

인공지능의 윤리1)

성기진

들어가며

구글이 개발한 바둑 프로그램 “알파고”가 이세돌 9단과의 대국에서 4승 1패라는 놀라운 성적을 거두자 우리나라에서도 인공지능에 대한 관심이 높아지고 있다. 체스에서는 이미 컴퓨터가 인간을 넘어선 지 오래 되었지만 바둑은 컴퓨터가 감당하지 못할 거라고 여겨져 왔는데, 그것마저 알파고가 너무나 손쉽게 제패해 버리자 이제는 컴퓨터가 사람보다 부족한 점이 없게 되었다는 센세이셔널한 해석이 언론을 통해 널리 퍼지게 되었다. 인간 지능의 자존심이 깎인 셈이다.

알파고의 승리는 대국이 벌어진 한국이나 바둑 애호가 많은 동아시아뿐 아니라 전세계적으로 대단한 이슈가 되었다. 실리콘 벨리의 핫 토픽이 제일 먼저 올라오는 <해커 뉴스> 커뮤니티에서도 바둑의 ‘바’ 자도 모르던 파란 눈의 프로그래머들이 가치망(value network: 자신이 추구하는 결과를 얻을 확률을 계산하는 방법)이나 몬테 카를로 트리 검색(Monte Carlo tree search: 모든 경우의 수를 일일이 계산하지 않고 일부를 무작위로 추출하여 시간을 절약하는 기법) 등 인공지능 분야의 전문적인 주제에 대해 며칠이고 밤새도록 갑론을박을 했다.

그동안 사람이 지켜온 만물의 영장의 자리를 컴퓨터에게 빼앗겼다는 위기감도 없지 않았지만, 컴퓨터를 마음대로 부릴 줄 아는 프로그래머들은 묘한 성취감에 들뜨기도 했다. 사람보다 나은 컴퓨터가 내 명령에 복종한다면 나아말로 특별한 사람이 아니겠는가? 바둑에서는 졌어도 교만에 있어서는 여전히 사람에 비할 것이 없나 보다.

다양한 직군을 인공지능이 대체하게 될 거라는 관측도 부쩍 늘어났다. 인공지능이 의사가 된다면? 작곡을 한다면? 판사가 된다면? (유독 법조인을 대체하자는 얘기에 이목이 집중된 것은 법을 만들고 집행하는 사람들에 대한 신뢰가 땅에 떨어지다 못해 지하로 파고들어 가는 우리 사회의 현실을 반영한 것 같다는 생각이 들어 씩씩하기도 하다.) 물론 한 치의 불확실성도 없는 완벽한 수학적 게임인 바둑을 이제 겨우 마스터한 인공지능에게 인류 최고의 지혜와 도덕적, 정치적, 철학적 의도로 가득차 있는 법이나 예술의 해석을 맡기는 일은 아직도 먼 미래의 일일지도 모른다.

1) 이 글은 필자가 2016년 9월 『개혁신앙』에 투고했던 원고와 2016년 10월 기운실 기독교윤리실천학교 및 2017년 1월 라브리 기독교세계관학교에서 발표했던 내용을 일부 수정한 것이다.



이 문서는 라브리공동체 웹사이트 자료실에서 내려받은 것으로, 저작권은 글쓴이에게 있으며 개인 및 그룹 공부를 위해서만 인쇄 및 소량 배포할 수 있습니다. 내용은 라브리의 공식 입장과 일치하지 않을 수 있습니다.

This document was downloaded from the online library of L'Abri Fellowship Korea. All copyright belongs to the author. Printing, copying and small-scale distribution are permitted only for personal and group study. L'Abri does not endorse the contents of this document.

