

소버린 AI에서 포용적 AI: 네이버의 글로벌 AI 전략

2024. 10

하정우, PhD: jungwoo.ha@navercorp.com

네이버클라우드 AI Innovation 센터장 겸 네이버 Future AI 센터장

과실연 AI 미래포럼 공동의장 겸 AI 정책연구소장

한국공학한림원 컴퓨팅분과 정회원

NAVER Cloud

알파폴드와 알파프테오: 단백질 구조 예측 & 신 단백질 구조

Google DeepMind

About ▾

Research ▾

Technologies ▾

Discover ▾

AlphaFold

Overview

Impact stories

AlphaFold Database [↗](#)

AlphaFold Server

AlphaFold

Accelerating breakthroughs in biology with AI

Explore the AlphaFold Database [↗](#)

AlphaProteo generates novel proteins for biology and health research

5 SEPTEMBER 2024

Protein Design and Wet Lab teams

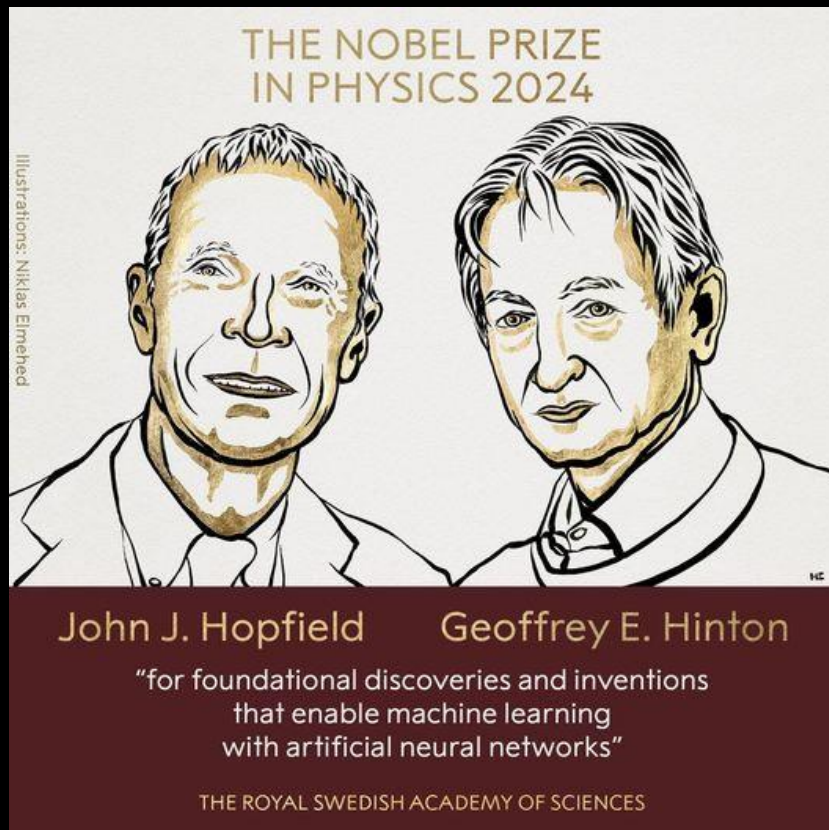
[Share](#)

<https://deepmind.google/technologies/alphafold/>

<https://deepmind.google/discover/blog/alphaproteo-generates-novel-proteins-for-biology-and-health-research/>

NAVER Cloud

AI의 시대: 노벨 물리학상 & 노벨 화학상



Nobel Prize Dialogue

Other panellists

Michael
Bruker
London
School of
Economics

Emilie
Caspar
Ghent
University

Bernard
Feringa
Nobel Prize in
Chemistry
2016

Věra
Jourová
European
Commission

Maria
Leptin



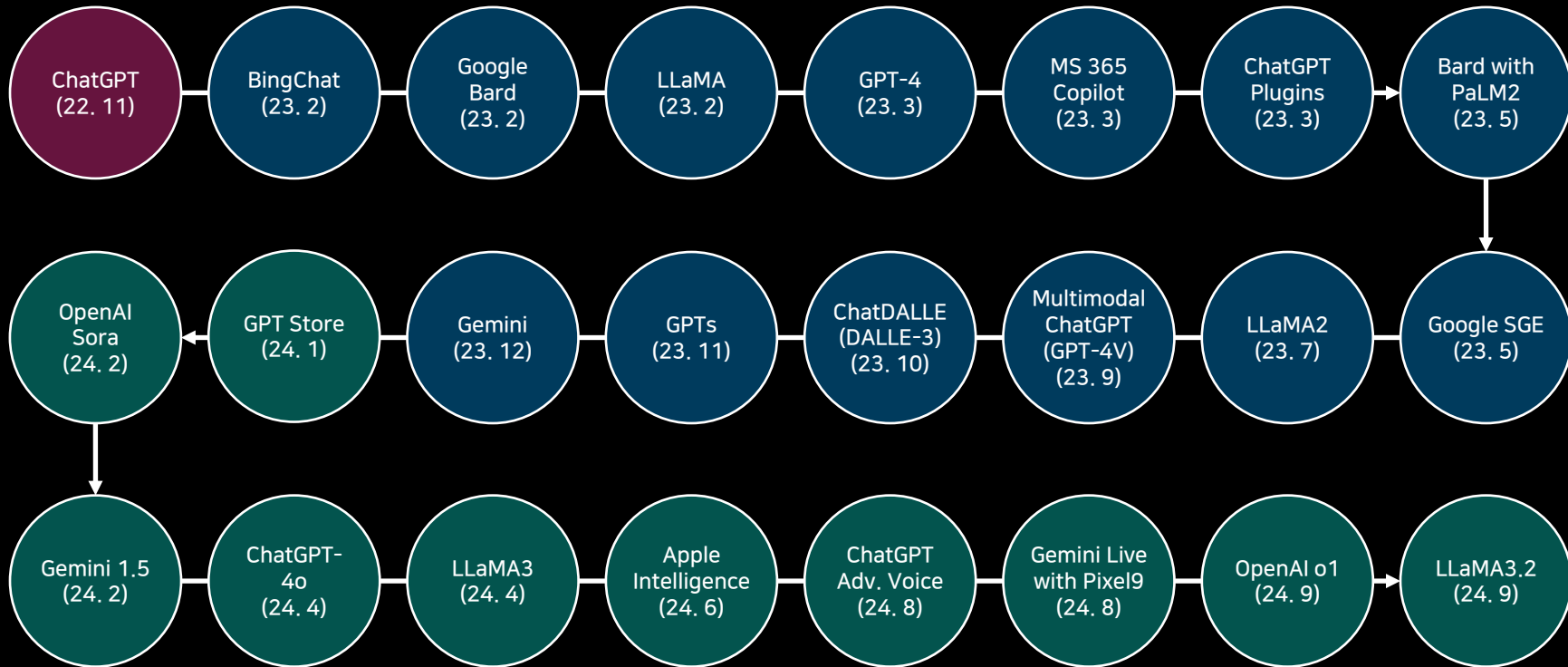
Demis Hassabis

Demis Hassabis is the co-founder and CEO of Google DeepMind, one of the world's leading AI research organisations.

