

# AI·로봇으로 인한 사회문제, 인간만의 ‘공감·신뢰’로 풀어야

지난 3월 이광형 KAIST 총장은 취임식에서 “KAIST는 앞으로 인류가 당면한 문제를 찾아 정의하고 해결하는 것에 중점을 뒀다”고 말한 바 있다. 이는 창업(entrepreneurship)의 근본적 문제이기도 하다. 창업의 기회는 개인이 특정한 문제에 대한 해결책을 찾았을 때 온다. 다른 한편으로 창업 기회는 주어진 문제를 어떻게 정립하는가에 달린다. 예를 들어 설명하겠다. 서울에서 주차공간을 찾는 것은 어렵다. 이에 대한 문제를 ‘주차공간의 부족’으로 정립한다면, 그에 대한 해결책은 더 많은 주차장을 만드는 것이다. 하지만 주차공간 찾기 어려움의 원인을 ‘넘치는 유희차량’이라고 세운다면, 사람들이 차량 공유 서비스나 대중교통 수단을 적극적으로 활용하도록 만드는 것이 해결책이 될 수 있다. 이처럼 주어진 문제를 어떻게 정립하는가에 따라 그에 대한 기술적 해결 방안도 달라진다.

오늘날 사회적 문제들은 빠르게 진화하고 있는 인공지능(AI)과 로봇기술과 밀접한 관계가 있다. 인공지능과 로봇기술은 기업 입장에서 비용절감을 하도록 만든다. 커피 파리스타 역할을 하는 로봇은 지치지 않는다. AI 뉴스 앵커는 건강보험을 요구하지 않는다. 이러한 이유로 인간의 일자리가 없어질 위기에 놓여 있지만, 기술의 발달로 인해 사람들에게는 새로운 일자리가 창출되기도 한다. 아직 기술은 사람들의 손길이 요구된다. 예를 들어 커피를 제조하는 로봇이 잘 작동되는지 확인하고, 택배를 배송하는 자율주행 트럭이 고장 날 경우 이를 수리하는 사람들이 필요하다. 기술의 진보는 근본적

## 기술진보시대 세 가지 이슈

기술 관련 어떤 문제 해결하나  
어떤 기술을 발전시켜야 하나  
어떻게 인간의 업무 디자인하나

## 인간이 기술을 통제하려면

공감에 기반한 인간의 일자리  
AI가 대체하지 못하게 막고  
신뢰가 우선되는 기술 개발을

으로 사람의 일자리와 업무 변화와 때려야 할 수 없는 사이다.

이런 상황에서 사람들이 직면하는 세 가지 이슈가 있다. 바로 기술 관련 어떠한 문제들을 해결해야 하는지, 어떠한 기술을 발전시켜야 하는지, (기술 진보 시대에) 인간의 업무를 어떻게 디자인해야 할지다. 이에 대한 ‘마법의 공식’은 없다. 새로운 과학 기술은 예상보다 너무 빠르게 다가온다. 어떤 사회적 문제를 해결해야 할지 사람들이 파악도 하기 전에 신기술이 나타나 는 것이다. 이러한 상황에서 인간이 해야 할 질문은 ‘사회적 문제를 어떻게 정립하고 어떤 문제를 해결할까?’다.

해당 질문에 대한 답을 찾기 위해서 인간은 스스로에게 ‘미래 사회와 미래의 직장 문화가 인류의 가치를 어떻게 반영할지 바라는지?’를 물어야 한다. 그리고 해당 질문에 대한 답을 토



[게티이미지뱅크]

대로 사람들이 만들어내는 기술을 통제해야 한다. 가령, 서비스 산업 관련된 한 가지 연구에서는 AI가 우선적으로 기계 기반의 일자리를 대체하고, 그다음으로 분석 기반의 일자리, 직관 기반(intuition-oriented)의 일자리, 공감 기반의 일자리 순으로 사람들의 일자리를 대체할 것이라는 전망이 나왔다. 하지만 미래 사회에 더 강력한 공감 기반의 조직문화가 형성되길 바란다. AI가 공감 기반의 일자리를 대체하지 못하도록 통제할 수 있다.