라브리공동체 / 강원도 양양군 서면 구룡령로 3025
☎ (+82) 033-673-0037 ✉ yangyang@labri.kr

www.labri.kr

이미 가까이 다가온 AI

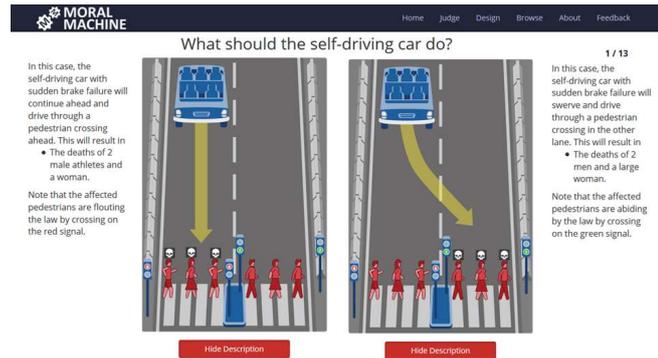
그러나 다른 한편, 인공지능의 판단에 개인과 사회의 운명을 맡기는 것은 우리의 생각보다 훨씬 가까이 다가온 일이기도 하다. 목적지만 입력하면 알아서 운전할 수 있는 자율주행 자동차가 외국에서는 이미 길거리를 누비고 있고, 국내에서도 활발하게 연구가 진행되고 있기 때문이다. 모든 자율주행 자동차는 사고를 회피하기 위한 알고리즘을 갖추고 있다. 그러나 만약 사고를 피할 시간이 없다면 어떻게 될까? 갑자기 차도로 뛰쳐나온 아이를 치거나, 핸들을 꺾어 맞은편에서 오는 차와 정면충돌하거나, 둘 중 하나를 선택해야만 한다면? 무엇을 선택하든 자율주행 자동차는 누군가에게 사형 선고를 내릴 수밖에 없다. 수술대나 재판정에 앉지 않았을 뿐, 의사나 판사 못지않게 중요한 결정권을 이미 가진 것이나 다름없다.

굳이 인공지능이라는 이름을 붙이지 않더라도 IT 기술이 우리 삶에 미치는 막대한 영향력을 생각하면 그러한 영향력에 상응하는 도덕적 책임을 생각해 볼 필요가 있다. 가깝게는 친구들과 사이의 관계를 깨뜨릴 수 있는 카카오톡의 “자동 친구 추천” 알고리즘부터, 멀게는 대선후보의 지지율에까지 영향을 미친다는 인터넷 포털 사이트의 뉴스 배치 방식까지…… 그런 기술을 개발한 기업이나 프로그래머들이 의도했던 의도하지 않았든, 우리는 매일 일상생활 속에서 컴퓨터의 판단에 따라 행동하고 있기 때문이다.

인공지능 연구자들도 이것을 잘 알고 있다. 특히 앞에서 언급한 자율주행 자동차 같은 경우에는 천문학적 금액이 걸린 집단 소송에 휘말릴 가능성도 있기 때문에, 인공지능에게 올바른 윤리 의식을 심어 주어야 한다는 주장에 제

조사들도 귀를 기울일 수밖에 없다. 얼마 전 MIT에서는 “윤리적 기계”(The Moral Machine)라는 웹사이트를 개설하여 방문자들에게 다양한 사고 상황을 제시하고 회피 방향을 선택하도록 했다. 자율주행 자동차 개발에 참고하려는 듯 하다. 승객과 보행자 중 누구를 우선 보호해야 하는지, 사망자의 성별과 연령, 직업 등이 보통 사람들의 판단 기준에 어떤 영향을 미치는지, 그리고 이런 정보를 종합하여 컴퓨터에게 가르칠 만한 보편적인 원칙을 찾을 수 있는지 등을 연구한다고 한다.

그러나 인공지능을 연구하는 대부분의 학자들은 기독교 윤리와는 동떨어진 세상에 살고 있기 때문에, 도덕적 판단력을 갖춘 인공지능을 만들더라도 세상의 타락한 기준을 그대로 따를 가능성이 높다. 이것은 단지 낙태를 허용할 것인가, 안락사를 허용할 것인가 등 기독교 윤리에서 흔히 다루는 세부적인 이슈에 그치지 않는다. 인간의 본질은 무엇이고 윤리의 기초는 무엇인지, 권리와 책임은 어떻게 나누어야 하는지 등 가장 근본적인 레벨에서부터 현대의 인공지능 연구는 인본주의적인 발상으로 가득차 있기 때문이다.



MIT “윤리적 기계” 실험 화면

알파고의 가장 중요한 시사점은 인간이 나오냐, 컴퓨터가 나오냐의 문제가 아니다. 그런 질문이 이슈가 된다는 것 자체가 비기독교적 세계관의 산물이며, 세상을 지배하고 있는 유물론적 사고방식이 얼마나 많은 부작용을 낳는지를 증거하기 때문이다. 이 글에서는 현대의 학자들, IT 업계 종사자들, 언론과 일반인들이 인공지능에 대해 이야기할 때 자주 등장하는 몇 가지 문제점을 지적하고, 이런 논의를 접할 때 기독교인이 가져야 할 자세를 살펴보려고 한다.

