

# 생성형 AI 시대와 기업의 준비

2024년 4월 4일(목)  
컨설팅팀 김공환 그룹장



# AGENDA

---

- 01 AI의 현재와 미래
- 02 미래를 준비하는 SDS Gen AI 솔루션
- 03 Gen AI 도입을 위한 기업의 준비

01

---

# AI의 현재와 미래

☆ 시가 그린 그림  
**QUIZ** ☆

★ **힌트** ★

○ ○ ○ 를  
만들고 있는  
○ ○ ○ ○

**1번 문제**



☆ 시가 그린 그림  
**QUIZ** ☆

★ **힌트** ★

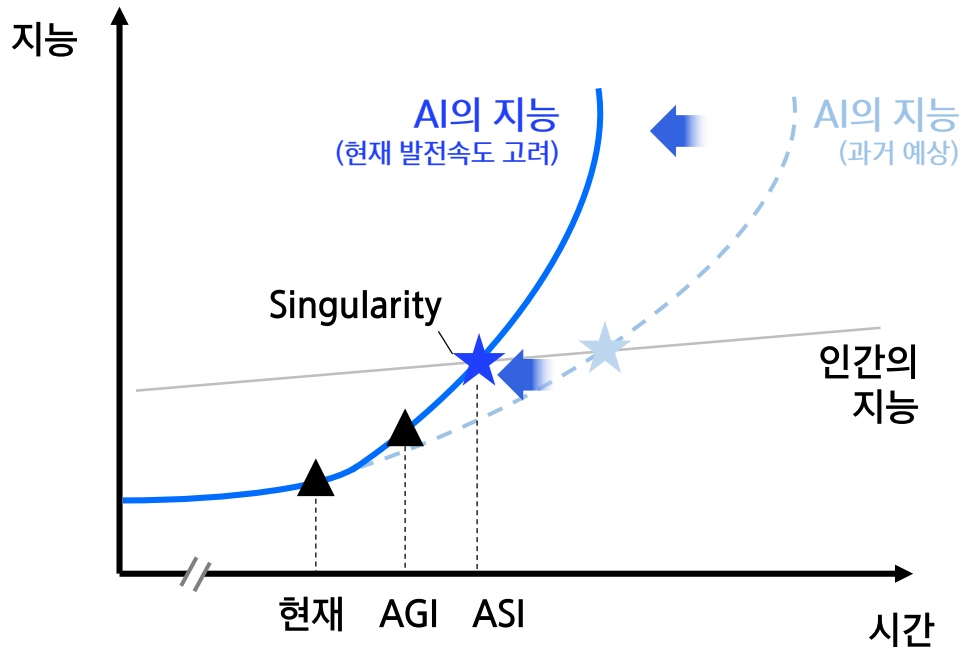
○ ○ 에서  
○ ○ 를 부르는  
○ ○ ○

**2번 문제**



# 우리가 마주할 AI의 미래

## AI지능 발전의 가속화



※ AGI(Artificial General Intelligence): 사람처럼 다양한 지적 작업을 수행할 수 있는 인공지능  
 ASI (Artificial Super Intelligence) : 인류의 전체 지능을 초월하는 인공지능

## AGI 5년내 등장, Computing 작업 방식의 변화

앞으로 5년 이내에 AGI가 등장할 것.

- NVIDIA CEO, 젠슨 황 (24.3월)

AI는 컴퓨터를 사용하는 방식을 완전히 바꿔놓을 것이다.

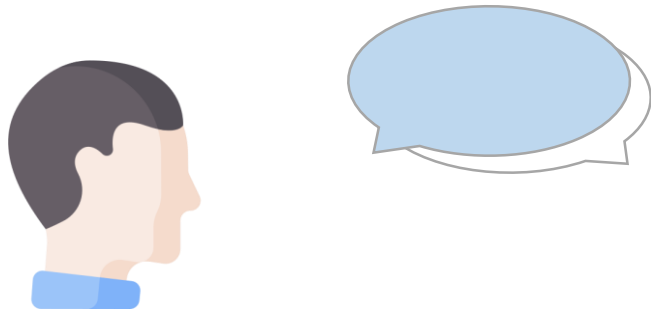
- MS Founder, 빌게이츠 (23.11월)

우리는 다시 출발점으로 돌아가...인류가 가진 독특한 능력들을 재검토해야한다.

- 사피엔스 저자, 유발 하라리 (22년 말)

# AI로 변화되는 일하는 방식

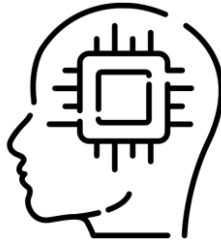
## 인간이 말하면



“경쟁사 M사의 주요 동향을 파악해서 메일로 보고해줘”

LUI (Language User Interface)

## AI는 이해하고



- STT (Speech To Text)
- TTS (Text To Speech)
- LLM (Large Language Models)
- Cognitive Search
- Multimodal Model

⋮

### 의도/목적 인지

- 경쟁사 M사의 다양한 정보를 수집, 분석 후 시사점을 도출한 보고서가 필요하구나...