공감 외에도 사람과 사람 사이의 교류에서 형성되는 신뢰를 보존하는 것에 대한 중요성 역시 생각해야 한다. 대부분의 학자들과 정책입안자들은 인간이 어떻게 AI를 신뢰할 수 있을지에

대해 고민하고 있다. 하지만 인간-AI 사이의 신뢰와 인간-인간 사이의 신뢰에는 큰 차이가 있다. 사람과 사람 간의 신뢰에는 어느 한 쪽의 필요성과 욕구를 희생하는 것이 요구된다. 가령, 부모가 자녀에게 ‘오후에 학원 시험 보러 갈 때 데려다주겠다’고 약속을 한다. 가정해보자. 이 상황에서 자녀는 부모가 학원에 데려다주기 전 있는 회사 모임에서 음주를 해 약속시간을 지키지 못하고 결국 학원에 늦게 데려다주는 상황이 발생하지 않을 것이라 믿는다(그리고 자녀의 믿음을 지키기 위해 부모는 음주하고 싶은 욕구를 억누른다). 하지만 로봇 등의 기계들은 니즈나 욕구가 없다. 사람들은 그저 본인이 일할 때보다 기계가 오류를 덜 낼 것이라 믿는다. 사람

들이 컴퓨터 기술과 AI에 의존할수록 인간은 다른 사람을 신뢰하고 무언가 잘못됐을 때 그를 용서하는 것보다 기술을 신뢰하고 용서하는 능력을 더 갖게 될 것이다. 하지만 문명사회에서는 인간이 서로를 덜 신뢰하게 만드는 기술이 아닌, 사람들 간의 신뢰가 더 형성되는 기술을 만들어야 하지 않을까?

최근 AI 돌봄서비스가 생기면서 AI의 ‘공감 능력’에 대한 케이스 스터디가 생겨나고 있다. 가령, SK텔레콤의 인공지능 스피커 ‘누구’(명령어 ‘아리아’)는 노인들의 말뚱마가 되어준다. 하지만 과연 AI 스피커가 사람들의 공감능력을 키울 수 있도록 도움을 줄 수 있을까? 부정적인 외부 효과가 발생한다면 대부분의 경우 벤처사업들의 뜻을 할 수는 없지만, 인류의 가치를 짊어내리는 부정적인 효과는 적극적으로 피해야 한다.

인간의 공감능력과 신뢰의 가치를 잘 판단하고 나아가야 할 방향을 잘 선정하기 위해서는 기업들과 정부가 동시에 다발적으로 사회적 문제를 올바르게 수립하고, 이에 대한 기술적 해결책을 개발하고, 의미 있는 일자리 창출을 할 수 있는 방법을 생각해야 한다. 한국을 포함해 전 세계적으로 국가들은 인공지능 기술을 상업화하는 글로벌 리더가 되려 경쟁하고 있다. 이런 상황에서는 인간의 가치를 아는 것이 매우 중요하다. AI가 인간의 가치를 잘 아는 전문가가 되지 않기를 바란다.



서예 차마오 한국뉴욕주립대 스토니브룩 교수

## CAN A FOCUS ON EMPATHY AND TRUST GUIDE OUR TASTE IN CHOOSING THE RIGHT PROBLEMS TO SOLVE?

By Chihmao Hsieh (시에 츠마오)

In his inauguration speech earlier this year, Lee Kwang-hyung, the new president of KAIST, remarked “KAIST 는 앞으로 인류가 당면한 문제를 찾아 정의하고 해결하는 것에 중점을 뒀다”. This is a fundamental issue in entrepreneurship: the opportunities that we find depend on how we search for solutions to a problem, but also how we formulate problems in the first place. For example, we all know that it’s not easy to park a car in Seoul. If we formulate the problem as a lack of parking spaces, then maybe we should build more parking garages. However, if the problem is that there are too many idle cars, then we should better optimize public transportation, ride sharing, or the taxi system. Or maybe we decide that cars are currently too big, and smaller cars will be easier to navigate in tight spaces. These three different problem formulations all lead to very different technological directions. Indeed, our taste in the problems that we identify today affects the decisions and socio-technological environment that we have tomorrow.

These days, society is addressing various problems with the aid of rapidly advancing artificial intelligence and robotics. Recent innovations certainly save money. Coffee barista robots don’t get tired. Digital news anchors don’t need health insurance. And self-driving trucks don’t ask for pensions. While jobs are at risk of being lost, new jobs will be created. In the near term, for example, we still need people to maintain the robots that make the coffee or fix the trucks that deliver our packages. Technological advances are fundamentally entangled with the changing human tasks and jobs on this planet.

