11월 11일 AI 아침 브리핑

국내 소식을 먼저, 건설·BIM 관련 AI 동향을 함께 정리했습니다. 둘째 항목은 건산연 브리핑을 인용한 일본 건설 AI 데이터 전략 분석으로, BIM 실무와 직접 맞닿아 있습니다.

1. 엔비디아·KAIST 공동연구소, 우리말 AI 학습 데이터 100만 건 구축·공개 계획

구분: 국내 AI 생태계 / 2026-11-10, 연합뉴스 공동인터뷰

- 김현우 엔비디아-KAIST AI 공동연구소장이 11월 10일 인터뷰에서, AI 모델 학습에 쓸 우리말 데이터를 내년까지 문서 단위로 100만 건 이상 구축해 공개하겠다고 밝혔습니다. 합성 데이터 생성 기술과 적은 데이터로 학습 효과를 높이는 방법론 연구에 주력합니다.
- 공동연구소에는 엔비디아가 초기 5년간 총 3억 달러(약 4,000억 원)를 투입하고, 매년 5,000만 달러 상당의 AI 컴퓨팅 자원을 제공합니다. KAIST 학생 대상 연간 10명 이상 인턴십과 국내 연구원 채용도 진행 중입니다.
- 한국어를 출발점으로 삼은 이유는 영어·중국어와 극소수 언어 사이의 '중간 자원 언어'라 다른 언어로 확장하기 좋다는 설명입니다. 지난 4월 공개된 '네모트론 페르소나 코리아' 데이터 세트가 허깅페이스 세계 다운로드 1위를 기록했고, LG AI연구원의 엑사원 2.5도 이를 토대로 학습한 것으로 소장은 알고 있다고 말했습니다.
- 주의: 100만 건 구축·공개는 내년 목표이며 현재 완료된 결과물이 아닙니다. 엑사원 관련 언급은 LG 측 공식 확인이 아닌 소장의 인터뷰 발언입니다.

제안 (BIM 연계): 공개 우리말 데이터가 쌓이면 건설 기준·시방서 같은 국내 전문 문서를 다루는 한국어 모델의 바닥 성능에도 영향을 줄 수 있습니다. 공개 시점에 건설 도메인 적용 실험을 검토해 볼 만합니다.

인터뷰하는 김현우 엔비디아-KAIST AI 공동연구소장

이미지: 연합뉴스 인터뷰 기사에 실린 김현우 공동연구소장 사진. 출처: https://www.yna.co.kr/view/AKR20261110107900017

2. 일본은 건설 AI 데이터 체계부터 재편, 한국은 BIM·기준 디지털화 중심

구분: 건설·BIM AI / 2026-11-10, 건산연 건설동향브리핑 인용 뉴시스 보도

- 한국건설산업연구원의 건설동향브리핑에 따르면, 일본 국토교통성이 9월 '인프라 분야 AI 실장 추진 방침'과 '건설 분야 피지컬 AI 전략(중간 정리)'을 공개하고 모델 개발보다 데이터 기반 정비에 속도를 내고 있습니다.
- 구체적 조치: 토목공사 공통사양서를 PDF에서 HTML+JSON-LD로 전환해 AI가 기준을 정확히 검색하게 하고, 5월부터 직할 토목 설계·조사 용역에 생성형 AI 활용 계획서 제출과 AI 작성 성과품 표시를 의무화했습니다. 2025년 11월에는 국토교통 데이터 플랫폼에 MCP 서버를 열어 35개 시스템 약 336만 건 데이터를 AI가 자연어로 검색할 수 있게 했고, 제조사가 다른 건설기계를 공통 제어신호로 연동하는 자율시공 기반 'OPERA'도 14개 기관 공동연구로 진행 중입니다. 2029년 봄부터는 건축확인·검사에서 IFC(BIM) 데이터 직접 심사를 추진합니다.
- 한국은 건설기준 디지털화와 BIM 기반 자동검토에 정책이 집중돼 있어, 데이터 생산-납품-유통-활용 전 과정을 AI 친화적으로 바꿔야 한다는 지적을 브리핑은 담고 있습니다.
- 주의: 건산연 브리핑을 인용한 보도로, 일본 각 조치의 실제 현장 정착도는 별도 확인이 필요합니다. 2029년 IFC 심사 등은 추진 계획 단계입니다.

제안 (BIM 연계): 납품 성과품의 '데이터 자산화'와 MCP 기반 공공 데이터 연결은 국내 BIM 납품·유통 체계 논의에 바로 참고할 수 있는 설계입니다. 특히 마크다운 등 구조화된 형식으로 보고서를 납품하는 방식은 지금 브리핑 자료 관리에도 적용해 볼 여지가 있습니다.

아파트 공사 현장의 타워크레인

이미지: 기사에 실린 공사 현장 자료 사진(2024년 촬영분)으로, 일본 사례 현장은 아닙니다. 출처:
https://www.fnnews.com/news/202611101344091826