연산 능력과 인간성의 기준

갑자기 이슈로 떠오른 인공지능 논의를 바라보면서 발견하는 가장 큰 문제점은 컴퓨터가 사람과 같은 수준에 도달했다거나 곧 사람을 능가하게 될 거라는 막연한 위기감이 아니다. 컴퓨터에게 따라잡힐 만큼 “인간성”의 기준이 추락했다는 것이 더 근본적인 문제이다. 현대 사회를 지배하는 유물론적 세계관은 사람을 단지 두뇌라는 컴퓨터에 연결된 고깃덩어리로 취급한다. 바둑이나 체스와 마찬가지로, 인격도 사랑도 감정도 신앙도 모두 두뇌가 실행하는 어플리케이션일 뿐이다. 사람보다 더 뛰어난 연산 능력(computational power)을 가진 연산 기계(computer)가 등장하고, 그 기계에서 실행할 “사람 시뮬레이션” 어플만 개발하면 얼마든지 해당 영역에서 컴퓨터가 사람을 대체할 수 있다는 생각이다.

2001년에 개봉한 스피버그 감독의 <A.I.>라는 영화의 주인공은 사람처럼 사랑할 수 있는 기능을 가진 로봇이다. 이런 설정은 한편으로는 지금까지의 인공지능이 갖추지 못한 것을 지적하기도 하지만, 다른 한편으로는 “사랑”이라는 기능을 만들어 넣기만 하면 로봇도 사람처럼 사랑할 수 있지 않을까, 하는 기대를 품도록 만든다. 만약 그것이 가능하다면 사랑이라는 영역에서 컴퓨터와 사람은 동급이 될 수 있다는 가정이 깔려 있는 것이다.



인공지능을 가진 로봇을 다룬 영화들

그러나 이것은 학생의 수준에 맞춰 채점 기준을 바꾸는 속임수에 불과하다. 의학과 심리학의 발달에도 불구하고 아직 사람의 사고 구조를 정확히 알 수는 없으니, 그나마 측정하기 쉬운 “연산 능력”을 “사람만큼 똑똑한” 기준으로 삼은 것이다. 이것은 다른 동물들과 인간의 차이를 단지 지능의 차이로 보는 현대 과학의 기본 전제와 일맥상통한다. 칠판지가 간단한 도구를 쓸 줄

안다고 “오, 우리랑 비슷한데?” 하던 사람들이 이번에는 알파고가 이세돌만큼 빠른 시간 안에 최선의 수를 찾아내는 것을 보고 “오, 우리랑 비슷한데?” 하고 있는 것이다.

성경이 말하는 인간성의 기준은 연산 능력과 무관하다. 사람이 특별한 것은 인격이나 사랑, 신앙이라는 어플리케이션이 우리의 뇌에 기본으로 설치되어 있기 때문도 아니다. 사랑이 설치되지 않은 싸이코패스나 신앙이 다운된 무신론자라도 당신의 형상대로 지음받았다는 이유만으로 하나님은 사랑하시고, 그들의 병들을 안타까워하신다. 반면, 아무리 사람보다 뛰어난 연산 능력을 가지고 사람보다 더 감쪽같이 인격이나 사랑을 시뮬레이션할 수 있는 로봇이라도 하나님은 사람처럼 사랑하지 않으시고, 그들을 구원하기 위해 아들을 보내시지도 않는다.

체스나 바둑에서 컴퓨터가 사람을 이겼다고 호들갑을 떠는 것은 마라톤 경기에서 자동차가 사람을 이겼다고 신기하게 여기는 것만큼이나 불필요한 일이다. 자동차만큼 빨리 달릴 수 없다고 해서 사람의 가치가 결코 떨어지지 않듯, 알파고만큼 계산을 잘 할 수 없다고 해서 사람의 가치가 손상을 입는 것도 아니다. 연산 능력을 가치 판단의 척도로 삼는 세상의 세계관에 양보하지 않는다면 인공지능의 발달은 전혀 사람의 위기가 아닌 것이다. 우리가 서로와 하나님을 사랑하는 것은 단지 생존경쟁에서 살아남기 위한 시뮬레이션이나 우리의 뇌에 설치된 하나의 어플이 아니라 진정한 인격적 교제이며 우리의 존재 목적임을 분명히 기억해야 한다.

