

# 검색부터 구매·예약까지... '실행형 AI' 구현

네이버, 테크 딥독서 AI랩 기술 소개 '프로덕트 네이티브 LLM' 정의 역할별 SLM 분업으로 효율화 검색 다음단계 '멀티모달 AI' 제시

네이버가 생성형 인공지능(AI) 경쟁의 초점을 '더 똑똑한 모델'에서 '더 일 잘하는 서비스'로 옮겼다. 검색 결과를 제시하는 수준을 넘어 쇼핑과 예약, 정보 탐색까지 이어지는 '실행형 AI' 구현을 위해 AI랩 전용 언어모델과 AI 에이전트 운영 기술, 멀티모달 AI를 공개했다.

네이버는 지난 2일 서울 강남 네이버D2SF에서 '탐색에서 실행까지, 차세대 AI 기술이 만드는 네이버 AI 검색'을 주제로 테크 딥독을 열고 AI랩 핵심 기술을 소개했다. 이날 발표에는 이기창 네이버클라우드 하이퍼스케일 AI 모델 2 이사, 한승균 네이버 AI 검색 서비스 리더, 윤상두 네이버 퓨처 AI센터 리더가 각각 모델과 서비스, 멀티모달 기술을 설명했다.

세 발표자의 설명은 서로 다른 기술을 다뤘지만 공통된 방향은 명확했다. AI 모델 자체의 성능보다 실제 서비스에서 얼마나 빠르고 정확하게 사용자의 일을 대신 수행할 수 있는지가 차세대 AI 검색 경쟁력이라는 것이다.

## ◆ "벤치마크보다 서비스"... 검색·예약에 맞춘 AI 모델

첫 발표에 나선 이기창 이사는 AI랩에 적용된 차세대 모델을 '프로덕트 네이티브 LLM'이라고 정의했다.

그는 "우리가 지향하는 것은 모든 벤치마크에서 1등을 하는 모델이 아니다"라며



네이버 AI 검색 테크 딥독서 이기창 이사가 AI랩에 적용된 차세대 모델을 '프로덕트 네이티브 LLM'이라고 설명하고 있다.

"사용자가 검색하고 구매하고 예약하는 실제 서비스 순간에 가장 잘 작동하는 모델을 만드는 것이 목표"라고 말했다.

이를 위해 네이버는 기존 '하이퍼클로바X'를 기반으로 검색과 쇼핑, 플레이스, 생활정보 데이터를 학습 초기부터 반영했다. 모델 구조도 대규모 서비스에 맞게 개선했다. 전문가 혼합(MoE) 구조를 적용해 긴 문맥에서도 응답 속도를 유지 하도록 했고, 사용자 경험과 직결되는 E2E 레이턴시를 줄여 처리 효율을 높였다.

강화학습 방식 역시 서비스 중심으로 바뀌었다. 실제 검색과 예약 서비스를 학습 환경에 연결해 AI가 여러 도구를 활용해 사용자의 과업을 끝까지 수행하도록 설계했다. 질문이 모호하면 먼저 추가 질문을 하도록 학습시키는 '명료성 강화 학습(Clarify RL)'도 적용했다.

네이버 자체 평가에서는 검색·구매·예약 등 실제 서비스 수행 능력인 '서비스 역량'이 글로벌 동급 모델 평균(100점)보다 높은 108점을 기록했고, 지시 이행과

도구 호출 등 기본 역량도 104점을 기록했다.

## ◆ "좋은 LLM 하나로 부족"... AI의 '일머리' 설계

한승균 리더는 AI 검색 서비스에서 중요한 것은 모델보다 '일머리'라고 표현했다.

그는 "좋은 LLM 하나만 있으면 AI 서비스가 완성될 것이라고 생각하기 쉽지만 실제 서비스를 만들어 보면 그렇지 않다"고 말했다.

네이버는 이를 위해 AI가 어떤 모델을 사용할지, 어떤 도구를 호출할지, 답변을 어떻게 구성할지를 설계하는 '하네스 엔지니어링'을 구축했다.

AI랩은 질문이 입력되면 안전성 검증부터 사용자 의도 분석, 검색과 플레이스·쇼핑 등 서비스 호출, 출처 제공, 액션 카드 생성까지 여러 단계를 거친다. 이 과정에서 하나의 거대 언어모델이 모든 작업을 처리하는 대신 역할별 소형언어모델(SLM)이 분업하는 구조를 채택했다.

그 결과 일부 시스템은 운영 비용을 최대 3배 줄이고 응답 속도는 2배 이상 개선했다. 또한 새로운 소형 모델이 개발되면 필요한 부분만 교체할 수 있어 서비스 중단 없이 성능을 높일 수 있다.

## ◆ "AI의 눈도 진화"... 이미지 이해 넘어 예약까지

윤상두 리더는 AI 검색의 다음 단계로 멀티모달 AI를 제시했다.

그는 "앞으로 AI 에이전트는 텍스트뿐 아니라 이미지를 통해서도 사용자의 의도를 이해하고 실제 행동까지 연결하는 방향으로 진화할 것"이라며 "스마트렌즈를 통해 10년 가까이 축적한 시각 검색 기술이 AI 에이전트의 '눈' 역할을 하게 될 것"이라고 말했다.

