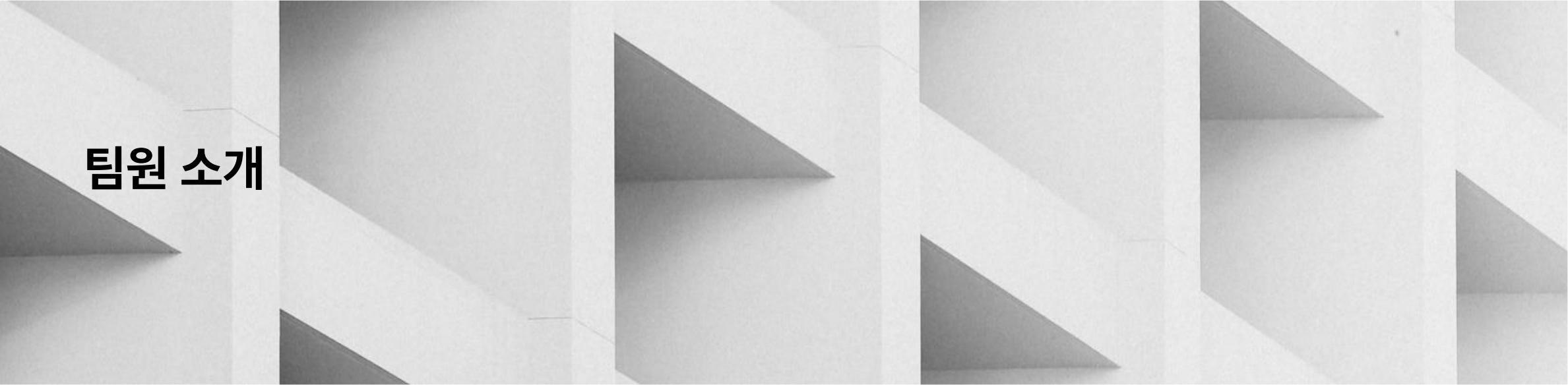


탈중앙화 AI로 구축하는 지속가능한 에이전트 시스템



DAILY: 조민주, 김정현, 우지훈, 이재은



팀원 소개



조민주

팀장 (경영 22)

- 인공지능경영 심화전공



우지훈

팀원 (컴과 19)

- 인공지능경영 심화전공



김정현

팀원 (경영 22)

- 인공지능경영 심화전공



이재은

팀원 (인공지능 22)

목차(Index)

Situation

AI 에이전트
시스템의 도입

·
·
·

인프라적
AI 에이전트 시스템

Problem

중양화 된
AI 에이전트:
데이터 권력 독점

·
·
·

AI 에이전트
시스템에 의한
구체적 문제 상황

Main Solution

탈중양화 AI
(DAI)

·
·
·

AI+블록체인
탈중양화 AI 도입

Domain Solution

국제·정치·법·교육

·
·
·

탈중양화 AI 온보딩 위한
각 영역별 임무

Expected Results

DAI 도입 이후
사회적 영향

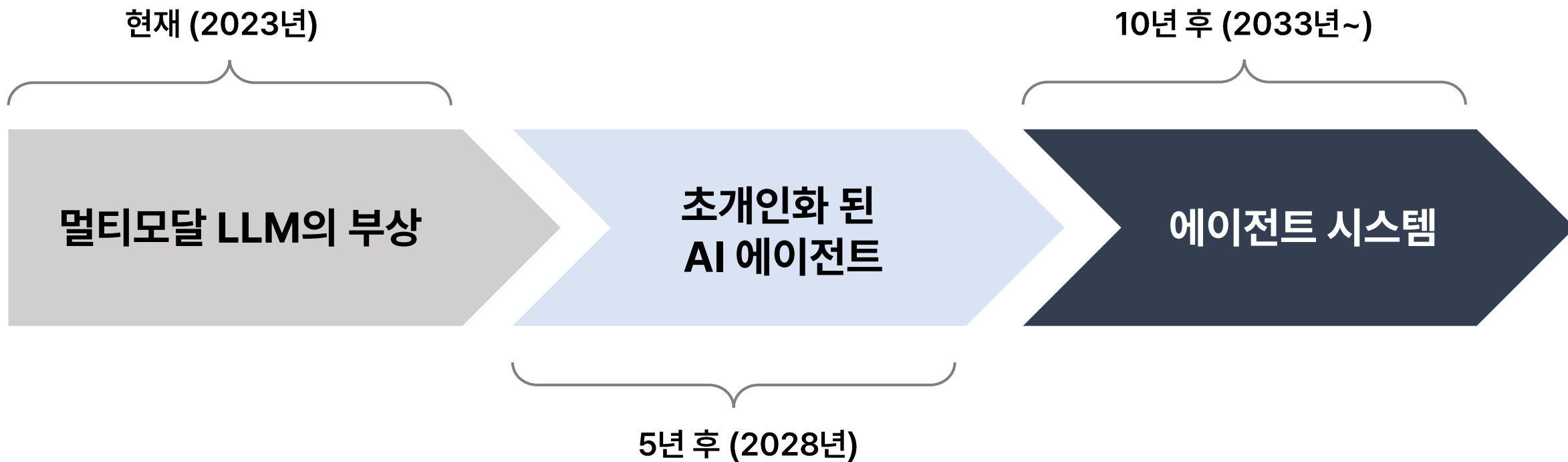
·
·
·

탈중양화 AI 도입 이후
예상 결과 및 사회에
미치는 영향

Part 1

Situation: AI 에이전트 시스템의 도입

10년 후, 인류는 멀티모달 LLM과 AI 비서 시대를 넘어 **에이전트 시스템 인프라**를 사회 전반에 구축했을 것이다.



(1) 2023년 현재, **멀티모달 LLM의 부상**과 함께 초개인화 된 AI 비서 “에이전트”에 대한 관심이 높아지고 있다.

1



Chat GPT

2



Gemini

Multi-Modal LLM

3



AI Agents

(2) 5년 후 2028년에는 AI 에이전트가 개개인 맞춤형 비서로 활용되며 일상 생활 곳곳의 생산성이 극대화될 것이다.



Bill Gates
Former CEO of Microsoft

In the next five years, this will change completely. You won't have to use different apps for different tasks. You'll simply tell your device, in everyday language, what you want to do. And depending on how much information you choose to share with it, the software will be able to respond personally because it will have a rich understanding of your life. In the near future, anyone who's online will be able to have a personal assistant powered by artificial intelligence that's far beyond today's technology.

This type of software—something that responds to natural language and can accomplish many different tasks based on its knowledge of the user—is called an agent. I've been thinking about agents for nearly 30 years and wrote about them in my 1995 book *The Road Ahead*, but they've only recently become practical because of advances in AI.

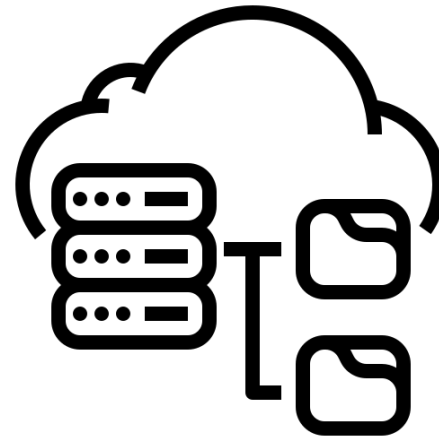
GatesNotes 중 발췌

“ AI 에이전트가 5년 내 컴퓨팅 완전 변경... 비즈니스/사회까지 혁신 ”

(3) 10년 후 2033년에는 비서를 넘어 AI 에이전트 기반의 통합 시스템 및 인프라가 사회 전반에 도입 되었을 것이다.



단순 개인 비서



사회 전반에 도입

Problem

중앙화된 AI 에이전트 시스템의 데이터 권력 독점

1

GPU Memory의 부족으로 인한
거대 모델의 작동 중지 문제

해킹의 가능성 배제 불가

2

현 중앙 집중형 AI의 한계

3

최근 AI 모델 개발에 사용되는
'강화학습' 로 AI 편향성 악화

거대 기업의 AI 작동원리 관련
불투명한 정보공개

4

특정 국가/기업이 중앙화된 AI 에이전트 시스템을 통해 모든 유저의 초개인화된 데이터를 수집한다면 **권력 독점 현상**이 발생할 수 있다.