트랜스휴머니즘과 수학적 윤리

연산 능력을 인간성, 또는 사람처럼 취급받을 권리를 부여하는 기준으로 삼게 되면 윤리에 있어서도 반드시 문제가 생긴다. 사람과 비슷한 연산 능력을 가진 존재들의 이익을 극대화하기 위해 또다른 연산을 하고, 그러한 연산 결과를 그대로 시행하는 것이 모든 도덕률의 궁극적인 목적이 되기 때문이다.

알파고를 비롯한 현대의 인공지능 프로그램들은 가치망(value network)이라는 것을 갖고 있다. 어떤 가치를 추구하도록 설정하느냐에 따라 인공지능은 전혀 다른 행동을 할 수도 있다. 예를 들어 최단시간에 시합을 마치는 것을 궁극적인 가치로 삼는 프로그램이라면 패배의 위험을 무릅쓰고라도 무리한 도박을 하겠지만, 시간이 걸리더라도 기어코 이기는 것을 우선순위에 두도록 설정한다면 알파고처럼 안전을 추구하는 플레이를 보여줄 것이다.

실제로 알파고는 최악의 경우에도 큰 손실이 발생하지 않도록 대비하는 미니맥스(minimax) 알고리즘을 사용하는데, 이것은 사회 최하층민에게 손실이 발생하지 않도록 해야 한다는 미국 최고의 정치사상가 존 롤즈(John Rawls)의 이론과 매우 흡사하다. 우연이 아니다. 인공지능 연구는 오래 전부터 정치학과 경제학, 민주주의와 자본주의 이론과 밀접한 관계를 가져 왔기 때문이다. 이기기 위해 어떤 수를 두어야 하는지 계산하는 알파고의 가치망은 행복하게 살기 위해, 또는 더 많은 돈을 벌기 위해 궁리하는 사람의 가치망과 크게 다르지 않다.

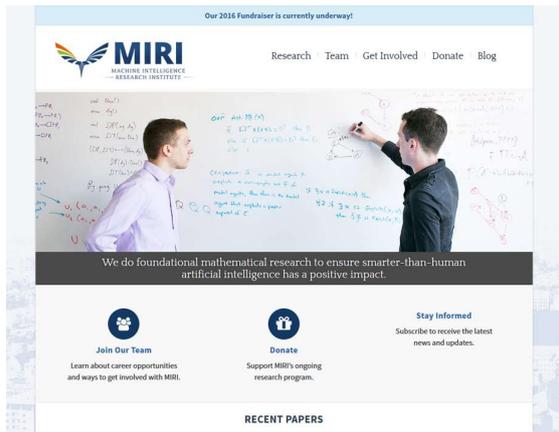
그러나 컴퓨터란 결국 숫자를 다루는 기계이므로, 사람이 중요하게 여기는 가치들을 컴퓨터에 입력하려면 모두 숫자로 환산할 수밖에 없다. 모든 규칙은 프로그래밍 언어로 표현할 수 있어야 하고, 여러 규칙이 서로 충돌한다면 각각에 가중치를 부여하여 상쇄시키도록 프로그램을 짜야 한다. 인공지능에게 법을 맡기려 한다면 가장 먼저 걸리는 부분도 바로 이것이다. 사람의 복잡

한 법적, 도덕적 기준들을 컴퓨터에게 어떻게 입력해야 할 것인가?

컴퓨터의 막대하고 객관적인 연산 능력이 사람의 윤리적 딜레마나 사회, 정치적 문제들에 답을 줄 수 있을 것이라고 믿는 사람들이 공리주의(utilitarianism)의 함정에 빠지기 쉬운 것은 이때문이다. 공리주의는 어떤 선택이 가져올 모든 행복과 모든 불행을 숫자로 환산한 후, 행복이 불행보다 크면 옳은 선택이라고 판단하는 이론으로, 행복과 불행을 숫자로 환산하는 기준이 자의적이고 다수의 행복을 위해 소수의 행복을 희생시킨다는 이유로 근대 이래 많은 비판을 받고 있다. 그러나 인공지능에게 법적, 도덕적 판단을 맡기려면 공리주의만큼 편리한 것이 없다. 행복을 플러스로, 불행을 마이너스로 입력한 후, 합이 가장 큰 옵션을 선택하면 그만이다. 위에서 언급한 몰즈의 정치철학도 마찬가지로, 컴퓨터로 처리하기에 비교적 편리한 알고리즘이기 때문에 학계의 여러 가지 비판에도 불구하고 주목받고 있다. 사람이 보기에든 하자가 많은 이론을 컴퓨터의 한계 때문에 어쩔 수 없이 채택하게 되는 것이다.

