# 기계가 만든 소프트웨어가 가져오는 불평등과 데이터 과학의 역할

이광춘

KPMG Korea 상무

1. 인간과 기계의 경쟁관계와 불평등의 발생

기계(Machine)라 하면 기계장치를 떠올릴 수 있지만, 영어로 머신(Machine)은 인공지능을 탑재한 컴퓨터를 의미한다. 사람들이 하는 작업의 자동화 수준을 여러 단계로 나눌 수 있는데, 기계의 도움 없이 모든 결정과 행동을 사람이 취하는 수준부터 인간을 배제하고 기계가 모든 의사결정을 내리고 자율적으로 운전, 판결, 세금계산 등 완전한 자동화 수준이 있다 [1].  이렇게 작업 자동화 수준을 구분할 수 있는 이유는 과거 인간과 기계가 서로 잘하는 영역에 뚜렷한 차이가 있어 가능했으나 급격한 자동화 기술의 발전으로 이러한 구분도 의미가 없어지고 있다.

생산성과 임금격차, 보울리의 법칙, 노동인력 참여율을 통해 확인되는 공통된 사항은 1980년 이후 일자리에 구조적인 변화가 생겼다는 점이다[2].  과거 일자리와 관련하여 국가 내에 정규직과 비정규직 프레임 혹은 외국인 노동자로 대표되는 국외노동자와 국내 일자리 프레임에 추가하여 사람과 기계 프레임[[1]](#footnote-2)이 추가되었다.

반명에 소프트웨어를 만들기 때문에,인간이 만든 소프트웨어보다 만든가한 바와 같이현재 하고 있다.[[2]](#footnote-3)

2. 데이터 과학의 시대

과거에는 사람이 직접 설계한 소프트웨어를 연구실에서 꺼내 실생활에 적용함으로써 큰 변화를 이끌어 냈다면, 현 시대는 빅데이터 기술을 통해 축적한 데이터를 활용하여 기계로 하여금 소프트웨어를 만들게 하는 시대이다. 이런 변화는 학문분야 뿐 아니라 실생활에까지 폭넓게 영향을 미치고 있으면 그 영향의 정도는 혁명적이라고 할 수 있다. 이러한 혁명적 변화를 가능하게 한 기술의 중심에는 요즘 자주 언급되는 기계학습(Machine Learning)과 딥러닝(Deep Learning)이 자리하고 있다. 하지만, 기계로 하여금 소프트웨어를 작성하게 하고, 기계가 작성한 소프트웨어를 검증하고, 실생활에 적용시키는 것 또한 아직까지는 인간이 큰 부분을 담당하고 있으며 이러한 추세는 한동안 유지될 것으로 예측되고 있으며, 이와 관련된 업무를 과거 빅데이터라고 통칭했다면 현재는 **데이터 과학(Data Science)**이라고 부른다[6].

데이터 과학은, 석유화학 산업이 석유를 자원으로 여러 단계의 공정을 거쳐 플라스틱, 섬유, 고무와 같은 상품으로 재탄생시켜 경제적 가치를 창출하듯이, 데이터 자원을 여러 단계의 과정을 거치면서 기계가 기존에 없던 가치를 창출한다[5]. 기존의 통계학도 데이터를 원자재로 사용하고 데이터에서 유용한 가치를 찾아낸다는 점에서 유사하지만, 통계학이 통계이론과 GUI 통계 팩키지를 결합하여 통계기반 의사결정을 지원하는 보고서 작성에 집중한 반면, 데이터 과학은 데이터를 원재료로 삼아 데이터 속 패턴을 찾는 알고리즘을 만들고, 대쉬보드를 통해 인간과 커뮤니케이션은 물론 의사결정도 지원하고, 예측모형을 제작하여 기계간의 자동화 업무도 포괄한다. 데이터 과학은 수학적 가정을 바탕으로 데이터를 해석하고 설명하는데 중점을 둔 반면, 데이터 과학은 데이터를 주어진 상태로 두고 유용한 정보를 추출하고 활용하는데 관심을 둔다.