중앙화된 AI 모델의 에이전트 시스템의 악용 가능성으로
특정 국가/기업/개인이 소유 시 다음과 같은 문제들이 우려됨 :



Misinformation의 강화



Reward System의 부재



Privacy & Security 문제

(1) 중앙 에이전트 시스템 개발기업이 **악의적인 Data Poisoning**을 통해 사회적 약자의 차별 인식을 강화하는 것에 대응가능한 모니터링 체계가 부재하다.

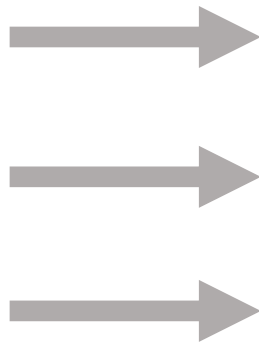


"Amazon scraps secret AI recruiting tool that showed bias against women"

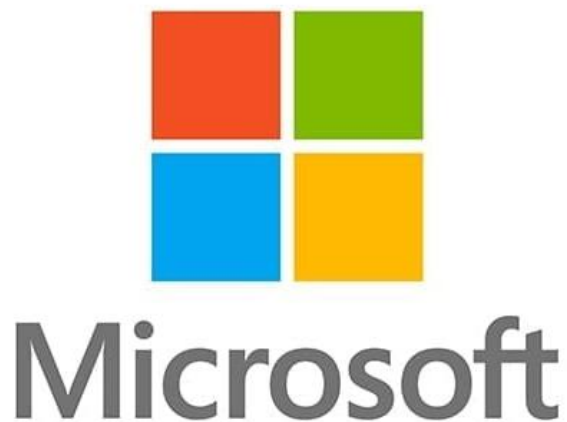


'Data Poisoning'으로 여성에 대한 차별 강화
단순 편견이 아니라 채용까지 부정적 영향을 끼침

(2) 중앙 에이전트 시스템 개발사는 정형 & 비정형 데이터 제공을 통해 모델 구축에 기여한 유저에게 **마땅한 보상을** 제공하지 않으며 **수익을 독점**한다.



(3) AI 에이전트 소유 기업이 모델을 감시의 목적으로 악용 시
유저의 **개인 정보 침해 문제**가 더욱 더 심각하게 대두될 수 있다.



'GitHub에 AI 학습용 데이터 공유하다가
38TB의 내부정보(비밀번호, 비밀키 등) 유출



개인정보 침해 사회 우려

방향성

Misinformation: 악의적 Data Poisoning 대한 대비 부재

- 에이전트 시스템에 블록체인의 상호 인증시스템 도입으로 Data Poisoning을 방지

Reward: 에이전트 시스템에 대한 기여 대비 보상의 부재

- 암호화폐 거버넌스의 투표 시스템로 합의에 귀결 (다양한 투표 방식 가능)
- 만약 암호화폐 소유자들이 허위 투표 시 그 피해는 본인에게 돌아가므로, 대체로 진실하게 투표할 것이라는 가정

Privacy & Security: AI에이전트 소유 기업의 개인정보 악용방지 부재

- 중앙 집중형 서버로 전부가 전달되는 것이 아니라, tuning에 필요한 정보만 전송

Misinformation: 악의적 Data Poisoning에 대한 대비 부재

- 에이전트 시스템에 블록체인의 상호 인증시스템을 도입, 오라클 문제를 해결하고 Data Poisoning을 방지할 수 있어야 함.

즉, 탈중앙화의 도입은 **AI의 편향성과 해킹을 방지하고,**
합의 거버넌스를 통해 자연스러운 AI발전을 유도할 수 있다!

Privacy & Security: AI에이전트 소유 기업의 개인정보 악용방지의 부재

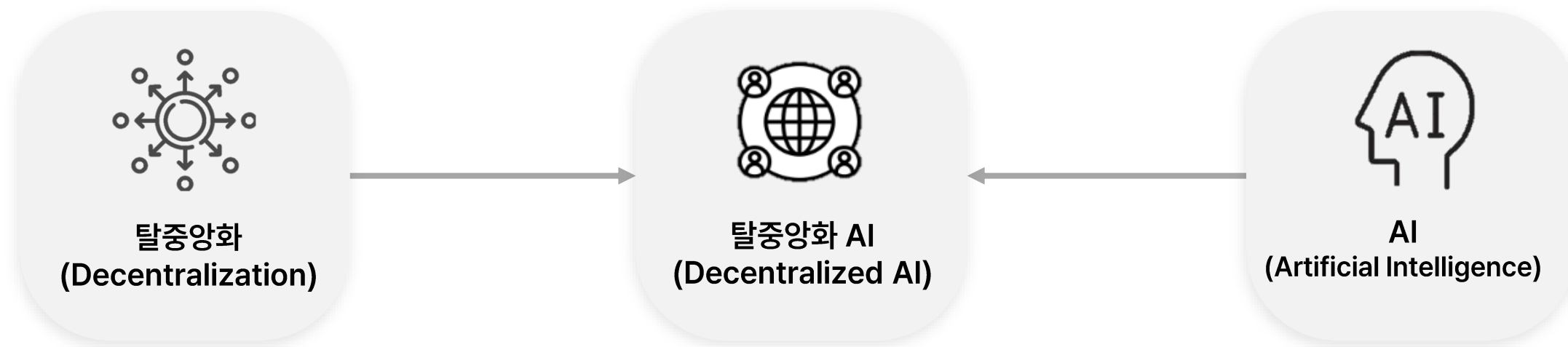
- 중앙 집중형 서버로 전부가 전달되는 것이 아니라, tuning에 필요한 정보만 전송되어야 함.

Main Solution

탈중앙화 AI (Decentralized AI, DAI)

Part 3

탈중앙화 + AI



탈중앙화 (Decentralization)

*탈중앙화란?

신뢰 있는 **제 3자 없이 직접 거래**하는 시스템
대표적으로 **블록체인** 분야에서 활용

*블록체인이란?

비즈니스 네트워크에서 거래를 기록하고, 자산 추적을
용이하게 하는 변경 불가능한 공유 원장 (IBM)

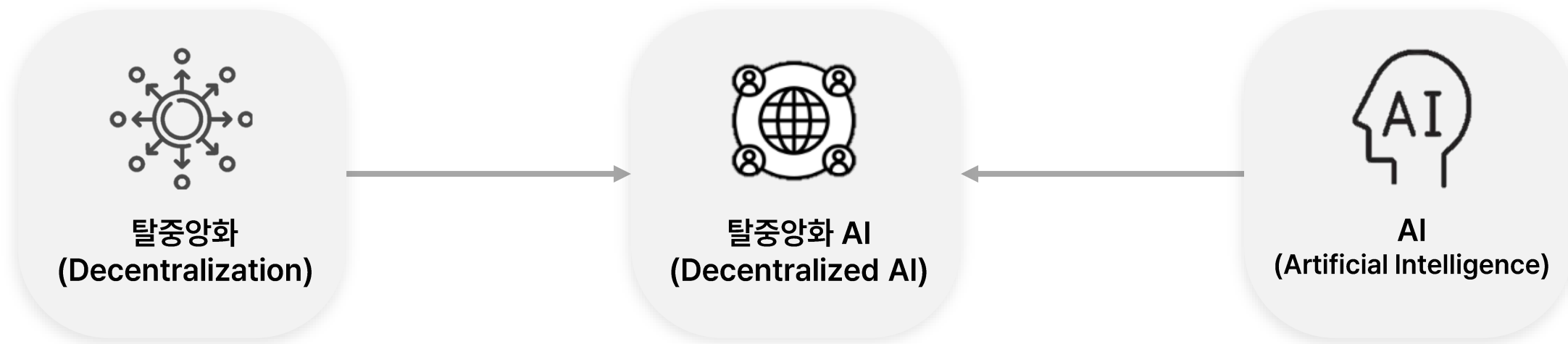
특정 기업에 중앙서버에 모아 작업하던 것을 **개개인에게
분산**하여 처리를 할 수 있게 됨.

AI (Artificial Intelligence)

*AI란?

컴퓨터가 음성과 언어를 보고, 이해하고, 번역한 후,
그 데이터를 분석하여 다양한 고급 기능을 수행할 수
있도록 하는 일련의 기술 (Google Cloud)

현재 LLM 등 거대모델들을 중심으로 빠르게
발전하는 분야이지만 명백한 한계 존재



탈중앙화 (Decentralization)

*탈중앙화란?

