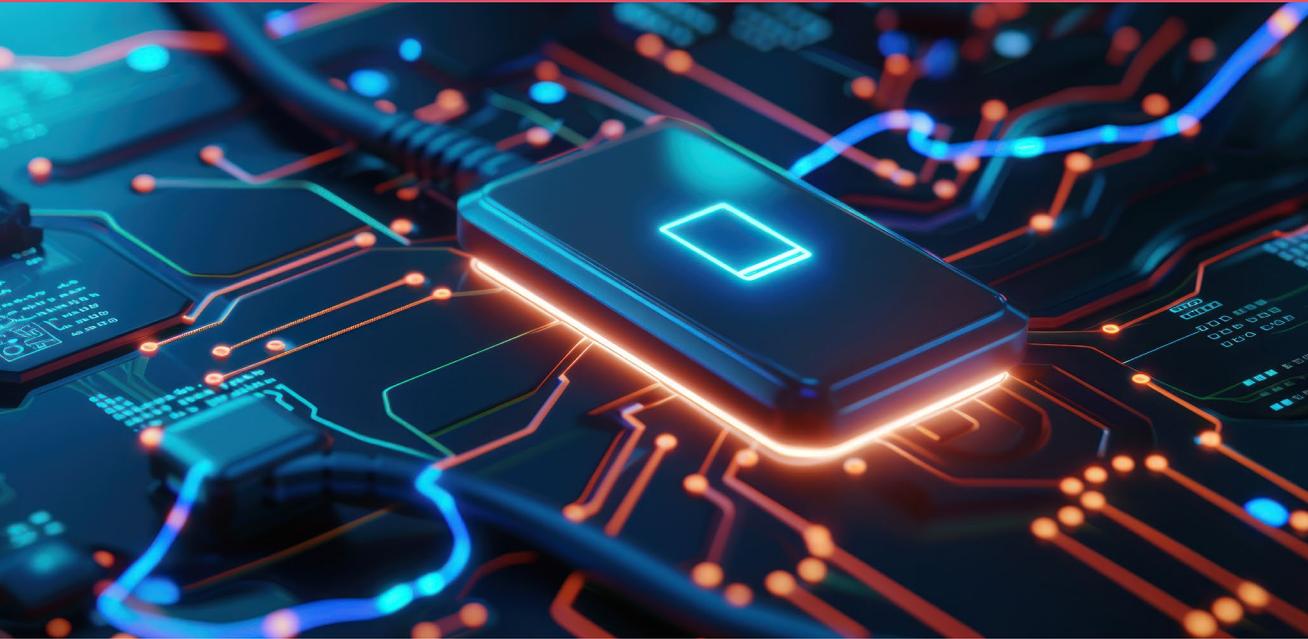


Issue Brief

AI는 전기를 먹고 자란다

AI시대, 전력 소비 증가 원인 및 시사점

삼일PwC경영연구원
October 2024



삼일회계법인



Contents

- 1 AI 시대, 전력 수요 폭증
- 2 학습과 추론 과정에서 전력 다소비
- 3 데이터센터 냉각을 위한 전력 필요
- 4 대응 방안 및 시사점

Summary



1 AI 시대, 전력 수요 폭증

- AI 시스템은 학습과 추론 시 대규모 데이터 처리와 고성능 컴퓨팅 능력이 요구되어 데이터센터 내에 탑재됨. 따라서 **AI 확대는 궁극적으로 데이터센터의 전력소모 증가**로 이어짐
- IEA에 따르면 전 세계 데이터센터 전력 사용량은 22년 기준 460TWh에서 26년에는 1,050TWh로 **2배 이상 증가**할 것으로 전망