### 할일 정의

- 경쟁사 홈페이지 접속 후 정보 추출
- 경쟁사 공시자료 수집 후 분석
- 경쟁사 최근 뉴스 검색 후 요약

## AI가 스스로 실행한다.



- AI Analytics
- Robotics
- VR/AR
- Autonomous Driving
- Biometrics

⋮

### 경쟁사 정보 수집

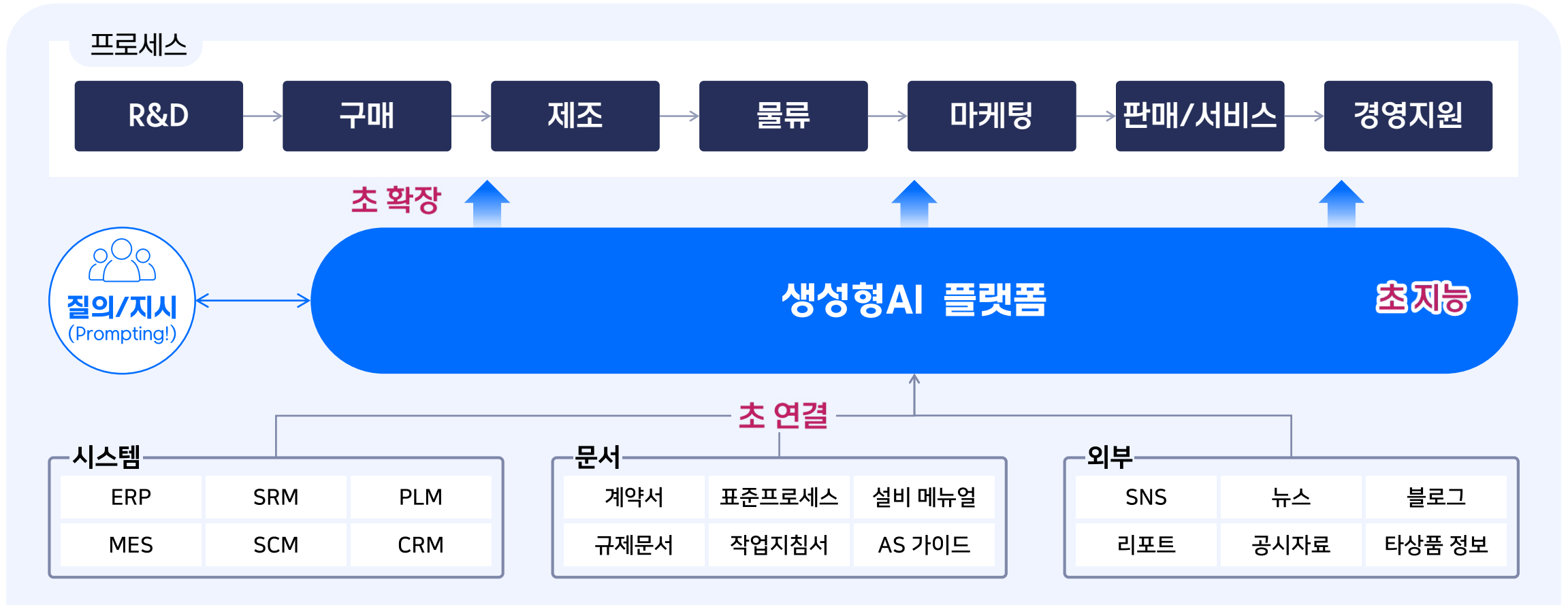
수집된 정보 요약/분석

동향 보고서 작성

메일 전송

# 기업 업무방식의 미래

사람은 질의/지시만! 생성형 AI가 스스로 연결, 생성, 분석 수행





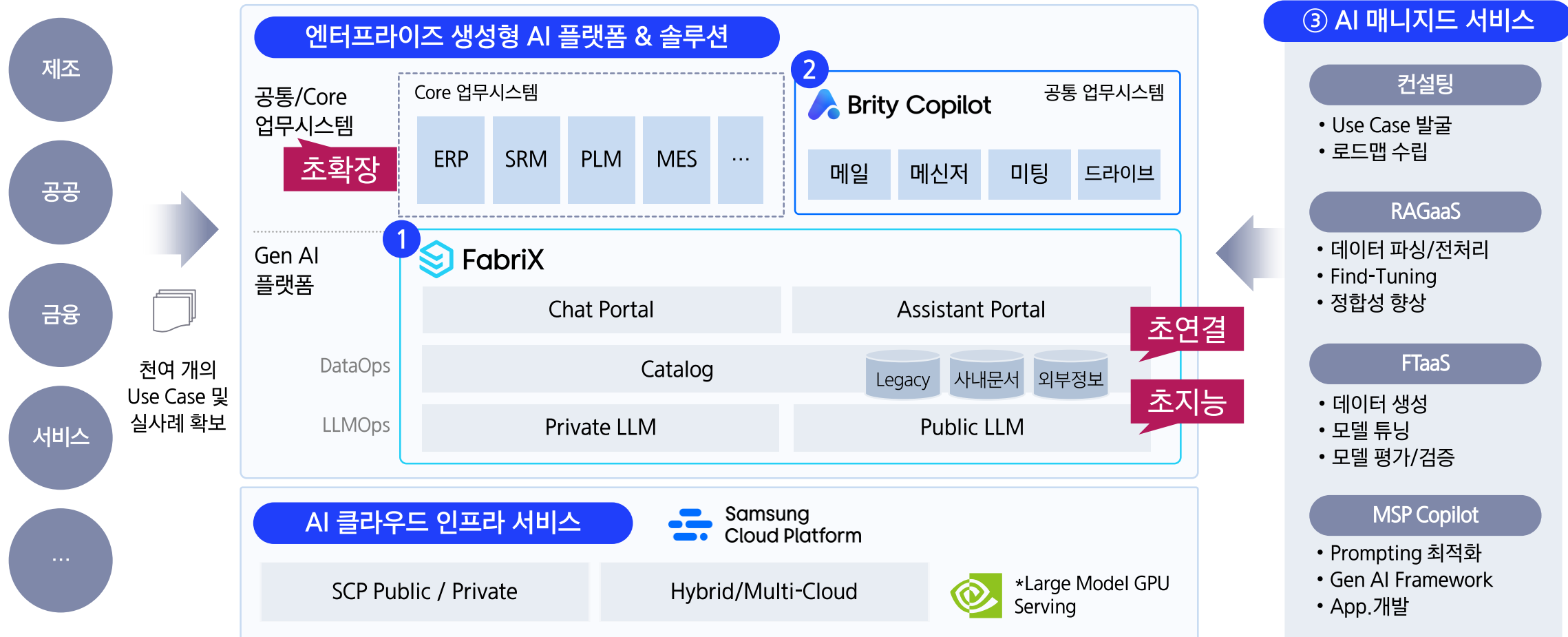
02

---

## **미래를 준비하는 SDS Gen AI 솔루션 (Live Demo)**

# Enterprise AI로 변화된 