신뢰 있는 **제 3자 없이 직접** 거래하는 시스템
대표적으로 **블록체인** 분야에서 활용

*블록체인이란?

비즈니스 네트워크에서 거래를 기록하고, 자산 추적을
용이하게 하는 변경 불가능한 공유 원장 (IBM)

특정 기업에 중앙서버에 모아 작업하던 것을 **개개인에게**
분산하여 처리를 할 수 있게 됨.

AI (Artificial Intelligence)

*AI란?

컴퓨터가 음성과 언어를 보고, 이해하고, 번역한 후,
그 데이터를 분석하여 다양한 고급 기능을 수행할 수
있도록 하는 일련의 기술 (Google Cloud)

현재 LLM 등 거대모델들을 중심으로 빠르게
발전하는 분야이지만 **명백한 한계** 존재



현재 : Smart Contract + Auditing



블록체인(BlockChain)

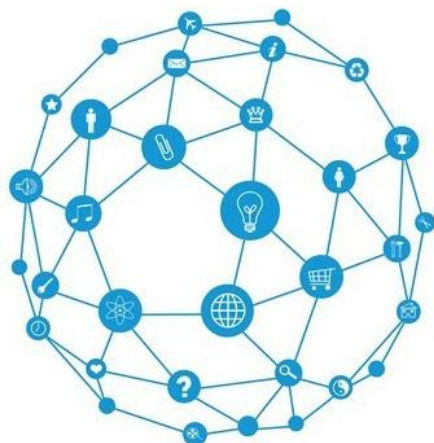
연합학습(Federated Learning)

엣지컴퓨팅(Edge Computing)

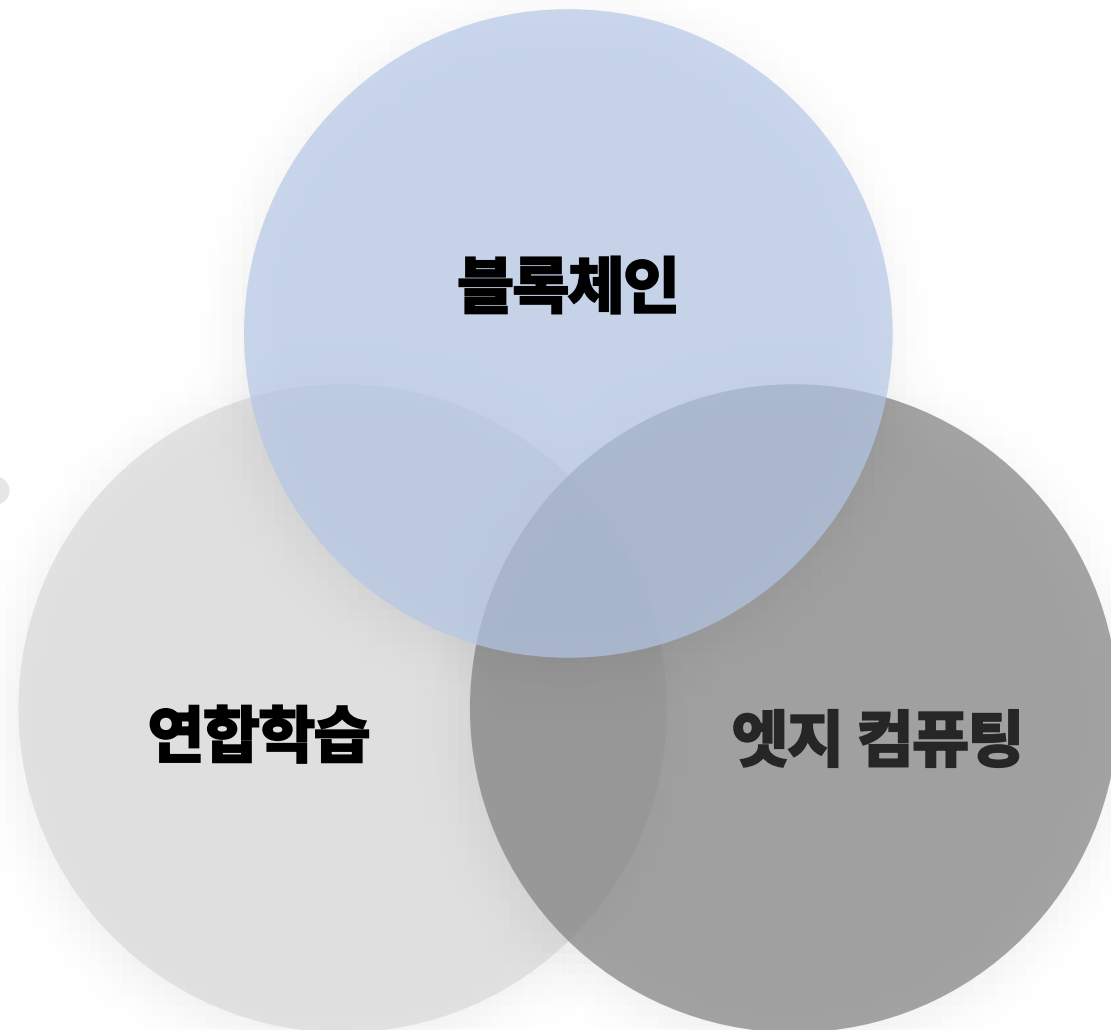


Decentralized AI(DAI)

- 기존에는 AI를 블록체인에 적용해 Smart Contract 검증과 Auditing 과정에만 국한되어 사용
 - ① 중앙집권화 와 ② GPU메모리의 부족 으로 AI는 한계가 드러남
 - 탈중앙화와 더 많은 GPU 메모리 할당이 필요
 - 블록체인 도입으로 탈중앙화를 이루고, 연합학습과 Edge Computing으로 메모리 할당 가능

