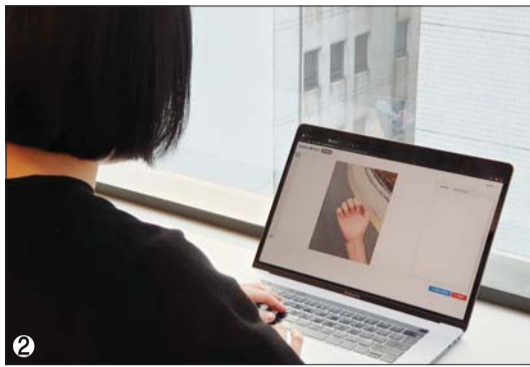


“데이터 가공하고 돈 벌어요” 클라우드 소싱 플랫폼 뜬다

정부, D·N·A 분야에 5조원 투입
AI 학습용 데이터 시장 성장 기대
클라우드웍스 작업자 12만명 보유
네이버·카카오 등 100개사 확보
셀렉트스타 어플 ‘캐시미션’ 주목
모바일로 편리하게 작업 ‘인기’
답내츄럴 3분기 만에 매출 10억
카이스트·서울대학교 등 고객사



1 답내츄럴은 AI 개발에 필요한 텍스트, 음성, 영상 학습 데이터를 생산하는 플랫폼인 ‘답내츄럴에이아이’ 서비스를 운영하고 있다
2 클라우드소싱 방식의 데이터 플랫폼에서 일을 하는 한 작업자가 클라우드웍스의 데이터 가공 작업을 진행하고 있다.
3 셀렉트스타는 모바일 데이터 가공 작업을 중심으로 하는 작업자용 플랫폼 앱 ‘캐시미션’을 운영하고 있다.



/각사

일반인들이 회원 가입 후 인공지능(AI) 학습을 위한 데이터 가공에 참여하고 작업 비용을 받는 클라우드소싱 데이터 플랫폼이 큰 인기를 얻고 있다. 클라우드웍스·셀렉트스타·답내츄럴 등 클라우드소싱 데이터 가공기업들은 국내의 데이터를 가공하는 작업자를 수만명에서 최대 10만명 이상 보유하고 있으며, 매주 수백명씩 국내의 데이터를 가공하는 작업자가 늘어나고 있다. 특히 최근 정부가 AI 일등국가로 도약한다는 비전을 내세우고, DNA(데이터·네트워크·AI) 분야에 올해 5조원을 투입하기로 하는 등 AI 활성화를 위해 힘을 쏟고 있으며, 다양한 산업에 AI 기술이 접목되면서 AI 학습용 데이터 시장도 큰 폭으로 성장할 것으로 전망된다. 이에 따라 클라우드소싱 데이터 가공기업들은 최근 고객사를 확대하면서 수십개에서 100여개사를 고객으로 확보하고 있다. 또 데이터 가공 프로젝트도 늘리고 있어 재택근무로 맡게는 수백만원까지 벌 수 있는 데이터 가공 작업

은 재택으로 할 수 있는 알바나 퇴근 후 일하는 ‘투잡’으로 인기를 끌고 있다. 국내 최대 데이터 가공 기업인 클라우드웍스는 유튜브 등에서 재택근무로 고소득을 벌 수 있는 알바로 유명세를 타면서 3월 현재 작업자수가 약 12만명에 달하고 있다. 이 회사는 2017년 4월 설립 후 클라우드소싱 데이터 플랫폼 서비스를 같은 해 10월 오픈했다. 현재 네이버, 카카오, 삼성전자, KT, SK텔레콤, LG CNS, SK C&C 등 대기업과 마인즈랩 등 스타트업, 카이스트, 포항공대 등 대학 등 100여개사를 고객으로 확보하고 있다.

김대영 클라우드웍스 최고운영책임자(COO)는 “많은 기업들에서 AI에 대한 수요가 폭발적으로 증가하면서, 딥러닝 학습에 필요한 많은 데이터는 기존 방식으로 감당하기 어려워 온라인 클라우드소싱 방식을 활용하고 있다”며 “AI 개발기업들이 우리와 작업하면서 데이터 준비 시간과 노력을 크게 절감할 수 있어 인기를 얻고 있다”고 설명했다. 클라우드웍스의 플랫폼 사업이 활기를 띠에 따라 지난해 한국투자파트너스, HB인베스트먼트 등 8개 투자사로부터 총 100억원 규모의 시리즈 B 투자를 유치하는 성과도 거뒀다.

셀렉트스타는 모바일 데이터 가공 작업을 중심으로 하는 플랫폼 앱 ‘캐시미션’을 운영해 주목을 받고 있다. 모바일 기반 작업자 플랫폼은 세계에서 최초라고 회사측은 설명했다. 황민영 셀렉트스타 영업 부문 대표는 “일반적인 데이터 가공 작업이 인터넷으로 진행되는 것과 달리, 모바일로 언제 어디서나 편리하게 작업을 해 돈을 벌 수 있다는 점에서 인기를 얻고 있다”며 “사용자 수만 우리나라와 필리핀을 합쳐 약 2만5000명이며, 매주 200명 이상 늘고 있다”고 밝혔다. 회사측은 일거리의 양과 작업자 수의 밸런스를 맞추