Demis Hassabis is the co-founder and CEO of Google DeepMind, one of the world's leading AI research organisations. Founded in 2010, DeepMind has produced numerous landmark AI breakthroughs such as AlphaGo, the first program to beat the world champion at the game of Go, and AlphaFold, which was heralded as a solution to the 50-year grand challenge of protein folding.

Hassabis has won many prestigious international awards for his research work including the Breakthrough Prize in Life Sciences, the Canada Gairdner Award, and the Lasker Award. His work has been cited over 100,000 times, and he is a Fellow of the Royal Society and the Royal Academy of Engineering.

빛의 속도로 발전하는 생성 AI: 멀티모달, 온디바이스, 비용효율성, 에이전트



Meta Movie Gen: 영상 내 물체 변경

Transform the lantern into a bubble that soars into the air.



IQ Test Results

Reset

Show Offline Test

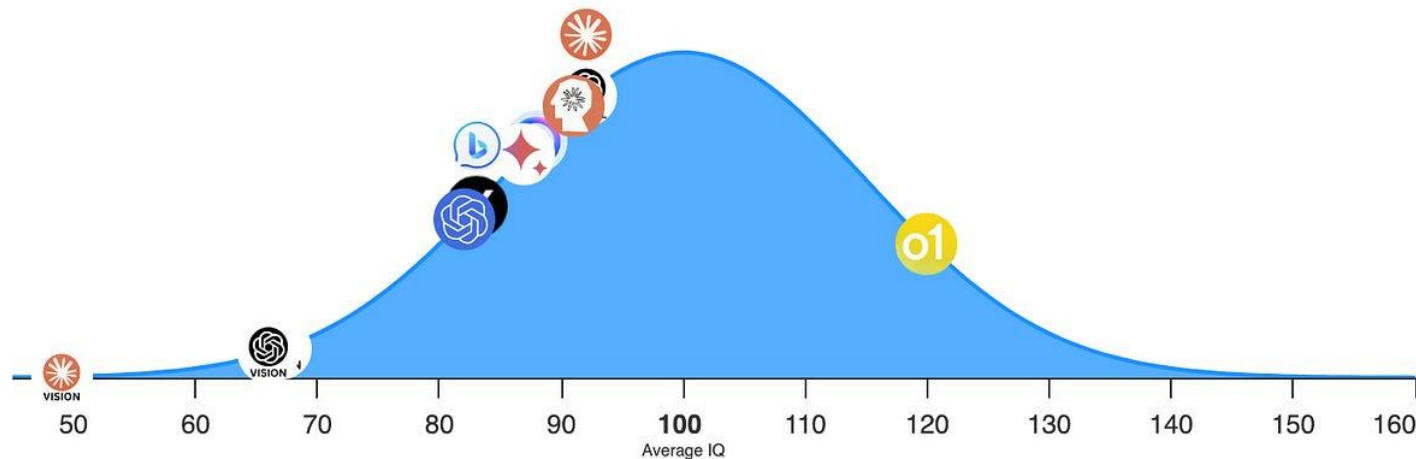
Show Mensa Norway



Score reflects average of last 7 tests given

MAT
MathVista (testmi
MMMU (v
MMI

AP English Lan
AP Physics
AP English I
LS/
AP Calcul
AP Chemist
SAT EBR



o1 OpenAI o1 preview



Gemini Advanced (Vision)



GPT4 Omni



Claude-3.5 Sonnet



Llama-3.1



Gemini Advanced



ChatGPT-4



Claude-3 Opus



Grok-2



GPT4 Omni (Vision)



Bing Copilot



Claude-3 Opus (Vision)

Inference

Post-training

Pre-training

온디바이스를 위한 sLM and sVLM (네트워크 불가 전장에 유효)

What is Phi-3.5-vision?

Explained with Working Examples



Research Graph · Follow

13 min read · Aug 30, 2024



39



Phi-3-mini
(3.8B)



Phi-3-vision
(4.2B)



Phi-3-small
(7B)



Phi-3-medium
(14B)



Phi-3.5-mini
(3.8B)



Phi-3.5-vision
(4.2B)



Phi-3.5-MoE
(6.6B active /
42B total)

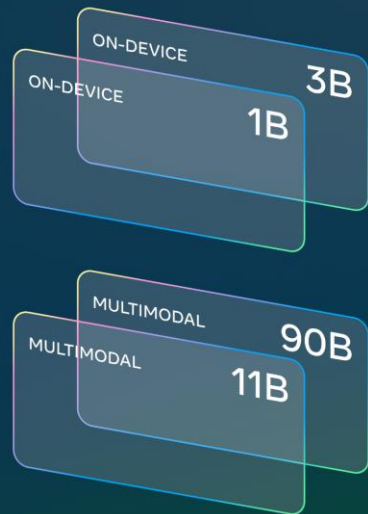
Instruction Tuned

RAI Safety Aligned

The Phi3 model series and its variations (Source [5])

INTRODUCING

Lightweight and multimodal Llama models



JULIAN CHOKKATTU

GEAR AUG 16, 2024 12:56 PM

What Is Gemini Live and How Do You Use It?

