“AI 기술 주도권, 모델보다 ‘고품질 데이터’로 결정된다”

‘네이버 벤처스’ 기자간담회
이해진 “데이터 차별화 중요해져”
UGC 기반 플랫폼 재구축 필요
빅테크 기업, ‘데이터 라벨링’ 집중

8일(현지 시간) 미국 실리콘밸리에서 네이버 이해진 네이버 이사회 의장이 네이버 벤처스 네트워킹 행사에서 발표를 진행하는 모습. 이날 이 의장은 소버린 AI 구축을 위해 확보 중인 데이터의 중요성을 설명했다. /네이버

“검색도 처음에는 알고리즘 싸움이었지만 결국 다 비슷해지고, 데이터를 갖고 차별화하는 것이 중요해졌습니다. 인공지능(AI) 또한 비슷한 일이 생길 것입니다.”

10일 〈메트로경제신문〉 취재에 따르면 이해진 네이버 이사회 의장은 지난 5일(현지시간) 미국 실리콘밸리에서 진행 된 네이버의 첫 해외 투자법인 ‘네이버 벤처스’ 기자간담회에서 이같이 말했다. 이 의장의 발언은 AI 기술의 주도권이 이제 ‘누가 더 똑똑한가’가 아니라 ‘누가 더 고품질의 데이터를 가졌는가’로 이동했음을 시사한다.

AI 개발 기업의 데이터 싸움은 오래 전부터 이어졌다. GPT-4나 제미나이(Gemini), 클로드(Claude) 등 거대언어모델(LLM)은 모델 구조 자체보다는 얼마나 방대한·다양한·정제된 데이터를 학습했는가가 성능을 좌우한다. 특히 최근 AI 개발의 초점은, 다른 LLM이 갖지 못한 고품질 데이터를 선점해 학습하는 쪽으로 옮겨가고 있다.

이 의장은 “AI 시대에도 차별화의 핵심은 결국 데이터”라고 강조하며, 사용자 제작 콘텐츠(UGC)를 통한 네이버의 검색 경쟁력을 사례로 들었다.

그는 AI 생태계에서도 블로그·카페 같은 UGC 기반 플랫폼의 재구축이 중요하다고 보고, 중고거래 플랫폼 ‘포쉬마크’ 인수 또한 상거래 데이터 확보를 위한 전략의 일환이라고 설명했다. 기초 AI 모델은 미국·중국이 선점했지만, 한국 내수 특화 AI 영역에서는 네이버가 데이터 경쟁력을 기반으로 충분히 승부를 걸 수 있다는 판단이다.

특화 데이터 확보와 함께 글로벌 빅테크 기업들은 최근 학습 데이터에 정확한 이름을 붙이는 ‘데이터 라벨링(DATA labeling)’ 기술 확보에 집중하고 있다. 데이터 라벨링은 인공지능(AI)에 학습시키는 데이터를 정확히 어떤 사물과 객체인지 분류하는 작업을 말한다.

가령, 거리에서 자전거를 탄 소녀를 찍은 사진이 있다면 ‘거리’ ‘자전거’ ‘소녀’ 등으로 분류해 라벨링 하는 식이다. 최근까지도 해당 작업은 사람의 수작업으로 이뤄졌는데, 주요 빅테크 기업이 비용 절감을 위해 폭력적이고 음란한 콘텐츠까지 포함해 개발도상국에 하청을 준 사실이 드러나 논란이 일기도 했다. 최근에는 이같은 작업 또한 AI를 통해 자동화 하는 추세다.

블룸버그 통신이 지난 8일(현지시간) 보도한 바에 따르면 최근 메타는 인공지능(AI) 학습 데이터 정제 기업 스케일AI에 최대 100억달러(약 13조6000억원)의 투자를 추진 중이다. 스케일AI는 MS·아마존·엔비디아 등으로부터 투자받은 비상장 기업으로, 데이터 라벨링 기술에 특화했다. 과거 데이터 라벨링 기업에 데이터를 구매했으나 더 나아가 원천 기술을 확보하려는 전략으로 풀이된다.

AI 학습 데이터 가공 업체 관계자는 “최근 학계에서도 AI 학습 데이터가 점차 고갈되고 있다는 연구 결과가 잇따르고 있다”며 “이제는 다른 LLM들이 확보하지 못한 차별화된 데이터나, 정밀하고 고품질의 데이터를 얼마나 갖고 있느냐가 핵심이 될 것”이라고 설명했다. 이어 “LLM 모델 성능이 상향평준화된 지금, 결국 미래 경쟁은 데이터 확보 전쟁으로 갈 수밖에 없다”고 밝혔다.

