

두산인문극장 2026: 신분류학 New Taxonomy

인공지능과 미래 예측: 판단하는 인간, 예측하는 기계 전준

2026년 6월 22일 저녁 7:30~9:30 @두산아트센터 연강홀

* 본 강연은 촬영할 수 없습니다.

주요 논의 내용

- 인간은 왜, 어떻게 미래를 예측하는가
- 인공지능이 수행하는 예측의 유혹
- 예측을 실현하거나 폐기하는 인간
- 인간의 선택, 인간의 미래

예측이란 왜 강력한 것일까? 예측은 사실상 개인의 삶과 조직의 흥망성쇠를 매 순간 결정짓는 것과 다름없는 강력한 행위이다. 우리 모두는 의식적, 무의식적으로 수많은 선택을 하고, 이에 대한 결과로 무한한 복잡성을 가진 새로운 미래를 맞이한다. 인류 문명의 역사는 한편으로는 올바른 예측을 해왔던 조상들의 생존의 역사이기도 하다. 다가올 겨울을 예측하고, 가을에 식량을 저장하지 않았다면 지금의 우리는 없었을 것이다. 예측에 실패했던 집단들의 후손은 이제 우리 주위에 없다. 이때 개인과 조직의 공통점은 이러한 '예측'을 위해 지배적으로 많은 자원을 소비하면서도 최선의 예측을 해내기보다는, 거의 언제나 '최적'의 예측만을 간신히 해낸다는 점이다. 우리의 오감과 지능은 일정 수준 한계가 있기에 예측의 정확성이 획기적으로 올라가지는 않기 때문이다. 자연과 사회의 불확실성은 때사에 존재할 수밖에 없기 때문에, 우리는 과도한 에너지 투자 대신, 적절한 양의 에너지를 소모하여 적절한 정도의 정확성을 갖는 예측을 해내는 방식으로 진화했다고 볼 수 있다.

인공지능은 정말로 예측을 잘해내는 강력한 도구일까? 인공지능을 매개로 한 알고리즘적 예측이 각광받는 이유를 살펴본다. 첫째, 인공지능의 예측은 인간보다 훨씬 저렴하다. 인간의 노동과 비교한다면 인공지능이 지출하는 비용이 크지 않다는 점을 알 수 있다. 예컨대 인공지능을 활용해 수백 권의 책을 읽고 이를 바탕으로 타당한 예측을 하게 하는 데에는 몇 번의 대화면 충분하다. 같은 과업을 인간이 수행하기 위해서는 어떻게 해야 할까? 현대 교육 시스템 속에서 살아가는 인간은 20년 내지는 30년가량의 장기적인 교육을 받고서야 개인의 전문성을 가까스로 획득한다. 게다가 고도로 훈련된 인간이라고 하더라도 혼자서는 충분히 전문성을 발휘할 수 없다.



둘째, 인공지능의 예측은 감각을 외주화한다. 인간이 시각기관과 지능을 최대한 활용해서 책을 ‘읽어야’ 하는 반면에, 인공지능은 책을 ‘읽지’ 않는다. 인공지능은 책에 쓰인 인간의 언어를 최소 단위로 쪼개고, 숫자를 부여하고, 이 숫자들 사이에 규칙적인 패턴을 ‘계산’한다. 셀 수 없이 많은 인쇄물과 인터넷상의 문서들을 학습해 계산값을 도출하는 인공지능은 마치 인류의 축적된 지적 자산인 것으로 인식되고는 한다. 또한 감각의 외주화는 더 나은 절대다수의 조연을 얻게 되리라는 왜곡된 신념으로 우리를 이끈다.

최근 현실에 존재하지 않는 인간을 데이터의 세계에 구현해서 인공지능을 활용해 다양한 실험을 하는 방법이 도입되고 있다. 현실 세계와 동일한 성질을 갖는 가상 세계를 구축하고, 가상 세계에서 인위적 조작과 개입을 통해 현실 세계의 미래를 예측하기도 한다. 이러한 접근이 정확하게 작동한다면, 어쩌면 인간은 미래 예측의 불확실성을 획기적으로 극복하고 인공지능의 예측을 기반으로 한 미래를 살게 될지도 모른다. 그러나 이러한 낙관론의 가장 중대한 모순은 인공지능이 어떤 미래를 예측하든 그러한 미래를 실제로 구현하는 주체는 인간이라는 점이다. 인공지능 모델이 A라는 예측을 발표했을 때 A를 지지하는 사람들은 굳건하게 결집해서 실제로 옳은 예측으로 공고히 하는 효과가 벌어질 수 있다. 그렇다면 이는 자기 실현적 예언이 된다. 반대로 A를 반대하는 사람들은 예측대로 되도록 내버려둘 수 없기 때문에 적극적인 대응에 나서서 예측과 반대되는 미래가 발생할 수 있다. 예측 행위로 인해 예측 스스로가 부정되는 역설적인 효과가 나타나는 것이다.

인공지능이 구축한 ‘가상’의 세계가 우리가 살아가는 현실을 최대한 모사하고 이를 바탕으로 미래를 예측한다고 할 때, 우리는 현실과 가상 사이의 경계가 쉽게 무너질 것이라고 상상하기 쉽다. 그러나 가상의 세계는 현실의 복잡성을 완벽하게 반영하여 구현되지 않을뿐더러, 가장 잘 작동하는 인공지능 시스템조차 다양한 형태의 일관적인 오류를 발생시킨다. 인공지능을 만들고 사용하고, 이를 특정 행동의 결과로 구현하는 주체는 우리 자신이라는 본질적인 전제로 돌아가 보자. 저렴하고 성능 좋은, 체계적 오류 가능성을 내포한, 어쩌면 썩 믿을 만하지 않은 인공지능을 우리는 어떻게 받아들여야 할 것인가? 어쩌면 ‘받아들인다’는 수동적인 태도보다는 ‘길들여야 할 것인가?’라고 질문을 수정하는 게 더 바람직할 수도 있다. 인공지능은 예측을 생성하지만 그 예측에도 불구하고 미래를 생성하는 것은 인간 고유의 몫이기 때문이다.

전준

한국과학기술원(KAIST) 디지털인문사회과학부 교수이자 KAIST 김재철AI대학원 겸임교수이다. 과학기술사회학자로 현대 사회 이론을 폭넓게 연구하면서 인공지능, 기후위기, 미래 기술에 대해 분석하고 있다. 저서로는 『인공지능 시대의 사회과학』(공저)이 있다.