

이찬규 교수, 인공지능에 ‘인문학’을 묻다 “결국은 인간의 문제, 협력해야”

✎ 박수연 기자 | Ⓞ 승인 2023.01.12 11:15



이찬규 중앙대학교 인문콘텐츠연구소 소장(국어국문학과 교수)가 11일 오후 서울 동작구 중앙대학교 인문콘텐츠연구소에서 한스경제와 인터뷰하고 있다. /사진=김근현 기자

[한스경제=박수연 기자] 중앙대학교 국어국문학과 교수이자 인문콘텐츠연구소 소장을 맡고 있는 이찬규 교수는 지난 2017년 국가과제로 주어진 ‘HK+’ 인공지능인문학 사업을 수행하고 있다.

그는 인공지능에 대한 관심이 커진 것이 오래되지 않았다고 설명한다. 2017년 인공지능과 이세돌의 바둑대회를 계기로 ‘앞으로 인공지능이 어떻게 나아갈 것인가’ 하는 고민이 본격화 됐고 2018년부터 국가에서 본격적으로 인공지능에 대한 심도 깊은 연구를 진행하게 됐다고 한다.

이에 따라 이찬규 교수도 인공지능인문학 구축을 위해 △인공지능 관계·소통학 △인공지능 인문데이터 해석학 △인공지능 윤리·규범학 △인공지능 기술 비평학 △인공지능 사회문화학 등 5가지 연구분과에 집중하고 있다.

인공지능은 수많은 방식으로 인간의 삶과 사회전반에 영향을 미치고 있다. 그래서 이 교수는 인공지능은 결국 ‘인간의 문제’, ‘사회의 문제’라고 말한다. 인공지능이 인문학적 성찰을 거쳐야 하는 이유이기도 하다.

하지만 여전히 인공지능과 인문학의 결합은 일반인들에게 낯설다. 이철규 교수에게 인공지능인문학이란 무엇일까.

“산업혁명 이후 기술의 시대가 도래했다. 기술은 1차, 2차, 3차 산업혁명을 거치면서 인간들이 필요로 하는 것을 달성하도록 했다. 가령 자동차는 공간적 제약을 보완해줬다”

“하지만 인공지능은 양상이 조금 다르다. 자동차가 인간을 대신해주지는 않지만 인공지능은 인간이 하는 것들 중 일정부분을 대체한다. 그런 점에서 인공지능은 인간의 문제라고 생각한다”

“인공지능의 발전이 결국 인간의 문제라면 아무리 기술이라 하더라도 인문학적 관점에서 관심을 가져야 하며 앞으로 인공지능이 사회에 큰 영향을 미칠 수 있는 만큼 우리가 선제적으로 검토해봐야 한다고 생각해 연구를 시작하게 됐다.”

이찬규 교수는 특히 인공지능의 ‘개발단계’에서 ‘설명 가능한 인공지능’을 만들기 위해 인문학과 윤리학이 강조되어야 한다고 말했다.

“인공지능은 크게 세 가지 영역으로 나뉜다. 하나는 데이터고 두 번째는 알고리즘, 세 번째는 컴퓨팅 파워 즉 하드웨어인데, 초기에는 알고리즘 쪽에 집중했다. 그러나 시간이 지나면서 알고리즘이 만들어낸 결과물들 중 굉장히 비인간적인 부분들이 발견됐다. 그렇다보니 설명이 필요한 인공지능을 만들어야 한다는 논의가 활발해졌다.”

“최소한 인간 사회에서 살아가면서 지켜야하는 윤리적이고 인문학적인 관점들이 기본적으로 인공지능에 반영이 돼야 한다는 것이다. 또 학습의 결과물들이 인간에게 해가 되지 않도록 하는 장치가 필요하다”

흔히들 인공지능은 감정이 없다고들 한다. 사람을 위로하는 로봇이나 인공지능은 어딘가 어색하기만 하다. 하지만 인문학에서 말하는 ‘감정’이란 굉장히 다채로우면서도 복잡한 영역이다. 인공지능이 인간의 감정을 구현할 수 있을까. 인공지능인문학연구소에서는 인공지능이 다채로운 감정 데이터를 구축할 수 있도록 하는 연구도 진행되고 있다.