예를 들어 사용자가 영상 속 카페를 보여주며 "이런 분위기의 카페를 우리 동네에서 예약해 달라"고 요청하면 AI는 이미지 속 공간 분위기와 인테리어를 이해한 뒤 플레이스와 예약 서비스를 연계해 실제 예약 가능한 장소를 추천하는 방식이다.

이를 가능하게 하는 핵심 기술은 멀티모달 임베딩이다. 이미지와 텍스트를 하나의 의미 공간으로 변환해 AI가 사진과 문장을 함께 이해하도록 하는 기술이다.

윤 리더는 최근 세계 최고 권위 컴퓨터 비전 학회 CVPR에서 발표한 'MuCo' 연구도 소개했다. 하나의 이미지를 반복 처리하지 않고 여러 차례 이어지는 질문의 문맥까지 이해하도록 설계한 기술로, 향후 실행형 멀티모달 AI의 핵심 기반이 될 것으로 기대된다고 설명했다.



vitna@metroseoul.co.kr metro

## SKT 15GW급 AIDC 구축 2Q 인프라 허브 도약

영남권에 2GW 규모 클러스터 구축



SK텔레콤이 아시아 지역 내 인공지능(AI) 인프라 허브 도약에 나선다. SK텔레콤은 15GW 규모의 AI 데이터 센터 구축(AIDC)을 추진한다고 5일 밝혔다.

이번 사업은 정부가 추진하는 AI 3강 전략과 지역 균형 발전 정책에 발맞춰 국내에 15GW급 AI 데이터 센터를 추진하고, 아시아 지역에서 AI 인프라 허브로 도약하기 위한 기반을 마련하는 게 목표다.

이에 대해 정재현 SKT CEO(사진)는 "이번 AI 데이터센터 구축은 글로벌 AI 생태계가 필요로 하는 컴퓨팅 인프라를 선제적으로 준비하기 위한 것"이라며 "정부·산업계·지역사회와 긴밀히 협의해 대한민국이 아시아의 핵심 AI 인프라 허브로 성장하는 데 기여하겠다"고 했다.

먼저, 울산에 짓고 있는 1호 AI 데이터 센터를 시작으로 영남권 전체에 2GW 이상 규모의 클러스터를 구축한다. 이 곳은 글로벌 빅테크의 AI 인프라 수요를 한국으로 유치하는 거점으로 활용한다. 서남권에는 1GW를 추가로 구축해 국내에 총 5GW 규모의 AI 데이터 센터를 2029년부터 단계적으로 여는 것을 추진한다. 2035년에는 총 15GW 규모의 데이터 센터가 순차적으로 가동할 수 있도록 한다는 계획이다.

AI 데이터센터 구축에 약 70조원에 이르는 대규모 사업비가 투입될 전망이다. 회사는 전략적 파트너 투자, 고객사 장기 계약, 프로젝트 파이낸싱 등을 통해 조달한다. 다만, 초기 투자 부담 및 사업 위험 최소화를 위해 수요와 투자 여건을 고려할 계획이다.

## LG U+ '익시오' 이용 고객은 해외서도 통화 무료

'익시오 로밍콜' 선포

LG유플러스는 해외 무료 통화 서비스 '익시오로밍콜'을 출시한다고 5일 밝혔다.

이번 서비스는 AI 통화앱 '익시오'를 이용하는 고객이 대상이다. 로밍 요금제 가입 또는 와이파이 환경에서 음성 통화를 이용할 경우 국제 통화요금 부담 없이 사용할 수 있다.

'익시오로밍콜'은 해외 여행이나 출장 시 발생하는 통화 요금 부담을 줄이고 국내와 유사한 통화 환경을 제공하기 위해 기획했다. 특히 해외 체류 고객들은 국내에 있는 가족, 지인, 업무 관계자들과 부담 없이 통화할 수 있도록 했다.

LG유플러스는 이번 출시를 위해 해외 통신사와 품질 점검을 진행했다. 로밍 환경에서도 안정적인 통화 품질을 확보하고 익시오 주요 기능이 원활하게 작동하는지 사전 검증했다. 특히 국내 통화 품질과 유사한 수준의 이용 경험을 제공하는 데 집중했다.

이 서비스는 먼저 일본에 선보이고 하반기 중 동남아·유럽 등 약 100개국으로 순차적으로 확대할 계획이다.

## 통신업계, '5G 과장광고' 후속판결에 촉각

공정위, 표시광고법 위반 과징금 부과 LG U+, 취소소송 항소심 패소 KT·SKT 행정소송 결과 영향 '주목'

통신업계가 5G 과장광고에 대한 후속 판결을 예의주시하고 있다. LG유플러스가 공정거래위원회의 과징금 부과 결정에 불복해 제기한 행정소송에서 패소하면서다.

5일 통신업계에 따르면, LG유플러스는 최근 공정거래위원회를 상대로 제기한 시정명령 및 과징금 납부명령 등 취소소송 항소심에서 패소했다. 법원은 5G 서비스 속도를 거짓·과장하고 기만적으로 광고했다는 공정위 결정이 적법하다고 봤다.

