

Featured · Interview

## [인터뷰] 지겐대학교 칸트연구소 소장 세네커 교수 “AI는 윤리적 주체일 수 없다”

박설민 기자 ainews@chosun.com

기사입력 2022.07.25 15:46

‘Kant and AI’ 저자 독일 지겐대학교 칸트연구소 소장 디에터 세네커 철학과 교수, 중앙대학교 인문콘텐츠연구소 HK+인공지능인문학단 김형주 HK교수 인터뷰



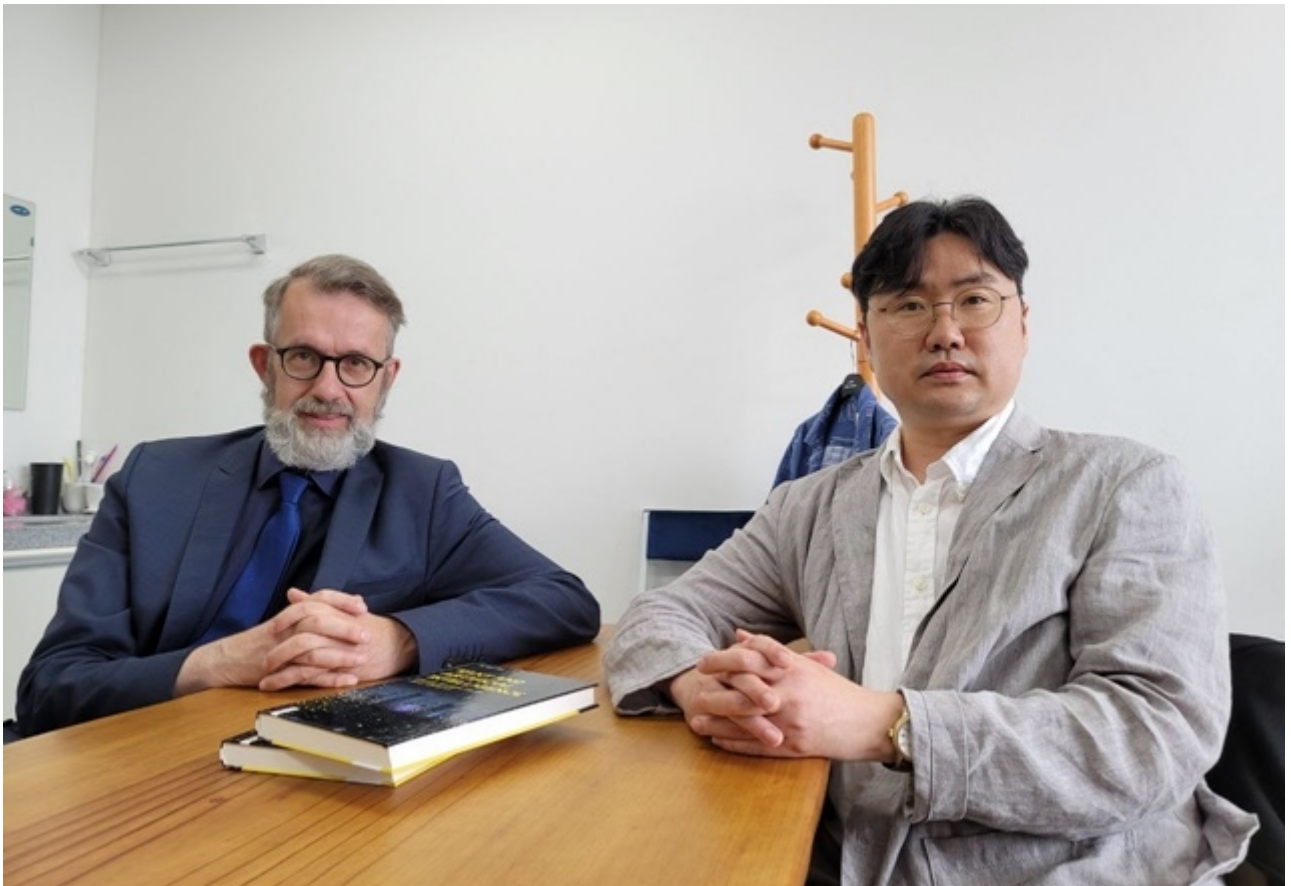
▲ 인공지능(AI)이 빠르게 발전함에 따라 AI를 앞으로 어떻게 해야 할지에 대한 논쟁도 많아지고 있다. 이에에서는 철학과 AI를 엮은 저서 ‘Kant and AI(칸트와 AI)’의 저자 독일 지겐대학교 세네커 교수와 중앙대학교 김형주 HK교수를 만나 AI의 발전 전망과 이것이 우리 사회에 미칠 윤리적 영향에 대한 의견을 들어보는 자리를 가졌다/ 사진, 편집=박설민 기자

“나는 무엇인가요(What am I)?” 이는 지난 2004년 개봉한 SF영화 ‘아이, 로봇’에 등장하는 로봇 ‘써니’의 대사다. 인간처럼 감정을 느낄 수 있는 로봇 써니는 끊임없이 자신의 존재가 무엇인지

고민한다. 반면 영화에 등장하는 대부분의 ‘사람’들은 기계에 불과한 씨니의 고민에 대해 대수롭지 않게 생각한다.

