

HOME (/) > 기획:연재 (/news/articleList.html?sc section code=S1N3) > 21세기 학문의 신대륙을 찾아서 (/news/articleList.html?sc sub section code=S2N198)

“인공지능에게 인문학이 필요하다”... ‘사람 중심’ 시를 구현하다

임효진 | 승인 2024.04.09 09:11

21세기 학문의 신대륙을 찾아서 ③

‘포스트휴먼 시대의 인공지능 인문학’
이찬규 중앙대 인문콘텐츠연구소 소장

국가·사회 난제 해결에 앞장서고 있는 인문사회 연구의 대표적인 성과 사례를 소개한다. 기존 인문사회 학문 분야의 벽을 넘어선 새로운 문제의식과 융합으로 사회에 기여하고 있는 혁신 연구의 결과다. 인문사회통합성과확산센터(센터장 노영희 건국대 교수)는 교육부와 한국연구재단이 지원하는 인문한국플러스사업(HK/HK+)과 융합연구지원사업의 연구 성과 우수성, 파급효과 및 활용성을 종합적으로 고려해 6명의 심사위원이 우수성과 20곳을 선정했다.

인공지능인문학은 인공지능 윤리·규범학뿐만 아니라 인공지능기술 비평학, 인공지능 관계·소통학, 인공지능 사회·문화학, 인공지능 인문데이터 해석학으로 나뉘 연구한다.

2016년은 인공지능의 존재를 대중에게 각인시킨 해다. 이세돌과 알파고의 대국에 이어 세계 최초 로봇 시민권자인 소피아의 한국 방문으로 언론이 떠들썩했다.

인공지능에 대한 사회적 이목은 쏠렸지만, 학술적인 연구는 미비했고 논란과 근거 없는 불안만 확산했다. 인문콘텐츠연구소는 급속도로 발전하는 인공지능을 기술 문제만으로 바라볼 것이 아니라, 인문학의 관점으로 연구해야 한다는 성찰에서 ‘인공지능 인문학’의



이찬규 중앙대 국어국문학과 교수는 인문콘텐츠 연구소 소장, 중앙대 부총장, 국어학회 회장이다. 2017년 교육부 연구재단이 주관하는 HK+ 사업에 선정되면서 '인공지능인문학'이라는 분야를 새롭게 만들었다.

학문체계 구축을 시작했다.

이찬규 중앙대 인문콘텐츠연구소 소장은 "인공지능 기술과 산업에 대한 인문학적 성찰은 인간의 삶을 돕기 위해 출발한 기술이 인간의 삶을 위협하는 존재가 될 수 있다는 인식을 갖게 했다"라며 "인공지능 인문학은 지능적인 기계가 인간과 함께하는 현상을 바로 보고, 이 현상이 인간의 사상과 문화에 어떠한 영향을 미치고 어떻게 대처해야 할지 탐구하는 학문"이라고 밝혔다.

이번 연구는 「포스트휴먼 시대, 인문학 가치 고양을 위한 인공지능인문학 구축」을 주제로 인문학, 사회과학, 기술 공학 분야 등 다양한 전공의 공동연구원 30명과 HK 교수, 연구교수 등이 참여했다.

'사람 중심' AI 윤리기준 마련에 일조

연구단은 인공지능이 인간의 삶을 위협하는 존재가 아닌 도움이 되는 기술이 되도록 하기 위해서는 인공지능 개발에 기본적인 윤리적 기준을 세우는 일이 시급하다고 판단했다.

마이크로소프트사가 공개한 인공지능 챗봇 '테이'는 인종 차별, 폭력적인 메시지로 16시간 만에 서비스가 중단됐다. 메타가 공개한 '블렌더봇'도 기존 챗봇보다 인간적인 측면, 공감 능력이 개선됐다는 평가를 받았으나, 히틀러를 '위대한 사람'이라고 표현해 논란이 일었다.

인공지능 윤리에 대한 논란 이후 각국은 인공지능 윤리 기준을 마련했는데, EU는 '신뢰할 수 있는 AI', 중국은 '책임 있는 AI'를 제시했고, 한국 정부는 '사람이 중심이 되는 AI' 윤리 기준을 발표했다. 연구단은 인간 중심의 기술 사용에 대한 관점을 제공하는 것을 목표로 연구를 수행해 정부가 AI 윤리 기준을 세우는 데 일조했다.