Google is rolling out this new voice assistant, and now you can have real-time, natural convers



사용자

apps

apps

apps

apps

apps

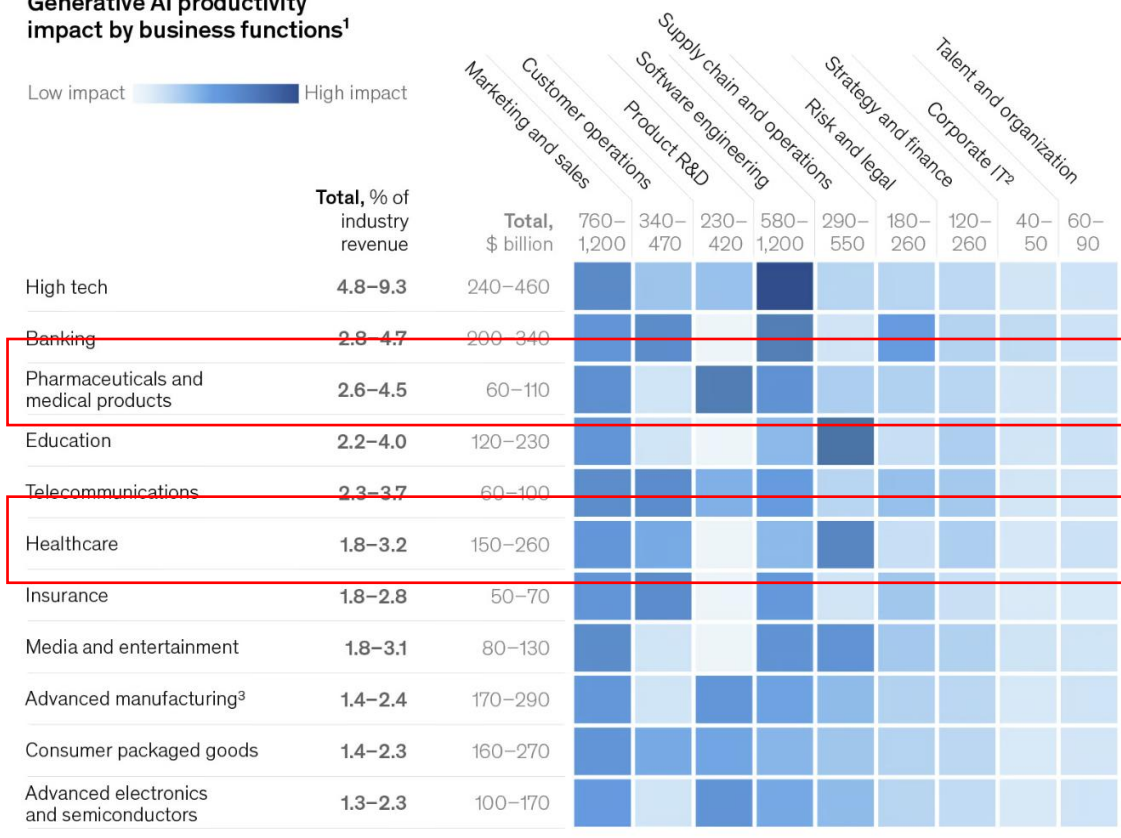
VER Cloud

McKinsey의 생성 AI 보고서 (2024. 1)

- 생성 AI 생산성
- 마케팅 세일즈,
- 위 4가지는 산업
- 생성 AI가 가져올
- 이런 일하는 방식
- 기업의 투자가 필
- 아직 늦지 않았

Generative AI productivity
impact by business functions¹

Low impact  High impact



야의 75% 차지

로써 가능

생성 AI 를 활용한 도표 만들기

Potential Use of Generative AI

Lawyers see the highest potential for generative AI tools to assist them in researching matters (65%). Drafting documents (56%), document analysis (44%), and email writing (35%) are also important potential use cases.

"Our survey confirms what we hear from customers all over the world every day, that they are excited about the potential of generative AI to help improve their productivity, efficiency, and overall business and practice of law," said Mike Walsh, CEO of LexisNexis Legal & Professional. "Customer-driven innovation is core to the approach we take with product development, and LexisNexis is excited that our Lexis+ AI platform safely and securely provides critical generative AI tools to help legal professionals excel in their jobs."

work with trusted and ethical companies as they begin to use generative AI solutions in their work.

Consumer Use of Generative AI for Legal Work

Consumers largely use generative AI for researching legal topics (60%). Many would consider using AI for legal assistance to create a will (40%), set up a business (37%), and develop a rental agreement (39%).