Workplace

삼성SDS는 엔터프라이즈 생성형 AI 서비스로 초확장/초연결/초지능 기반 기업의 업무를 보다 빠르고 편리하게 변화시킵니다.



03

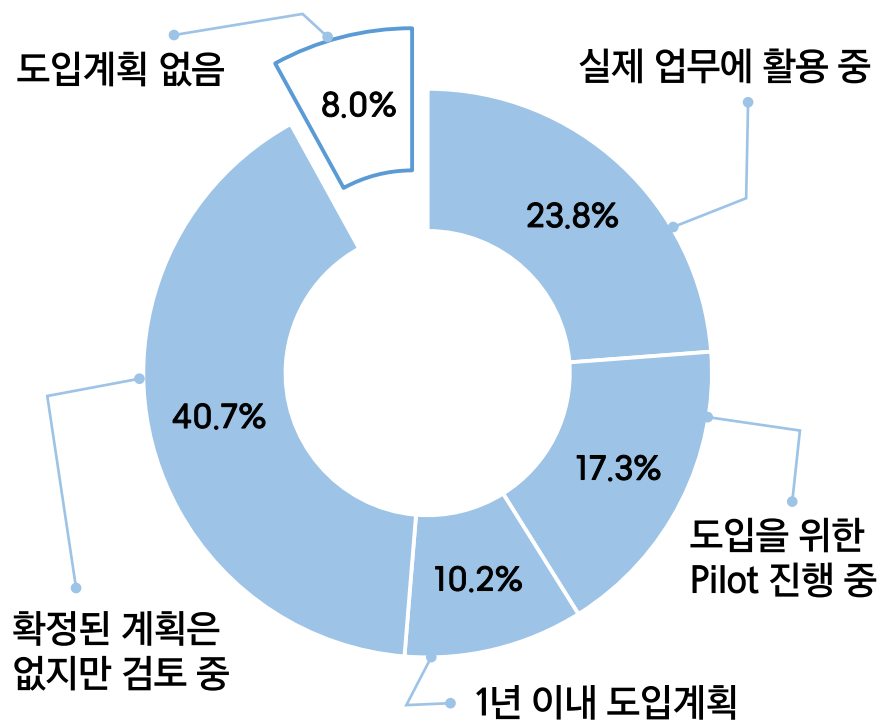
---

## Gen AI 도입을 위한 기업의 준비

# Gen AI을 도입하려는 기업의 주된 고민은?

대다수 기업이 도입 검토 중. 다만 어디에 우선 적용해야 할지, 어떤 솔루션과 인프라 구성으로 도입하면 좋을지 고민

## 90% 이상 기업이 도입 또는 검토 중



\* '23 국내 AI 도입 및 활용현황 조사 (IDG Korea, '23.7)

## 도입 시 주된 고민

Use Case	① 어디에 활용할 수 있지?
Architecture	② 어떻게 구현해야 되지?
Security	③ 기업데이터 유출 위험은 없을까?

