

국산 AI반도체 성능지표 'K-Perf', 공공 현장 도입 검증에 첫 활용

- 과기정통부, 한전-리벨리온-TTA 국산 AI반도체 전력분야 기술협력 협약식 참석
- K-Perf를 전력현장 AI 영상분석에 활용, 성능 평가·검증 영역 확대

【관련 국정과제】 22. 초격차 AI 선도기술·인재 확보

과학기술정보통신부(부총리 겸 장관 배경훈, 이하 '과기정통부')는 10월 2일 금요일 리벨리온 본사에서 열린 한국전력공사(이하 '한전'), 리벨리온, 한국정보통신기술협회(이하 'TTA') 간 「K-Perf^{활용} 국산 AI반도체(NPU) 기반 전력분야 AI 영상분석 기술협력 업무협약(MOU)」 체결식에 함께 참석했다고 밝혔다.

이번 협약은 국산 AI반도체를 실제 전력현장에 적용하고, 현장 운영 환경에서 성능과 신뢰성을 검증하기 위한 협력체계를 마련하기 위해 추진됐다. 협약에 따라 한전은 전력현장 시범운영과 활용모델 개발을, 리벨리온은 NPU 성능 최적화와 기술지원을, TTA는 전력현장 운영환경을 반영한 성능·신뢰성 검증을 맡는다.

한전은 기존 변전소별로 분산 운영하던 AI 영상분석 인프라를 NPU 기반 통합 관제 시스템으로 전환하고 실증을 추진할 예정이다. 이를 통해 칩 단위 성능을 넘어 실제 운영 환경에서의 국산 AI반도체 활용 가능성과 처리 성능, 비용 절감 효과 등을 검증할 계획이다.

특히 이번 협약은 국산 AI반도체 성능지표 'K-Perf'가 실제 도입 과정의 검증 수단으로 활용되는 첫 사례라는 점에서 의미가 크다.

'K-Perf'는 국산 AI반도체 성능에 대한 신뢰성을 높이고, 수요기업이 도입 판단에 필요한 성능 데이터를 객관적으로 제공하기 위해 마련된 지표다. 작년 12월 15개 수요·공급기관이 참여해 실제 사용 환경을 중심으로 측정

모델과 시험 조건, TTFT*·TPOT**·TPS/1user*** 등 세부 지표를 정해 발표했다.

* TTFT(Time to First Token) : 사용자 응답 지연 최소화 여부 확인

** TPOT(Time per Output Token) : 긴 답변 생성시 속도 안정성 확인

*** TPS/1user(Throughput per Second/1user) : 단일 사용자 시간당 처리량 확인

이어 올해 6월에는 과기정통부와 TTA, 정보통신기획평가원(IITP)이 국산 AI반도체 2종을 대상으로 AI 챗봇, 문서 검색 등 이용자들이 많이 사용하는 대표적인 생성형 AI 서비스 유형별 시범 검증을 진행해, 수요 기업들이 요구하는 서비스별 성능 기준(SLO)을 충분히 만족하는 것을 확인한 바 있다.

시범 검증을 마친 지 수개월 만에 수요기관인 한전이 한전이 이번 실증에서 K-Perf를 활용한 성능평가·검증을 추진하면서, 성능지표가 실험실을 벗어나 실제 도입 현장에서 쓰이게 됐다. 또한 그간 생성형 AI 서비스 위주로 시험해 온 K-Perf를 전력현장 AI 영상분석 환경에 적용한다는 점에서, 검증 영역을 넓히는 계기가 될 전망이다.

이도규 정보통신정책실장은 "K-Perf의 검증 결과가 도입 판단의 근거가 된다는 것은 수요기관과 공급기업이 하나의 기준으로 성능을 논의하는 체계가 현장에서 작동하기 시작했다는 의미"라며, "국가 전력망을 책임지는 한전이변전소 AI 영상분석 실증에 K-Perf를 활용하는 만큼, 다른 수요 기관으로 확산되는 데 강력한 레퍼런스가 될 것"이라고 말했다.

이어 "정부도 국산 AI반도체의 기술개발과 성능 검증이 실제 시장 활용으로 이어지는 생태계 구축에 정책 역량을 집중하겠다"며, "K-Perf를 지속 고도화해 수요기관이 믿고 도입할 수 있는 환경을 만들고, 이번 협력의 성과를 공공부문 전반으로 넓혀갈 수 있도록 적극 지원하겠다"고 밝혔다.

담당 부서	정보통신정책실 정보통신방송기술정책과	책임자	과장 이강용 (044-202-6230)
		담당자	사무관 손희수 (044-202-6232)
관계 기관	한국정보통신기술협회 AI 인프라단	책임자	단장 신일진 (sssjin12@tta.or.kr)
		담당자	팀장 한주연 (hanjy@tta.or.kr)