이 같은 결정에 법원 선고를 앞둔 KT·SK텔레콤의 행정소송 결과에도 관심이 쏠리고 있다. 재판부가 이론상으로만 설명 가능한 5G 서비스의 최고 속도를 실제 소비자들이 이용 가능한 속도처럼 광고하는 것은 과장광고에 해당한다는 취지로 공정위 판단을 인정해서다.

공정위는 지난 2023년 7월 통신 3사가 광고에 사용한 표현이 표시광고법을 위반한 것으로 보고 총 336억 1000만원의 과징금을 부과했다. 한 달 뒤 KT가 이 같은 처분을 취소하는 행정소송에 나섰고, 이어 SK텔레콤과 LG유플러스가 각각 같은 취지로 처분 취소를 청구했다.

통신 3사는 5G가 본격적으로 상용화되기 전부터 'LTE보다 20배 빠른 5G'를 강



/AI로 생성한 이미지

조하며 5G가 궁극적으로 지향하는 최대 속도를 20Gbps로 제시했다.

이후 상용화가 시작된 2019년에는 최대 지원속도를 강조하며 가입자 유치에 나섰다. 통신사가 제시한 최대 지원속도는 SK텔레콤 2.7Gbps, KT 2.5Gbps, LG유플러스 2.1Gbps 등이다. 이는 1GB 크기의 일반 화질 영화를 약 3초 만에 내려받을 수 있는 수준이다. 그러나 공정위는 5G 평균 속도가 각사가 내세운 최대 지원속도의 약 3% 수준인 800Mbps에 불과하다고 봤다.

통신 3사는 최대 속도 20Gbps에 대해서는 이론상 최고 속도에 불과하며 광고에는 정부의 행정지도에 따라 실제 속도가 사용 환경에 따라 달라질 수 있다는 점을 표시했다고 주장했으나 받아들여지지 않았다.

이번 LG유플러스 판결에서 법원은 5G를 구현할 환경이 갖춰지지 않은 상태에서 현재형 광고 문구를 사용했다는 점을 근거로 들어 거짓·과장 광고 및 기만적 판

고에 해당한다고 봤다. 또 경쟁사 대비 속도가 빠르다는 내용의 비교광고에 대해서도 적절하지 않다고 판단했다.

LG유플러스는 당시 '5G 속도측정 또 다시 1등'이라는 문구를 내걸었으나 정부 조사 결과는 정반대였다. 2020년 8월 정부가 진행한 5G 품질평가에서 전속속도 측정 결과 LG유플러스는 SK텔레콤과 KT의 뒤를 이었다.

이번 판결은 SK텔레콤과 KT가 제기한 행정소송 결과에도 적지 않은 영향을 미칠 전망이다. 쟁점이 이론상 최고 속도와 최대 지원속도를 실제 환경에서 구현 가능한 성능처럼 표현했는지 여부에 집중해서다.

다만 이번 판결은 항소심으로, 절차상 대법원의 최종 판단이 남아 있다. 이에 따라 공정위 처분의 적법성이 인정되면 과징금 처분도 최종 확정된다. 과징금 규모는 SK텔레콤이 168억 2900만원으로 가장 크다. LG유플러스가 28억 5000만원으로 가장 적고, KT는 139억 3100만원이다.

한편 이번 행정소송과 별개로 법원은 지난 5월 740여 명의 소비자들이 통신 3사를 상대로 제기한 손해배상청구 1심에서는 회사 측 손을 들어줬다. 원고는 5G 커버리지에 대한 설명의무를 위반해 계약 무효 등을 주장했지만, 재판부는 커버리지 제한 등은 초기 상용화 단계에서 발생하는 기술적 한계로 판단했다.

/조민선 기자 msjo@



지난 3일 기상청 서울청사에서 열린 '위험기상 등 정보 확산 및 취약계층 지원을 위한 업무협약식'에서 권대열 카카오 지속가능경영 총괄리더(왼쪽)와 이미션 기상청장이 협약 체결 후 기념 촬영을 하고 있다. /카카오

## 카카오-기상청 카톡으로 폭염 알린다

폭염 취약계층 모금 캠페인도 진행

카카오가 기상청과 손잡고 위험기상 정보 확산과 폭염 취약계층 지원에 나선다. 카카오는 지난 3일 서울 기상청 청사에서 기상청과 '위험기상 등 정보 확산 및 취약계층 지원을 위한 업무협약(MOU)'을 체결했다고 5일 밝혔다. 협약식에는 이미션 기상청장과 권대열 카카오 지속가능경영 총괄리더가 참석했다.

이번 협약에 따라 카카오는 카카오톡 채널과 카카오프렌즈, 카카오페이지 등을 활용해 위험기상 정보를 신속하게 전달한다.

사회공헌 플랫폼 카카오킴이 가치를 통해 탑상공동체복지재단과 함께 폭염 취약계층 지원을 위한 모금 캠페인도 진행한다. 무더위에 취약한 이웃들의 건강과 안전을 위한 물품 지원 등에 활용될 예정이다.

/최빛나 기자