# 어디에 활용할 수 있지?

## 고객사 대상 Gen AI 적용 사례 (200건↑) + SDS 임직원 Use Case BP (1,190건)

적용영역	기술유형	Use-Case
내/외부 지식기반 질의응답	RAG	<ul style="list-style-type: none"> <li>[ R&amp;D ] 특허동향 보고서 생성, 특허 침해여부 검토</li> <li>[ 제 조 ] 규제기관 Inspection 대응 위한 내부 문서검색</li> <li>[서비스] 정비 엔지니어를 위한 업무 매뉴얼 검색</li> </ul>
자연어 질의로 정형데이터 분석	Text2 Query	<ul style="list-style-type: none"> <li>[경영관리] 매출/채권데이터 분석</li> <li>[서비스] 상담사 고객대응 지원 위한 고객정보 분석</li> </ul>
코딩 지원	Text2 Code	<ul style="list-style-type: none"> <li>[SW개발] SW 개발자를 위한 AI Code Assistant 등</li> <li>[ 품 질 ] 수율 분석 프로그램 코딩</li> </ul>
전문용어기반 통역/콘텐츠 생성	Fine Tuning	<ul style="list-style-type: none"> <li>[마케팅] 광고카피 생성 등</li> <li>[마케팅] 온라인몰 내 상품설명 콘텐츠 생성</li> </ul>
고객서비스	STT/TTS 연계	<ul style="list-style-type: none"> <li>[서비스] 매장 방문객 대상 안내로봇에 Gen AI 적용</li> <li>[서비스] 기 접수된 VoC를 분석하여 고객서비스 응대 지원</li> </ul>

# 어디에 활용할 수 있지? - 사례 1

## [R&D] 특허동향 보고서 작성

### Pain Points

- ✓ 연구 개발에 앞서 관련 특허 동향을 파악하는데 방대한 양의 정보 검색 및 검토가 필요
- ✓ 수작업으로 특허를 검색 시, 누락되는 특허 발생