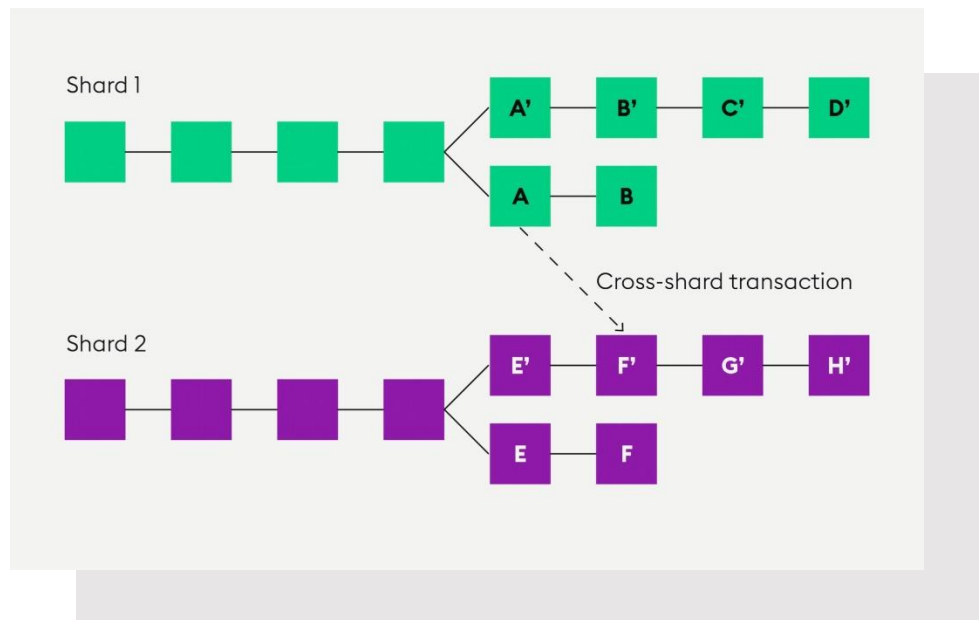


**Decentralized
AI**



Part 3

Solution (1): 블록체인 - 원리

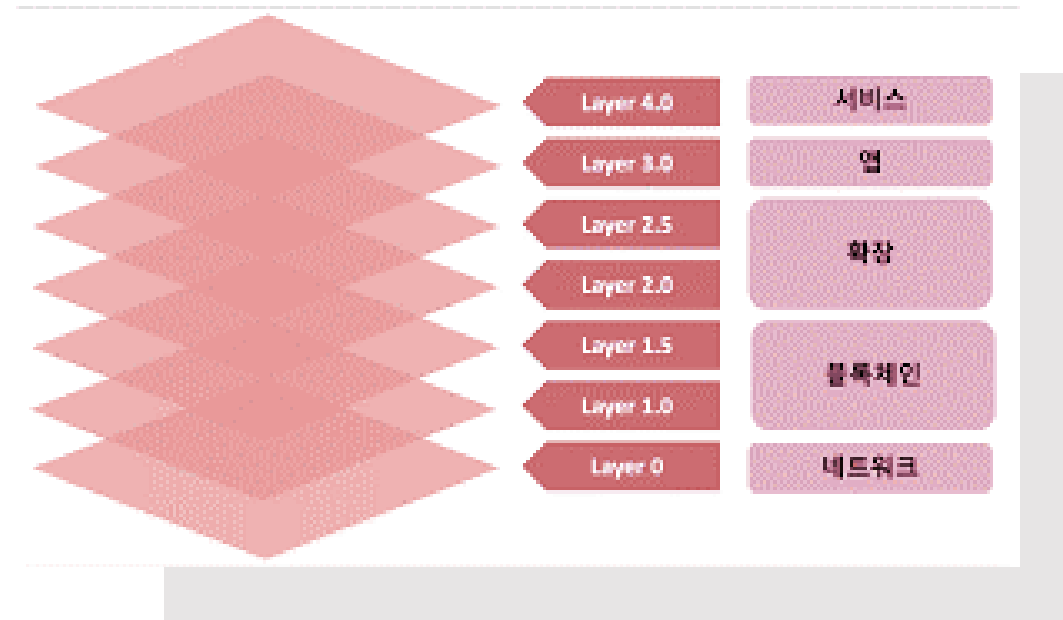


<Sharding>

각 Edge 노드로 분산(Sharding)하여 연산

*로컬(OffChain, Edge): 학습 및 데이터 처리

*서버(Onchain): 결과물 수신, 내용 기록

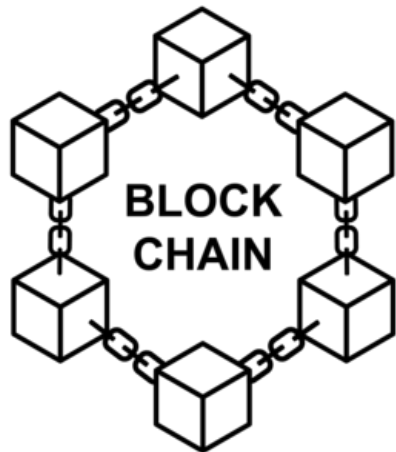


<Multi Layer>

다수 Server, 다수 Layer

같은 AI, 다른 결과

(각자만의 데이터로 학습을 진행하기 때문)



상호 인증시스템을 도입하여
Data Poisoning 방지



변경 불가능한 계약으로 체인상 내용은 수정 불가

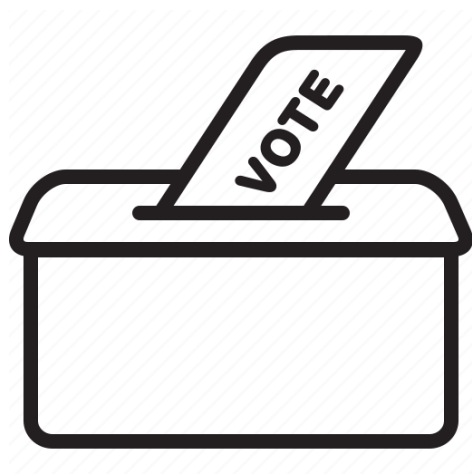


다수 참여자 간 통일적 의사결정



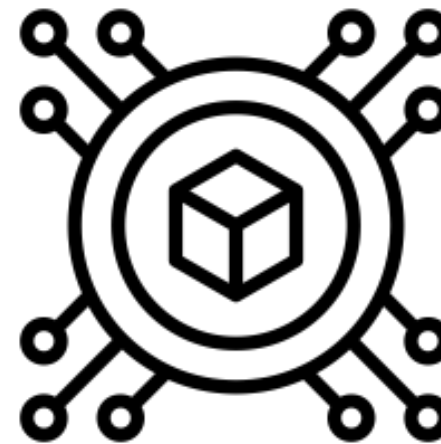
Poisoning 발생 시, 과반 이상이 받아들이지 X

→ Web3를 통한 오픈소스의 투명성 강조



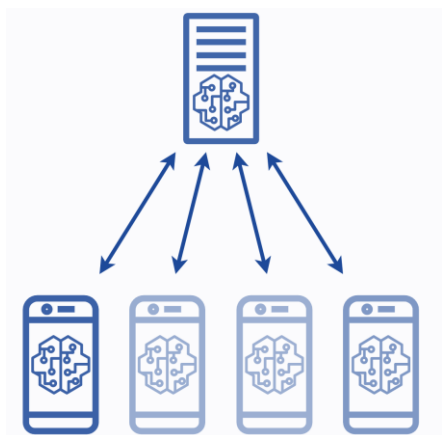
DAO는 탈중앙화된 블록체인 위 **자율적 조직**

(주식회사의 이사회와 유사하게
모든 사용자가 투표권을 가지고 의결에 참여)



투명한 과정의 공개와 토큰 가치 상승으로 보상을
재분배받아 **경제적 유인 증가**

→ 집단 거버넌스(DAO) + 블록체인 프로토콜 도입, 발전적인 커뮤니티 생성



<기본 구조>

정의

분산형 모델을 전제로 한 기계 학습에 대한 혁신적 접근 방식
장치와 중앙 서버 간 지속적인 데이터 교환 없이 Edge에서 기계 학습
모델을 실행하는 것을 전제로 함

핵심 원칙

개인 정보 보호
데이터 지역성
분산 intelligence

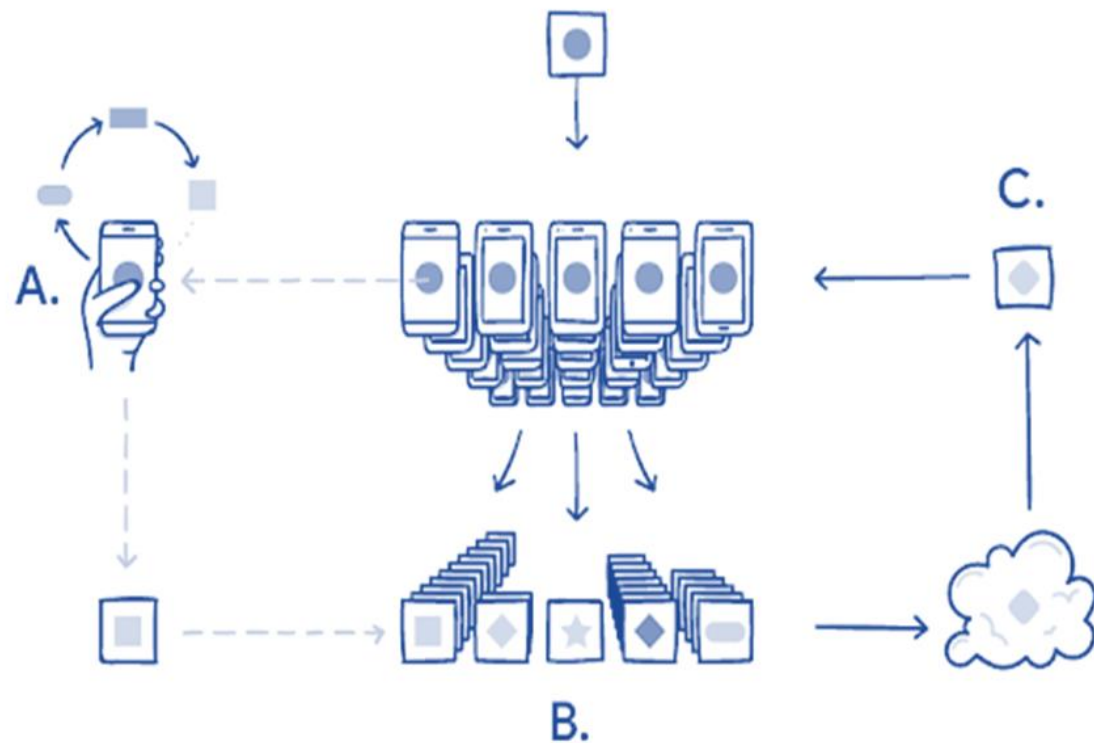
도입 배경

분산형 데이터의 증가
민감 정보 보호
모델 개선에 기여

Part 3

Solution (2): 중앙 집중학습 vs 연합학습





* 초기화

학습 모델을 중앙 서버에서 Edge로 전송 (모델은 오픈 소스로 공개된 범용 모델).