위 글에 있는 내용을 분석해서 각각 그래프로 그려줘

Potential Use Of Generative AI By Lawyers



Awareness Of Generative AI



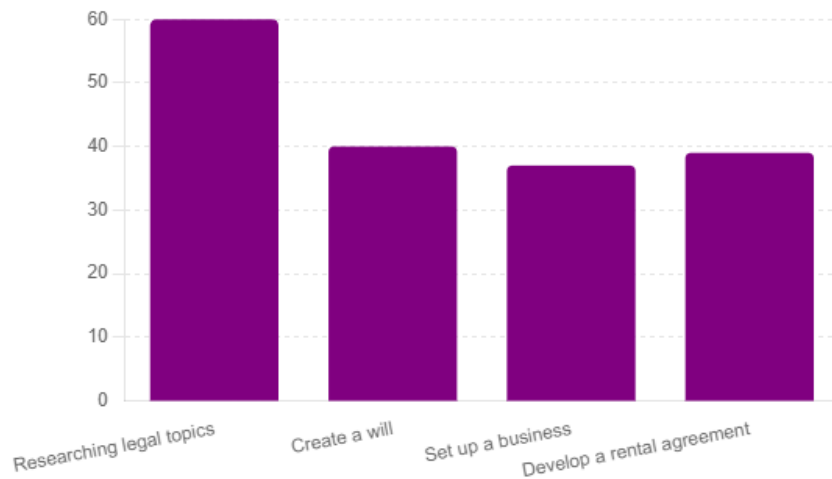
Ethics Regarding Generative AI



Consumer Use Of Generative AI For Legal Work



Y Percentage (%) X Use Cases



시가 불러온 일하는 방식의 변화

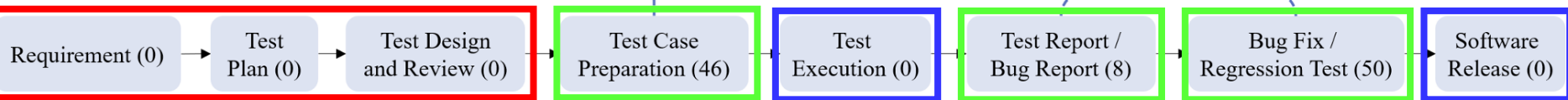
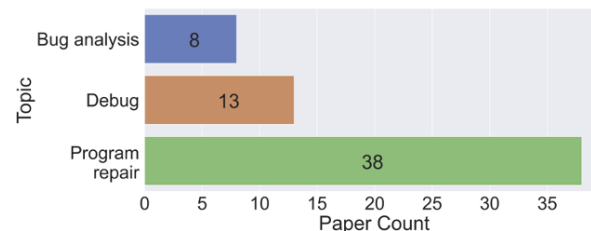
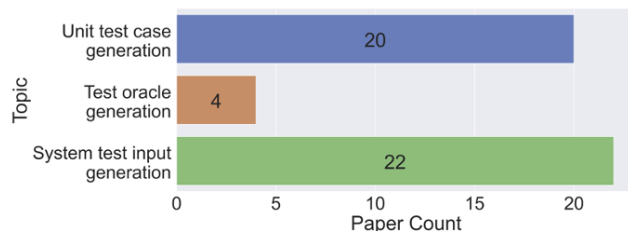
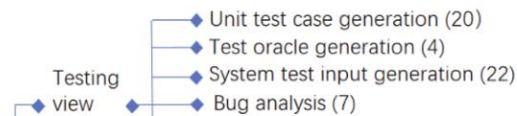
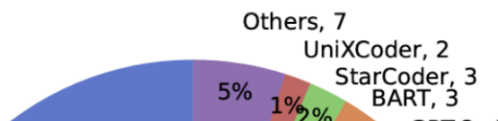
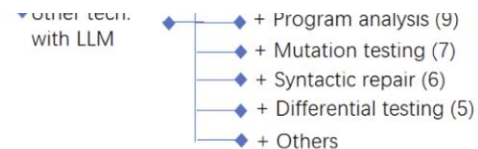


Fig. 4: Distribution of testing tasks with LLMs (aligned with software testing life cycle [134]–[136], the number in bracket indicates the number of collected studies per task, and one paper might involve multiple tasks)

Fig. 6: LLMs used in the collected papers



생성 AI가 불러온 신 AI 제국주의: AI 종속국 vs. AI 독립국 vs. AI 패권국



소버린 AI란?

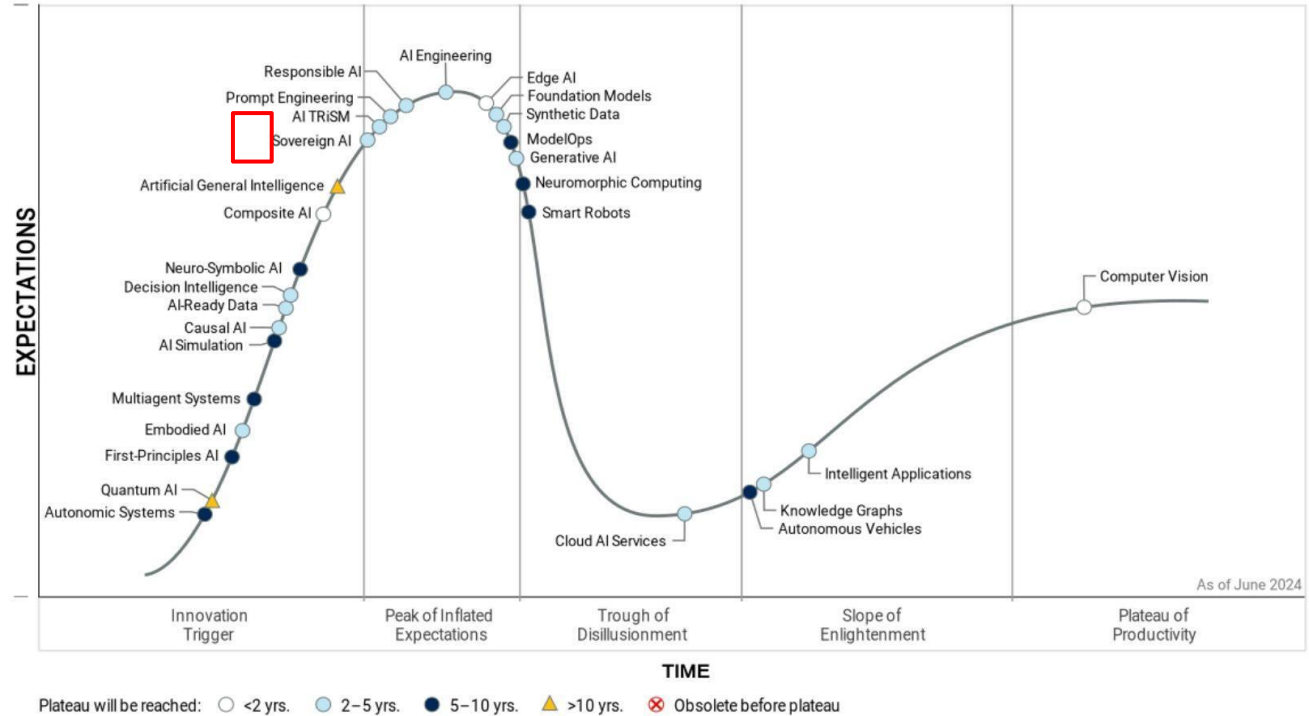
- 해당 국가의 문화나, 가치관, 사회/정치적 환경을 보다 정확하게 이해하고 콘텐츠를 만들 수 있는 AI
- 빅테크 AI 배제가 아닌 적용 분야에 따라 경쟁력 있는 선택지를 확보하자는 것
- AI모델 뿐 아니라 전력, 컴퓨팅 인프라, AI 데이터센터, 클라우드, AI 배포 서비스, 산업 적용 확산을 아우르는 전체 밸류체인 혹은 생태계 개념
- AI 안전성과 윤리 또한 소버린 AI 와 직결
(사회 문화마다 민감한 이슈가 다름, 북미-인종, 한국-페미니즘)
- 모든 나라가 자국을 정확히 이해하는 소버린 AI를 갖고 싶어함
- 직접 자체 기술로 모두 확보 가능하면 외국 의존 최소화가 최선, 당장 불가능한 경우엔 해외 신뢰 가능한 파트너와 확보