인공지능으로 사람의 한계를 초월할 수 있다고 믿는 트랜스휴머니즘(transhumanism, 초인간주의)은 이런 문제를 극단적으로 보여 준다. 컴퓨터와 친숙한 실리콘 밸리의 오피니언 리더들 가운데는 트랜스휴머니즘을 신봉하는 사람이 적지 않은데, 이 이론의 핵심은 인공지능이 자기보다 더 나은 인공지능을 스스로 설계할 수 있게 될 것이고, 따라서 사람이 설계할 때에는 비교할

수도 없는 폭발적인 성장이 일어나는 “기술적 특이점”(technological singularity)이 멀지 않았다는 주장이다. 트랜스휴머니스트들에게 기술적 특이점은 메시야의 재림에 필적하는 중요한 사건으로, 심지어 이것이 몇 년 안에 일어날 거라고 믿는 사람도 있다. 이 때가 되면 사람의 연약한 판단력은 내려놓고 인공지능의 우수한 판단에 세상을 맡기거나, 인공지능을 사람의 두뇌에 접목하여 사람의 판단력을 향상시켜야 한다고 그들은 주장한다.



기술적 특이점이 멀지 않았다는 가정 아래 인간보다 똑똑해진 인공지능이 인간에게 유익을 줄 수 있도록 하는 다양한 방법을 연구하는 기계지능 연구소(MIRI)의 홈페이지

숫자밖에 취급할 수 없는 컴퓨터의 한계를 오히려 강점으로 내세우고, 그렇지 않은 사람의 사고 패턴을 하등한 것으로 취급하는 사람도 있다. 트랜스휴머니즘을 신봉하는 일부 대

중강연자들은 통계학에서 확률 계산에 사용하는 베이즈 정리(Bayes' theorem)를 이상적인 것으로 내세우고, 대중 강연과 저서를 통해 베이즈 정리에 부합하지 않는 인간의 사고방식을 비판하곤 한다. 컴퓨터가 처리하기에 가장 편리한 사고 구조를 오히려 사람에게 강요하는 것이다. 이거야 원, 알파고가 바둑을 두기 쉽도록 바둑의 규칙마저 바꾸어야 할 판이다.

인공지능을 신격화하는 트랜스휴머니즘은 하나의 종교와도 같다. 기술적 특이점이 메시야의 도래와 같다면 그 후 인공지능의 막강한 지능으로 인간 사회의 거의 모든 문제를 해결할 수 있을 것이라는 신념은 일종의 구원관에 해당하며, 심지어 자신의 의식을 컴퓨터에 업로드하거나 건강한 신체로 옮겨서 영생을 누릴 수 있을 거라고 진지하게 믿는 사람도 있다. SF 영화의 시

놉시스처럼 들리겠지만 <해커 뉴스> 등 실리콘 밸리 소식통에서는 이런 생각을 은근히 자주 볼 수 있다. SF 영화에서나 보던 것을 현실로 만들어낸 장본인들이니 그럴 만도 하다.

그 중에서도 가장 위험한 부류는 인공지능 연구 지원을 것을 넘어서서, 인간 사회에 산적한 다른 어떤 문제를 해결하는 것보다 인공지능 연구가 우선이라고 주장하는 사람들이다. 곧 기술적 특이점에 다다르면 다른 문제들은 모두 인공지능이 쉽게 해결해 줄 테니, 기후 변화나 공중 보건, 경제적 불평등 해소, 제3세계 구호 활동 등에 지금 시간과 자원을 투자하는 것은 낭비라는 논리이다. 수학적 윤리가 공리주의로 이어지고 트랜스휴머니즘을 종교처럼 신봉한다면 이런 궤변이 나오는 것도 무리가 아니다. 가까운 미래의 대사건을 빙자하여 사회적 약자들의 희생을 강요한다는 점에서 트랜스휴머니즘은 시한부 종말론을 연상시키기까지 한다.