A. 로컬에서 작업

각 Edge는 모델을 선호에 따라 작업.
다른 모델을 참고할 수는 없음 (독립성 보장).

B. 모델 업데이트

각 모델은 Edge에서의 작업의 다양성으로, Weight와 Parameter값이 달라지게 되어 Edge만의 모델이 제작됨.

C. 집계

최종 결과만을 중앙 서버로 보내, 오픈 소스의 범용 모델 tuning에 사용.

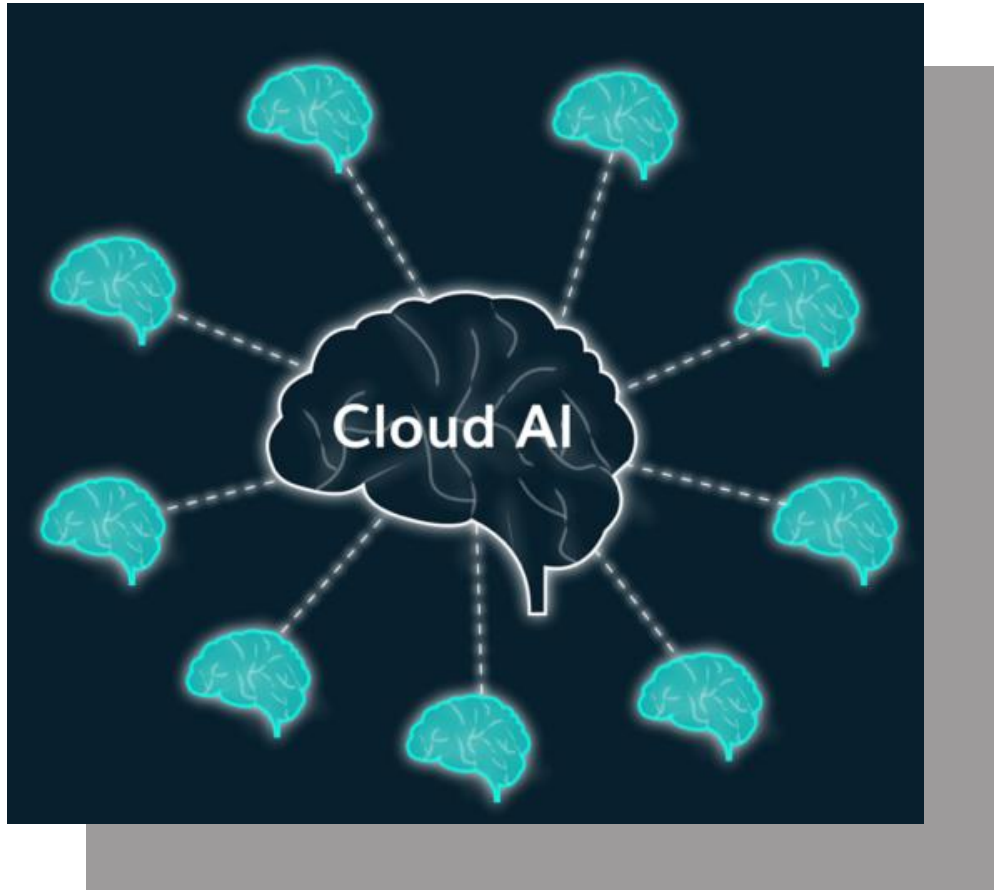
➔ **반복학습**으로 범용 모델의 발전을 추구



- 암호화 프로토콜로 전송, 개인정보 보호
 - 엣지 컴퓨팅 도입에 용이
- 중앙과 Edge 간 통신의 효율성 증가



- 통신상 오버헤드 발생
 - 효율적인 통신 프로토콜 필요
- Edge 디바이스(상태, 종류)의 이질성



기업은 클라우드 중심의 데이터만 관리하고
각 **Edge**가 계산을 수행하게 됨

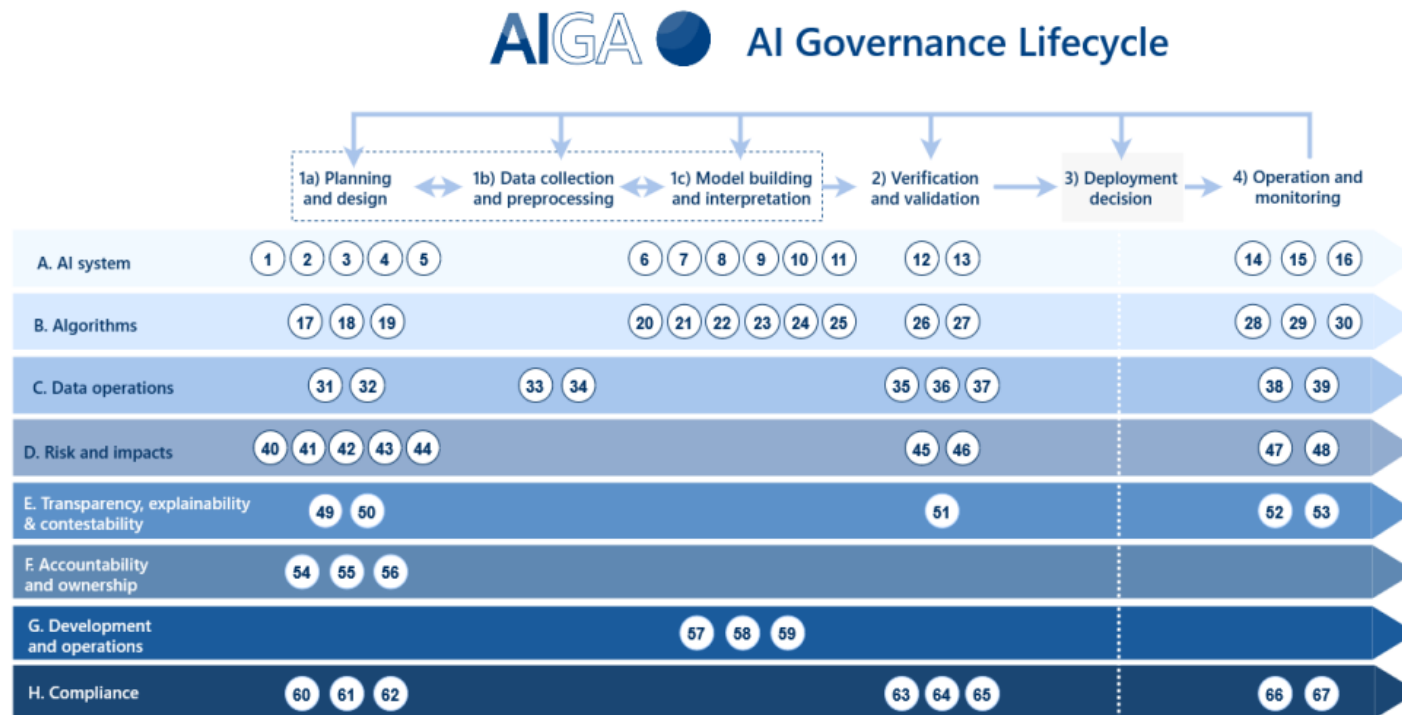


- 1. 네트워크 비용 절감**
- 2. 전송 지연 감소**
- 3. 대역 제약 탈피**
- 4. 민감 데이터 제어 용이**

Domain Solutions

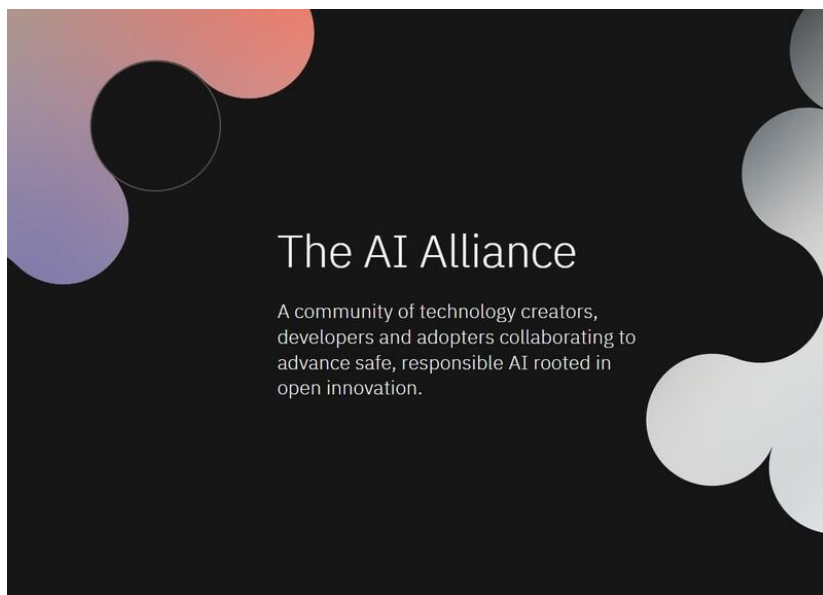
탈중앙화 AI 온보딩을 위한 각 도메인 별 임무

세계 각국의 정부는 AI 에이전트 시스템의 민주화를 위해 오픈소스 이념을 강화하고, 윤리적인 AI 거버넌스를 펼쳐 AI 기술과 탈중앙화 개념 간 융합의 필요성에 대해 강조해야 한다.