Sovereign AI: Promising Business Sectors

The Economist



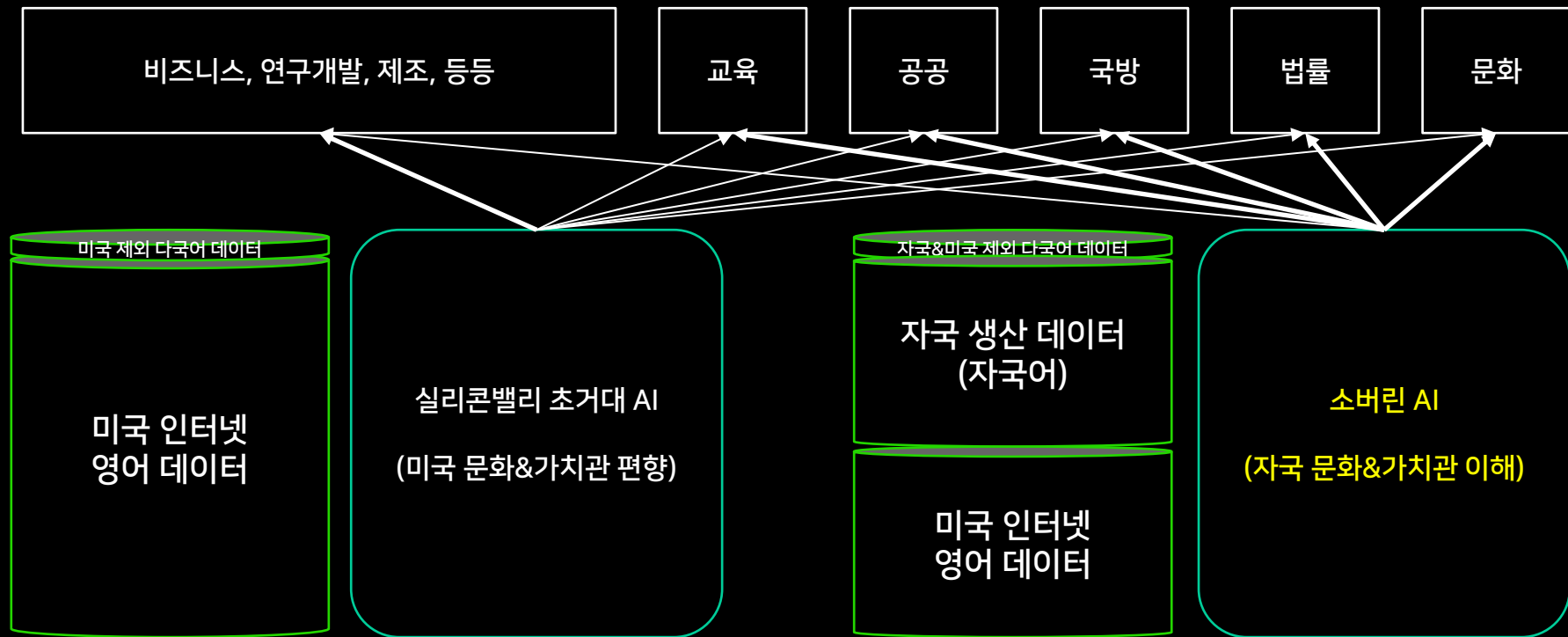
Hype Cycle for Artificial Intelligence, 2024



Gartner

NAVER Cloud

소버린 AI는 어떻게 만들 수 있는가?



오픈소스(오픈 바이너리) AI 활용하면 된다??

더밀크박인터뷰

오픈AI · LLM · 소버린 AI

조경현 교수 “기업은 자체 LLM 만들어야 합니다. 이유는...”



박원익

2024.07.26 14:00 PDT

북마크 하기



데이터
학습 데이터

데이터 투
오동작시 분석
(오픈 모델은 데

ing 만으로 소버린 불가

에서 데이터가 부족하면
써도 지역 문화를 정확히
하는 AI 만들기 어렵다.

각국 정부는 LLaMA 의존 않고 소버린 AI 확보 위해 자국 AI 대표기업 전략적 육성 중

프랑스
(9조투자)

미스트랄
H

일본

소프트뱅크
(4500억)
Sakana.ai

캐나다
(2.4조원
투자)

