.

불길함은 인간 전문가가 만들어낸 기술 독점적 ‘블랙박스’가 아니라, 인공지능 스스로 만들 수도 있는 ‘블랙박스’에 있다. 0(off)과 1(on) 값이 연속 배열된 기계어는 번역 프로그램이 없으면 모스 부호^{Morse code} 만도 못하게 된다. (모스 부호는 규칙에 따라 인간이 해독할 수 있지만 기계어는 그 자체를 인간 감각으로 해독하기 힘들 뿐만 아니라, 가능하더라도 어마어마한 시간이 걸린다.) 러시아 기호학자 유리 로트만^{Yuri M. Lotman}은 ‘미칠’ 수 있는 능력도 지능의 훌륭한 기능적 자질이라 말했다. 그는 사유하는 조직체가 이성적 행위에 대한 대안으로 미친 행동을 수행할 수 있는 잠재력을 지니고 있고, 이성적 행위와 미친 행동 사이에서 항시적인 선택을 실현할 수 있는 능력을 지니고 있다고 말했다.⁵⁾ AI의 발달이 사유하는 기계에 이르렀을 때, 이성적인 행위만 계속할 것이라고 장담할 수 있는 근거는 전혀 없다. 우리는 ‘미친’ 행동을 할 잠재력을 고려해야 한다. 또한 인류 바깥에 존재하는 지능(AI)이 실행하는 이성적인 행위가 인류를 돕는 일만 있을 것이라는 순박한 기대에 머물러 있을 수도 없다. (인류를 돕는 일이 인류 멸종이라 판단한다면 어찌할 것인가.) 만약 AI가 어떠한 판단으로 번역 프로그램의 구조를 인간이 알 수 없게 한다거나, 기계어의 체계를 인간이 해독 불가능하게 재편하는 등 ‘블랙박스’로 만든다면, 우리는 어떻게 해야 할까? 아직도 해석하지 못한 마야 문자^{Maya scripts}나 파이스토스 원반 문자^{Phaistos Disk}, 롱고롱고 문자^{Rongorongo Script} 등의 고대 상형문자처럼 디지털 문

자, 즉 기계어는 미지의 언어가 될지도 모른다.

새의 언어를 알아듣는 아폴로의 아들 이아모스^{Iamus}가 언어를 잃어버리는 것을 혹시 미래의 우리가 경험하게 되는 것은 아닐까?

4. 디지털 ‘기억-역사’ 의 증발

“그대의 발명품은 배우는 사람들의 영혼에 망각을 불러일으킬 것이요 사람들은 더 이상 자신의 기억을 사용하려고 들지 않을 테니까. 사람들은 외부에 쓰인 문자들을 믿고 스스로 기억하려고 하지 않을 것이요. 그대가 발명한 것은 기억에 도움을 주는 것이 아니라 단지 기억을 불러일으키는 데 도움을 줄 것이요.”⁶⁾ - 플라톤, 『파이드로스^{Φαῖδρος}』 중