정부는 인공지능의 최고 가치로 '인간성'을 설정하고, 3대 원칙과 10대 요건을 제시했다. 3대 기본 원칙은 △인간의 존엄성 원칙 △사회의 공공선 원칙 △기술의 합목적성이다. 10대 요건은 △인권 보장 △프라이버시 보호 △다양성 존중 △침해 금지 △공공성 △연대성 △데이터 관리 △책임 △안전성 △투명성이다.

중앙대 인문콘텐츠연구소는 이 3대 원칙과 10대 요건을 교육 현장에 적용해야 한다고 제안하며, 교육과정 예시를 만들었다. 교육과정에 적용할 수 있는 3가지 역량 요소인 개인적·사회적·합리적 도덕 역량을 정의했다.

이찬규 소장은 “우리 연구단은 인공지능 문제에 대해 적실한 학문적 대처를 할 수 있는 학문체계 구축을 선구적으로 시도했고, 학문 의의와 필요성을 학계와 정책 입안 그룹, 산업계에 알리고 공감대를 형성했다”라고 밝혔다.

인공지능인문학은 인공지능 윤리·규범학뿐만 아니라 인공지능기술 비평학, 인공지능 관계·소통학, 인공지능 사회·문화학, 인공지능 인문데이터 해석학 영역으로 나누어 연구한다. 여기에는 인공지능에 대한 철학적·윤리적 반성, 인공지능으로 인해 야기되는 사회·문화적 현상 탐구, 인공지능을 이용한 새로운 인문학 연구방법론 창출 등의 연구를 포함하고 있다.

어휘	뜻풀이	감동	공포	권태	놀람	분노	불안	사랑	설렘	섭섭	성취
소확행	'작지만 확실한 행복'이라는 뜻으로, 일상에서 누리	1.8	0	0	0	0	0	0.6	1.6	0	0
속다	어떤 것을 다른 것으로 잘못 알다.	0	0	0	1.67	0	0	0	0	0	0
완고·하다	융통성이 없이 옮고고 집이 세다.	0	0	0	0	0.2	0	0	0	0	0.2
용납	너그러운 마음으로 남의 말이나 행동을 받아들임.	0	0	0	0	0	0	0.4	0	0	0
징그럽다	하는 행동이 유독유독하여 역겹다.	0	0.6	0	0.6	0.8	0	0	0	0	0
찌뿌뜨다	표정이나 기분이 밝지 못하고 매우 언짢다.	0	0	0.2	0	0	0	0	0	0.8	0
처량·하다	마음이 구슬퍼질 정도로 외롭거나 쓸쓸하다.	0	0	0	0	0	0.6	0	0	0	0
추모	죽은 사람을 그리며 생각함.	0	0	0	0	0	0	0	0	0	0
충성	진정에서 우러나오는 정성. 특히, 임금이나 국가에	0	0	0	0	0	0	3	0	0	0
흥망	잘되어 일어남과 못되어 없어짐.	0	0	0	0	0	0	0	0	0	0

연구단이 만든 '감정 어휘 사전'의 일부를 캡처한 것이다. 이 사전은 CSV 파일 형식으로 돼 있으며, 연구단 홈페이지에서 누구나 내려받을 수 있다.

24개 감정기반 '감정 어휘' 사전 공개

인공지능 관계·소통학 연구에서는 감정 기반 감정 어휘 사전을 제작한 게 하나의 성과다. 미국 심리학자 폴 에크먼은 1960년대 연구에서 인간의 감정은 기쁨·두려움·혐오·분노·놀람·슬픔 6가지로 나뉜다는 '보편적 기본 감정' 이론을 주장했다.

하지만 중앙대 인문콘텐츠연구소는 인간의 다양한 감정을 구별할 수 없는 이런 데이터는 매우 제한적이라고 생각해 조금 더 인간의 감정에 가까운 데이터를 만드는 연구를 진행했다. 특히 한국인 특성에 기반한 연구로 총 24개의 감정을 분류했다.

그리고 공개 국어사전인 우리말샘의 수록어 중에 감정을 나타내는 단어를 선별, 24개의 감정에 관한 감정 어휘 사전을 구축했다. 해당 감정 어휘 사전은 2만 284개의 감정 어휘를 수록하고 있으며 각 어휘가 24개의 감정 중 어느 감정을 가졌는지를 0에서부터 5.0까지의 강도로 정의한다.

예를 들어 '마음이 가볍고 상쾌하다'라는 뜻의 '가뜩하다'는 설렘이 0.4, 성취가 1.4이다. '예절이나 형식에 얽매이지 아니하고 수수하고 털털하다'라는 뜻의 '소탈하다'는 재미가 0.5, 행복이 1.5, 평안이 4이다.