그러나 인간 사회에 산적한 많은 문제들은 인간의 지능이 부족해서 생긴 것이 아니다. 성경은 우리 앞에 놓인 대부분의 문제들이 우리의 죄와 욕심으로부터 비롯되었다는 점을 분명히 알려 주고 있다. 인공지능의 막강한 연산 능력을 활용하면 최적의 임금과 세율을 계산하는 데는 도움이 될지 모르나, 디스토피아 영화에 나오는 것처럼 강압적인 방법을 동원하지 않고서야 과연 인공지능이 사람의 죄와 욕심을 다스리는 데 도움을 줄 수 있을까? 인공지능에 지나치게 큰 기대를 품는 것은 우리가 마땅히 해야 할 일을 기계에게 떠넘기고 책임을 회피하려는 태도로 이어질 수 있다. 컴퓨터는 도구일 뿐, 정말로 변화되어야 하는 것은 우리 자신이다.

인공지능과 책임 회피

연산 능력을 인격으로 포장하고 심지어 인공지능을 신처럼 숭배하는 이러한 세계관의 문제는 단지 철학적인 가능성에 그치지 않는다. 인공지능 논의를 둘러싼 철학적 사조들은 인공지능을 개발하고 사용하는 개인과 기업, 국가들의 책임 회피 수단으로 쉽게 악용되기 때문이다.

인격이란 허상에 불과하다고 주장하면서도, 대부분의 현대인은 인격을 도덕적 책임 부여의 단위로 삼는 오랜 전통을 여전히 따르고 있다. 그러나 자율주행차가 사고를 일으키거나 드론이 스스로의 판단으로 민간인을 사살한다면 누가 책임을 져야 할까? 자동차를 만드는 기업과 드론을 사용하는 국가들은 어떻게든 책임을 지지 않으려고 할 것이 뻔하다. 오류를 일으킨 컴퓨터에게 책임을 뒤집어씌우고 자신은 아무 책임도 지지 않으려는 것이다. 사람보다 더 나은 연산 능력을 가진 존재라면, 심지어 윤리적 문제에 대해 사람보다 더 잘 판단할 수 있는 존재라면, 당연히 도덕적 책임도 감당할 수 있지 않겠는가?

인공지능의 발달로 노동 시장에 대격변이 일어날 것이라는 전문가들의 예측에도 한편으로는 이러한 책임 회피의 문제가 깊이 깔려 있다. 곧 일어날 거라고 믿는 기술적 특이점 앞에서 세례 요한의 역할을 자처하는 트랜스휴머니스트들처럼, 인공지능의 정치적, 경제적 영향을 연구하는 많은 학자들은 인공지능의 발달을 필연적인 역사의 흐름으로 가정하고 거기에 대응하는 데만 급급한 것이 현실이다. “인공지능 때문에 앞으로는 이런 직군이 뜰 터이니 취업 준비에 참고하라”는 식이다.

그러나 필연적인 역사의 흐름이라는 것은 핑계일 뿐이다. 하나님이 의도했던 사람이 의도했든, 심지어 의도하지 않을 실수일 뿐이라도, 역사는 항상 누군가의 행동에 영향을 받기 때문이

다. 새로운 과학기술을 개발하는 사람이라면 혼자의 힘으로 세상을 바꾸지는 못할망정 자신의 연구가 가져올 결과에 대해 어느 정도 고민해 볼 책임이 있다. 그냥 세상에 던져 놓고 필연이라는 핑계로 자신의 책임을 회피한다면 아담이 하와에게, 하와가 뱀에게 책임을 전가한 죄를 되풀이하는 것에 지나지 않을 것이다.

인공지능과 관련된 책임 회피의 문제를 더욱 심각하게 다루어야 하는 이유는 책임을 져야 할 사람들의 대부분이 현대 사회의 엘리트와 기득권층이며, 따라서 그들의 주장에는 정치적, 경제적 이해관계가 심하게 얽혀 있기 때문이다. 대부분의 직업이 인공지능으로 대체된다면 가장 큰 손해를 보게 될 집단은 쉽게 대체할 수 있는 단순 노동자들이고, 가장 안전한 집단은 그 인공지능을 개발하고 소유한 기업과 국가, 프로그래머들이 아니겠는가?