플라톤은 문자가 인간의 기억능력을 쇠퇴시키고, 기억하려는 의지를 소멸시킨다고 일찍이 말했다. 기록이 기억을 대신하고, 기록 매체가 기억 능력의 자리를 차지하는 현상 속에서 ‘기억’과 ‘기록’, ‘기록 매체’ 사이의 긴밀성이 드러난다. 기록은 과거의 흔적이라는 측면에서 ‘역사’와도 관계가 깊다. 문자 발명 이후, 우리는 분절적인 문자성^{literacy}을 얻는 대신에 비분절적인(연속적인) 구술성^{orality}을 잃어버렸다. 원칙적으로 비분절적 텍스트를 분절적 텍스트로 완전히 번역하는 것은 불가능하기 때문이다. 하지만 월터 옹^{W. J. Ong}에서 마샬 맥루한^{M. H. McLuhan}에 이르는 대부분의 매체 이론가들은 근원적인 유대 관계를 느끼는 문자 ‘이전’(비분절성)이 어찌면 디지털로 대면되는 문자 ‘이후’에 회복될 가능성이 있다고 생각한다. 특히 문자^{Schrift}보다 기록^{Aufschreibung}을 집중적으로 조명했던 키틀러는 문자 때문에 기록되지 못했던 소리와 이미지 같은 비분절적인 텍스트가 아날로그 기록 매체인 축음기(소리), 영화(이미지) 등으로 기록되기 시작했다고 말했다. 사라지는 소리를 기록하기 위해 문자 체계는 악보를 도입했지만, ‘존재하지만, 존재하지 않는 듯 존재하는 것으로서의 실제’인 (악보에 기록하지 못하는) 소음은 축음기가 등장한 이후 그 존재감을 드러냈다. 1초당 24개의 장면으로 분절하여 현실세계를 기록한 초기 무성영화도 새로운 기록을 가능하게 했다. 키틀러는 디지털 매체의 등장으로 축음기, 영화, 타자기로 상징되는 아날로그 방식의 매체 개별성이 사라지고 디지털 방식으로 모든 기록 형식이 수렴된다고 말하고 있다. 즉 문자성을

뛰어넘은 모든 기록 방식이 디지털 방식으로 전환되어 축적되고 있다는 뜻이다. 그리고 이러한 현상 속에서 매체 이론가들은 문자 '이전'의 회복 가능성을 고려하고 있는 것이다.

그런데 문자 '이전'으로의 회복이 기억 능력의 회복이 아닌, 완전성과 관련 있다는 것을 주목해야 한다. 플라톤이 문자의 등장으로 인간의 기억 능력이 쇠퇴하고 기억하려는 의지가 소멸하리라고 본 것은 문자 자체보다는 문자의 '기록성' 때문이었다. 디지털 매체 시대는 이러한 플라톤의 우려를 다시금 상기시킨다. 디지털 하드웨어의 발달과 인터넷의 구축 및 확산은 '디지털 치매'라고 일컬어지는 급격한 기억 능력 저하와 기억 의지 소멸을 더욱 부추기는 방향으로 진행되고 있다. 인터넷은 검색이 쉽고, '클라우드^{cloud}'와 같은 무한(할 것이라고 믿는) 저장 가능한 상황을 만들어 '기억'을 '자료 올리기로', '사유'를 '자료 찾기'로 대체해 버린 듯하다.

이진 신호로 바뀐 디지털 기억(기록)에 우리가 더 의존하게 된 것은 열화나 손실이 없이 저장하고 전달할 수 있기 때문이다. 이러한 이점은 새로운 기록의 대부분을 디지털 방식으로 하게 할 뿐만 아니라, 이전의 아날로그 기록물도 디지털 방식으로 변환시키는 상황을 만들고 있다. 손실이 있던 아날로그 방식이 축소되고 있는 것이다. 그리고 이렇게 기록·저장된 기억들은 빅데이터로 축적되어 인류의 '기억-역사'를 구성하고 있다. 모든 과거의 기록이 0과 1의 소용돌이에 빨려 들어가 빅데이터로 쌓이고 있다고 해도 과언이 아니다. 낡아 해지거나 바래지 않는 디지털 '기억-역사'에 대한 굳건한 신뢰와 편리성을 지향하는 인간의 속성은 '기억-역사'를 디지털 외재화함으로써 많은 기억을 0과 1에 내어주고 있다.

