

<기술을 통해서 보는 인류의 모습>

책의 마지막 장에서는 컴퓨터를 통해 인간의 지능, 본능과 현실이 확장된다고 하지만 확장을 하기 전에 컴퓨터는 거울의 역할을 한다. 컴퓨터를 통해서 시간과 공간을 뛰어넘고 우리의 현실을 확장하기에 앞서서 이런 컴퓨터를 구현하기 위해서 우리는 현실을 직면하고 분석해야 한다. 우리 세상은 어떤 가정을 바탕으로 논리로 구성되어있는가? 우리는 어떤 편견과 가치관을 갖고 사는가? 이는 우리의 말과 행동과 정책에서 어떻게 드러나고 있는가? 이는 과연 옳은건가?

기술의 발전에는 크게 두 가지 위험이 존재한다. 첫째로 컴퓨터는 기계이기 때문에 “객관적”이라는 착각이다. 컴퓨터가 고작 계산기 수준의 작업을 했을 때는 사실이였겠지만 오늘날의 컴퓨터와 앞으로 발전할 컴퓨터는 절대로 객관적일 수 없다. 기술은 현실과 유기적으로 소통하며 발전하고 있기 때문에 우리 현실에 존재하는 온갖 주관적 요소가 모두 반영된다. 둘째로 컴퓨터는 현실을 일반화시키고 단순화시킨다. 책에서도 설명했듯이 세상에는 아직 기계로도 풀리지 않는 문제들이 많다. 따라서 모든 경우에 옳은 답을 찾는 방법보다는 적당히 옳은 답을 찾는 방법을 선택 할 수밖에 없다. 이런 식의 알고리즘은 일반화를 통해서 답을 찾는 것을 의미한다. 이는 사회적으로 매우 위험한 현상이다. 우리가 현실에서 인종이나 소득계층을 기준으로 사람들을 일반화시키기 시작할 때 사회적 갈등과 오해가 발생한다.

데이터는 의견에 비하면 상대적으로 객관적이지만 절대로 절대적으로 객관적이지는 않다. 그 데이터가 어떻게 측정되었는가? 그 데이터를 어떻게 분석했는가? 이런 질문을 던져보면 데이터의 밑바탕에 많은 가정들이 존재한다는 사실을 쉽게 확인 할 수 있다. 뉴욕타임스의 <When Algorithms Discriminate>라는 기사는 우리가 일상에서 알고리즘에 의해 차별 받고 있는 사례들을 보여준다. 카네기 멜론 대학의 연구에 의하면 구글은 여성보다는 남성에게 더 높은 연봉의 직종 관련 광고를 보여준다. 하버드 대학의 연구에 의하면 구글은 흑인 관련 검색을 할 때 범죄 기록 조사에 대한 광고를 더 많이 보여준다. 이는 알고리즘이 남성보다는 여성이 돈을 더 많이 벌고 흑인은 범죄자일 가능성이 높다는 가정을 바탕으로 실행되고 있기 때문이다. 우리는 기계, 숫자와 데이터는 객관적이고 사실을 보여준다고 말하지만 이 데이터를 해석하고 사용하는건 인간이기 때문에 필연적으로 결과에는 우리의 편견이 주입되기 마련이다. 또한, 머신러닝을 이용하는 기계들은 우리의 행동에 의한 영향을 더 많이 받는다. 우리의 행동이 바로 데이터가 되는 것이다.

현실은 매우 복잡하다. 현실은 흑백 논리가 성립하지 않는 공간이다. 도둑질은 범죄라고 배우면서도 굶어 죽어가는 아기를 위해서 우유를 훔친 여성을 심판해야할 때 우리는 혼란스럽다. 살인은 최악이라고 배우지만 영화의 주인공이 아버지를 위해 복수를 하는 모습을 보면 우리는 한 편으로 만족감을 느낀다. 현실은 이성과 감성이 교차하는 공간이기 때문에 흑백 논리로 판단하기 어려운 상황들이 너무나도 많다. 따라서 현실에서의 우리의 평가기준도 그만큼 복잡하다. 법, 대학입시, 기업 공채 등의 과정이 점점 복잡해지는 이유도 현실의 복잡성을 반영하기 위해서이다.

