

‘백인남성’ 자아 답습한 AI?… IT업계, 편향성 개선 주력

훈련 데이터 속 사회적 편견 학습
불평등 심화·결과물 품질 영향

카카오, 리스크 관리체계 ASI 구축
네이버, 잠재적 위험 인식·평가·관리
NIA-TTA, 유해표현 데이터 공개

지브리풍 이미지 생성으로 화제를 모은 챗GPT 등 생성형 AI는 성별이나 인종 정보를 명시하지 않아도 대부분 ‘백인 남성’을 주인공으로 설정한다. 이는 반복 실험에서도 동일한 결과를 보인다. AI 모델들이 훈련 데이터에 내재된 사회적 편견을 학습해 성별, 인종, 나이 등에 따른 차별을 재생산하는 구조 때문이다.

21일 <메트로경제> 취재 결과 AI 모델의 차별, 혐오표현, 편향성 등이 여전히 문제시 되는 가운데 이를 해결하기 위한 다양한 대책이 쏟아지고 있다.

AI 모델의 편향성은 훈련 데이터에 내재된 사회적 편견을 그대로 학습해 성별, 인종, 나이 등에 따른 차별을 학습, 재생산하는 현상을 뜻한다. 데이터의 불균형과 알고리즘 설계자의 무의식적 편향이 주요 원인으로 꼽힌다.



거의 모든 AI 모델은 인구통계학적 정보가 없는 명령어를 제시해도 항상 ‘백인·남성’을 그린다.
/캡처

AI 모델의 편향성은 생성형 AI 대중화 이전부터 문제로 지적돼 왔다. 중국 난징항공우주대 천지성 연구원은 AI 채용 시스템이 효율성을 높이면서도 비백인 남성에 불리한 차별적 결과를 낳았다고 밝혔다. 그는 “AI의 편향은 데이터셋과 알고리즘 설계자의 편견에서 비롯되며, 기존 사회적 불평등을 강화할 수 있다”고 경고했다.

헬스케어 분야에서도 단일 인종·성

별 중심의 학습은 특정 집단에 대한 과잉 해석 등 왜곡을 유발할 수 있다는 우려가 제기된다.

편향으로 인한 각종 문제는 사회적 불평등을 심화시키고, AI 모델이 생성하는 결과물의 품질에도 영향을 미친다. 이에 따라 AI 모델을 개발하는 기업들은 적극적으로 편향을 해결하기 위해 노력하고 있다.

카카오는 지난해 10월 AI 기술 개발

과 운영 과정에서 발생할 수 있는 리스크를 사전에 식별하고 관리하는 리스크 관리 체계인 ‘카카오 AI 세이프티 이니셔티브(카카오ASI)’를 구축했다. 카카오 ASI는 ▲Kakao AI 윤리 원칙 ▲리스크 관리 사이클 ▲AI 리스크 거버넌스라는 3가지 핵심요소로 구성된다.

카카오 관계자는 “카카오 ASI는 지난해 5월 개최된 ‘AI 서울 서밋’에서 선언한 ‘서울 AI 기업 서약’을 위한 실천에 해당한다”며 “AI 윤리와 안전을 위해 다양한 선제적 노력을 이어갈 예정”이라고 설명했다.

네이버 또한 AI 시스템을 개발하고 배포하는 과정의 모든 단계에서 AI의 잠재적 위험을 인식·평가·관리하기 위한 대응 체계를 갖고 있다.

네이버의 AI 모델 하이퍼클로바X는 레드팀 테스트, 전문 데이터 확보, 강화 학습 알고리즘 개선 등을 통해 생성 정보의 정확성, 편향성, 안전성 등을 개선하고 있다. 또한 타 문화권과의 협업을 통해 특정 문화의 특징과 사회적 맥락을 학습하는 데도 주력하고 있다.

정부 또한 AI 모델의 편향성을 개선하기 위한 노력을 이어가고 있다. 한국

지능정보사회진흥원(NIA)와한국정보통신기술협회(TTA)는 대규모 언어모델(LLM) 학습용 데이터 내 ‘유해 표현 검출 인공지능 모델 및 유해 표현 학습용 데이터’를 공개하기도 했다.

공개된 인공지능 모델은 주어진 말뭉치 데이터 내에서 유해 표현을 검출하고 다양한 유해 표현 카테고리로 분류하는 기능을 수행할 수 있다. 과학기술정보통신부의 초거대 AI 확산 생태계 조성 사업의 일환으로, LLM 데이터의 품질 검증 과정에서 ‘유해성’을 측정하기 위해 개발됐다. 자체 성능 검증 결과, 해당 모델은 80~90% 수준의 높은 정확도를 기록하며 유해 표현 검출의 실효성을 입증했다.