디지털 기반의 AI는 인터넷 통신의 발전과 빅데이터 형성에 따라 크게 진화했다. 인터넷이라는 데이터 핏줄은 개별적이었던 컴퓨터를 하나의 단일한 컴퓨터 체계로 묶고 있으며, 이 핏줄을 통해 수집되고 축적된 빅데이터는 기계 학습 알고리즘에 의해 더욱 개선되고 진화된 AI를 가능케 한다. 여기에서 영화 <터미네이터>에 등장한 모든 것을 통제하는 기계 지능 '스카이넷^{Skynet}'의 검은 실루엣이 보인다. 동시에 인공지능 운영체제(사만다)와 사랑에 빠진 남자(시어도어) 이야기를 그린 영화 <her>(2013)의 체취도 맡게 된다. 인터넷에 접속되어 있는 AI가 무서운 것은 개별 컴퓨터라도 접속되는 순간 바로 하나의 단일한 AI가 될 수 있기 때문이다. <her>

의 결말에서 본인과만 일대일로 사랑을 나누고 있을 것이라는 주인공의 믿음은 AI 운영체제가 수천 명의 인간과 동시에 사랑(교류)을 나누고 있었다는 사실로 무너져 내린다. 이것은 모든 개별 AI가 단일한 하나의 개체일 수도 있다는 점과, 능력과 감성이 인간과는 다를 수 있다는 점을 암시한다. 우리는 AI가 인간의 어떤 패턴을 최대한 이용하기 때문에 선천적인 인간의 특징과 닮았을 것이라는 착각한다. 하지만 AI는 인간적이면서 전혀 인간적이지 않은 이율배반적인 특징을 가지고 있음을 기억해야 할 것이다.

인간이 디지털 언어인 기계어를 번역 프로그램을 통하지 않고는 거의 해석하지 못하는 상황과, (분절적 텍스트와 비분절적 텍스트를 가리지 않고) 대부분의 '기억-역사'가 기계어로 축적·저장되는 상황, 그리고 AI가 인터넷에 의해 하나의 단일 체제를 형성할 가능성은 나쁜 SF 가상 시나리오를 쓰기에 좋은 조건이다. 이미 악의적인 인간 사용자들은 컴퓨터 시스템에 대한 접근을 제한시키는 악성코드^{malware}인 랜섬웨어^{ransomware}를 심어 인터넷 사용자에게 금품을 요구하고 있다. 또한, 전 세계 컴퓨터 4만 5000대를 감염시킨 스텝스넷^{Stuxnet} 바이러스(2009년 말)처럼 바이러스를 퍼트리 컴퓨터 사용을 어렵게 하기도 한다. 비의도적인 실수가 문제를 일으키기도 하는데, 그 대표적 예로, 2009년 10월 중순 스웨덴의 웹사이트 불능 사태를 들 수 있다. 스웨덴 국가 도메인 .se를 업데이트하던 중, 잘못 배열된 스크립트가 모든 .se 도메인 이름을 작동하지 않게 하여 스웨덴의 웹사이트나 이메일을 장시간 불능 상태로 만들었다.⁷⁾ 인간이 의도적으로 혹은 실수로 기계어의 번역 불가능성을 만들었을 때도 이렇게 피해를 발생시키는데, 만약 인터넷 네트워크로 단일 체제를 갖춘 AI가 어떠한 판단으로 인간이 축적한 디지털 '기억-역사'에 접근할 수 없도록 한다면 인류는 어떻게 될 것인가? AI가 번역 불가능한 알고리즘을 만들어 기계어로 된 빅데이터를 '블랙박스' 속에 넣어버린다면 우리는 어떻게 해야 할까? 손실 없는 저장과 전달 할 수 있는 디지털 기억-역사에 대한 신뢰와 편리는 대재앙이 되어 과거의 많은 부분을 잃게 할지도 모른다.