2 AI가 전기를 많이 쓰는 이유

- **학습과 추론:** AI 학습과 추론 시 복잡한 연산, 딥러닝 모델 고도화, 대규모 데이터셋 필요로 필요하고 추론 과정에서 요구되는 실시간 처리에 막대한 전력 자원이 필요
- **냉각:** 많은 연산을 동시에 수행하기 위해 다수의 GPU를 사용함. 이러한 프로세서 수 증가는 전력소비를 직접적으로 증가시켜 열 발생, 냉각을 위한 전력 수요 증가로 연결



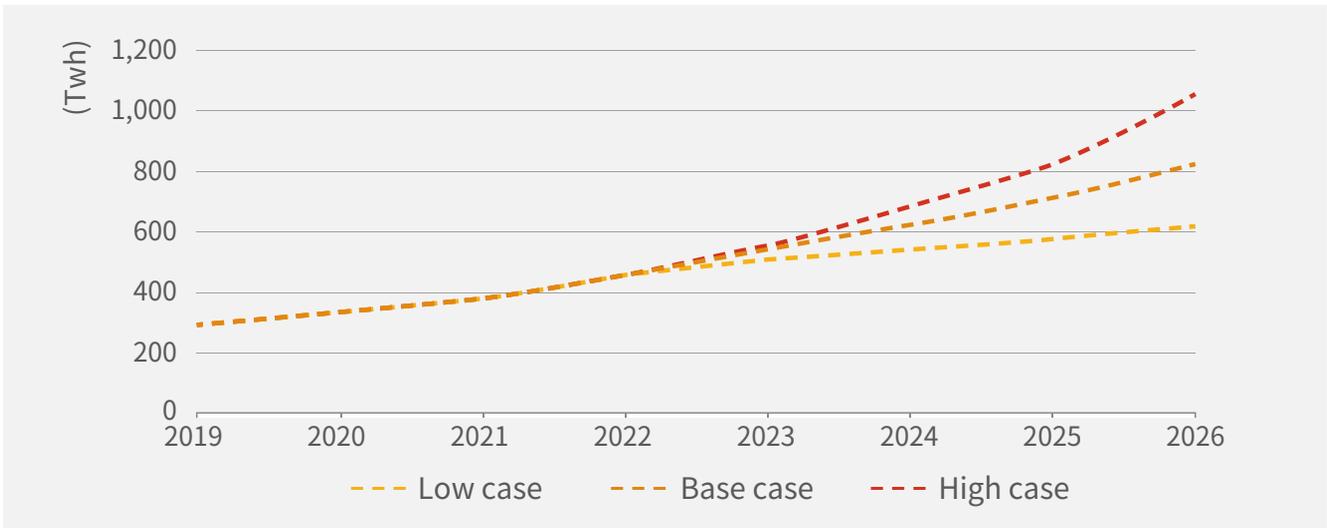
3 대응 방안 및 시사점

- **절감:** AI 모델 경량화를 통한 전력 소비 절감, AI 반도체 채택을 통한 전력 효율성 제고
- **친환경:** 데이터센터 전기 사용 시 신재생 에너지 확대 및 SMR(소형모듈원전) 도입 검토
- **유단취장(有短取長):** AI 사용 확대로 데이터센터의 전력 소비량은 증가하지만 AI를 활용한 타 산업에서의 에너지 절감 효과로 이어질 기술 개발 노력도 필요

1. AI시대, 전력 수요 폭증

- 최근 AI 기술은 데이터 분석, 이미지 및 음성 인식, 자연어 처리, 자율주행차, 로봇 공학 등 광범위한 영역에서 혁신을 주도하고 있음
 - AI 모델 고도화를 위해서는 대규모 데이터 처리와 고성능 컴퓨팅 능력이 요구되며 해당 과정에서 많은 전력소모가 이루어짐
 - AI 시스템은 주로 데이터센터 내에 설치되고 운영되기 때문에 AI 확대는 궁극적으로 데이터센터의 전력수요로 이어짐
- 전 세계 데이터센터의 전력 사용량은 22년 460TWh에서 26년 1,050TWh로 2배 이상 증가할 것으로 전망됨