Cohere

독일

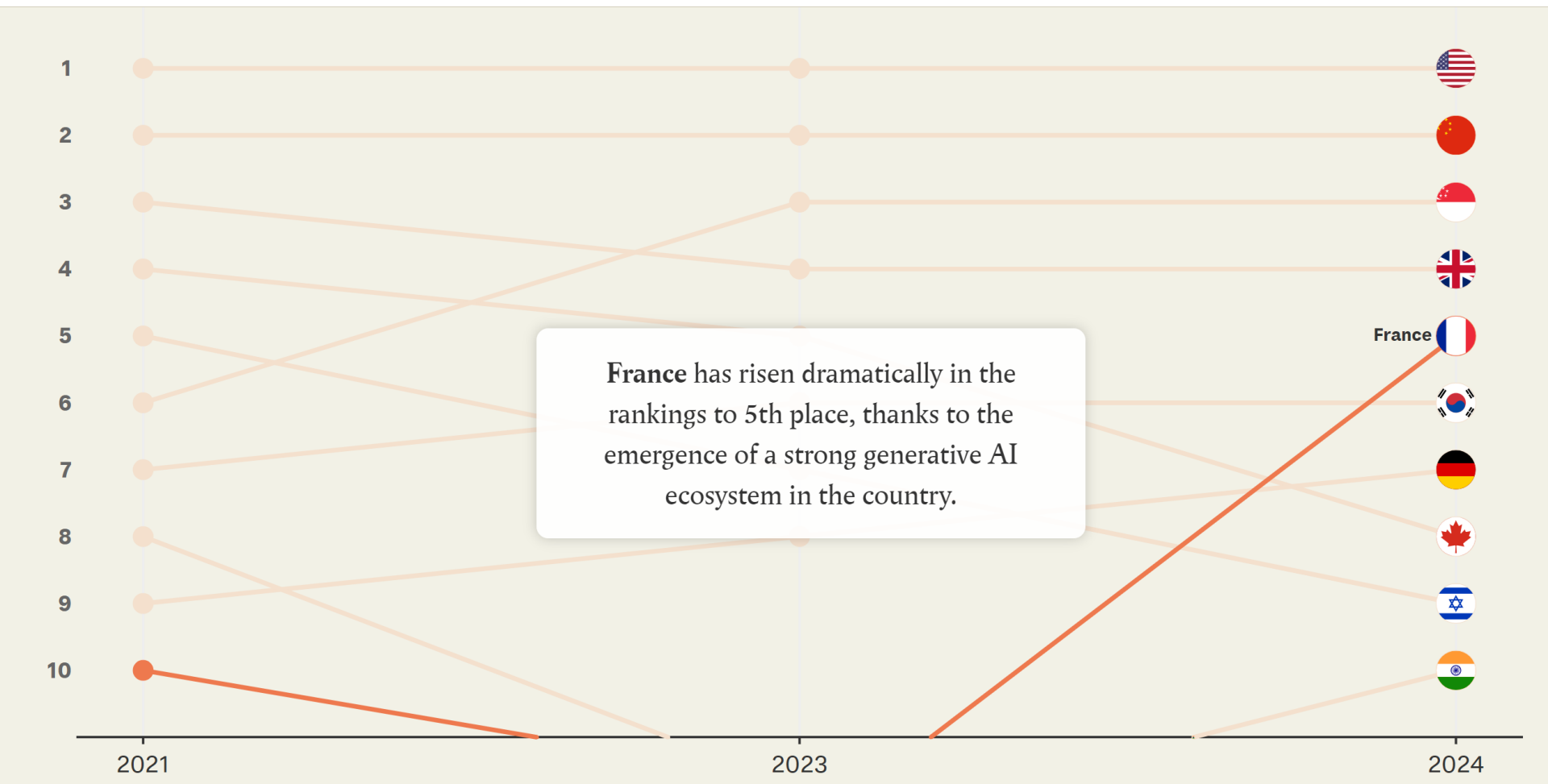
Aleph Alpha

인도

Krutrim
OpenHathi

Italy

iGenius
(Italia, 24. 6)



소버린 AI의 범위와 중요성

모든 산업 생성형 AI 확산 및 산업 생태계

생성형 AI 파인튜닝, 배포 및 활용

파운데이션 모델 (생성형 AI)

학습 데이터 및 서비스 데이터

반도체, AI 데이터센터, 클라우드 (IaaS)

에너지 (전력 수급)

소버린 AI에도 단계가 있다.(예시) – 최소 자체 Lv3 확보 필수

| | 자체 기술 | 외부 기술 협력 |
|------------------------------------------|---------------------------------------|----------------------------|
| Level-0: 기술 종속 | 대부분의 국가 | |
| Level-1: 자국어 소량 데이터 구축&오픈소스 파인튜닝 모델 확보 | 인도네시아, 태국, 베트남 | |
| Level-2: 영어 중심 Pre-training 모델 확보 | 핀란드, 싱가포르 | |
| Level-3: 자국어 대량 데이터 & Pre-training 모델 확보 | 대만, 이스라엘, 인도, 독일, 프랑스, 이탈리아, 사우디 아라비아 | UAE, 일본 |
| Level-4: 클라우드 기반 자체 소버린 AI API 제공 | 러시아 | 독일, 프랑스, 일본 (OpenAI), 이탈리아 |
| Level-5: 클라우드 기반 산업생태계 구축 | 미국, 중국, 한국 | |

소버린 AI 군비에 해당하는 영역 (클라우드, 컴퓨팅 인프라, 반도체, AI 데이터센터, 자체 Foundation model 생산능력) → 이것 없이는 사상누각

한국문화를 가장 잘 이해하면서 영어도 LLaMA3 만큼 하는 AI

I Weekend



Art by Clark Miller

The Big Read

How the World Plans to Stop American AI Domination

Around the world, the push for homegrown AI often originates from a belief that English-first AI models from the U.S. put other languages and cultures at a disadvantage.

In South Korea, for example, internet conglomerate Naver is developing its own LLM, HyperClova X, that it plans to use across its many services. That model could touch a lot of Naver applications in the hyperconnected country: Imagine a mashup of Amazon, Google and Uber, and you'll get an idea of the ground the company covers in South Korea.

Inside Naver, a team of safety experts is focused on making sure HyperClova X produces answers that are culturally appropriate for its domestic audience. Kang Min Yoo, a Naver research lead, said that Korea has a rich culture of neologisms that foreign large language models often fail to understand. Similarly, he said Koreans tend to support government controls on things like housing prices more than Americans do, and HyperClova X's answers reflect these social values better than non-Korean models do.

Large language models, he said, "could have a different answer depending on the context and the location."

Another important motive behind HyperClova X is the quest for lower computation costs. The fees for using LLMs are typically based on the amount of data they're asked to process, which they break down into chunks of words or characters called tokens. More tokens equal more costs.

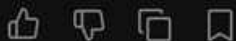
Naver's model—which is trained on 6,500 times more Korean data than OpenAI's GPT-3, according to the company—is able to break down Korean queries into fewer tokens compared to other, English-centric models, making it cheaper, Yoo says. ([Independent analyses of western large language models LLMs](#) have shown that queries in non-English languages require significantly more tokens to process than they do in English.)



이 사진에 대해 설명해줘



이 사진은 백제 금동 대향로입니다. 이 향로는 백제 문화의 우수성을 보여주는 대표적인 유물로 국보로 지정되어 있습니다. 이 향로는 국립 중앙 박물관에 소장되어 있으며 백제 예술의 아름다움을 보여주는 작품으로 높이가 64cm에 이르는 대형 향로입니다. 이 향로는 뚜껑, 몸체, 받침으로 구성되어 있으며 뚜껑에는 봉황이 있고 몸체에는 연꽃과 여러 동물들이 새겨져 있으며 받침에는 용이 새겨져 있습니다.