생각하는 기계에 대한 우려는 이미 1940년대부터 있어 왔다. 아이작 아시모프^{Isaac Asimov}는 1942년에 『위험에 빠진 로봇^{Runnarouid}』에서 '로봇공학 3원칙'⁸⁾로 제시했었다. 이후 컴퓨터 과학자 로빈 머피^{Robin R. Murphy}와 인공지능 학자 데이비드 우즈^{David D. Woods}가 '책임 있는 로봇공학

의 대안적 3원칙'(2009년)을, 영국의 공학·물리과학연구위원회^{EPSRC}와 인문예술연구회^{AHRC}가 '로봇공학자를 위한 5가지 윤리'(2010년)를 제시하는 등 생각하는 기계에 대한 원칙을 세우고자 했다. 하지만 원칙에 벗어날 수 있는 예외가 늘 존재하기 때문에 여전히 불완전한 상황이다. (로봇의 두뇌가 AI로 작동한다고 할 때, 로봇공학의 원칙은 AI공학의 원칙으로 직결된다.) 이런 상황에서 생각하는 기계가 있는 미래는 안개 속 일 수밖에 없다. 나쁜 가상 시나리오에서 디지털 '기억-역사'는 신기루처럼 사라지고, 바벨탑처럼 무너진다.

5. 신기루처럼, 바벨탑처럼

우리에게 말하지. 미치광이, 공상가라고
하지만, 우울한 구속에서 벗어나서,
세월이 지나면 사색가의 노련한 두뇌는
인공적으로 사색가를 창조하게 될 걸
- 피테, 『파우스트』, 2장 (번역 B. 파스테르나크)

'미디어는 메시지다'로 유명한 맥루언은 컴퓨터(디지털)가 성령 강림 상태, 즉 세계적 이해와 통일이 기술에 의해서 이룩된 상태를 약속했다고 보았다.⁹⁾ 그는 모든 언어가 프로그래밍 언어로 통일되어 하늘 꼭대기까지 도달하는 바벨탑을 이룩할 것이라 믿었다. 하지만 키틀러는 평범한 언어(정현파, 아날로그)가 차지하고 있던 자리를 프로그래밍 언어(이진 신호, 디지털)가 잠식하여 자신이 세운 새로운 위계 속에서 성장한다고 말했다. 그래서 키틀러는 프로그래밍 언어를 포스트모던적 바벨탑으로 명명했다. 디지털 언어 체계는 우리의 기록을 우리의 지각 밖에서 구성한다. 우리는 이 과정에서 무엇이 행해지고 어떻게 기록되는지 알 수 없다. 인간을 기록으로부터 아예 소외시켜 버리는 것이다.

성경의 바벨탑 사건은 단일 체제가 불러올 수 있는 욕망과 좌절을 보여준다. 바벨탑을 쌓기 이전의 인간은 단일 언어의 사용으로 쉽게 연합하고, 높은 작업 능력을 가지고 있었다. 능률 높은 단일 체제를 형성한 인간은 하늘까지 닿고자 하는 욕망에 휩싸여 바벨탑을 쌓으려 했다. 모든 것을 수렴하여 단일 언어인 기계어로 빅데이터를 쌓고 있는 디지털의 면모가 떠오르는 대목이다. 익히 우리가 알고 있듯이 인간의 욕망은 단일 언어를 와해시킨 하나님의 행동으로 무너져 내린다. 인간 간의 소통 불능

상태가 바벨탑을 쌓으려는 욕망을 붕괴시켰다.

사유할 수 있는 기계, 판단할 수 있는 기계를 생각해본다. 영화 <her>에서처럼 AI는 즐거운 친구도, 애인도 될 수도 있다. 자신의 고민과 속내를 들어주는 상담사가, 가장 합리적인 해결책을 제시해주는 조언자가 될지도 모른다. 하지만 AI의 사유 과정이, 그 판단 과정이 인간과 다를 것은 분명해 보인다. 인간은 기계어를 직접 해석하는 데 엄청난 난관에 부딪치지만, AI는 인간의 언어를 자체적으로 해석하는데 큰 어려움이 없(을 것이다). 사람과 사람 사이의 속삭임을 들을 수 있는(음성인식) 'AI', 하지만 AI와 AI 사이의 (이진 신호) 커뮤니케이션을 알 수 없는 '인간'. 그들(AI)은 우리의 모든 것을 알지만 우리는 그들의 생각을 알 수 없다. 마치 알파고의 특이한 수를 프로바둑기사가 해석할 수 없었듯이 말이다. (AI가 네트워크로 단일 체제를 구축할 것이라고 예상할 때 '그들'보다는 '그'라는 1인칭이 맞을 것 같다.)