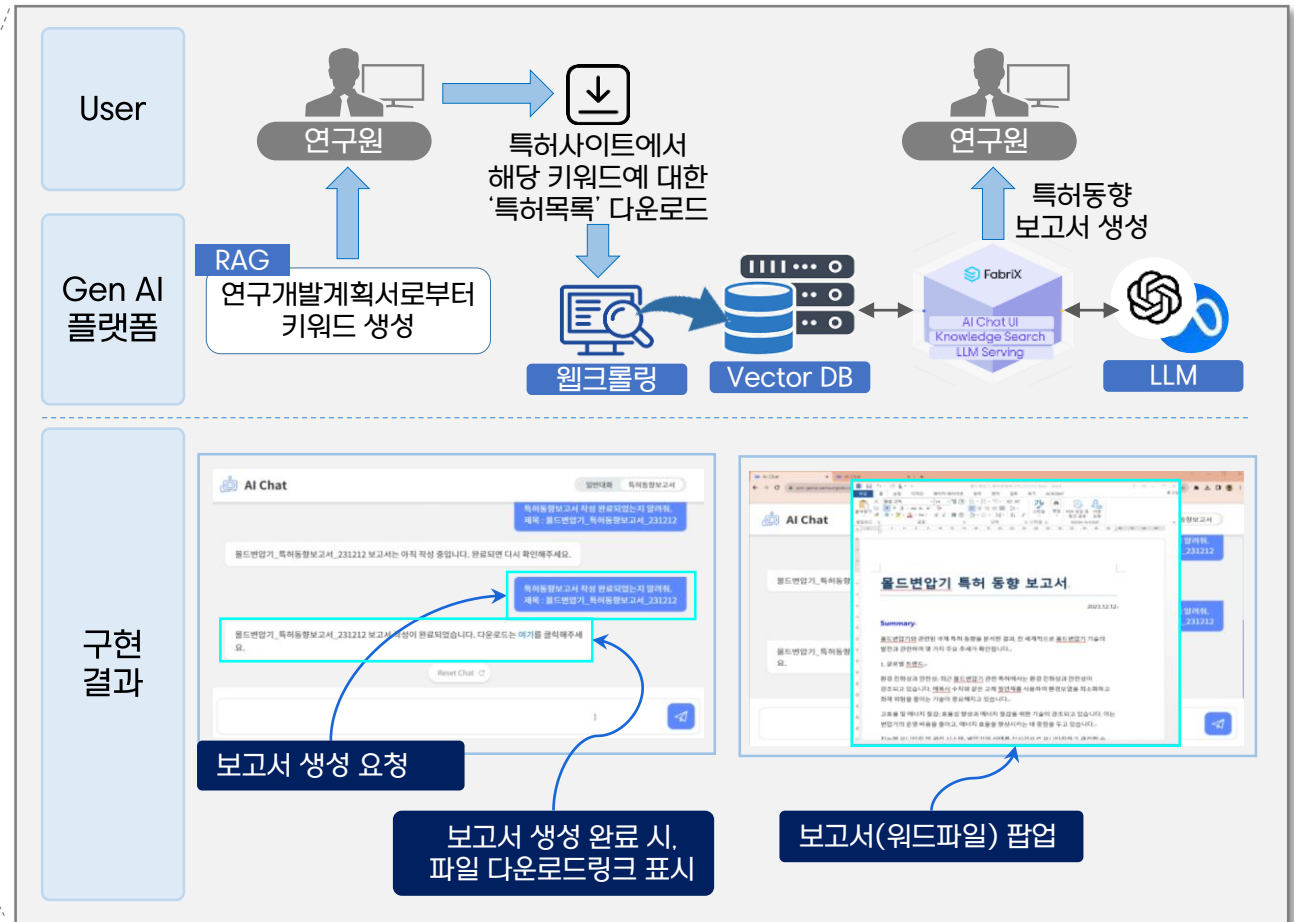
### Solution

보고서 작성을 위한 정보수집/요약/작성 업무 자동화

- ✓ 정보 수집을 위한 키워드 추출 자동화 및 정보수집
- ✓ 분류/요약/번역/분석 수행 후 동향 보고서 작성

### Benefits

- ✓ 기존 보고서 작성 업무의 50% 수준 자동화
- ✓ 방대한 양의 특허 내용을 자동 분석함에 따른 업무효율 향상



# 어디에 활용할 수 있지? - 사례 2

## [서비스] 오프라인 매장에 지능형 응대봇 적용

### Pain Points

- ✓ 붐비는 시간대, 소수의 직원이 모든 내방객 케어 어려움
- ✓ 고객은 직원 역할을 알기 어렵고, 직원별 수준 편차 존재, 비숙련 직원은 응대 시간 길어짐