데이터 과학은 원재료 생성, 처리, 축적부터 축적된 데이터에서 가치를 찾아내고 이를 활용하는 전체 과정에 소프트웨어 공학 기술이 적용되고 이를 “데이터 과학 소프트웨어”로 산업계에서 정의한다. 데이터를 원재료로 삼아 기계가 소프트웨어를 작성[[3]](#footnote-4)하게 되면서 기존 인간이 작성한 소프트웨어로 해결하지 못했던 번역 등 자연어 처리(NLP)분야와 이미지 인식 등 시각분야에서 가졌던 한계를 뛰어 넘는데 큰 기여를 했다.

3. 인간과 기계의 경쟁관계에서 데이터 과학이 할 수 있는 역할

과학자와 기술자들이 제품과 서비스를 만들고 이를 기업가가 확대시키는 과정이 시간의 흐름과 함께 선순환하여 확대발전하면 많은 사람들에게 큰 혜택을 주게 된다. 그러나 동전의 양면처럼 그 이면에는 앞에서 언급한 불평등과 독과점과 이라는 폐해가 따른다.

데이터 과학은 단기적으로 기계가 데이터로부터 소프트웨어를 제작하는 과정에서 남녀차별, 인종차별 등 비윤리적 알고리즘이 포함될 수 있는 블랙박스를 투명하게 만들어 차별과 불평등이 심화되는 것을 막는데 기여하고 있다. 또한, 기존 가내수공업 중심 데이터 과학 제품개발단계를 구글, AWS 서비스로 확인되는 주문형 대량생산체계로 넘어가는 소프트웨어 성숙도(Software Maturity)를 가속화하여 Commodity화 시킴으로써 독과점으로부터 발생되는 불평등을 방지하는데 기여할 것이다. 데이터 과학을 통해 기계가 만드는 소프트웨어가 축적되고 데이터 과학에 더 많은 사람들의 참여가 이뤄짐에 따라, 중장기적으로는 기계가 만든 소프트웨어와 사람이 만든 소프트웨어가 통합된 소프트웨어 제작 방식이 일반화 될 것이다. 데이터 과학은 기존 사람만이 소프트웨어를 제작하던 방식에서 벗어나 기계도 소프트웨어를 제작하는 방식에 길을 열어 둠으로써 문제 해결의 새로운 길을 제시하고 있다.

참고문헌

[1] Cummings MM, Man versus machine or man+machine, IEEE Intelligent Systems, 2014;29:62–9.

[2] Mishel L., Gould E., Bivens J., Wage stagnation in nine charts. Economic Policy Institute, 2015;6:2–13.

[3] Kaplan J., Humans need not apply: A guide to wealth and work in the age of artificial intelligence, Yale University Press 2015.

[4] Wing JM., Computational thinking, Communications of the ACM, 2006;49:33.

[5] 김기범, 이효정, 박도휘, 기업 운영 혁신을 위한 데이터 과학: 기업의 활용 방안, 삼정KPMG 경제연구원, 2020;제121호.

[6] F Provost, T Fawcett , Data science and its relationship to big data and data-driven decision making, Big data, 2013 - liebertpub.com

1. 인간은 기계가 장점이 있는 지루하고 반복된 문제를 해결하는 데 적합하지 않은 반면, 컴퓨터는 인간이 장점을 갖는 추상적이고 일반화하는 작업에 적합하지 못함. [↑](#footnote-ref-2)
2. 그 뿐 아니라 데이터를 원자재로 삼아 소프트웨어를 개발할 수 있느냐에 따라 개발자들 사이의 디지털 불평등도 심화시키고 있다. [↑](#footnote-ref-3)
3. 프로그래밍은 사람이 알고리즘을 작성하고 데이터를 넣어주면 컴퓨터가 출력결과를 만들어 주는 반면, 기계는 기계학습을 통해 데이터와 출력결과를 받아 알고리즘을 스스로 작성하게 된다. [↑](#footnote-ref-4)