현실을 있는 그대로 반영하는 것은 아마도 다항 비용의 알고리즘이 아닐거다. 또한, 우리도 현실에서 발생하는 문제들에 있어서 일관적으로 답을 찾을 수 있는 방법을 아직 모르기 때문에 이를 기계로 만드는건 불가능하다. 따라서 대부분의 알고리즘은 엄청난 일반화를 바탕으로 결과를 낸다. 앞의 기사의 예를 보면 구글의 알고리즘은 전 세계의 기업들을 봤을 때 여성이 CEO인 경우는 15.4%밖에 안 되기 때문에 여성에게는 높은 연봉의 직종을 보여주는 것이 마케팅상 비효율적이라고 판단하고 있다. 또한, 이 알고리즘은 미국에서 흑인은 백인보다 범죄를 저지를 확률이 3배 이상이기 때문에 흑인의 이름을 검색하는 사람은 아마도 범죄 기록 확인에 관심이 있을거라고 판단하고 있다. 컴퓨터에 앉은 여성이 야후의 CEO Marissa Mayer 이고 검색하는 흑인의 이름이 오바마 대통령이라도 이런 가정이 성립하는가? 물론 답은 '아니다.' 일반화는 이래서 무서운거다. 인종차별, 종교차별, 성차별 등의 사회의 거의 모든 갈등은 일반화의 산출물이다.

Frank Pasquale의 <The Black Box Society>에서도 우리의 검색정보가 기업의 마케팅의 악용되거나 신용등급에 영향을 주는 사례들을 보여준다. 사실 우리가 인터넷에서 클릭하는 링크들의 큰 일부는 실수로 클릭하거나 아무 생각 없이 클릭하는 경우들이다. 그냥 친구가 페이스북에 공유해서 클릭하거나 그냥 앱에서 보여주거나 클릭하거나 그냥 사진이 재밌어서 클릭해 본다. 이런 유저 데이터를 바탕으로 한 사람을 평가하기에 앞서서 데이터에 내제된 정보와 사용용도에 대한 충분한 고민이 필요하지만 인간의 오류와 알고리즘의 한계에 의해서 우리는 거의 무작위로 데이터를 사용한다.

고등학교 Mission Statement에 이런 말이 있었다. 학교를 설립한 John Phillips는 "Goodness without knowledge is weak and feeble, yet knowledge without goodness is dangerous."라고 말했고 3년 내내 이 말을 들으면서 자랐다. Phillips가 1781년에 학교를 설립할 때는 엘리트층의 사회적인 책임을 떠올리면 이런 말을 했겠지만 대부분의 단순 노동이 기계로 대체되고 점점 인류의 지적 수준이 높아지게 되면 우리 모두가 이런 사회적 책임에 대해서 생각해 볼 필요가 있는 날이 올 거다.

최근 인공지능 성과를 보면서 사람들은 인간을 뛰어넘은 기계의 탄생을 두려워하지만 기계는 어디까지 기계이다. 인간보다 더 잘하는 것이 훨씬 많은 기계일지라도 기계이고 하나의 도구이다. 이런 기계를 통해서 인류의 삶은 더 풍부해 질 수 있다. 지금까지 풀지 못 한 문제의 답을 찾고 새로운 사실을 발견하고 더 넓은 세상을 탐험 할 수 있게 될 것이다. 그러나 이 정도로 강력한 도구의 탄생은 인류에게 그 만큼 더 큰 책임을 요구하게 된다. 이 도구가 잘 못 된 방향으로 발전하지 않기 위해서는 맹목적으로 개발을 위해 노력하기 전에 컴퓨터라는 거울을 자세히 볼 필요가 있다. 우리의 현실에서 발생하는 문제들에 대한 고민을 하지 않은 상태로 오로지 기술의 발전만 추구하면 구글의 차별이나 신용등급의 하락보다 더 심각한 문제들이 발생하게 될 것이다.