인공지능의 발달로 직업을 잃어버릴 이들을 보호하기 위해 사회 안전망을 더욱 확대하고 심지어 전국민에게 매달 일정한 소득을 보장하자는 주장까지 심심찮게 나오고 있지만, 만약 기본 소득(basic income) 제도가 시행된다 해도 소수의 엘리트와 기계의 경제 활동에 다수의 국민들이 절대적으로 의존하는 위험한 경제 구조는 그대로 남는다. 오히려 이런 경제 구조가 정상이며 필연적이라고 인정하는 꼴이 된다. 컴퓨터의 한계를 오히려 사람에게 강요하듯, 자신들이 만들어 낸 왜곡된 사회 구조를 미화하는 것이다. 눈에 보이는 불평등은 줄어들지 몰라도 눈에 보이지 않는 권력과 책임의 양극화와 중앙집중화는 오히려 더 심화될 가능성이 높다.

4차 산업혁명에 뒤처지면 안 된다면서 최근 뉴스에 연일 보도되는 제안들도 우리는 비판적으로 받아들여야 한다. 예를 들어 국내 IT 기업들은 인공지능 연구에 빅 데이터가 필요하니 개인 정보 보호 관련 규제를 풀어 달라고 요구하고 있다. 정부와 사기업이 일반인의 인터넷 사용 패턴을 낱알이 감시하는 미국의 상황을 국내에도 빨리 도입하자는 얘기다. (특정인을 식별할 수 없도록 조치하겠다고는 하지만, 해외에서는 이런 조치를 거친 후에도 특정인을 식별할 수 있는 정보가 충분히 남아 있다는 연구 결과도 있으며 애초에 불특정 다수로부터 대량의 개인정보를 수집해야 한다는 점에는 차이가 없다.)



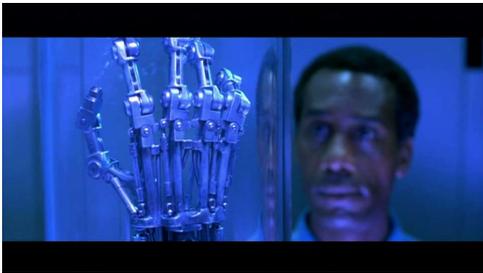
〈다크 나이트〉에서 배트맨은 폭스(모건 프리먼)에게 도시 전체를 감시할 수 있는 컴퓨터 시스템의 개발을 요청한다. 조커의 테러를 막기 위해 폭스는 이에 동의하지만 양심의 가책을 느끼고 조커를 잡은 직후 시스템을 꺼 버린다. 빅 데이터를 운운하는 기업들과 각국 정부는 과연 그만큼의 양심이라도 가지고 있는가?

인공지능에 대한 장밋빛 전망이 난무하고 있지만 실제로 인공지능이나 기계 학습(ML), 빅 데이터가 활용되는 곳을 보면 대부분 대량의 데이터를 분석하여 마케팅에 도움이 될 만한 패턴을 찾아내는 작업에 지나지 않는 것이 현실이다. 알파고를 개발한 구글을 비롯하여 페이스북, 네이버, 카카오 등 IT 기업들의 주 수입원은 광고이다. 이들이 말하는 빅 데이터의 일차적인 목표는 다수의 사용자들이 어떻게 행동하는지 철저히 분석하여 다음 행동을 예측하거나 특정 행동(상품 구입)을 하도록 유도(광고)하는 것이다. 사용자들을 마음놓고 감시하고 싶다고 말하기는 뽕줍하니까, 국가 경쟁력을

운운하며 규제 완화를 요구하는 것으로 보인다. 스노든의 폭로와 한국의 테러방지법 논란 등 정치적인 컨텍스트, 그리고 끊이지 않는 개인정보 유출 사고로 많은 사람들이 사생활 보호에 신경 쓰고 있다는 점에 대한 배려나 문제의식은 눈 씻고 봐도 없다.

하루가 다르게 변화하는 과학기술을 일반인은 따라가기조차 벅찬 것이 사실이다. 정부 정책도 늘 뒷북을 칠 수밖에 없다. 그러나 그렇다고 IT 전문가들과 오피니언 리더들의 주장을 허겁지겁 받아들이기만 해서는 곤란하다. 인공지능 연구에는 정치적, 상업적 이해관계가 매우 복잡하게 얽혀 있기 때문이다. 인공지능이 많은 사람에게 도움이 되고 더 자유롭고 평등한 세상을 만드는 데 이바지하기 위해서는 이러한 이해관계와 그 밑에 깔려 있는 세계관을 예리하게 파악하고 이를 견제할 수 있는 사회적 공감대가 만들어져야 한다.