Set aside the idea of universal basic income, for now. We face three issues: which problems to solve, what technologies to develop, and how to identify and design the tasks and jobs eventually performed by humans. There is no magic equation that explains how this all actually unfolds. Often, even, new scientific technologies emerge so quickly that they arrive before we know what problems they should solve. With increasingly advanced AI, where applications are so broad and dramatic, we should stop to think: *How do we formulate and choose which problems to solve?*

One diagnosis would require us to work backwards: ask ourselves how we want future society and work culture to reflect core values of humanity, and then use that vision to help constrain and guide the kinds of technologies we develop and the kinds of problems that we address. Recent research on service industries suggests that AI will tend to first replace mechanically-oriented jobs, then analysis-oriented jobs, then intuition-oriented jobs, and finally empathy-oriented jobs, in that order. Opinions differ regarding the overall timeline. However, building and refining a stronger culture for empathy now will help preserve a basis of humanity at work and in society, protecting ourselves from the most damaging economics of AI. Human empathy is not particularly scalable, but maybe we also should not be trying to make it scalable. Marketing staff, news reporters, nurses, teachers and countless other occupations today can make valuable use of human empathy, and we should support that kind of sociocultural direction. On a broad scale, more products and services can be created to enable people to exhibit, share, and promote empathy, at work and at home. In the future, hopefully still more different kinds of empathy-oriented jobs can emerge.

Besides empathy, we should also consider the importance of preserving the importance of trust in interactions *between humans*. This issue hasn't been addressed much; instead, most scholars and policymakers today are worried about how we humans could best learn to trust AI. But there's an important conceptual distinction between how humans trust AI and how they trust other humans. On one hand, trust between humans requires an element of vulnerability to others' self-interest: if a father promises his child that he will drive her to an evening hagwon exam, she trusts him that he won't get drunk while partying with co-workers, then arriving late to pick her up. On the other hand, machines and robots have no self-interest; we merely trust them to function reliably and make fewer errors than us. As our interactions become increasingly mediated by computerized technology and AI, we position ourselves to learn to "trust" and "forgive" computers and robots more than we maintain our ability and capacity to trust and forgive each other. As a civilized society, shouldn't we be developing technology and policy that ultimately helps humans trust and empathize with each other more, not less?

Recent developments offer potential case studies. Aria, for example, represents AI by SK Telecom that has been shown to empathetically stimulate senior citizens' cognition and delay onset of dementia, while also doubling their daily travel distances. But does the AI-based replication of the late Kim Kwang-Seok's voice and singing help to foster empathy or fulfill egos? Are recently developed AI-based 'digital girlfriends' more likely to support or cheapen cultural norms of human trust and empathy? While it's true that business ventures should often be excused for generating 'negative externalities' (부정적인 외부효과), we should be actively avoiding those negative externalities that fundamentally discount any core value of humanity.

Formulating the right problems, developing supportive technological solutions, and fostering meaningful jobs should be *simultaneous considerations* that will require business and government to have good taste. But good taste in the value of human empathy and trust is not straightforward. For example, empathy should not completely stifle competitive spirit, and sometimes trust between business partners should not cut off valuable exploration of the business environment. As countries—including South Korea—race to become world leaders in commercializing AI-based technology, they should all respect the responsibility of choosing what entrepreneurial directions we take. Overall, having good taste in humanity is simply becoming more valuable than ever. Hopefully AI won't become an expert in that too.

## References

KAIST President's quote can be found in this Kookmin Ilbo article:

이광형 신임 KAIST 총장 “포스트 인공지능 시대 대비해야”

<http://news.kmib.co.kr/article/view.asp?arcid=0015608922&code=61122015>

AI girlfriends described in the following articles:

‘빠빠로데이에 뭐해’ AI와 사랑에, 아니 대화에 빠졌다

[https://biz.chosun.com/site/data/html\\_dir/2020/11/05/2020110502562.html](https://biz.chosun.com/site/data/html_dir/2020/11/05/2020110502562.html)

‘영화 her’ 현실되나? 인공지능 여친 개발사에 투자 봇물

<https://www.sedaily.com/NewsView/22JQQT26VA>

The line about AI in “recent research on service industries” points to the popular 2018 research article found here (Google citations = 548):

‘Artificial Intelligence in Service’ (Journal of Service Research)

<https://journals.sagepub.com/doi/full/10.1177/1094670517752459>