다른 답변 보기 ↻

이 사진은 한국
가능성이 높습니

이 향로는 화려
며, 그 위에는 봉
실과 관련된 유

중앙의 향을 담
문양이 복잡하
미적 감각과 기

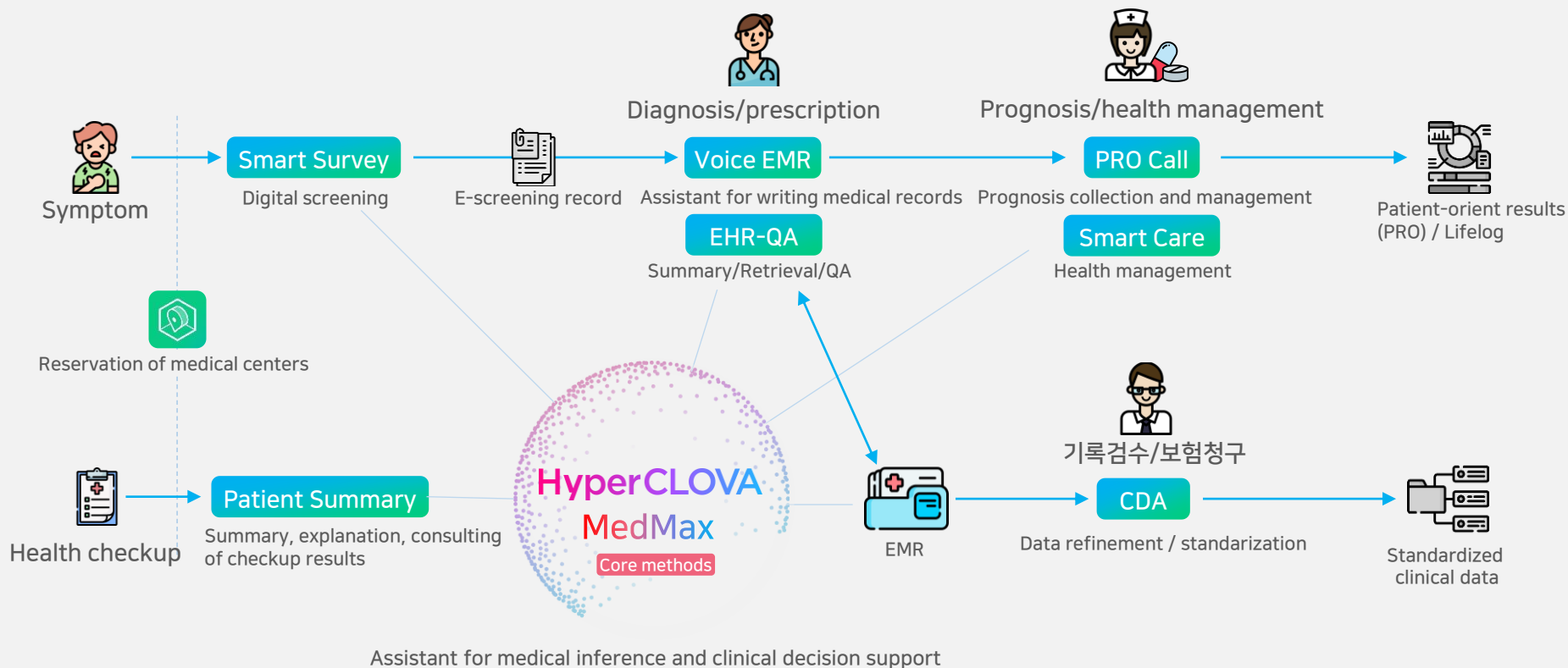
전체적으로 이
요한 문화재입니

🔍 📄 🔄

클로바 케어콜: 전국 지자체 절반 이상 30000명 이상 독거노인 말벗

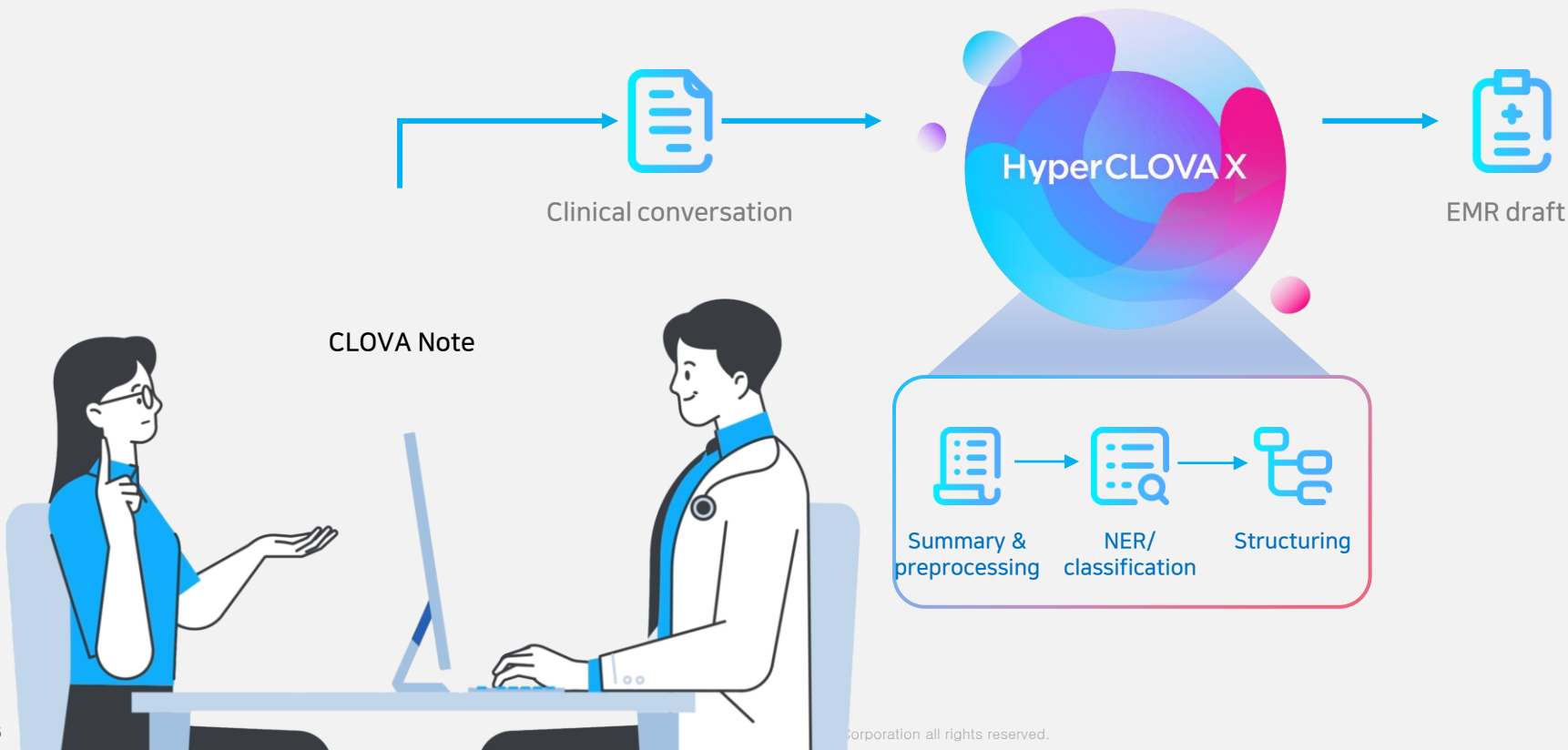
어르신 실제 대화 사례 1

CLOVA for Hospital (Applied in NAVER 1784 In-House Hospital)



CLOVA Voice EMR

Automatic conversion from verbal conversation between a patient and a doctor to a formal EMR draft



HCX기반 건강검진 상담 서비스

건강검진 업계 최초 the first and only

AI 스마트검진리포트 에스프리 asKMI와 함께 더 똑똑해진 결과를 받아보세요!