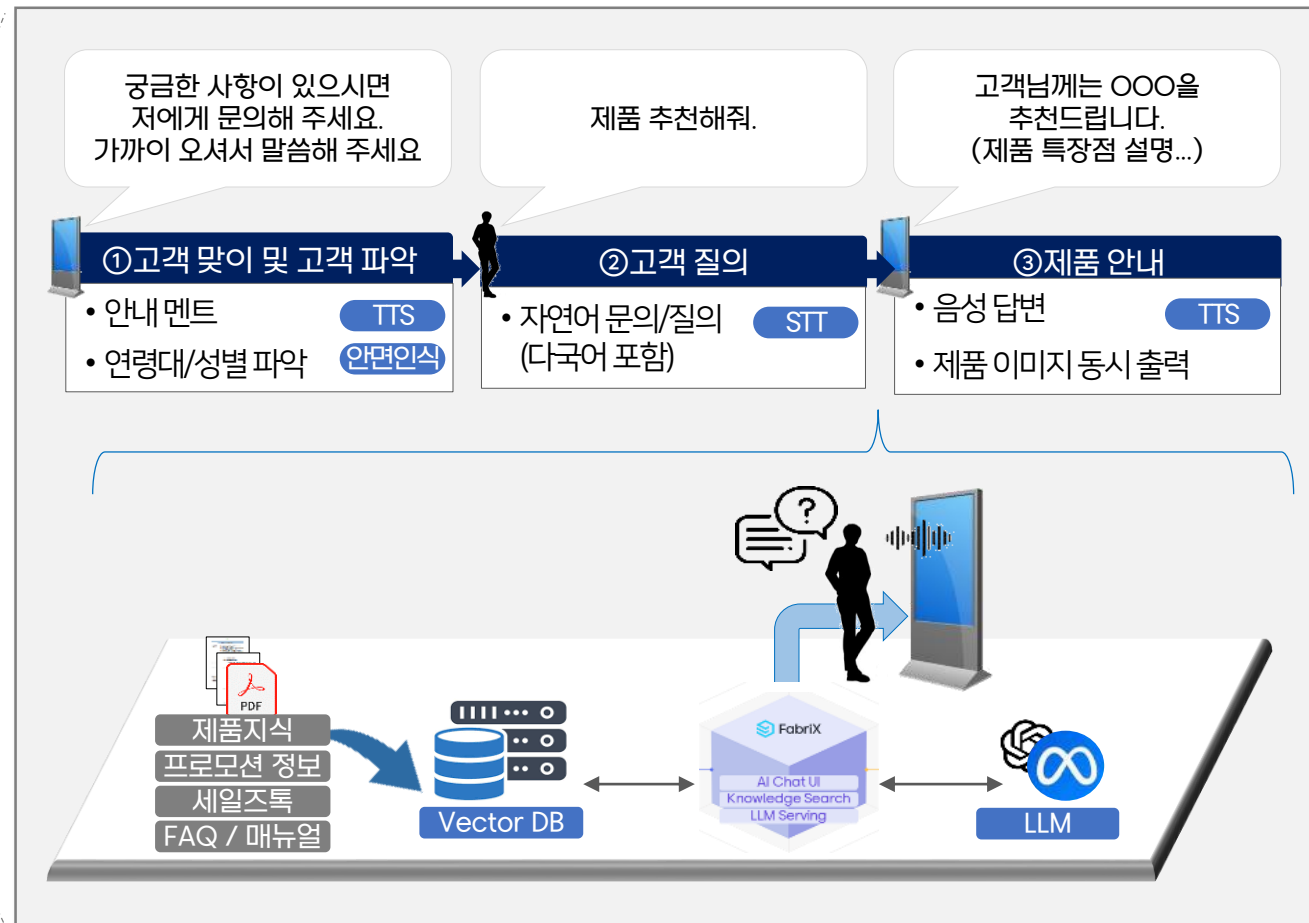
### Solution

#### 오프라인 매장에 고객 응대봇 적용

- ✓ 최초 접근 시 안면인식을 통해 고객 파악(연령/성별)하여 맞춤형 상담 진행
- ✓ 고객 질의에 따라 제품간 기능 및 가격 비교 설명, 일반 질문 등 자연스러운 대화 가능

### Benefits

- ✓ 직원의 부족한 응대로 인한 고객 불만/이탈 최소화
- ✓ 고객별 맞춤형 대응이 가능해져 고객 대기 절감



# 어디에 활용할 수 있지?

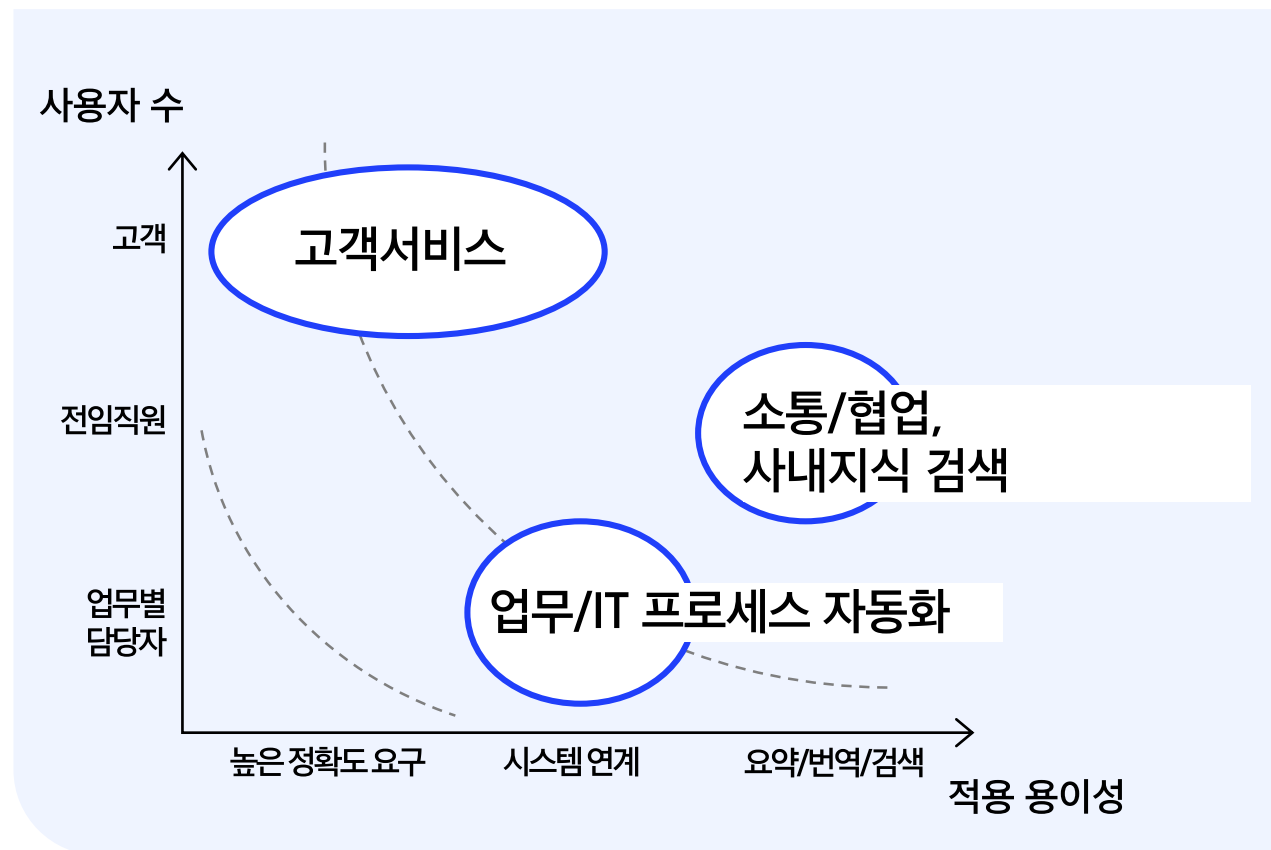
도메인별 비즈니스 Impact 고려하되, 시작은 많은 사용자가 사용할 수 있고 구현 난이도가 높지않은 Use Case를 최우선 적용 필요