/김서현 기자 seoh@metroseoul.co.kr

좋은 질문이 좋은 검색을 만든다

김서현 기자의 AI시대 적응하기

검색창에 “어떤 집이 좋을까”라고 묻고선 속으론 ‘어떻게든 지금 사는 집이 나쁘지 않다’는 결과를 바란 적 있다. 검색은 정보를 찾는 게 아니라, 위로받는 의식일 때가 있다. 그리고 인공지능(AI) 시대, 알고리즘은 당신보다 앞서가 위로를 전한다.

“힘들었죠? 이 결과를 봐요. 당신은 틀리지 않았어요.”

책으로 정보를 찾던 때를 지나 처음 온라인 세상이 펼쳐지고 검색엔진이 등장했을 때, 우리는 정보의 바다에 던져졌다. 그때 우리는 몇 가지 키워드를 조합함으로써 검색 결과를 획득했다. 내가 만든 홈페이지가 검색결과 최상단에 뜨기도 하던 시기였다.

시간이 흘러 AI 시대에 이른 현재, 우리는 더 이상 ‘검색어를 입력하는 사람’이 아니라 알고리즘이 조합한 정보의 결과물 속을 걷는 사람이 되었다. 정보를 거르는 데에 필요한 것은 여전히 검색이지만, 검색에 앞서 한 마디의 질문이 더 필요해졌다. 중요한 건 키워드가 아니다. ‘질문’을 던질 수 있느냐다.

검색은 기술이지만 동시에 습관이고, 더 나아가 권력이다. 우리는 늘 검색한다고 믿지만, 사실은 검색하는 방식을 배운 적이 없다. 그래서 우리는 깊은 생각을 않고 검색어를 떠올려 검색하곤 한다. 그러나 현대 온라인 환경은 ‘어떤 키워드로 묻느냐’에 따라 완전히 다른 결과가 나오고, 묻는 방식에 따라 사람의 확신과 불안, 때론 행동까지 바뀐다. 이제는 정보를 어디서 찾느냐보다 어떤 질문이 필요한지 돌아볼 때가 된 것이다.

‘왜?’를 떠올렸다면 검색창에 질문하는 법은 어렵지 않다. 검색창에 키워드를 넣기 전, 이렇게 자문해보자. “내가 이걸 왜 궁금해하게 됐지?” “이 정보는 나에게 어떤 결론을 밀어줄까?” 그리고 가장 중요한 질문, “이게 진짜 내가 알고 싶은 건가?”

질문은 검색보다 먼저 와야 한다. 검색은 답을 주지만, 질문은 방향을 바꾼다. 검색창 앞에서 멈춰 서는 몇 초가, 판단력의 출발점이 될 수 있다. 정보가 넘치는 시대, 결국 우리에게 필요한 건 더 나은 검색어가 아니라, 더 나은 질문이다. /seoh@

팀네이버, 사우디 3개 도시 디지털 트윈 구축

스마트도시 기반 마련

팀네이버가 사우디아라비아 메카·메디나·제다 3개 도시를 대상으로 디지털 트윈 플랫폼을 구축했다고 10일 밝혔다. 이로써 총 6800㎢, 건물 수 92만 동에 달하는 대규모 도시 데이터를 기반으로 스마트시티 구현의 기반이 마련됐다.

이번 프로젝트는 사우디 자치행정주택부 산하의 발라디(Balady)와 NHC 이노베이션이 주도하며, 팀네이버가 기술을 지원하고 있다.

디지털 트윈 플랫폼은 3D 모델링, 지형·일조량·경관 분석, 건축 사전 검토, 홍수 시뮬레이션 등 도시 계획에 필요한 다양한 기능을 포함하고 있다. 고해상도 이미지와 실시간 데이터 통합으로 현지 도시 관계자들의 의사결정과 운영 효율성을 높이고 있다.

발라디는 이를 국가 차원의 통합 플랫폼으로 확장할 계획이다.

야세르 알로바이단 발라디 CEO는 “이 플랫폼은 스마트하고 지속가능한 도시를 위한 사우디의 비전을 보여주는 사례”라고 밝혔다.

한상영 네이버클라우드 글로벌 DX&이노베이션 부문장은 “현지 파트너와 함께 기술 고도화와 적용 범위를 넓혀 갈 것”이라고 말했다.

이번 프로젝트는 2023년 10월 계약을 체결하고 2024년 7월 본격 착수됐으며, 네이버클라우드와 NHC 이노베이션은 ‘네이버 이노베이션’이라는 합작법인을 설립해 관련 사업을 전개 중이다.