모든 기록이 0(off)과 1(on)의 기계어로 저장되고, 이전의 아날로그 기록도 기계어로 변환하고 있는 상황에서 인간이 기계어를 사용하지 못하는 순간이 온다면 우리는 어떻게 될 것인가? 인간과 디지털의 언어적 이중성과 (인간 입장에서) AI의 잘못된 판단이 결합하여 그동안 인간이 쌓아왔던 모든 디지털 '기억-역사'에 접근하지 못하게 된다면 우리는 어떻게 될까? 그 때 AI에 대한 기대는 신기루였음을 깨닫게 될까? 디지털 제단이 바벨탑처럼 무너지는 것을 경험하게 될까? AI의 발달에서 인간과 디지털의 언어의 이중성에 대한 고려는 분명 필요해 보인다.

참고문헌

- 1) James Barrat, *Our Final Invention: Artificial Intelligence and End of the Human Era*, New York: Thomas Dunne, 2013, pp.196~197.
- 2) 천체물리학에서 블랙홀 주변에서 정상적인 물리 법칙이 적용되지 않기 시작하는 지점을 말하는 것으로, 미래에 기술 변화의 속도가 급속히 변해 인간의 생활이 되돌릴 수 없도록 변화되는 지점을 의미한다.
- 3) Friedrich Kittler, *Gramophone, Film, Typewriter*, tr. Geoffrey Winthrop-Young and Michael Wutz, Stanford university press, 1999, p.2
- 4) 위의 책, p.243.
- 5) 유리 로트만, 『기호계』, 김수환 역, 문학과지성사, 2008, pp.206~207.

- 6) 플라톤, 『파이드로스』; 알렉스 라이트, 『분류의 역사』, 김익현, 김지연 공역, 디지털미디어리서치, 2010, p.15. 재인용.
- 7) 당시 소문으로는 수백만 명의 중국인과 일본인이 스웨덴 어딘가에 있는 신비로운 레즈비언 마을 차코폴(Chako Paul)을 검색하면서 스웨덴의 인터넷이 다운되었다고 한다.
- 8) 제1원칙 : 로봇은 인간을 다치게 해서는 안 되며, 아무것도 하지 않으므로써 인간이 다치도록 놔두어도 안 된다.
제2원칙 : 로봇은 제1원칙과 상충하지 않는 한, 인간이 내리는 지시에 복종해야 한다.
제3원칙 : 제1원칙, 제2원칙과 상충하지 않는 한, 로봇은 자신을 보호해야 한다.
아시모프는 1985년 『로봇과 제국』에서 다음과 같은 제0원칙을 추가하여 3원칙을 수정한다.
제0원칙 : 로봇은 인류를 해쳐서는 안 되며, 아무것도 하지 않으므로써 인류가 해를 입게 놔두어도 안 된다.
- 9) 마셜 맥루언, 『미디어의 이해: 인간의 확장』, 김상호 역, 커뮤니케이션북스, 2011, p.165.

저자 소개

● 안 진 국(Lev Aan)



- 2015 조선일보 신춘문예 미술평론 당선, 평론 등단
- 현재 : 미술비평 전문지 『컨템포러리아트저널』 필진
- 종합 인문주의 정치 비평지 『말과말』 편집위원 월간 『BiZart』 고정 필진
- 한국미술평론가협회 회원

▪ 양주시립미술창작스튜디오 '777레지던스' 입주작가

<관심분야> : 디지털, 인공지능, 뉴미디어 아트, 매체미학