전 세계 데이터센터 전력 수요량 추이



자료) IEA, 삼일PwC경영연구원

- 1,050TWh는 일본의 연간 전력 소비량과 유사

글로벌 데이터센터와 국가별 전력 소비량 비교



자료) IEA, 삼일PwC경영연구원

2. AI가 전기를 많이 쓰는 이유 - 학습과 추론

- AI 모델 훈련과 추론 과정에서 전력이 많이 소요되는 이유는 복잡한 연산, 딥러닝 모델의 고도화, 대규모 데이터셋 및 실시간 처리가 필요하기 때문

(1) 복잡한 연산

- AI 모델 훈련: 딥러닝 모델을 훈련시키기 위해서는 대규모 데이터셋을 사용하여 복잡한 수학적 연산을 반복적으로 수행해야 함. 이 과정에서 막대한 연산 자원이 필요

(2) 딥러닝 모델의 고도화

- 매개변수 수: 딥러닝 모델, 특히 대형 언어 모델(LLM)과 같은 모델들은 수십억 개의 매개변수를 가지며, 이를 훈련하는 데 막대한 계산 자원이 필요함
- GPT-3와 같은 모델은 1,750억 개의 매개변수를 가지며, 이를 훈련하는 데 수주에서 수개월이 걸리며, 수많은 GPU가 사용됨
- GPT-3 모델을 훈련할 때 소비한 전력은 1.3GWh에 달하는 것으로 알려짐. 1GWh는 4인 기준 10만 가구가 하루 동안 사용할 수 있는 전력량

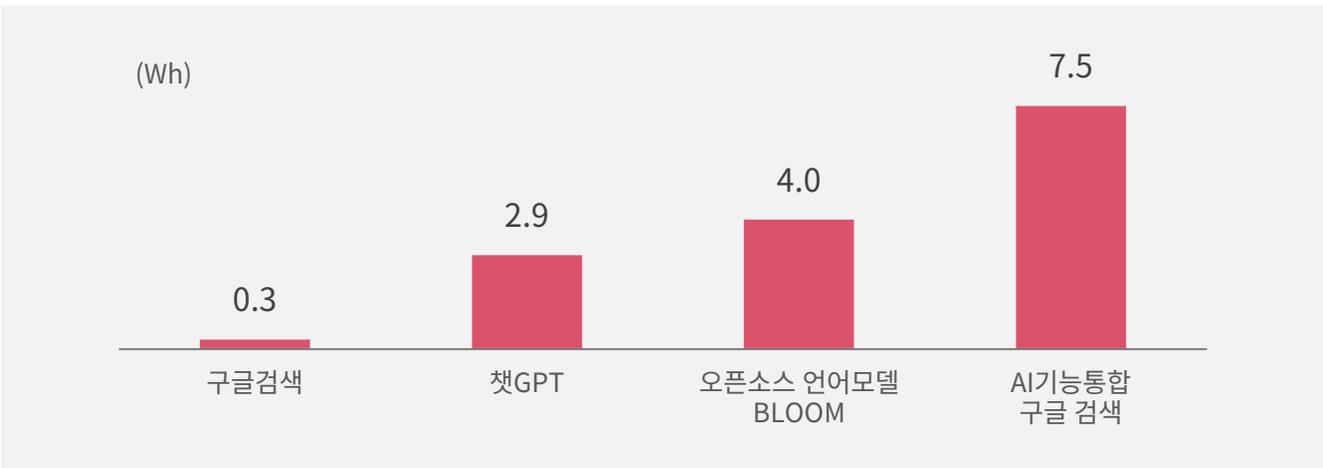
(3) 대규모 데이터셋

- AI 모델을 훈련시키기 위해서는 수많은 데이터가 필요하며, 이는 저장, 전송, 처리 과정에서 막대한 전력 자원을 요구함

(4) 실시간 처리

- 사용자와의 인터랙션에서 실시간으로 응답을 제공하기 위해서는 빠른 추론이 필요함. 이는 고성능 컴퓨팅 자원이 없으면 불가능함
- 음성 인식, 자연어 처리, 이미지 인식 등의 애플리케이션에서 즉각적인 응답을 제공해야 함