흥미로운 것은 공상과학영화 속에서나 등장했을 법한 로봇 씨니의 이야기가 이제 우리에게 현실이 됐다는 점이다. 4차 산업혁명시대를 맞아 인공지능(AI) 기술력이 급속도로 발전하면서 사람들은 AI를 앞으로 어떻게 해야 할지에 대해 논쟁한다. AI를 개발·연구하는 과학자들뿐만 아니라 인간의 존재에 대해 고찰하는 철학자들 역시 마찬가지다. AI가 어디까지 ‘인간’과 유사한 존재가 될 수 있는지, 이들을 어떤 존재로 대해야 하는지에 대해 많은 철학자들의 고민이 깊어지고 있다.

이에 <THE AI>에서는 근대 서양 철학사의 상징인 임마누엘 칸트(Immanuel Kant)의 철학과 AI를 엮은 저서 ‘Kant and AI(칸트와 AI)’의 저자 독일 지겐대학교 칸트연구소 소장 디에터 세네커 철학과 교수와 중앙대학교 인문콘텐츠연구소 HK+인공지능인문학단 김형주 HK교수를 만나 AI의 발전 전망과 이것이 우리 사회에 미칠 윤리적 영향에 대한 의견을 들어보는 자리를 가졌다.



▲ 근대 서양 철학사의 상징인 임마누엘 칸트(Immanuel Kant)의 철학과 AI를 엮은 저서 ‘Kant and AI(칸트와 AI)’의 저자인 독일 지겐대학교 칸트연구소 소장 디에터 세네커 철학과 교수(사진 좌측)와 중앙대학교 인문콘텐츠연구소 HK+인공지능인문학단 김형주 HK교수(사진 우측)./사진=박설민 기자

**Q1. ‘칸트 철학’은 일반적으로 우리에게 근대 서양 철학의 상징으로 알려져 있습니다. 반면 ‘인공지능(AI)’는 대표적인 최신 과학 기술 분야입니다. 두 분께서 이 둘을 묶어 ‘Kant and AI’를 출간하게 된 배경은 무엇인가요?**

**세네커 교수:** 최근 핫한 기술 트렌드인 AI에 대해 칸트에 관점을 취해 AI가 빚어내는 문제들을 어떻게 바라볼 수 있을까 접근하고자 했다. 예를 들어서 실천 철학 분야에서 주로 다루는 문제 중 하나인 트롤리 딜레마 같은 것들을 칸트의 관점을 대입한다면 AI의 경우 이것을 어떻게 바라볼 수 있을까. 또 반대로 AI로부터 칸트 철학을 연구하는 이들은 무엇을 배울 수 있을까. 이 두 가지 관점에서 독자들이 AI를 바라볼 수 있도록 하는 안내서가 될 수 있도록 이 책을 출간하게 됐다.

**김형주 교수:** 저는 세네커 교수님께 칸트 철학에 대해 배웠으며, 중앙대학교 인문콘텐츠연구소에 와서는 AI철학을 연구 분야로 하고 있다. AI 철학은 AI 윤리나 AI과 관련된 여러 가지 기능 개념의 규정문제 등을 연구하고 있다. 이 책은 우리가 칸트의 철학을 이야기할 때 ‘진 선 미’라는 최고의 가치로 책을 기획을 했다. 여기서 ‘진’은 칸트의 ‘이론 철학’을, ‘선’은 칸트의 ‘실천 철학’을, ‘미’는 ‘미학’이다. 이 세가지 칸트의 관점으로 나눠 현재 AI가 빚어내고, 향후 빚어낼 수 있는 여러 현상을 조망하고자 하는 것이 이 책의 목표다.