기 위해 마케팅을 자제하고 있음에도 입소문으로 가입하는 사용자들이 많다고 설명했다. 플랫폼 런칭 후 1년이 지난 현재 네이버, LG CNS, 롯데정보통신, 한국정보화진흥원 등 43여개의 고객사와 프로젝트를 진행했고, 총 1250만 건의 학습데이터를 생산했다. 플랫폼에 올라가는 프로젝트 수도 상시 3~4개에서 많을 때는 20개 정도가 올라가며, 동시 작업자수도 수백명에서 천 여명까지 참여하고 있다. 셀렉트스타는 이 같은 성과를 거두면서 카카오벤처스로부터 4억원의 투자도 유치했다. 답내츄럴은 지난해 9월 ‘답내츄럴 AI’ 플랫폼 오픈 베타 서비스를 시작한 후 3개월 만에 회원수 4000명을 돌파하는 성과를 거뒀다. 또 지난해 AI 데이터 수집 및 가공 서비스를 시작한 지 3분기 만에 매출 10억원을 달성했다. 이 회사는 카이스트, 서울대학교 등 학교 및 국립국립원 등 기관, 대상홀딩스와 같은 대기업 및 스타트업을 고객사로 확보하고 있다. 지난해에는 본엔젤스투자파트너스로부터 5억원의 투자를 유치하는 등 관심을 받고 있다. 박상원 대표는 “작업, 검수도구를 비롯한 플랫폼 전반에서 사용성을 높이기 위해 유저인터페이스(UI) 및 사용자경험(UX) 디자인에 많이 신경을 쓴 결과, UI/UX가 좋다는 피드백이 이어지고 있다”며 “AI 기술에 대한 이해와 활용 능력을 바탕으로 작업 효율을 높일 수 있으며 고품질 데이터 구축으로 이어지고 있다”고 설명했다. /채윤정 AI전문기자 echo@metroseoul.co.kr

“국내 최대규모 식물 유전체 DB구축 AI기반 신약 후보물질 개발 총력전”

AI 기업인
박 종 선 인포보스 대표

“
국내 자생 70종 유전체 DB구축 해외 638종 공개정보 구축 완료
1개 식물 분석에 최대 1억 소요 인포보스, 3000만원으로 줄여
현재 프리A 투자 유치 진행 중
완료면 자생식물 100종 추가분석



인포보스는 국내 자생식물에 대해 유전체를 해독하는 등 연구를 진행해 큰 성과를 거두고 있다. /인포보스

국내 최대 규모로 자생식물 유전체 데이터베이스(DB)를 구축하고 해외 식물 600여종에 대해서도 DB 구축을 완료한 인공지능(AI) 바이오 스타트업이 주목을 받고 있다. 주인공은 인포보스다. 이 회사는 식물에 대한 전장 유전체 정보DB를 활용해 AI 기술로 신약 후보물질을 발굴하는데 주력하고 있다. 박종선 인포보스 대표는 “전 세계에

서 해독이 안 된 국내 자생식물 70종에 대해 유전체를 해독해 DB로 구축했다”며 “해외 식물 638종에 대해서도 공개된 염기서열 등 유전체 정보에 대한 DB 구축을 완료했다”고 밝혔다. 그는 1개의 식물에 대해 유전체를 분석하기 위해서는 8000만~1억원 정도가 소요되는데, 인포보스는 이 비용을 3000만원 정도로 1/3로 줄였다고 강조했다.

그는 “국내 자생식물은 4500종에 달하지만 비용이 많이 들고 분석이 쉽지 않아 환경부, 농림부, 산림청 등 정부기관이 유전체를 해독한 국내 자생식물은 10여종 내외에 그쳐요. 우리는 현재 프리A 투자를 유치하고 있는데, 투자가 완료되면 추가로 국내 자생식물 100종에 대한 유전체 분석을 진행할 계획이에요.” 박 대표는 서울대학교에서 식물병리

학과 컴퓨터 공학을 전공한 후 유전체 연구로 서울대에서 박사학위를 취득했다. 이후 미국에서 암유전체 관련 벤처기업의 생물정보학 부사장을 맡았으며, 한국으로 돌아와 한 커피 로스팅 회사에서 연구소장을 지내면서 커피 유전체를 연구했다. 2015년 인포보스를 설립하고 환경부에서 국가생물종목록 시스템 관리용역을 수행하고, 여러 대학 및 기업과 함께 유전체 해독 프로젝트를 진행했다. 박대표는 창업 이후 유전체 DB 분석과 생물 다양성 자료 분석, 식물이 사는 환경을 모니터링하는 기후모니터링 시스템을 개발해왔다. 그러다가 기초과학탐사연구소를 운영하던 손장혁 대표를 공동 대표로 영입하면서 AI 기반의 신약 개발 사업으로 확대했다. “식물의 종 정보와 유용정보의 관계를 정리해 관계형DB를 만들었어요. 종 정보를 2800종 이상 수집했고 동의보감, 본초강목 등 문헌과 논문 등에서 ‘감기에 좋다’, ‘혈압에 좋다’와 같은 유용정보를 47만8000건을 얻어 관계형 DB로 구축·분석했어요.” 이 회사가 개발한 AI 유전체 정보분



석 시스템인 ‘메타프리-AI’는 유전체 정보에서 생화학 합성과정 정보를 기반으로 자생식물이 생산하는 2차 대사산물 후보를 정밀하게 예측하는 AI 솔루션이다. 1차 대사산물이 생존에 꼭 필요한 물질이라면, 2차 대사산물은 생존에 필수적이지는 않지만 필요에 의해 만들어지는 물질로, 다양한 효소가 관여한다. 하지만 현재는 2차 대사산물 관련 정보가 산재돼 데이터 활용에 어려움이 있어 신물질 발굴에 한계를 지닌다. “식물 유전체를 바탕으로 식물이 만들어내는 2차 대사산물이 무엇인지를 확인해요. 생화학합성 과정에 효소가 관여를 하는 데, AI로 관련된 효소 및 인자들의 존재여부를 예측하고, 실제 그 기능을 하는지 추정할 수 있어요.” /채윤정 AI전문기자