“인간의 감정은 굉장히 다양하다. 한 문장과 한 단어가 가지고 있는 의미들이 복합적이기 때문이다. 가령 ‘슬프지만 후련하다’던가 ‘기쁘지만 짹짹하다’와 같이 인간 감정은 대부분 복합적으로 작용하고 있다. 그렇기 때문에 인공지능이 단순히 긍정감정, 부정감정 등 6가지 감정으로 나눠 표현한다면 기계적인 효율

성과 편의성은 높겠지만 그것은 인간의 감정이라고 보기는 어렵다.”

이찬규 교수는 인공지능의 정확한 감정 구축을 위해 6가지의 각 모델이 더해져 ‘멀티모달’을 구축해야 한다고 말한다.

“인공지능이 모델을 구축하기 위해서는 ‘좋다’, ‘싫다’ 등 단어만 가지고 표현할 수 있는 방법과 사람의 표정을 보고 학습하는 방법, 말을 할 때 나오는 어조를 파악하는 방법, 전체적인 상황과 맥락을 파악하고 문장단위로 데이터화하는 방법, 실제 사람들에 대한 심리검사 후 심리적인 반응을 구축하는 방법 등 6가지가 있다. 이를 모두 더한 것이 ‘멀티모달’이다. 멀티모달이 구축될수록 더 정확한 감정데이터를 구축할 수 있다.”



이찬규 중앙대학교 인문콘텐츠연구소 소장(국어국문학과 교수)가 11일 오후 서울 동작구 중앙대학교 인문콘텐츠연구소에서 한스경제와 인터뷰하고 있다. /사진=김근현 기자

이철규 교수는 ‘인공지능인문학’이라는 새로운 학문적 장르를 실현시킨 장본인이다. 하지만 여전히 인공지능은 나아가야 할 길이 멀다. 우리는 인공지능을 얼마나 신뢰할 수 있을까.

“인공지능의 오류 이미지에 관해 다룬 책이 있다. 인공지능이 이미지를 학습하는 중간에 오류 이미지를 만들어 내는 경우가 있는데, 이 때 얼굴이나 손이 구부러지는 등 이상한 이미지가 형성된다. 이러한 이미지가 잘못됐다고 생각하는 경우에는 사람이 바로잡아줘야 한다. 사실 이게 인공지능의 현주소다. 우리가

아직 인공지능을 완전히 신뢰하기 어렵다는 것이고 이를 고치고 바로잡는 것도 인간이기 때문에 결국 아직 모든 것을 인공지능에게 맡겨 놓기는 어렵다.”

이 교수는 현재 우리나라가 인공지능 연구를 주도하기 위해서는 갈 길이 멀다고 했다. 전 세계적으로도 ‘인공지능 시대’에 도래했다고는 보기 어렵다. 현재 인공지능은 데이터 구축과 알고리즘 기술을 얼마나 잘 구현하는가가 관건이라 볼 수 있다.

“2021년 미국의 Open AI가 ‘GPT-3’를 공개하면서 전 세계인들이 깜짝 놀랐다. GPT-3는 엄청난 양의 데이터들을 모으고 이를 학습시켜 사람들과 대화할 수 있도록 만든 자연어 처리 거대모델이라고 할 수 있다. 현재로서는 이 모델이 최첨단 인공지능이다. 그런데 곧 인간의 언어를 다 습득한 GPT-4가 나올 것이다. 인공지능이 한 개의 언어가 아니라 영어와 불어 중국어 한국어를 다 할 수 있다고 생각해보라. 어머어마하지 않은가. GPT-4가 인공지능 산업에 미칠 파장이 엄청 클 것이다”

“국내 기업들 중에서는 네이버가 ‘하이퍼 클로바’라는 거대 모델을 만들었지만 완전 공개되지는 않고 있다. GPT-3보다 한국어 구현은 더 잘된다고들 하지만 조금 제한적으로 사용되고 있다. 향후 GPT-4가 공개되면 국내 기업들도 이를 활용할 것이다. 현재 우리 인공지능 기술은 거의 응용기술이라고 보면 된다.”