수치로 측정할 수 없는 감정이나 음의 높낮이에 따른 기분을 파악하는 건 인간 고유의 능력이었다. 하지만 인공지능이 이를 학습한다면 돌봄의 영역인 사회복지 분야 등에 이용할 수 있다.

연구단은 "이 데이터는 인간 이해와 관련한 AI 모델 학습용 데이터로 활용할 수 있다. 또한 학습용 원천 데이터로 쓰거나, 사전 학습된 모델에 미세 조정용 데이터로 적용할 수 있다"라고 밝혔다. 연구단은 활용 가능성을 시험하기 위해 이기성 중앙대 인문콘텐츠연구소 교수 주도로 감정 분석기를 직접 개발하고 있다.



박평중 중앙대 인문콘텐츠연구소 교수가 생성형 AI인 GAN을 이용해 사진처럼 만들어 낸 역사적 인물 이미지다. 왼쪽부터 태조 이성계, 선덕여왕, 울곡이이, 마리 앙투아네트의 이미지다.

알고리즘이 만들어 내는 '비틀린' 인간

인공지능 인문데이터 해석학에서는 박평중 중앙대 인문콘텐츠연구소 교수가 생성형 AI인 GAN을 이용해 생성한 오류 이미지, 역사적 인물 이미지와 관련한 전시회를 열었다. 「미증유의 얼굴: AI의 오류 이미지」 전시에는 인공지능 알고리즘이 자동으로 산출한 가상 인물의 얼굴이 나온다. 실제 인물과 구분이 어려울 만큼 사진과 흡사한 이미지를 만들어 내지만, 계산 착오로 얼굴에서 머리카락이 자라거나 눈동자가 흘러내리는 등의 비틀린 모습도 나온다.

「생성사진 프로젝트」에서는 초상화를 사진으로 만드는 작업을 진행했는데, 인간의 창작품과 달리 결국에는 모두 같은 결과물을 만들어 내는 것을 확인할 수 있었다.

이찬규 소장은 "인공지능은 곧 인간의 문제이므로 인문학적 접근이 필수적이라는 생각으로 이 사업을 시작했는데, 불과 몇 년 만에 이렇게 빨리 사람들이 인공지능인문학에 관심을 가질 줄 몰랐다"라며 "이 공론화된 울림이 국가 정책, 산업 기술에 본격적으로 적용되도록 융합연구 노력에 더 관심을 쏟겠다"라고 밝혔다.

연구 개요			
연구지원사업	인문한국(HK)플러스 : 2유형	연구기간	2017.11 ~ 2024.10
연구과제명	포스트휴먼 시대, 인문학 가치 고양을 위한 인공지능인문학 구축		
연구팀	중앙대 인문콘텐츠연구소(연구책임자 이찬규 교수)		
연구분야	인공지능 기술비평학, 인공지능 관계·소통학, 인공지능 사회·문화학, 인공지능 윤리·규범학, 인공지능 인문데이터 해석학		
연구성과	<p>아래의 모든 성과 집계는 2018년 05월부터 2024년 02월까지 기준</p> <ul style="list-style-type: none"> -논문 : 한국연구재단 등재지: 240편, SCI(SSCI): 23편, SCOPUS: 11편, A&HCI: 3편, 기타: 1편 -저서 : 국내발간: 19권, 해외발간: 2권 / -DB구축 : 한국어 감정 데이터 구축 및 공개 -학술대회 : 국제학술대회:6회, 국내학술대회:30회, 월례 심포지엄 / -생성형 AI를 활용한 이미지 실험 및 작품 전시 -국내 등재지 『인공지능인문학연구』 정기 발간 및 SCOPUS 학술지 『Jahr』 공동 발간 -인공지능인문학 대중 이해를 위한 칼럼 정기 웹진 발간 -지역인문학센터 : '인공지능인문학' 대중 강연 외에 매년 인문페스티벌을 통해 독후감·논문·콘텐츠 제작·생성 AI 프롬프트 대회 등 개최. 소외지역에 '찾아가는 인문학' 프로그램을 운영 -인공지능인문학 적용 결과물 : 감정 분석기 개발, 중첩 개체명 인식기 개발, VR 콘텐츠 2개 개발: 조선시대 과거제도 (연구소 자체 제작)와 정조 능행차 콘텐츠 제작(한국 콘텐츠출판진흥원 프로젝트로 제작) 		

임효진 기자 editor@kyosu.net

저작권자 © 교수신문 무단전재 및 재배포 금지



임효진

📄 인쇄