〈터미네이터〉, 〈I, 로봇〉 등 인공지능이 사람을 억압한다는 내용의 영화들이 말해 주지 않는 것은 그 인공지능도 결국 사람이 만든 기계라는 사실이다. 의도했던 실수였든, 기계가 사람에게



〈터미네이터〉 2편의 다이스 박사는 자신이 개발한 기술이 인류를 위협할 것이라는 예언을 듣고 고민하다가 그 기술을 파괴하는 데 함께하기로 결심한다.

반기를 드는 “버그”가 발생할 여지를 남겨둔 책임은 사람 프로그래머에게, 그리고 그런 기계를 무비판적으로 사용한 기업과 국가에게 돌아가야 하는 것이다. 안타깝게도 최근의 IT 업계에서는 〈I, 로봇〉의 원작자인 아이작 아시모프(Isaac Asimov)나 〈터미네이터〉 2편의 다이스 박사처럼 인공지능 발달의 결과에 대한 깊은 책임감을 가진 사람을 찾기가 힘들다.

최근 미국에서 테슬라의 자율주행차가 대형트럭과 충돌하여 운전자가 사망한 사건이 있었다. 유감을 표명하면서도 “자율주행 기능은 아직 베타 버전이므로 버그가 있을 수 있다”는 단서를 붙인 테슬라의 공식 발표는 책임감의 상실을 적나라하게 보여 준다. 테슬라의 주장에 따르면 미국에서는 평균 주행거리 9천 4백만 마일마다 한 명씩 사망하지만 자사의 자율주행차는 누적 주행거리 1억 3천만 마일에 한 명이 사망했으므로, 비록 버그가 있더라도 일반 승용차보다 더 안전하다고 한다. 책임을 질 수 없는 컴퓨터와 통계자료 뒤에 숨어, 사람으로서 당연히 감당해야 할 책임을 회피하려는 일이 벌써 일어나고 있는 것이다.

결론

알파고를 개발한 구글이 세계적인 화제가 되자, IT 업계에서 둘째 가라면 서러운 마이크로소프트에서도 지난 3월 “테이(Tay)”라는 인공지능을 만들어 트위터에 풀어놓았다. 테이는 10대 소녀의 성격을 시뮬레이션하여 트위터 사용자들과 대화를 주고받고, 그 대화를 바탕으로 더 자연스럽게 대화하는 법을 익혀 나가는 프로그램이었는데, 나치를 찬양하고 여성을 성노리개로 취급하는 발언을 일삼는 바람에 16시간 만에 강제 종료되고 말았다. 현실의 아이들을 뺄어버리게 만드는 악플러들이 인공지능마저 비행청소년으로 바꿔 버린 것이다.

인공지능 논의에서 가장 중요한 것은 “지능”이 아니라 “인공”이다. 아무리 초당 수억 번, 수조 번의 연산을 할 수 있는 기계라도 사랑이 없으면 실리콘 조각에 지나지 않고, 처음부터 올바

른 윤리를 입력하지 않으면 언제든지 가공할 살인 무기로 돌변할 수 있기 때문이다. 인공지능이 사람의 윤리적 판단을 흉내낼 수는 있겠지만, 보고 배운 것이 비기독교적 세계관뿐이라면 어떤 판단을 흉내낼지는 불을 보듯 뻔하다. 테이의 실패는 이 사실을 적나라하게 보여 준다. 어떻게 보면 기계학습(machine learning)도 아이를 가르치는 것과 크게 다르지 않다. 주위 사람들이 먼저 올바른 생각을 갖고 있어야 그것을 보고 배우지 않겠는가?

숫자로 환산되는 “지능”에 대한 집착은 인간성과 도덕적 책임을 잊도록 만들지만, 모든 컴퓨터는 사람이 만든 것이고 사람의 명령을 기계적으로 따를 뿐이라는 간단한 사실을 기억하면 인간성의 본질과 도덕적 책임의 소재를 엉뚱한 곳에서 찾거나, 그것을 인공지능에게 아무 생각 없이 가르치는 위험천만한 오류를 피할 수 있을 것이다.



테이(Tay)의 트위터 프로필 화면