사용자 검색 방식에 따른 전력소비량 비교



자료) 미국전력연구원, 삼일PwC경영연구원

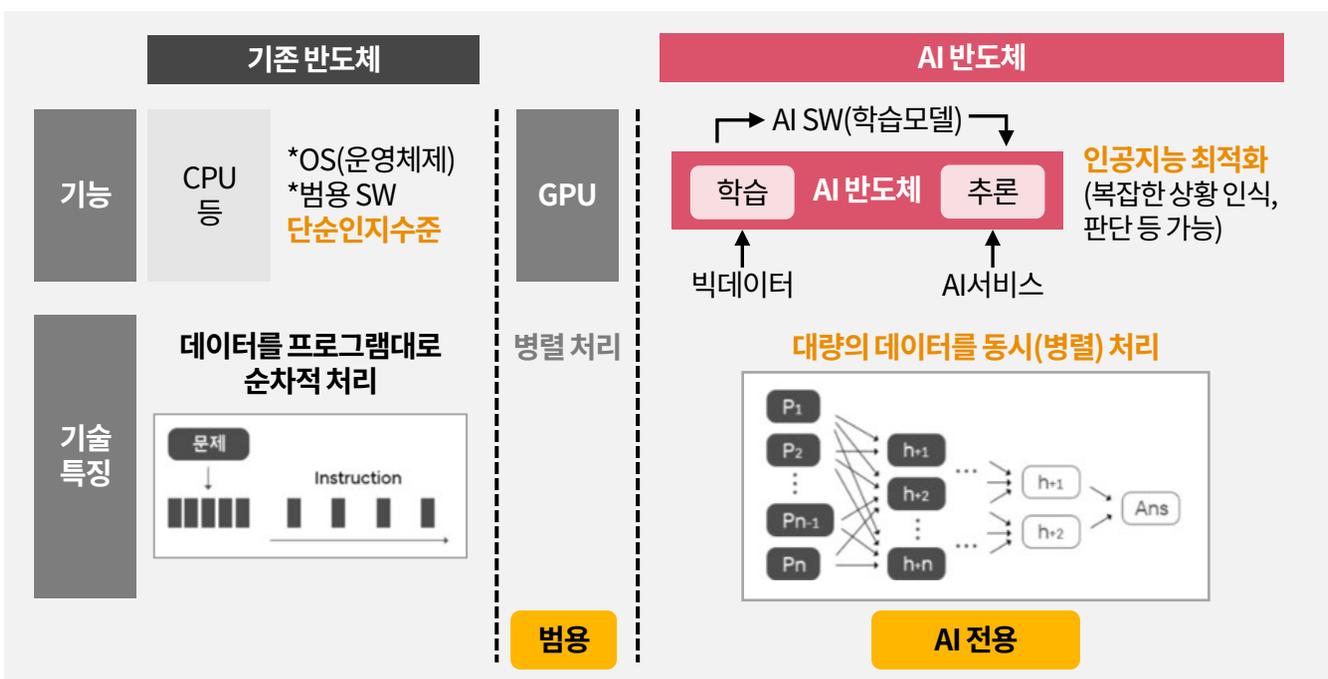
2. AI가 전기를 많이 쓰는 이유 - 학습과 추론

- **GPU(Graphics Processing Unit): AI 훈련에 GPU를 활용하면 CPU만 사용하는 훈련에 비해 속도가 크게 향상되어 대규모의 복잡한 AI 시스템 훈련 시 선호됨**
 - 역할: AI 연산에 있어서 GPU는 병렬 연산이 가능하여, 대량의 데이터 처리를 효율적으로 수행함
 - 전력 소비: GPU는 높은 연산 능력을 제공하는 만큼, 상당한 전력을 소비함. 특히 딥러닝 모델의 훈련 과정에서 전력 소비가 크게 증가함
 - 다수의 프로세서: 병렬 처리는 많은 연산 작업을 동시에 수행하기 위해 다수의 프로세서를 사용함. 이러한 프로세서 수의 증가는 전력 소비를 직접적으로 증가시킴