이렇듯 인공지능은 나날이 발전해가고 있지만 이에 우려를 표하는 사람들도 적지 않다. 앞서 2020년에는 국내 AI 챗봇의 혐오발언이 도마 위에 오른 바 있다. 인공지능이 아이들의 교육에도 관여하고 있는 지금, 사람들은 인공지능이 ‘혐오’를 부추길 수 있다고 우려한다. 이러한 문제에 우리는 어떻게 대응해야 할까.

“데이터 수집 단계서부터 거칠게 모든 데이터를 다 수집해 인공지능을 학습시키는 것은 문제의 해결책이 될 수 없다. 당장 사람들이 쓴 글만 봐도 얼마나 편향적이고 혐오적인 발언들이 많이 들어가 있는가. 이런 것들을 다 인공지능에 때려 넣으면 인공지능도 이를 학습하게 된다.”

“그렇기 때문에 데이터 수집단계에서부터 ‘전처리 과정’이 중요하다. 온갖 비윤리적인 것들을 미리 걸러내서 양질의 데이터 전처리가 가능할 수 있도록 해야 한다.”

“작년에는 연구소에서 여러 기업들, 타 대학들과 함께 ‘비윤리데이터’라는 것을 구축했다. 비윤리데이터란 욕설과는 다르다. 오히려 단어로 이뤄진 욕은 수집해 걸러내기 쉽다. 하지만 ‘비윤리적인 표현’은 비하적인 표현부터 혐오표현까지 그 범위가 광범위하다. 수집한 비윤리데이터는 인공지능에서 삭제하는 데 사용된다.”

그럼 비윤리적 데이터는 어떻게 구축할까? 사람들은 저마다 ‘비윤리적’이라고 생각하는 부분에 차이가 있기 마련이다.

“가령 ‘멜로영화는 여자애들이나 보는 것’이라는 문장을 예시로 들었을 때, 누군가는 이 문장을 비윤리적인 문장으로 볼 수도, 누군가는 아닐 수도 있다. 우리는 여러 명의 사람들을 대상으로 투표를 진행해 ‘강도 측정’을 한다. 이 표현이 얼마나 비윤리적인지 평균값을 매겨 비윤리적데이터를 구축했다.”



이찬규 중앙대학교 인문콘텐츠연구소 소장(국어국문학과 교수)가 11일 오후 서울 동작구 중앙대학교 인문콘텐츠연구소에서 한스경제와 인터뷰하고 있다. /사진=김근현 기자

인공지능의 발전을 막연하게 두려워할 필요도 없지만 또 마냥 '도구'로만 볼 수도 없다. 인공지능을 기본적인 지식으로 이해하고 어느정도 활용할 수 있는 'AI 리터러시'를 갖춰야 한다. 다만 이 교수는 '협력의 상실'은 경계해야 한다고 말한다. 기계 기술 문명이 고도화될수록 '협력할 필요가 없는 사회'가 될 수 있음을 우려했다.

“인간이 유인원 시절, 나무에서 살다가 700만년에 걸쳐 지구의 주인이 되기까지 인간의 삶은 협력의 모델을 계속 만들어왔다. 하지만 기술이 고도화될수록 협력할 일이 없어져가고 있다. 그렇기 때문에 협력을 통해 만들어 온 인류 문명이 어떤 방향으로 가게 될지는 항상 고민해봐야 할 문제다”

“사람들은 접촉을 하지 않는다. 협력 없이는 경쟁만 남게 되는데, 인공지능이 발전할수록 인간 사이 거리는 더 멀어질 것이다. 인간 사이에 인공지능이 들어오는 순간 내가 누군지도 모르는 것과 그냥 이야기를 할 뿐 감정의 교류, 협력, 갈등의 조정은 상실될 수밖에 없다”

“인공지능이 시키는 대로 하는 것이 지배되는 것이 아니다. 인공지능이 우리 삶에 들어옴으로써 인간관계가 멀어지고 홀로 고립되는 것. 이것이 지배당하는 것이다”



박수연 기자 ddunip@sporbiz.co.kr

저작권자 © 한스경제(한국스포츠경제) 무단전재 및 재배포 금지