

## 글로벌 ICT

## 주간동향리포트

정책/규제

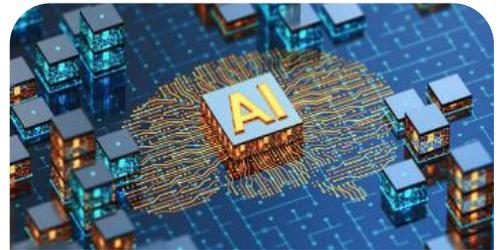
프로젝트/투자

인기기술/인기제품

신기술/제품

ICT 기업

## AI, 데이터센터 확장 및 엣지 컴퓨팅 혁신 촉발



### 인공지능(AI) 기술 발전에 따른 데이터센터 확장 필요

최근 AI 기술의 급속한 발전으로 인해 데이터 처리 및 학습에 필요한 연산 자원이 급격히 증가하고 있는데, 이에 따라 증강된 전력 및 냉각밀도(cooling density)가 추가된 대규모 데이터센터의 확장 필요성이 제기되고 있음. IT 인프라 운영 기업 하이퍼스케일 클라우드 서비스(hyperscale cloud service)가 운영중인 대규모 데이터센터는 약 50MW에서 200MW 이상의 전력을 소비하고 있으며, 향후 AI 확산에 따른 전력 수요를 충족시키기 위해서는 최소 1GW 이상의 전력 충족이 가능한 데이터센터가 필요할 것으로 전망됨

이러한 데이터센터 확장 필요성의 증가는 데이터센터 설계 및 운영에 큰 영향을 미치며, 특히 인프라 기술, 전력 생산, 냉각 시스템 등에 있어 중대한 변화를 초래할 것으로 보임. 가령 발전된 AI 기술 수준을 충족시키기 위한 ‘고성능 컴퓨팅’을 지원하기 위해 고급 액체 냉각시스템(advanced liquid cooling) 및 실리콘 아키텍처(Silicon Architecture)의 혁신이 필요할 것으로 관찰됨.

한편, 대규모 데이터센터 설립 비용은 기술 수준의 발전에 따라 현재의 최대 약 40억 달러(약 5조 2,000억 원) 수준에서 향후 최대 약 250억 달러(약 32조 5,000억 원) 수준으로 약 6배 이상 증가할 것으로 전망되며, 약 6,000-7,000명의 전문 인력이 필요한 것으로 확인됨

### 엣지 컴퓨팅(Edge Computing)의 변화와 혁신

AI의 발전은 사물인터넷(IoT: Internet of Things) 기기, 자율주행차, 스마트 시티 등 실시간 데이터 처리가 필요한 분야에 사용되고 있는 ‘엣지 컴퓨팅’의 성격을 변화시키고 있음. 특히, 특정한 목적에 맞춰진 AI 모델인 ‘도메인 특화 언어 모델’(Domain-specific language models)이 중요해지고 있으며, 이러한 언어 모델은 크기가 작고, 신속히 응답할 수 있도록 설계되어 엣지 컴퓨팅 환경에 최적화되어 있음

자율주행 기술의 경우 도로위에서 실시간 혹은 순간적으로 결정을 내려야 하는 경우가 많아 신속한 데이터 처리가 필요한 분야인데, 엣지 컴퓨팅은 자율 주행차량 자체에서 필요한 AI 연산을 처리하도록 하여 자율주행 기술의 확산에도 핵심적인 역할을 수행중임

한편, 엣지 컴퓨팅의 혁신은 엣지 AI 서버(edge AI server), AI PC, 로봇, 웨어러블 기기 등 새로운 형태의 기기 발전에 기여하고 있으며, 이러한 기기들은 AI와 인간과의 상호작용을 더욱 원활하게 할 것임

[ 엣지 컴퓨팅 기술의 주요 적용 사례 ]

적용 분야	적용 사례	엣지 컴퓨팅의 역할
자율주행	센서를 통해 수집된 데이터를 차량 내부에서 실시간 처리하여 신속한 의사결정 가능	차량이 서버와 상시 연결되지 않더라도 주변 환경을 분석하여 자율적으로 주행 가능
스마트 시티	CCTV 등을 통해 교통 관리, 공공 안전, 에너지 관리 등의 실시간 데이터를 처리하여 도시 관리 효율성 증대	현장에서 실시간으로 데이터 처리 및 의사 결정을 하여 서버로의 데이터 전송을 줄이고, 도시 내 인프라 관리 최적화
헬스케어	웨어러블 기기 혹은 IoT 의료 기기가 환자의 생체 신호를 실시간으로 분석하여 긴급 상황에서 신속한 대응	의료기기에서 수집한 데이터를 클라우드에 전송하지 않고, 기기 자체에서 처리하여 빠른 응답과 신속한 알림 제공
산업 자동화	공장 내 로봇 및 기계들이 수집한 데이터를 즉각적으로 분석하고 생산성 향상 및 주요 문제를 실시간으로 해결	산업 현장에서 데이터 수집과 분석이 빠르게 이루어져 기기의 고장, 이상 징후 등을 사전에 감지하고 대응
영상분석	드론 혹은 CCTV가 촬영한 영상을 실시간으로 분석하여 이상 행동이나 사건을 감지	대규모 영상 데이터를 서버로 보내지 않고 현장에서 처리하여 신속한 응답 제공 및 이상 행동 탐지 가능