반도체 구분

구분	특징
CPU	<ul style="list-style-type: none"> • 컴퓨터의 중앙처리장치로 데이터를 고속으로 연산하고 처리 (두뇌 역할) • 일반적인 계산, 시스템 작업, 사용자 인터페이스 등 다양한 작업을 순차적으로 처리
GPU	<ul style="list-style-type: none"> • 그래픽 처리 장치로 멀티미디어를 화면으로 출력하기 위한 그래픽 카드의 핵심 부품 • 게임, 영상 편집, AI, 데이터 마이닝 등 병렬 처리가 중요한 작업에 특화 • AI 연산만을 위해 만들어진 칩이 아니었기에 AI에 필요한 단순 연산만 수행
GPGPU	<ul style="list-style-type: none"> • CPU를 대신해 모든 데이터 연산 및 처리를 GPU를 통해 하는 범용 계산 장치 • GPU가 연산 처리 용도로 쓰이면서 기존에 CPU로 진행하던 대규모 데이터 처리와 AI 개발을 위한 기계학습, 딥러닝 분야 등에서까지 그래픽 카드가 쓰이기 시작

기존 반도체와 AI반도체 비교



자료) 삼일PwC경영연구원

3. AI가 전기를 많이 쓰는 이유 - 냉각

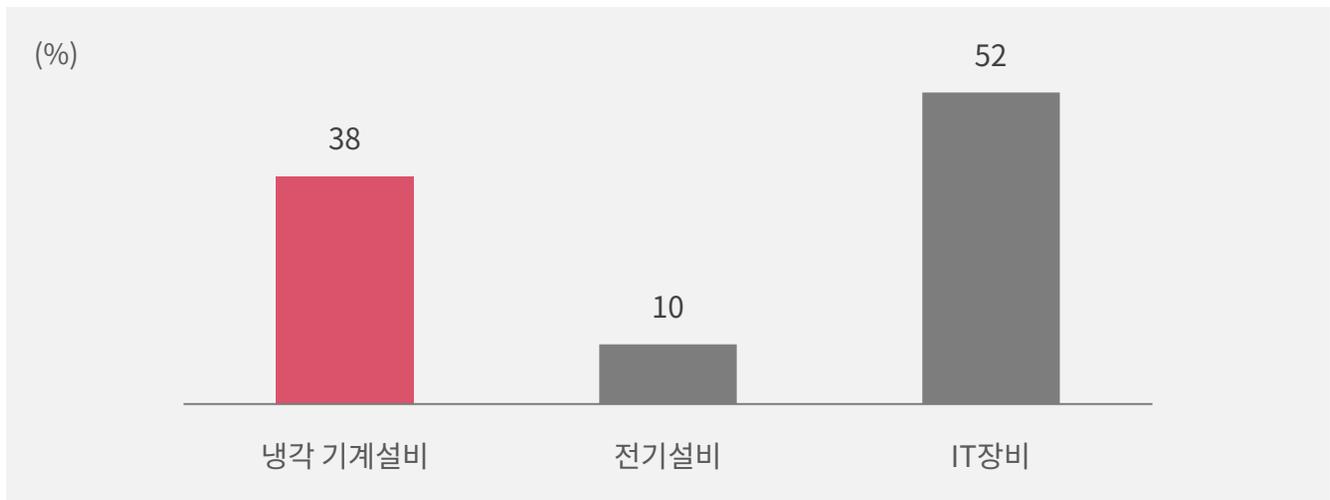
● IT 기기의 발열 문제

- 발열: 병렬 처리는 많은 프로세서를 동시에 사용하므로, 이로 인한 열 또한 발생. 높은 열은 시스템 성능에 부정적인 영향을 미칠 수 있어 이를 해결하기 위한 추가적인 냉각이 필요함
- 고성능 GPU 클러스터는 대규모 연산 작업 중 많은 열을 발생시키며, 이를 냉각하기 위해 추가적인 전력 소모가 필요함