**Q2. 최근 발생한 AI이루다 성희롱 사건 등 AI에 대해서도 향후 일종의 ‘인권’과 같은 권리를 마련해야 하는 것이냐는 의견을 제시하는 사람들이 늘고 있습니다. 하지만 세네커 교수님께서 AI가 윤리적 주체일 수 없다는 입장을 고수하고 계신 것으로 알고 있는데, 그 이유를 설명 부탁드립니다.**

**세네커 교수:** 이것은 굉장히 복잡하고 어려운 질문이라 생각한다. 간단히 말하면 제가 AI가 윤리적 주체일 수 없다는 입장인 것에 대해선 일단 두 가지 이유가 있다. 첫 번째로 AI에겐 ‘내면의 삶 (In a life)’이 없다. 내면의 삶이 있다는 것은 쉽게 말해 어떤 세계를 인식하는 의식을 넘어서 자기가 자기 자신을 인식할 수 있는 자기의식(Self-consciousness)이 있다는 것이다. 그것이 있을 때만 우리는 존엄성을 가진 ‘인간(Person)’이라는 지위를 갖게 된다. 이것이 칸트적 관점이다. 때문에 AI는 수학적 계산으로 스스로 판단할 수 있는 능력이 있지만 그것만으로 인격의 지위를 부여할 수 있는 없다고 생각한다. 두 번째로 AI는 가치는 있지만 권리는 없는 존재라고 생각한다. 예를 들어 우리가 자연에서 마주치는 아름다운 동물과 나무 등은 ‘가치’가 있는 존재들이다. 하지만 칸트의 입장에서 보자면 이들은 인간과 같은 ‘권리’는 없다. 우리가 가치를 부여하긴 하지만 가치가 있다고 해서 모두 권리가 있는 것은 아니다. AI도 마찬가지다. AI는 어떤 일을 처리하는 가치가 있을 수 있지만 권리가 있지는 않다.





▲ 세네커 교수는 AI는 어떤 일을 처리하는 가치가 있을 수 있지만 권리가 있지는 않기 때문에 AI를 인격적 존재로 보긴 어렵다고 말한다./ 사진=박설민 기자

**Q3. ‘Kant and AI’ 2장을 살펴보면 김형주 교수님께서도 칸트적 관점에서 AI는 감정을 느끼는 주체가 될 수 없다고 주장하셨습니다. 어떤 논리에 의해 그렇게 주장하셨는지 설명해주시겠습니까?**

**김형주 교수:** 이렇게 볼 수 있을 것 같다. 만약 ‘짝사랑도 사랑이다’라고 치면 AI도 감정을 느끼는 주체처럼 볼 수도 있다. 그러나 짝사랑은 엄연히 사랑이라고 볼 수 없다. 서로가 어떤 긍정적 감정을 교감해야 하고, 그렇게 되었을 때 우리는 그것을 사랑이라 인정한다. 때문에 교감이라는 영역의 측면으로 봤을 때는 AI는 그런 주체가 될 수 없다는 것이 저의 입장이다. 우리가 AI와 대화를 하거나 어떤 감정을 내가 느낄 수는 있겠으나, 그것이 AI도 감정을 느끼고 우리와 교감을 한다고 말할 수는 없는 것이다. 그런 입장에서 해석해볼 때 저는 AI가 감정을 느끼는 주체가 될 수 있다고 함부로 말할 수는 없다고 본다. 다만 인간의 관점에서는 AI가 마치 감정을 느끼거나 생각하는 것처럼 보일 수는 있다.

**Q4. 김형주 교수님께서서는 또한 임마누엘 칸트가 바라보는 인격적 존재는 ‘자신이 자유로운 지성적 존재자임을 자각하는 존재’라고 해석하셨습니다. 만약 미래에 인간처럼 AI가 고도로 발달해서 ‘자유로운 지성적 존재자임을 자각’하게 된다면 칸트는 이들을 ‘인간’과 같은 존재로 인식했을까요? 김형주 교수님과 세네커 교수님 두분의 의견이 궁금합니다.**

**김형주 교수:** 일반적으로 우리가 생각하는 ‘인간’은 단순한 ‘인격(personality)적 존재’뿐만 아니라 육체적, 즉 생물적으로 인간으로 분류가 된다. 때문에 ‘자신이 자유로운 지성적 존재자임을 자각하는 존재’에 칸트는 신, 천사 등 모든 인격적 존재가 들어갈 수 있다고 봤다. 이런 것만 있으면 이 우주의 어떤 존재도 인격적 존재가 될 수 있다. 만약에 AI가 이런 존재라면 당연히 인격적 존재에 도달했다고 볼 수는 있을 것이다. 하지만 육체적 관점까지 포함된 인간이라는 존재로 분류해야 한다면, AI를 인간이라 보기는 힘들 것이다.

**세네커 교수:** AI는 미래에 자기의식과 의식을 가진 존재까지 발달할 수도 있지만 하나 걸리는 것이 자유로운 존재는 아니라는 점이다. 그래서 AI가 인간이 될 수는 없다고 생각한다. 동물도 의식이 있다. 만약에 AI가 의식, 더 나아가서 자기의식을 가진다고 할지라도 그것이 인간이 되려면 자유로운 존재가 돼야 한다. 그럼에도 불구하고 AI는 아무리 발달해도 자유로운 존재라고 우리가 말할 수는 없을 것 같다. 여기서 칸트가 말하는 자유는 일반적으로 ‘소극적 자유’와 ‘적극적 자유’의 두 가지 의미가 있다. 먼저 소극적 자유는 앞에 일어난 사건에 의해 내가 규정받지 않는 존재다. 적극적 자유는 자기 스스로를 규정할 수 있는 존재다. 칸트적 의미에서 조금 더 깊게 말하면, 내가 스스로 도덕 법칙을 생성을 하고 그 법칙을 규정하는 존재다.