## Domain별 Biz. Impact

- **영업마케팅/고객서비스/R&D/SW엔지니어링 높은 영향도**

	Impact total* (억\$)	Marketing & Sales	Customer Operations	R&D	S/W Engineering	SCM	Risk & Legal	Strategy & Finance	IT
Consumer Service	2,400-3,900	High	High	Low	High	Medium	Medium	Low	Low
Banking, Invest	2,000-3,400	High	High	Low	High	Medium	Medium	Low	Low
Manufacturing	1,700-2,900	High	Low	High	Medium	Medium	Medium	Low	Low
Engineering	900-1,500	Low	Low	Medium	Medium	Medium	Medium	Low	Low
Bio, Health care	600-1,100	High	Medium	High	Medium	Medium	Medium	Low	Low
Insurance	500-700	High	High	Low	High	Medium	Medium	Low	Low

## Use Case 선정 우선순위



\* The economic potential of Generative AI (McKinsey, 2023.06)



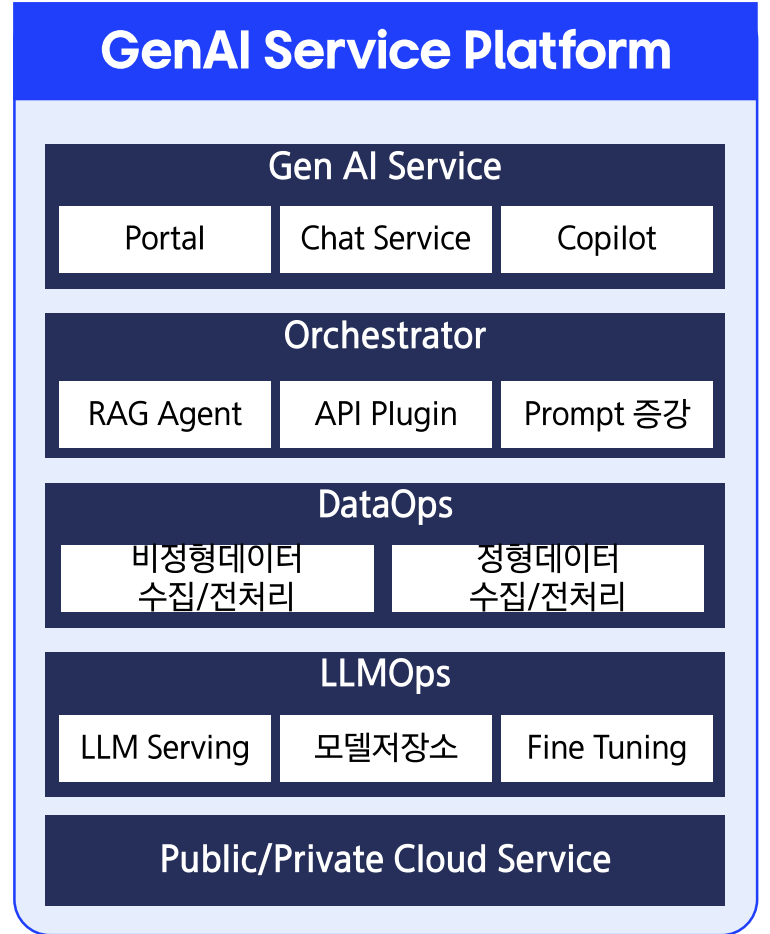
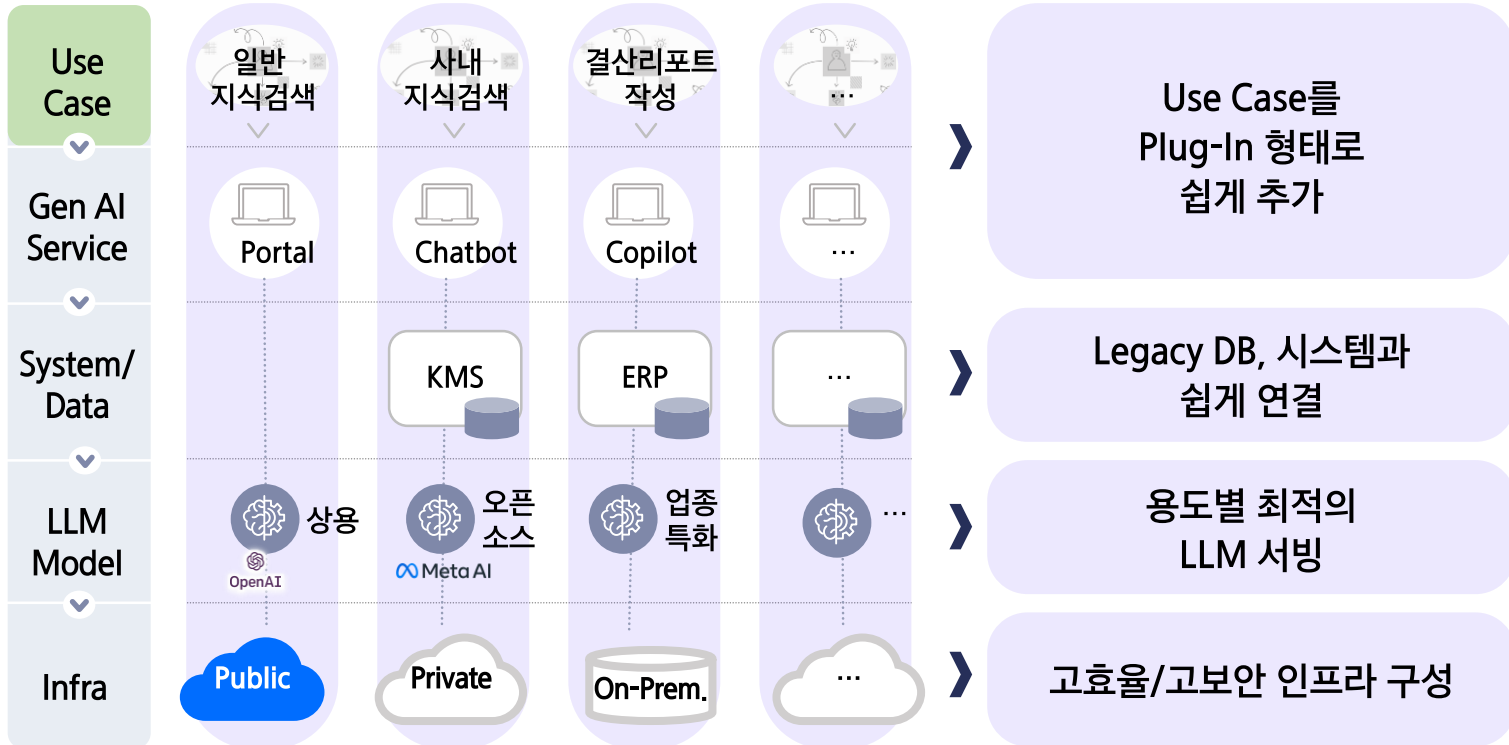
# 어떻게 구현해야 하지?