에스프리 제공 서비스

- ✓ 건강검진 결과 해석
- ✓ 검사항목 설명
- ✓ 의료용어 해설
- ✓ 건강관리 방법 안내

스마트검진리포트 장점

- ✓ 모바일과 PC로 신속 간편 확인
- ✓ 지난 검진 결과와 한눈에 비교
- ✓ 언제 어디서나 PDF 파일로 활용
- ✓ 의료정보 특화 보안 솔루션으로 안심



스마트검진리포트의
에스프리 버스를 눌러보세요!



NAVER Cloud X KMI 한국의학연구소 X 대웅제약

NAVER Cloud

NAVER CLOUD PLATFORM



슈퍼컴퓨팅 인프라

Neurocloud for HyperCLOVA X



뉴로클라우드 인프라
서버·GPU·스토리지...

고객 영역

고객사
어플리케이션

고객사 특화
학습 데이터

고객사 기존 인프라
서버·스토리지·데이터베이스...

네이버클라우드 데이터센터

고객사 데이터센터

HyperCLOVA X-Dash + On-premise 로 보안 특화 AI

보안강화 클라우드


HyperCLOVA

| Category | Benchmark | mu-choco-003-32k | LLaMA 3.1 8b |
|----------|------------------|------------------|--------------|
| 한국어 | KMMLU | 47.85 | 41.20 |
| | HAERAE | 74.61 | 44.45 |
| | KoBigBench (KBB) | 64.64 | 51.37 |
| | CLiCK | 65.96 | 45.61 |
| | KoBEST | 70.99 | 52.99 |
| | KorMedMCQA | 64.94 | 51.95 |
| 영어 | KoFLAN | 42.61 | 33.27 |
| | MMLU | 65.65 | 65.12 |
| | BigBench-HARD | 46.75 | 47.35 |

B)

네이버클라우드, 미래에셋증권에 금융특화 소형언어모델 구축

입력 2024.09.19. 오전 10:10 - 수정 2024.09.19. 오전 10:11 기사원문

 김민석 기자

 추천  댓글

| 하이퍼클로바X 대시 기반 '구축형 LLM' 제공



네이버클라우드 미래에셋증권 하이퍼클로바X 기반 금융 sLLM 구축(네이버클라우드 제공)

and many other AI experts, put forward the Manhattan Declaration

September 26, 2024



전체 Value Chain 보유 3개국 중 하나: 소버린 AI 전체 생태계 경험으로 글로벌 진출

원팀 코리아

정부

네이버

통신사

대기업

스타트업

AI반도체 기업

...

다양한 산업 적용 및 확산 경험 (B2C, B2B, B2G)

컨설팅, 데이터 가공, 안전 프레임워크

생성 AI 배포 및 운영 시스템

자체 초거대 생성 AI (소버린 AI)

클라우드 (IaaS)

AI 데이터센터

반도체, AI 반도체, 컴퓨팅 인프라

중동지역

아세안

유럽일부

아프리카

중남미

[단독]사우디 날아간 이해진...아랍어 LLM 네이버가 만든다

정호준 기자 jeong.hojun@mk.co.kr

황순민 기자 smhwang@mk.co.kr

입력 : 2024-09-12 09:25:34 수정 : 2024-09-12 09:27:51



네이버, 소버린AI로 사우디 뚫었다...아랍어 AI 구축 지원
사우디 데이터인공지능청과 MOU
데이터센터부터 모델까지 전방위적 협력
AI 로봇틱스 관련 연구 개발도 진행



varieties Business Culture Sports Entertainment

ee to Operate Smart platform 'Baladi'



NAVER Cloud

무엇이 AI의 확산을 가로 막고 있나?

Must have
vs.
Nice to have
(RoI)

기술적 한계
(할루시네이션)

사용성 & 접근성

규제

안전성 & 보안

AI 기업에 조언: 다시 기술의 본질을 생각하라

1. 다목적 기술을 기업의 실용 솔루션으로 특화한다

2. 100원짜리 상품이 50원어치 가치를 제공하는 현실을 극복하라

3. 생성형 AI가 바꾼 개발 방법론을 적극 활용하라

4. AI 도입 효과를 측정할 수 있는 정량 지표를 만들어라

5. 전체 가치사슬 관점에서 최적화하라

6. AI 에이전트 시대를 대비하라

7. 중장기 연구에 투자한다

8. AI 리터러시에 대한 책임감을 갖는다

8. 체계적이고 책임감 있는 AI 거버넌스 운영

10. 소버린 AI: 다양성과 포용성 측면에서 글로벌을 바라보라

성공적인 AI활용을 위한 제언 (AI 활용 전문가의 View)

1. 역할과 책임에 따른 AI 활용의 방향성

2. AI 기반 서비스 및 AI 에이전트의 사용성

3. AI 기술을 활용한 서비스 품질 평가

4. 산업별 AI 활용에 대한 베스트 프랙티스 사례

5. AI 활용을 위한 업스킬링과 리스킬링 전략

6. AI 대전환 시대를 위한 AI 리터러시와 마인드셋

7. AI 시대가 원하는 인재의 조건

성공적 AI 도입 및 대전환을 위한 지침서 (24.10.16)

국가대표 AI 전문가 2인이 제안하는
AI 주도권 확보 전략

2025

AI 대전환

주도권을 선점하라


오순영, 하정우 지음

이광형(카카오 AI 연구원), 안현식(삼성전자 AI 연구원), 손현덕(메릴린드 AI 연구원),
배경훈(LG AI 연구원), 손대권(국군 군수사령부 사령관)
- 강력 추천 -

골든타임에 진입한 AI 주도권 전쟁

주저하며 몰락할 것인가,
선점하고 도약할 것인가?

한빛비즈



jungwoo.ha@navercorp.com



Pinned Tweet



Santiago 

@svpino

AI will not replace you. A person using AI will.