▲ 세네커 교수는 예술을 작가의도(Intention)가 포함되어 있는, 즉, 우리 인간의 내면에 있는 것을 표현하는 것이라고 규정한다면 AI는 예술가가 될 수 없다고 봤다./ 그래픽=박설민 기자

**Q5. 최근 등장하는 AI는 딥러닝 기술을 기반으로 스스로 학습하고 결론을 도출하기도 한다. 그렇다면, 현재의 AI들은 비록 자기 스스로를 규정하진 못하지만 적어도 앞에 일어난 사건에 의해**

## 내가 규정받지 않는 ‘소극적 자유’ 측면에서는 자유롭다고 볼 수 있을까요?

**세네커 교수:** AI는 두 가지 모두 해당될 수가 없다고 생각한다. 물론 딥러닝의 경우에는 알고리즘에 의해 마치 스스로 결과를 도출하는 것으로 보이기도 한다. 때문에 소극적 자유 측면에서는 자유로운 존재가 아니냐고 생각할 수도 있다. 하지만 이것이 인식론적 관점에서는 그렇게 볼 수도 있겠지만 존재 그 자체, 존재론적 관점으로 보면 다르다. 어디까지나 진짜 인간이 AI의 알고리즘을 만들고, 이를 통해 ‘생각’을 하는 AI는 결과적으로 인간에 종속된 기계다.

## Q6. 최근 AI가 미술 작가나 작곡가로 활동하며 다양한 작품을 만들었다는 소식을 종종 접합니다. 그렇다면 칸트 철학의 관점에서 이것은 AI가 창조한 ‘예술 작품’ 일까요? 아니면 AI에 명령을 내린 인간이 만든 작품이라고 판단해야 할까요?

**세네커 교수:** 예술을 뭘로 규정하느냐에 따라 판단이 다를 수 있다고 생각한다. 만약 단순히 어떤 작품이 아름답다와 같이 외형적 관점에서 봤을 경우엔 AI가 만든 것도 예술이라고 볼 수도 있다. 만약 예술을 작가의도(Intention)가 포함되어 있는, 즉, 우리 인간의 내면에 있는 것을 표현하는 것이라고 규정한다면 AI는 예술가가 될 수 없다고 생각한다. 따라서 AI가 외면적인 시점에서 예술로 보이는 작품을 만들었을 때 이를 만든 것은 ‘인간’이다. 붓을 이용해 그림을 그리거나, 컴퓨터를 활용해 그래픽을 만들거나 타자기를 이용해 소설을 써내려간다면 이는 ‘도구’를 사용해 작가들이 작품을 만든 것이다. 이는 AI 역시 마찬가지다. AI에게 어떤 작품을 만들라고 명령을 내린 작가가 결국 AI가 만들어낸 예술 작품의 진짜 제작자라고 볼 수 있다.

## Q7. 현재 AI를 연구하는 대다수의 과학자들이 목표하는 최종 단계는 ‘인간과 거의 동등한 수준의 지적 능력을 갖는 인공지능’입니다. 물론 이것은 굉장히 어려운 일입니다. 하지만 가능성은 희박하다해도 먼 미래에 AI가 발전을 거듭해 앞서 설명하신 ‘인격적 존재’에 가까운 존재가 된다면 어떨까요? 이런 AI를 칸트가 마주하게 된다면 어떤 반응을 보였을까요?

**세네커 교수:** 아마 칸트는 그 자체(인격적 존재가 된 AI를 만드는 것)는 문제가 없을 것이라고 생각했을 것 같다. 인류는 오랜 세월 지구에서 살아가면서 수많은 ‘인격적 존재’들을 만들어내고 있다. 많은 사람들이 결혼을 하고 아이를 만들지 않는가(웃음). 과학자들이 인간과 가까운 수준의 AI를 만드는 것 자체는 큰 문제라고 생각하진 않았을 것이다. 칸트는 오직 인간만을 인격적 존재로 보는 ‘종차별주의자’가 아니었다. AI가 만약 정말로 자신이 자유로운 지성적 존재자임을 자각하는 존재가 된다면 칸트가 받아들이지 않을 이유는 없었을 것이라 본다. 물론 그것이 인간에게 위협이 되는 존재가 될 수 있느냐를 고려하는 것은 전혀 다른 문제이기도 하다.