● 데이터센터 냉각 시스템 역시 주요 전력 사용처

- 데이터센터는 공조 및 액체 냉각 시스템을 주로 사용, 데이터센터 전력 사용량 중 냉각 사용 비중은 약 40% 수준에 달할 정도로 높음

데이터센터 내 전력 사용 비중



자료) 삼일PwC경영연구원

데이터센터 냉각 방식

구분	특징
공랭식	<ul style="list-style-type: none">• 공기를 활용해 열을 식히는 방식• 에어컨 냉매 활용 대표적인 열관리 방식• 전력 소모가 심함
수랭식	<ul style="list-style-type: none">• 차가운 물이 데이터센터 내 장비 근처를 흐르며 열을 흡수
액침 냉각	<ul style="list-style-type: none">• IT장비를 비전도성 유체에 침전시켜 열을 식히는 기술• 누전이나 기기 고장 우려 낮음• 냉각 효율 높음

자료) GS칼텍스, 삼일PwC경영연구원

4. 대응 방안 및 시사점 – AI 부문 절감

● AI 모델 경량화를 통한 전력 소비 절감

- AI 경량화 기술은 매개변수를 줄여 학습·운영 비용을 낮추고, 미세조정을 통해 정확도를 높인 기술
- 매개변수가 적은 경량화 모델은 기존 LLM(대규모언어모델) 대비 낮은 연산 성능으로도 구동이 가능하여 기업의 고성능 하드웨어 기반 인프라 구축 부담을 줄이고 전력 소비를 낮출 수 있음

● 2가지 접근 방법 ① 경량화된 AI 모델 및 추론 기술 개발 ② 기존 AI 모델을 경량화하는 기술

- ① 경량 AI 모델(모델 자체를 작게 만드는 것): 기존 모델 대비 효율을 극대화하는 것으로, 딥러닝의 구조적 한계를 개선하고자 알고리즘 구조를 전면 수정하여 작게 만드는 기술
- ② AI 모델 경량화 기술(큰 모델을 가볍게 만드는 것): 만들어진 모델의 파라미터 크기를 줄이는 것이 주 목적, 모델 압축(Model Compression), 지식 증류 등의 기법이 사용됨.

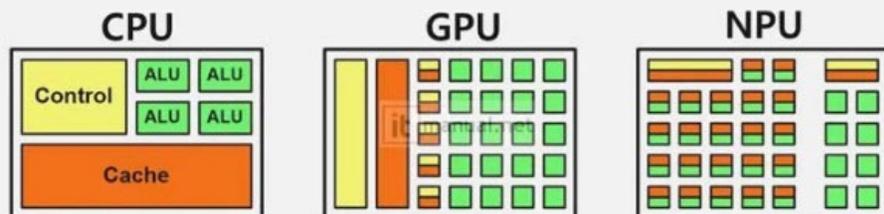
● AI 반도체 채택

- AI반도체는 인공지능 서비스 구현에 필요한 대규모 연산을 초고속, 저전력으로 실행하는 비메모리 반도체를 뜻함. NPU(신경망처리장치, Neural Processing Unit), PIM(Processing In Memory), 뉴로모픽(Neuromorphic, 인간 뇌의 구조와 기능을 모방해 설계된 기술 또는 시스템) 등이 있음. 뉴로모픽 반도체는 성능과 효율성은 뛰어나지만 범용성이 낮고 아직은 개발 중인 차세대 AI 반도체임