학습용 데이터는 ▲유해 표현 검출용 데이터 20만 건 ▲유해 표현 카테고리 구분용 데이터 21만 건으로 구성돼 있다.

황종성 NIA 원장은 “인공지능 기술의 발전은 기술적 진보뿐만 아니라 사회적 책임을 담보할 수 있는 윤리적 고려가 반드시 병행돼야 한다”며 “이는 지속 가능한 인공지능 생태계 조성을 위한 핵심 요소”라고 강조했다.

/김서현 기자 seoh@metroseoul.co.kr

네이버, SME 기반 로컬커머스 생태계 확장

(중소상공인)

日 법인 통해 ‘야놀자솔루션’ 인수
외식업 시장서 경쟁력 확보 기대
로컬커머서도 UX 구축 움직임

네이버가 외식업 디지털 솔루션 기업 ‘야놀자에프앤비솔루션’을 인수하며 로컬 커머스 생태계 확장에 속도를 낸다.

21일 IT 업계에 따르면 네이버가 일본 현지 법인인 ‘네이버 J.Hub 코퍼레이션’을 통해 야놀자의 자회사인 야놀자에프앤비솔루션을 인수했다. 거래금액은 약 100억원 규모로 알려졌다.

야놀자에프앤비솔루션은 외식업 매장을 위한 줄서기 관리 시스템 ‘나우웨이팅’, 모바일 주문 서비스 ‘야오더’, 고객 마케팅 솔루션 ‘도도포인트’ 등 다양한 디지털 솔루션을 운영하고 있다.

초기에는 중소형 매장 중심의 디지털 전환 기업으로 주목받으며 카카오가 주요 주주로 참여했던 바 있으나, 2020년 야놀자가 경영권을 인수한 이후 현재의 사명으로 변경됐다.

그러나 이후 적자가 지속되면서 전략적 투자 유지 또는 매각 가능성이 꾸준히



야놀자에프앤비솔루션이 모바일 기반 주문 솔루션 ‘ya(야)오더’ 홍보 웹사이트.

히 제기돼왔다. 2023년 기준 매출은 63억 원, 당기순손실은 83억 원에 달했으며, 지난해 야놀자가 매물로 내놓으면서 매각이 본격화됐다. 이미 시장 점유율 60%를 차지한 티오더 등 기존 사업자와 하이오더(KT), 배민오더(우아한형제들), 토스플레이스(비바리퍼블리카) 등 새로운 플레이어들이 테이블오더 시장 경쟁에 뛰어들면서 경쟁이 격화됐다. 그러나 결과적으로 네이버가 인수에 나서면서, 양사는 새로운 전략적 전환점을 맞게 됐다.

이번 인수는 네이버의 SME(중소상공인) 기반 로컬 커머스 사업 확장의 일환으로 해석된다. 네이버는 ‘네이버플

레이스’, ‘네이버 스마트주문’, ‘네이버페이’ 등 다양한 플랫폼과 솔루션을 이미 운영 중이며, 여기에 야놀자에프앤비솔루션의 기술과 인프라를 더해 외식업 시장에서 보다 견고한 경쟁력을 확보할 수 있을 것으로 기대된다.

특히 최근 네이버는 ‘네이버플러스스토어’ 같은 별도 쇼핑앱을 출시하며 초개인화 기반의 커머스 전략을 강화하고 있으며, 로컬 커머스 영역에서도 ‘예약·주문·결제·리뷰’로 이어지는 통합된 사용자 경험(UX)을 구축하려는 움직임이 뚜렷하다. 야놀자에프앤비솔루션의 웨이팅 및 주문 기능은 이와 같은 전략에 적합하다는 평이다.

인수 주체가 일본 법인이라는 점도 눈에 띈다. 단순 국내 서비스 연계를 넘어, 일본 시장을 겨냥한 포석으로 보인다. 일본 외식업계 역시 테이블오더, 웨이팅 관리 등 디지털 솔루션 도입 수요가 높은 만큼, 야놀자에프앤비솔루션의 기술과 네이버의 플랫폼 역량이 결합할 경우 현지 시장에서 새로운 기회를 창출할 수 있다는 분석이다.

/김서현 기자

LG U+, ‘캐치! 티니핑’ 한정판 굿즈 판매

A16 케이스, 그림톡 세트 선택

LG유플러스가 27일까지 유플러스닷컴에서 SAMG엔터의 인기 IP ‘캐치! 티니핑’ 한정판 굿즈 세트를 단독판매 한다.