AI 기술 발전과 데이터센터 확장 등이 가져올 미래 변화

**(클라우드 및 데이터센터 서비스 제공업체의 경쟁 증가)** 클라우드 및 데이터센터 서비스 제공업체들은 AI의 미래 수요를 충족시키기 위한 전략적 도전과제를 해결해야 할 필요성이 증가하고 있는데, 가령 메타(Meta)는 오픈 AI, 구글(Google) 등과 경쟁하여 대형언어모델 분야에서 리더십을 확보하기 위해 노력하고 있으며, 이를 위해 지난 2년 동안 컴퓨팅 용량을 대폭 확장시켜 오고 있음. 아울러, 고객의 요구를 충족시키기 위해 어떤 분야의 서비스를 제공할 지 결정하고, 언어모델 개발 가속화, 대규모 및 분산 데이터센터 구축에 중점을 둘 필요가 있음

**(인프라 및 소프트웨어 제공업체의 역할 강화)** 인프라 제공업체는 AI 기술 발전을 위한 높은 수준의 전문성을 요구받고 있음. 서버 네트워크, 스토리지, 냉각기기, 케이블 등 데이터센터를 구축하는데 필요한 모든 요소를 생산하는 기업은 AI 기술 수준을 충족시키기 위한 전문적인 제품 개발의 필요성이 증가하고 있음. 소프트웨어 제공업체는 자사의 경쟁력을유지하기 위해 핵심 제품에 AI 기술을 지속적으로 적용할 것이며, 언어 모델을 최적화하여 고객에게 보다 개선된 정보 및 결과를 제공해야 할 필요가 증가하고 있음

**(데이터센터 공급망 제공업체의 시장 재편 기회)** 데이터센터 공급망 제공업체는 대규모 데이터센터 및 엣지 컴퓨팅 기술에 대한 수요 증가에 따라 시장에서의 역할을 재편할 기회를 갖고 있음. 이들 업체는 규모를 확장하고, 인프라 건설 업체와의 파트너십을 구축하여 대규모 데이터센터 및 엣지 컴퓨팅 기술에 대한 글로벌 수요 충족을 위해 노력해야 할 것임. 예를 들어, 고성능 컴퓨팅 기능을 지원하기 위한 고급 액체 냉각 시스템과 실리콘 아키텍처의 혁신을 추진할 수 있으며, 이는 데이터센터의 성능을 최적화하고 AI 소프트웨어의 성능을 향상시키는데 기여할 수 있음

## 데이터센터 및 엣지 컴퓨팅 기술 확장에 따른 주요 이슈

**(전력소비 증가와 전력망에 대한 압박)** 데이터센터 및 엣지 컴퓨팅의 확장은 전력 소비 증가를 초래할 것인데, 이러한 전력 수요 증가에 대처하기 위한 전력망을 확장하고 강화하기 위한 투자가 필요한 상황임. 특히 현장전력발전(on-site power generation) 및 신재생에너지를 추가하여 전력망을 강화할 수 있으며, 이는 데이터센터의 전력 소비를 줄이고, 환경 영향을 최소화하는데도 기여할 수 있을 것임

**(전문 인력 부족)** 앞서 언급한 바와 같이, 대규모 데이터센터 확장에는 다수의 특화된 전문 인력(약 6,000-7,000명)이 필요한데, 현재 전기 및 냉각 분야에 대한 인력 부족 현상이 심각한 것으로 확인됨. 이러한 인력 부족 문제를 해결하기 위해 교육 및 훈련 프로그램을 강화하고, 주요국과의 국제협력을 통해 인력 부족 문제를 해결할 수 있는 방안을 모색해야 할 것임

**(네트워크 안전성 및 보안 취약성)** 엣지 컴퓨팅은 중앙 서버가 아닌 다수의 하부 서버에서 데이터를 처리하고 있는 바, 네트워크 연결이 불안정하거나 제한된 환경에서도 작동에 문제가 없어야 함. 특히, 자율주행차, 원격 의료 같은 분야에서는 네트워크 지연이 치명적인 문제로 이어질 수 있음. 아울러, 분산된 하부 서버에서의 데이터 처리 과정에서 해킹, 데이터 도난, 무단 접근 등 보안 취약성도 극복해야 할 과제로 지적됨

### 참고문헌

- Bein & Company, AI Changes Big and Small Computing, 2024년 9월 25일
- Medium, How AI is Transforming Data Centers: Can They Keep Up in the Next Decade?, 2024년 9월 21일
- Data Center Frontier, AI Driving Data Centers to the Edge, 2024년 9월 20일