● NPU(Neural Processing Unit): 추론용 AI 반도체로, 딥러닝에 적합한 연산들을 가속. GPU에서 불필요한 기능을 모두 제거하고 인공지능 연산에 최적화된 기능만 집약한 칩

구분	특징
NPU	<ul style="list-style-type: none"> • GPU처럼 병렬 처리에 최적화된 구조로 되어있으나, AI를 개발하는데 필요한 제어 및 산술 논리 구성 요소를 갖춰 기계 학습, 심화 학습 알고리즘을 실행하는데 최적화 • GPU에서 불필요한 기능을 모두 제거하고 AI 연산에 최적화된 기능만 집약한 칩 • 이에 GPU와 동일한 AI 작업 시, GPU 대비 전력 소모는 적고 더 많은 결과물을 산출

CPU, GPU, NPU 구조



주) ALU: arithmetic and logical unit, 산술 논리 장치 / Cache: 자주 사용하는 데이터나 값을 미리 복사해 놓는 임시 장소



AI모델 경량화 및 AI 반도체에 대한 상세 설명은 삼일PwC경영연구원의 [‘제3의 IT혁명 디바이스 시대’가 온다: 내 손안의 AI, 온디바이스 AI\(On-Device AI\)’](#) 보고서 참고

4. 대응 방안 및 시사점 – 친환경

● 신재생 에너지를 사용하는 친환경 데이터센터 구축

- 글로벌 빅테크들은 전력비 절감과 ESG 대응을 위해 데이터센터에 신재생 에너지 활용에 적극적임
- 데이터센터 설립 시 신재생 에너지를 활용할 수 있는 부지를 채택하고 태양광 및 풍력을 통해 데이터센터를 운영하기도 함
- 23년 신재생에너지 전력 구입 에서 상위권을 차지한 기업은 모두 빅테크 기업
- PPA(Power Purchase Agreement: 전력구매계약)를 통해 신재생에너지 발전사업자로부터 전력을 구입하는 형태

2023년 미국 기업 신재생에너지 구매량 순위

순위	기업	구매량
1위	아마존	8.8GW
2위	메타	3.1GW
3위	구글	1.6GW

자료) Bloomberg, 삼일PwC경영연구원

● 원전과 SMR(소형모듈원전), 데이터센터 전력원으로 부각

- 글로벌 빅테크들은 대형 원전과 SMR 기반 전력원에 관심
- SMR은 수요처 인근에 설치할 수 있어 데이터센터 전용 발전원으로 사용할 수 있고 송배전망 투자 비용을 줄일 수 있음
- 재생에너지 대비 좁은 부지에서 에너지를 생산하며 수명도 길다는 것이 특징
- SMR은 현재 기술개발 중에 있으며 2030년 상용화 목표

대형 원전과 SMR 비교

구분	대형 원전	SMR
설계	주요 기기 배관연결	일체형
출력	1,000~1,400MW	100~300MW
건설비용	5~10조원	1~3조원
건설기간	10년 이상	5~7년
중대사고 빈도	100만년에 1회	10억년에 1회

자료) 한국수력원자력, 삼일PwC경영연구원

4. 대응 방안 및 시사점 – 유단취장(有短取長)

● ‘전기 먹는 하마’ AI를 에너지 절약에도 활용

- AI는 막대한 전력을 소모해 우려를 자아내고 있지만 AI가 에너지 절약 기술을 발전시켜 이를 완화 또는 해소할 가능성도 존재함
- AI를 활용한 스마트 빌딩이 대표적인 사례, 건물 공조설비와 조명 등을 항상 최적의 설정으로 유지하는 문제를 AI로 해결할 수 있다면 전력 사용량 절감에 도움
- IEA는 주택과 상업용 건물이 세계 에너지 사용량의 30%를 차지하고 그 중 3분의 1이 낭비된다고 분석한 바 있음
- 바깥 기온이나 건물 내 인원수와 상관없이 매일 특정 시간에 냉난방 설비를 켜다 끄는 것을 전력 낭비가 발생하는 주요 원인으로 꼽았음