따뜻한 이야기로 어린이들 사이에서 큰 사랑을 받는 캐치! 티니핑은 유튜브와 OTT에서 누적 조회수 13억회를 넘긴 인기 애니메이션이다.

이번 세트는 휴대폰 케이스(A16 모델) 또는 그림톡, 레디백, 넥스트랩, 갤럭시 테마, 스티커 세트, 다이어리로 구성된다. 고객은 A16 케이스와 그림톡



LG유플러스 모델이 캐치! 티니핑 굿즈 세트를 소개하는 모습.
/LG유플러스

중 하나를 선택할 수 있다. 포토카드는 LG유플러스 고객에게만 증정된다.

/김서현 기자

LG CNS, 수학적최적화 기술확산 가속화

‘최적화 통한 경영혁신·미래전략’ 포럼
플랫폼 ‘MOP’, 광고업계 성과 공개

디지털 전환(DX) 전문 기업 LG CNS가 수학적최적화 기술로 다양한 산업군의 난제 해결에 나섰다.

LG CNS는 지난 18일 여의도 콘래드 호텔에서수학적최적화 기반의 혁신 사례와 기술 트렌드를 공유하는포럼을 진행했다.

수학적최적화는 가능한 모든 선택지를 검토해 최선의 해답을 도출하는 인공지능(AI) 핵심 기술이다. 의사결정 효율화와 비용 절감을 동시에 달성할 수 있다는 장점을 갖췄다. AI 기술의 진화 형태인 ‘에이전트 AI’ 구현에 필수 요소로 주목받으면서 최근 글로벌 빅테크 기업들도수학적최적화 기술을속속 도입하고 있다.

이번 포럼은 ‘최적화를 통한 경영 혁

신과 미래 전략’을 주제로, 수학적최적화 기술 확산을 가속화하기 위해 마련됐다. 제조, 교통, 금융, 물류등 각분야 산업군의 관계자가 참석해 기술 적용 방안과 효과를 논의했다.

LG CNS는 포럼에서 디지털 마케팅 플랫폼 ‘MOP’를 기반으로 광고 업계에서 낸 성과를 공개했다. 이 플랫폼은 검색 및 디스플레이 광고 전략을 수학적 최적화 알고리즘으로 분석해 셀러의 광고 효율을 극대화한다. 고객이 가장 활발히 검색하는 시간이나 최적의 광고 노출 위치, 경쟁사 입찰 정보 등을 종합적으로 살펴 광고 전략을 수립한다.

이날 LG CNS에 따르면, 2023년 5월 MOP 출시 후 2년 만에 2000여개 고객사를 확보하는 성과를 거뒀다. 기업 고객이 MOP를 통해 취급하는 광고 금액은 3000억원을 넘어섰다고 회사 측은 전했다.

/김현정 기자 hjk1@

SKT, 10년 우수고객 위한 ‘숲캉스’ 행사

에버랜드 포레스트 캠프서 진행
봄·가을 시즌, 총 7500명 규모 운영

SK텔레콤이 이달부터 장기 이용 고객을 위한 프리미엄 이벤트 ‘스페셜 T 숲캉스’ 운영을 시작했다고 21일 밝혔다. 숲캉스는 에버랜드 포레스트 캠프를 하루 동안 전용 공간으로 개방해, 10년 이상 SKT 이용 고객과 가족에게 자

연 속 힐링 경험을 제공하는 감사 프로그램이다.

올해 봄 시즌은 4월 19일부터 5월 25일까지 총 13회, 가을 시즌은 9~10월 진행된다. 한 회당 최대 5인까지 참여 가능하며, 총 7500명 규모로 운영된다. 응모는 T월드 내 ‘스페셜 T’ 페이지를 통해 가능하다.

숲캉스 현장에서는 도시락, 간식, 피

크닉 키트 등 전원에게 풀세트 지원이 제공되며, 숲 산책, 테라리움 클래스, 글라이더 날리기 등 맞춤형 체험 프로그램도 진행된다.

특히 지난해에는 130:1의 경쟁률을 기록하며 장기 고객 사이에서 큰 인기를 끌었다.

SKT는 2024년부터 ‘스페셜 T’를 정식 론칭하고, 가입 연수에 따라 데이터 지급, 30년 이상 고객에게는 T멤버십 V IP 자동 부여 등 다양한 장기 고객 우대 정책을 시행 중이다.

/김서현 기자