

CMC, 베트남어 법률 대형 언어 모델(LLM) AI 선도

(NIPA, 하노이IT지원센터)

- CMC 그룹 소속 CMC OpenAI(C-OpenAI), 베트남어 법률 대형 언어 모델 CMC-AI-Legal-32B 공개 및 자체 개발 VLegal-Bench 평가 기준 소개
 - AI는 2025년 6월 총리가 공포한 전략 기술 목록 11개 기술 그룹 중 하나로, 베트남어 LLM, 가상 비서, 전문 AI 등 제품군을 포함*
 - * 제1131/Qđ-TTg 결정, 전략 기술 목록 및 전략 기술 제품 공포(2025.06.12.)
 - 베트남 법적 틀에 따른 AI 표준화 및 국내화의 중요한 진전을 보여주며, 분야별 AI 표준화 및 평가에 대한 강력한 수요 반영

□ 베트남어 법률 대형 언어 모델(LLM) CMC-AI-Legal-32B

- 베트남 법률 시스템에 맞춰 기존 모델을 기반으로 한 심층 조정 구축
- 베트남 법률 맥락에서 일반 LLM(예: GPT-4o, Claude, Gemini)의 한계를 보완하는 것을 목적으로 함
- VLegal-Bench 기준, CMC-AI-Legal-32B가 최고 성능을 달성, 22개 과제 중 6개 선도, 복잡한 법률 추론 문제에서 특히 강점을 보임
- 핵심 역량
 - 법률 조항 및 규범 문서 이해 · 분석
 - 다단계 법률 추론
 - 구조화된 법률 내용 작성
 - 베트남 맥락에 따른 법률 상황 처리

□ VLegal-Bench*, 베트남어 법률용 대형 언어 모델 AI 평가 기준

- * 공식 웹사이트 : <https://vlegalbench.cmcai.vn/>
- 베트남 법률 특성에 맞춰 설계된 정답 포함 10,450개 데이터 샘플 구성
- VLegal-Bench는 C-OpenAI에서 다양한 법률 AI 모델을 비교 · 평가하기 위한 “경쟁의 장” 으로 활용, 오픈 소스 · 폐쇄형 소스 · 법률 분야 특화 모델
- arXiv에 프리프린트 형태로 공개된 기술 보고서로, 독일 국제 파트너 및 SRH 하이델베르크 대학교 Vu Duc Binh 교수 연구팀과 협력 · 지문을 통해 수행
- 22개 과제 범위로 구분

- 점진적 5단계 추론 수준 설계
 - 인식 및 회상
 - 이해 및 구조화
 - 추론 및 논리적 도출
 - 해석 및 생성
 - 윤리, 공정성 및 편향

□ 향후 개발 계획

- VLegal-Bench 소스 코드 및 데이터 공개
- AI 모델 비교 결과 웹사이트 확장
- 2026년 단계별 오픈 버전 배포
- 명확한 검증 체계와 표준 플랫폼 구축을 목표로, 전문 AI 응용 개발 지원 및 베트남 AI 생태계 지속 가능 발전에 기여

[참고자료]

■ 베트남어 법률 AI 대형 언어 모델 선도적 위치 확립 (2025.12)

<https://vjst.vn/xac-lap-vi-the-tien-phong-mo-hinh-ngon-ngu-lon-ai-phap-ly-tieng-viet-77872.html>