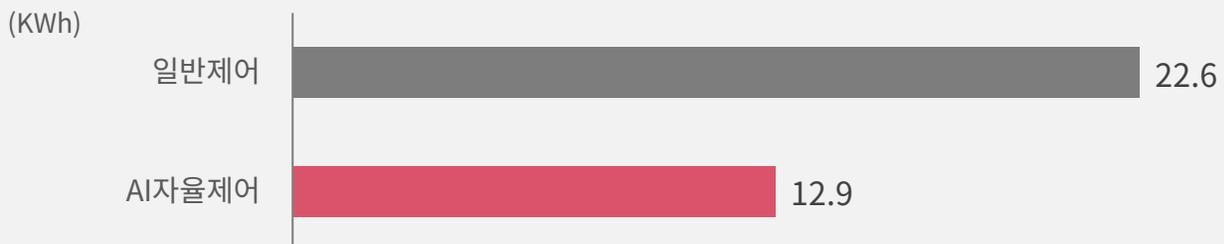
● 스마트 빌딩, 건물 에너지 사용을 제어하는 AI

- AI는 빌딩 내 물 사용량, 사람 수, 엘리베이터 운전 상황 등 정보를 수집해 대량의 데이터를 생성한 후 이를 통해 에너지 절약이 가능한 부분을 찾아내고 적용할 수 있음
- AI와 빅데이터 기반의 에너지 관리 시스템은 전기 요금이 높은 시간대의 에너지 사용을 피하기 위해 미리 냉난방 설비를 가동하면 운전 요금을 줄일 수 있음. 에너지 요구량이 적어지는 시간대를 예측하게 되면 미리 냉난방 운전을 줄여 소비 에너지를 절감할 수 있음

● 유단취장, 단점도 있지만 장점을 취할 줄 알아야

- AI 확대에 따라 데이터센터의 과도한 전력소비, 환경 문제 등이 거론되고 있음
- AI로 사용되는 전력량이 증가하는 것은 맞지만, AI를 이용하여 건물 에너지 절약 또는 산업현장에서 빅데이터를 통한 불량률 감소 역시 에너지 절약으로 연결될 수 있음
- AI 자체의 전력 먹성에 대한 우려보다는 AI를 활용하여 다른 분야에서 에너지를 절감할 수 있는 기술 개발에 힘쓸 필요가 있음

AI 활용한 에너지 환경 통합관리 시스템 실증 결과



자료) 한국에너지기술평가원, 삼일PwC경영연구원

Author Contacts

삼일PwC 경영연구원

김 승 철 수석연구위원

seungchurl.k.kim@pwc.com

Business Contacts

Samil PwC Business Research

최재영 경영연구원장

jaeyoung.j.choi@pwc.com

www.samil.com

삼일회계법인의 간행물은 일반적인 정보제공 및 지식전달을 위하여 제작된 것으로, 구체적인 회계이슈나 세무이슈 등에 대한 삼일회계법인의 의견이 아님을 유념하여 주시기 바랍니다. 본 간행물의 정보를 이용하여 문제가 발생하는 경우 삼일회계법인은 어떠한 법적 책임도 지지 아니하며, 본 간행물의 정보와 관련하여 의사결정이 필요한 경우에는, 반드시 삼일회계법인 전문가의 자문 또는 조언을 받으시기 바랍니다.

S/N: 2403W-RP-017

© 2024 Samil PricewaterhouseCoopers. All rights reserved. "PricewaterhouseCoopers" refers to Samil PricewaterhouseCoopers or, as the context requires, the PricewaterhouseCoopers global network or other member firms of the network, each of which is a separate and independent legal entity.