

테크

일반/정책

'데이터 댐'에 담은 AI 공공데이터 구축... 올해 2배로 늘려 지원

조선비즈 김윤수 기자

입력 2020.06.21 12:00

과기정통부, AI 학습용 데이터 구축 지원사업 20건 선정 추경예산 3000억원으로 하반기 150종 추가 계획

인공지능(AI)이 의료영상을 판독해 질병을 진단하거나 주변환경을 인식해 자율주행 성능을 높이도록 하는 학습용 공공데이터가 구축된다.

과학기술정보통신부는 올해 'AI 학습용 데이터 구축사업'의 20개 지원과제를 선정했다고 21일 밝혔다.

AI 학습용 데이터 구축사업은 정부가 AI 개발에 필요한 양질의 데이터를 대규모로 구축하고 개방하기 위해 2017년 시작한 사업이다. 현재까지 21종, 4650만건의 데이터를 구축해 AI허브(aihub.or.kr)에 공개하고 있다.

올해는 작년보다 선정 과제 수와 예산 모두 2배 늘어나 20개 과제에 총 390억원을 지원한다.

국립암센터는 유방암 진단을 위한 환자 의료영상 데이터를, 기업 '디노플러스'는 치매 진단을 위한 환자 뇌파 영상 데이터를 구축한다.

울산대는 자율주행 드론의 주변 인식 성능을 높이기 위해 관광지·도심지·산림지 등의 고화질(4K) 비행 촬영 데이터를 구축한다. '올포랜드'는 자율주행차에 활용할 수 있는 도로 파노라마 이미지 데이터 확보에 나선다.

딥페이크 영상을 잡아내기 위한 조작 영상, AI 문서요약 성능을 높이기 위한 텍스트 요약 샘플도 학습용 데이터로 구축된다.

과기정통부는 2925억원의 추경예산을 통해 올해 하반기 안에 학습용 데이터를 150종 추가로 구축해 총 191종까지 늘릴 계획이다.

날개의 데이터들을 모으고 가공하는 데 필요한 인력을 클라우드 소싱 방식으로 확보할 방침이다. 클라우드 소싱 방식은 모든 사람이 데이터 수집·가공에 참여할 수 있는 방식이다. 작년에 2개 과제에 시범 적용

한 데 이어 올해는 모든 과제에 확대 적용한다.

과기정통부 관계자는 "AI 학습용 데이터 구축을 대규모로 확대해 '데이터 댐'에 모으고 다양한 기술, 서비스 개발에 활용되도록 지원하겠다"며 "양질의 일자리 창출과 새로운 성장 원동력을 확보할 수 있을 것으로 기대한다"고 말했다.

2020년 AI 학습용 데이터 구축사업 지원과제 20종

지정공모 과제명		수행기관
1	대용량 동영상 콘텐츠 AI 데이터	KDX
2	자율주행드론비행영상 AI 데이터	울산대학교 산학협력단
3	시각정보 기반 질의응답 AI 데이터	유클리드소프트
4	수어 영상 AI 데이터	테스트웍스
5	한국인 대화음성 AI 데이터	슬루게이트
6	딥페이크 방지영상 AI 데이터	머니브레인
7	랜드마크 이미지 AI 데이터	피씨엔
8	사람 인체·자세 3D AI 데이터	스위트케이
9	문서요약 텍스트 AI 데이터	비플라이소프트
10	전문분야 한영 말뭉치 AI 데이터	플리토
자유공모 과제명		수행기관
1	질병진단 이미지 AI 데이터	국립암센터
2	도로환경 파노라마 이미지 AI 데이터	올포랜드
3	피트니스 자세 이미지 AI 데이터	슬릭코퍼레이션
4	K-Fashion 이미지 AI 데이터	오피니언라이브
5	한국인 재식별 이미지 AI 데이터	한국과학기술연구원
6	도로주행영상 AI 데이터	티큐에스코리아
7	치매진단 뇌파영상 AI 데이터	디노플러
8	감성 대화 말뭉치 AI 데이터	미디어젠
9	위성영상 객체판독 이미지 AI 데이터	한국항공우주연구원
10	구강악 2D·3D 이미지 AI 데이터	헬스허브

자료=과학기술정보통신부

AD [화제] 로또개발자曰 "이번주 1등 번호 꼭 나온다! 32, 43,!" 무조건 찍어라!

AD [추천]주식을 하고자 한다면 '이 규칙'만 지키면 月3천만원 수익 나온다.

Copyright © 조선비즈 & Chosun.com