/김서현 기자

카카오
브런치 작가 멤버십 참여자 사전 모집

카카오가 운영하는 브런치스토리가 오는 7월 정식 출시 하는 유료 콘텐츠 구독 서비스 ‘브런치 작가 멤버십’에 참여할 작가를 사전 모집한다고 10일 밝혔다.

브런치 작가 멤버십은 독자가 마음에 드는 작가를 구독하고, 작가가 제공하는 멤버십 프리미엄 콘텐츠를 월 3900원에 감상할 수 있는 유료 구독 기능이다.

브런치 작가 멤버십은 브런치 구독자 30명 이상, 3개월 내 글 3개 이상 발행한 작가라면 누구나 신청 페이지에서 신청하기를 눌러 신청할 수 있다.

멤버십 작가로 선정되면 유료 구독자 수에 따라 매월 구독료를 정산 받아 수익을 창출할 수 있다.

/김서현 기자

배민 “‘로드러너’ 전국확대 계획 없어”

“현재 베타 서비스 단계”

배달의민족(배민)이 성과제를 기반으로 한 ‘로드러너’ 시스템 전면 도입으로 라이더들에 대한 통제권을 강화하려 한다는 주장은 사실이 아니라고 10일 반박했다.

배민의 물류 서비스를 전담하는 우아한청년들은 이날 〈메트로경제신문〉에 로드러너 앱을 본격적으로 도입할 계획이 현재로서는 없다고 밝혔다.

로드러너는 독일계 모기업 딜리버리히어로가 개발한 라이더 전용 애플리케이션으로, 기존 자체 앱인 ‘배민 커넥트’와 가장 큰 차이점은 배달원에게 평가 등급을 부여해 경쟁을 유도한다는 것이다.

배민이 오산과 세종 등지에서 시범 운영해온 이 시스템은 시간당 평균 배달 건수, 주문 수락률 등의 평가 지표를 토대로 라이더에게 1부터 6까지 등급을 매긴다. 이로 인해 라이더들이 ‘더 빨리, 더 많이’ 배달해야 ‘똥콜(단가 낮고 난도 높은 호출)’을 피할 수 있다는 우려가 제기됐다.

우아한청년들 관계자는 “로드러너는 배달 서비스 개선을 위해 배민이 진행하고 있는 다양한 정책의 일환”이라며 “현재 베타 서비스 단계로 전국 단위 확대 등의 계획은 아직 구체적으로 없다”고 선을 그었다.

배민은 과거에도 유사한 방식의 ‘리워드’ 정책을 운영해왔다. 리워드 제도는 2년 전 ‘배달고수클럽’으로 개편됐다. 배달고수클럽은 목표 달성을 강제하는 것이 아닌 라이더가 선택해 참여하는 프로모션으로, 노조와 협의를 거쳐 합당한 보상을 제공한다고 회사 측은 강조했다. /김현정 기자 hjk1@

KT, 초고속 인터넷 가입자 1000만명 돌파

오는 8월 31일까지 고객감사 이벤트

KT는 국내에서 처음 초고속 인터넷 서비스 누적 가입자 1000만명을 돌파했다고 10일 밝혔다. 지난 1994년 상용 인터넷 서비스 ‘코넷(KONET)’ 시작 이후 31년 만이다.

한국지능정보사회진흥원의 지난해 인터넷 이용실태 조사에 따르면 우리나라는 전체 2227만가구 중 2226만가구가 유무선 인터넷을 이용하며, 가구 인터넷 접속률은 9.97%에 이른다.

KT는 1000만 가입자 달성을 기념해 오는 8월 31일까지 고객 감사 이벤트를 실시한다. 먼저 장기 고객을 대상 구형 기기 무상 업그레이드 혜택을 마련했다.

KT 초고속 인터넷을 3년 이상 이용한 고객 중 와이파이 공유기가 없거나 구형 기기를 사용하는 경우 베이직(500Mbps) 요금제 가입자에게는 기가

KT 초고속 인터넷 설치 기사와 고객이 ‘인터넷도 역시 KT 페스티벌’을 홍보하고 있다. /KT

와이파이 홈 공유기를 증정한다. 또 에센스(1Gbps) 이상 가입자는 와이파이 7D 공유기를 3년 약정시 무료로 받는다. 선착순 5만명에게 제공되는 혜택이다.

또 지니TV 3년 이상 이용 고객 중 구형 셋톱박스를 사용하고, 지니TV 슬림 이상 요금제 가입자에게는 안드로이드 기반 최신 셋톱박스인 지니TV 셋톱박스 A를 무상 교체해준다. /김서현 기자