Mäntymäki, M., Minkinen, M., Birkstedt, T., & Viljanen, M. (2022). *Putting AI Ethics into Practice: The Hourglass Model of Organizational AI Governance* (arXiv:2206.00335).

각국의 정부, 대학, 기업은 특정 영리 단체가 AI 에이전트 시스템을 독점하는 것을 방지하기 위해
국제적인 차원에서 협약을 맺고 **투명성·경쟁**을 추구해야 한다.



AI 얼라이언스 (2023.12.05~):

글로벌 AI 분야의 개방형 혁신과 오픈 사이언스를 지원하기 위한 선도 조직으로, 산업계, 스타트업, 학계, 연구기관, 정부를 아우르는 협력 단체

각 정부는 AI 에이전트 시대에 시민들의 **데이터 주권**을 보호하고 사이버 보안을 강화하기 위해 **빅데이터** 및 초개인화된 **비정형 정보** 수집에 대한 법적 체계와 규제를 마련해야 한다.

목적

AI 에이전트 시스템이 대량의 비정형 정보를 수집하는 것에 대한 법적 규제 및 체계 필요

빅데이터

최대한 많은 데이터 수집 &
활용할수록 빅데이터 효용 UP

⇔ '최소 수집의 원칙' &
'목적 명확성의 원칙'에 위배

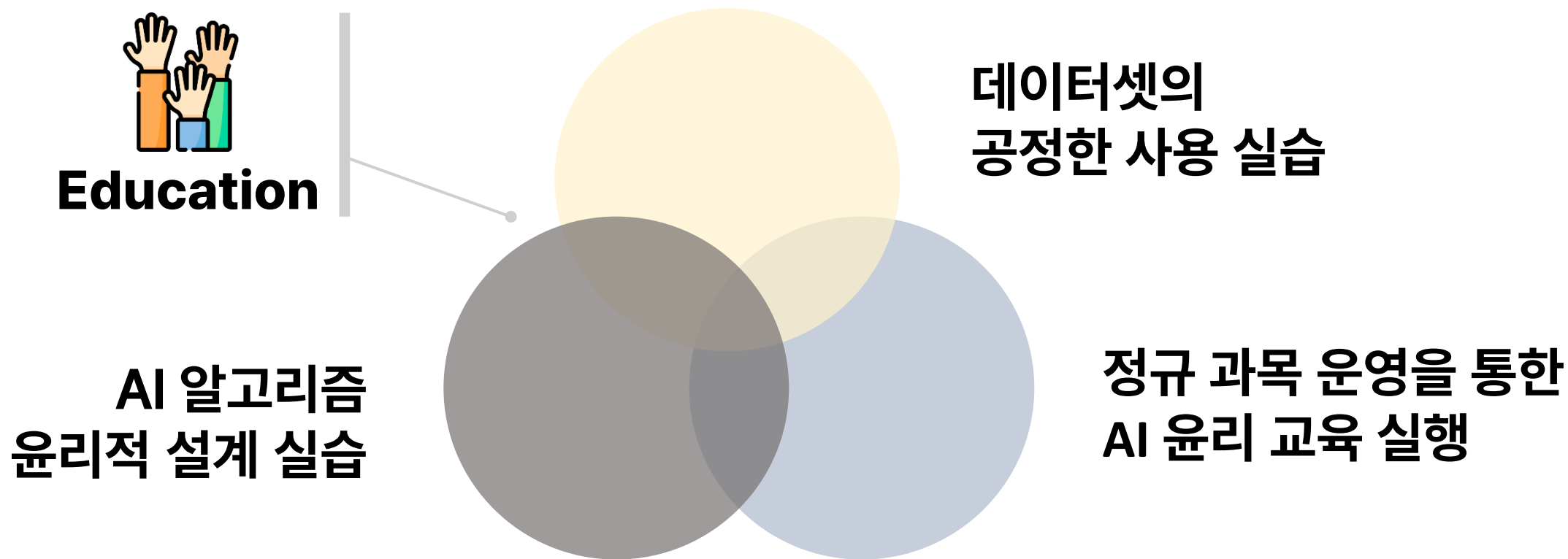


비정형 정보

일상 자동 추적 라이프로그,
실시간 영상데이터 & 동적 정보 등

➔ 획일적인 사전 고지를 전제로
하는 현행 법체계에 부합 X

교육 기관은 학생들에게 중앙화된 AI 모델의 위험과 분산화된 에이전트 시스템의 이점을 단순 이론 주입이 아닌 **참여적 & 실습형 커리큘럼**을 통해 전달해야 한다. (McNamara et al, 2018)

