

## LLM 특별세션

일자\_ 2025년 6월 19일(목) 11:10~12:10, 6월 20일(금) 09:50~10:50

장소\_ 제주 신화월드 랜딩 컨벤션 센터 LG층 랜딩볼룸A

## 프로그램

시간	발표주제	발표자(소속)
6월 19일(목)		
11:10~12:10	지능형 AI Agent의 구조와 진화: 자율성과 목표 기반 행동 설계	이경진 대표이사 (엑스투)
6월 20일(금)		
09:50~10:50	HyperAccel LPU: 생성 AI 시대 도약을 위한 AI 반도체	김주영 대표/교수 (하이퍼엑셀/카이스트)

## 강연 소개



## 지능형 AI Agent의 구조와 진화: 자율성과 목표 기반 행동 설계

이경진 대표이사

엑스투

- (現) 엑스투 대표이사
- (前) 카카오 엔터프라이즈 대표이사
- (前) 카카오 엔터프라이즈 부사장
- (前) 블릭스 소프트 CTO

지능형 AI Agent는 환경을 인식하고, 판단하며, 스스로 행동할 수 있는 자율적 시스템입니다. 초기 AI는 규칙 기반이었지만, 현대 AI는 학습과 적응 능력을 갖춘 구조로 진화해왔습니다. 구조적으로는 지각(Perception), 계획(Planning), 추론(Reasoning), 행동(Actuation) 모듈로 구성됩니다. 자율성은 외부의 지속적인 개입 없이 목표를 달성할 수 있는 능력을 의미합니다. 목표 기반 행동 설계는 Agent가 주어진 목적에 따라 상황을 분석하고 전략을 수립하게 합니다. 강화학습과 목표 기반 계획(goal-oriented planning)은 자율성과 유연성을 높이는 핵심 기술입니다. 인간-AI 협업을 고려한 설명 가능한 AI(XAI)도 Agent 설계의 중요한 방향으로 부상 중입니다. 멀티에이전트 시스템에서는 여러 AI들이 협력하거나 경쟁하면서 더 복잡한 문제를 해결할 수 있습니다. 최신 연구는 감정, 윤리, 사회적 규범 등 인간 중심의 요소를 AI Agent에 통합하는 방향으로 진행 중입니다. 궁극적으로 지능형 AI Agent는 다양한 산업과 일상생활에서 주도적인 역할을 할 것입니다.



## HyperAccel LPU: 생성 AI 시대 도약을 위한 AI 반도체

김주영 대표/교수

하이퍼엑셀/카이스트

- (現) 주식회사 하이퍼엑셀 대표이사 (2023.01 - 현재)
- (現) 카이스트 전기및전자공학부 교수 (2019.09 - 현재)
- (前) Microsoft Azure 엔지니어링 리더, 책임연구원 (2012.02 - 2019.08)
- IEEE/ACM DAC Engineering Best Presentation Award 수상 (2023)
- 인공지능반도체 산업유공 과학기술정보통신부 장관표창 (2022)

글로벌 인공지능(AI) 산업은 트랜스포머 기반 대규모 언어 모델(LLM)의 등장으로 인해 급격한 변화를 맞이하고 있으며, 이는 AI의 전체 판도를 새롭게 정의하고 있습니다. 이번 발표에서는 LLM 기반의 생성형 AI 추론 서비스를 위한 새로운 컴퓨팅 과제들을 살펴보고자 합니다. 아울러, 이러한 도전에 대응하기 위해 발전해 온 최신 AI 칩 기술의 동향도 함께 논의할 예정입니다. 또한, LLM 추론 연산에 최적화되어 설계된 첨단 반도체인 HyperAccel의 LPU(LLM Processing Unit)의 핵심 기술을 심도 있게 소개합니다. LPU의 혁신적인 기능과 설계 철학을 통해 LLM이 요구하는 성능 조건을 어떻게 충족하고 있는지를 살펴보고, 이를 통해 더 비용 효율적이고 확장 가능한 AI 솔루션의 가능성을 모색해보고자 합니다. 끝으로, 생성 AI 시대를 선도하기 위한 우리의 반도체 기술 자립화 전략도 함께 고민해 보고자 합니다.