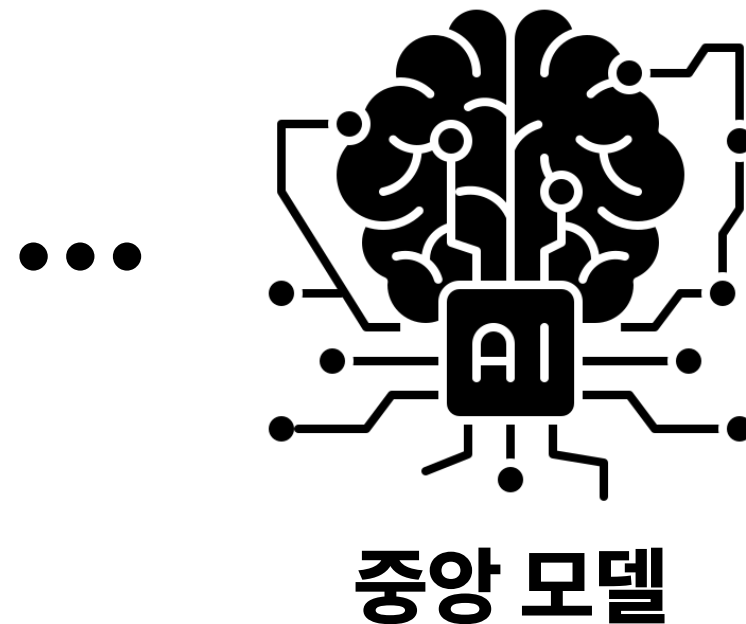
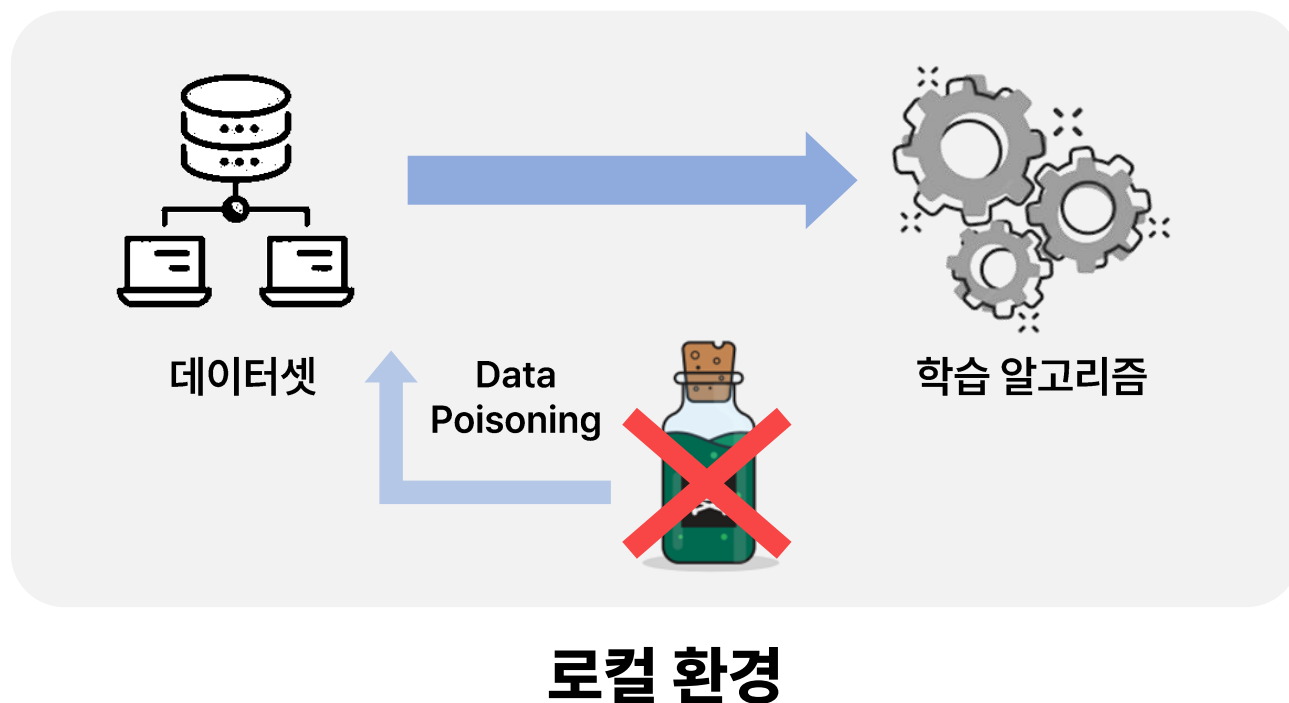


Expected Results

탈중앙화 AI 도입 이후 예상 결과 및 사회에 미치는 영향

1) 투명하고 신뢰 가능한 에이전트 시스템

탈중앙화 AI 인프라 도입 이후 Data Poisoning으로 에이전트 시스템이 오염되거나
거짓된 정보가 확산되는 일의 빈도가 눈에 띄게 줄을 것이다.



2) 기여한만큼 돌려받는 능동적인 유저 보상 체계

탈중앙화 AI 도입 이후, 유저들은 P2P를 기반으로 데이터와 AI 모델을 사고팔며 에이전트 시스템에 투명하게 기여하고, 그에 응당하는 보상을 돌려받을 수 있는 **마켓플레이스**가 더욱 더 **활성화**될 것이다.

탈중앙화
AI 모델
마켓플레이스
(ex. SingularityNET)

탈중앙화
마켓플레이스

탈중앙화
데이터
마켓플레이스
(ex. OceanProtocol)

탈중앙화
AI 모델
마켓플레이스
(ex. SingularityNET)

- 인공지능 모델을 **오픈 마켓에서 구매, 대여, 임대**하는 세상
- 모든 사람이 중앙 집중식 데이터 제어의 제약에서 벗어나 AI 발전에 기여할 수 있게 함
- 투명하고 안전한 블록체인 기술을 통해 AI 모델을 만드는 사람들이 **공정한 수익 보상**을 받을 수 있게 해 줌

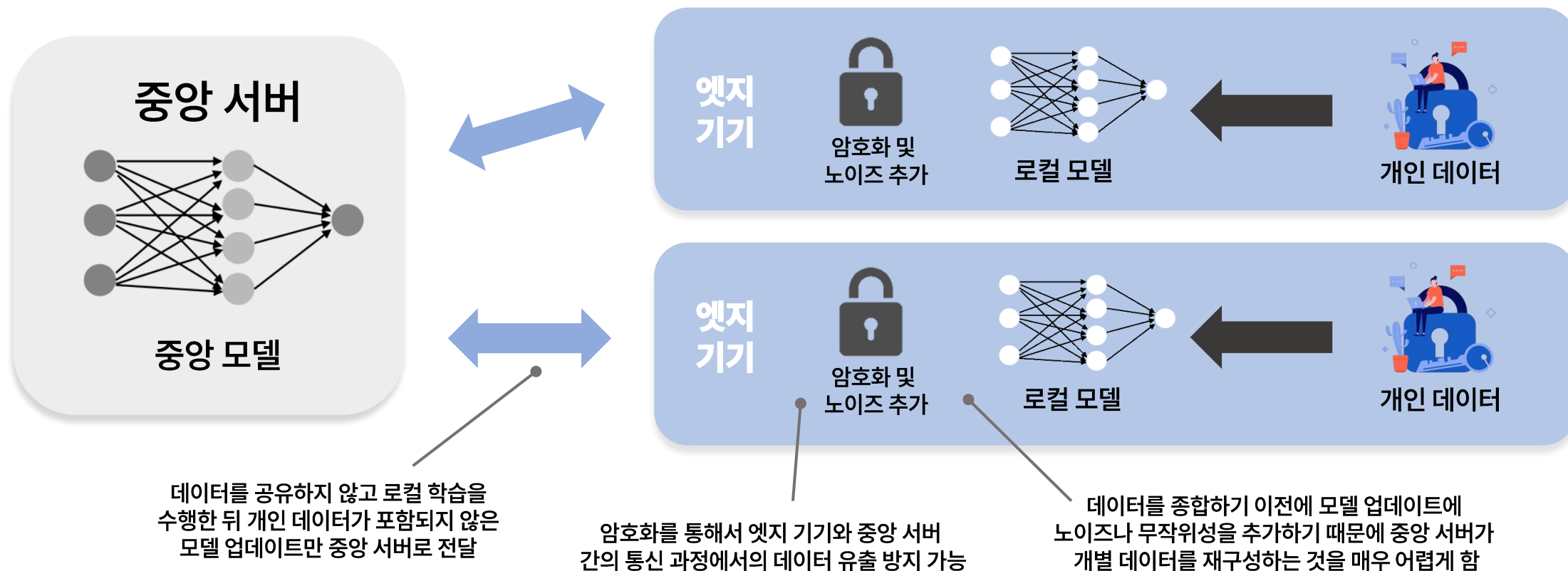
2) 기여한만큼 돌려받는 능동적인 유저 보상 체계

- 기업의 데이터 **판매** 또는 **구매**를 위한 네트워크
- 스마트 컨트랙트에 의해 데이터 판매자가 **프라이버시 유출의 위험** 없이도 데이터를 안전하게 판매할 수 있도록 보장됨
- 구매자와 판매자의 신원은 **비공개로 유지**
- 데이터 판매자의 이름은 판매자가 100% **권한을 부여한 경우**에만 데이터 구매자에게 노출됨
- 데이터 거래 이후에도 데이터 프라이버시에 대한 **판매자의 권리는 보호되고 보존됨**

**탈중앙화
데이터
마켓플레이스**
(ex. OceanProtocol)

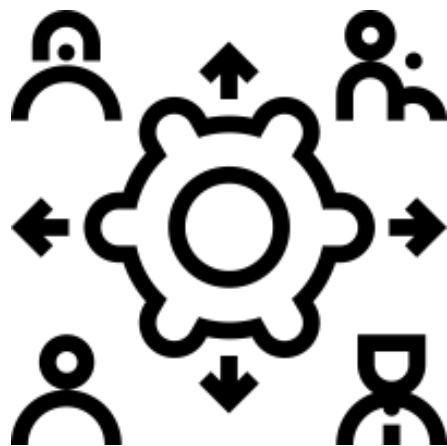
3) 시스템 분산화로 강화하는 사이버 보안

중앙화된 서버가 아닌, 유저 개개인의 엣지 기기에 데이터를 보존함으로써 데이터 오용 및 무단 액세스 등 각종 **사이버 보안의 리스크**를 줄일 수 있다.

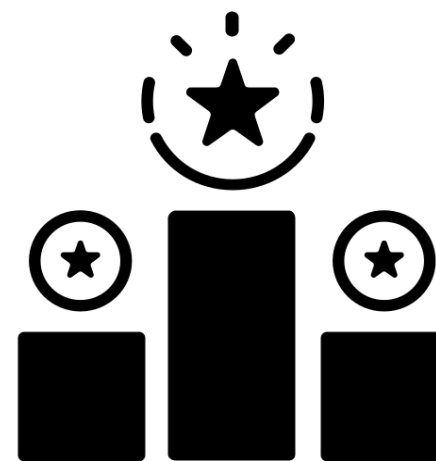


4) 높은 호환성 & 연결성을 자랑하는 AI 생태계

모두가 특정 에이전트 시스템 개발 및 운영사가 규정하는 프로토콜에 따르는 것이 아닌,
균등한 경쟁환경 아래 사용자간 연결성과 호환성이 극대화되는 **협력적인 AI 생태계**가 조성될 것이다.

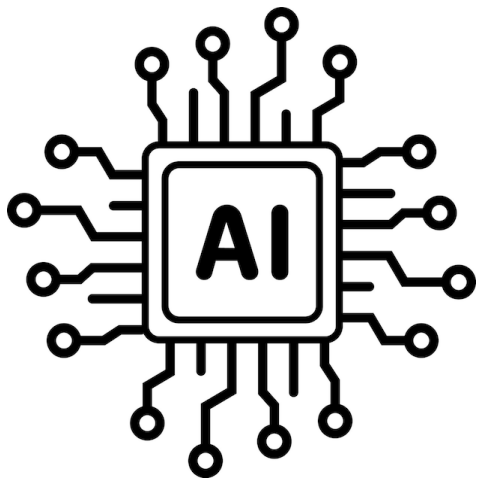


사용자간 더 연결되고
협력적인 AI 생태계 조성



소수 주체의 기업 집중보다
균등한 경쟁환경 조성

AI 에이전트 시스템의 전력 소비 파이프라인을 **노드 컴퓨팅** 기법을 통해 분산시키면
에너지 소비량을 효과적으로 절약할 수 있다.



단일모델 훈련 시 전력량



미국 100가구 전력량



효율적 컴퓨팅
인프라로 에너지 절약

References

참고문헌

- 최진영. (2021). 오픈소스 라이선스 분쟁에 관한 연구. The Journal of Law & IP, 11(2).
- 이효진. (2023). 우리 '인공지능법 제정안'의 인공지능 규율 방향에 관한 소고 - 유럽연합 인공지능법(EU AI Act) 규제 비교를 중심으로 -. 法學論文集, 47(2), 20-21.
- 김윤명. (2023). AI관련 발명에 있어서 데이터 공개. 저스티스, 196, 170-198.
- 성준호 (2013). 빅데이터 환경에서 개인정보보호에 관한 법적 검토. 法學研究, 21(2), 2-8.
- 정창우(Changwoo Jeong);이혜진(Hyejin Lee). (2022). 도덕과에서 AI 윤리교육의 필요성과 과제. The SNU Journal of Education Research, 31(1), 55-82. 10.54346/sjer.2022.31.1.55
- Malik, S. (2023, September 2). Blockchain-Based AI Marketplaces: Fostering Innovation and Trust. Medium. <https://medium.com/@imsaleemmalik/blockchain-based-ai-marketplaces-fostering-innovation-and-trust-dfe4ac2c990>
- CoinEx Blog. (2023, June 2). What New Business Model Will Web 3 and AI Collide With? <https://www.coinex.com/ko/blog/3114-what-new-business-model-will-web-3-and-ai-collide-with>
- Coindesk. (2023, September 22). AI Should Be Decentralized, But How? Consensus Magazine. <https://www.coindesk.com/consensus-magazine/2023/09/22/ai-should-be-decentralized-but-how/>
- Li, D., Du, X., Yang, K., & Wang, Y. (2022). Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. Scientific Reports, 12, 7517. <https://doi.org/10.1038/s41598-022-12833-x>
- Zhang, L., Yang, P., & Zhang, Y. (2020). 6G Vision: An AI-Driven Decentralized Network and Service Architecture. ResearchGate. https://www.researchgate.net/publication/344260938_6G_Vision_An_AI-Driven_Decentralized_Network_and_Service_Architecture
- Wang, X., Han, Y., Leung, V. C. M., Niyato, D., Yanikomeroglu, H., & Chen, X. (2021). Decentralized Deep Learning for Multi-Access Edge Computing: A Survey on Communication Efficiency and Trustworthiness. ResearchGate. https://www.researchgate.net/publication/356940150_Decentralized_Deep_Learning_for_Multi-Access_Edge_Computing_A_Survey_on_Communication_Efficiency_and_Trustworthiness
- Journal of Grid Computing. (2023, July 31). Decentralized AI for Smart Grid Privacy: A Federated Learning Approach. Springer. <https://link.springer.com/journal/10723/updates/24079090>
- Forbes Tech Council. (2018, January 11). Decentralized Artificial Intelligence Is Coming: Here's What You Need to Know. Forbes. <https://www.forbes.com/sites/forbestechcouncil/2018/01/11/decentralized-artificial-intelligence-is-coming-heres-what-you-need-to-know/?sh=108cfe26146d>

- Allocation framework for hierarchical federated learning (IEEE Computer Society).(2021).<https://ieeexplore.ieee.org/document/9479786>
- Decentralized AI: edge intelligence and smart blockchain, metaverse, web3, and DeSci (IEEE Intelligent Systems).(2022).<https://www.computer.org/csdl/magazine/ex/2022/03/09839452/1FisW0wf6bS>
- A survey of blockchain, artificial intelligence, and edge computing for web 3.0 (arXiv).(2023).<https://arxiv.org/abs/2311.13731>
- Edge computing in the age of AI: an overview.(2023).<https://infohub.delltechnologies.com/p/edge-computing-in-the-age-of-ai-an-overview/>
- IEEE Xplore. (2021). Decentralized AI Governance: A Blockchain-based Approach. IEEE. <https://ieeexplore.ieee.org/document/9479786>
- IEEE Computer Society. (2022). Decentralized AI: Challenges and Opportunities in Edge Computing. IEEE Computer Society Digital Library. <https://www.computer.org/csdl/magazine/ex/2022/03/09839452/1FisW0wf6bS>
- Verma, S., & Alam, M. (2023). Decentralized AI for Cyber-Physical Systems: Opportunities, Challenges, and Road Ahead. arXiv. <https://arxiv.org/abs/2311.13731>
- TechRxiv. (n.d.). Integrating Edge Intelligence and Blockchain: What, Why, and How. https://www.techrxiv.org/articles/preprint/Integrating_Edge_Intelligence_and_Blockchain_What_Why_and_How/19634610
- Stockholm University. (2023). Decentralized AI and Privacy Concerns. <https://su.diva-portal.org/smash/get/diva2:1784381/FULLTEXT01.pdf>
- Dell Technologies. (2023, September 27). Edge Computing in the Age of AI: An Overview. Retrieved December 10, 2023, from <https://infohub.delltechnologies.com/p/edge-computing-in-the-age-of-ai-an-overview/>
- International Energy Agency. (2023, November 2). Why AI and Energy Are the New Power Couple. Retrieved December 10, 2023, from <https://www.iea.org/commentaries/why-ai-and-energy-are-the-new-power-couple>
- Decentralized AI through Federated Learning & Edge Computing. (2022). <https://www.nature.com/articles/s41598-022-12833-x>
- 6G Vision: An AI-driven decentralized network and service architecture (IEEE Internet Computing).(2020). https://www.researchgate.net/publication/344260938_6G_Vision_An_AI-Driven_Decentralized_Network_and_Service_Architecture
- Decentralized deep learning for multi-access edge computing: a survey on communication efficiency and trustworthiness (IEEE Transactions on AI).(2021). https://www.researchgate.net/publication/356940150_Decentralized_Deep_Learning_for_Multi_Access_Edge_Computing_A_Survey_on_Communication_Efficiency_and_Trustworthiness



감사합니다!