Use Case별 개별 구축 보다는 다양한 Use Case를 공통 지원할 수 있는 Platform 적용 Approach 통해 효율성과 확장성 확보 가능

## 구현 시 필요 요소

## 고려사항

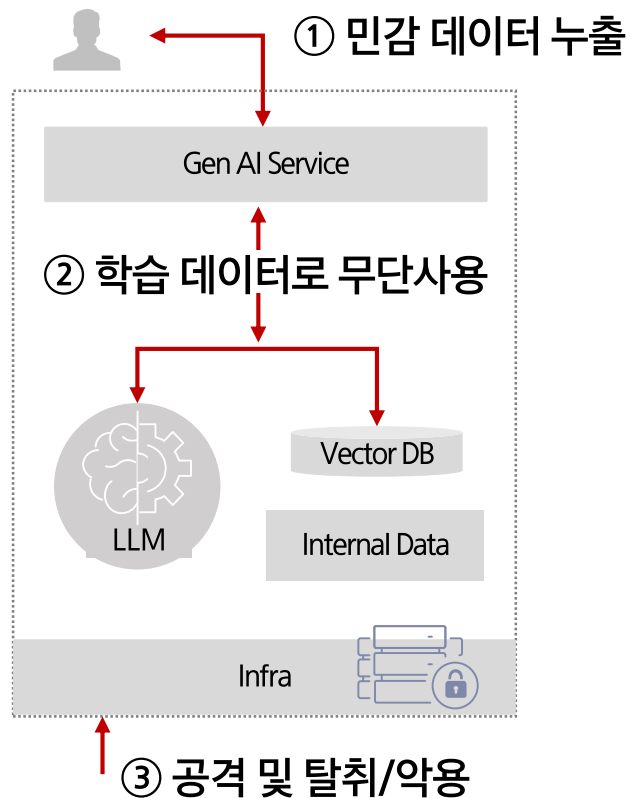
※ Use Case 개별 도입 시 구축 난이도 및 비용증가



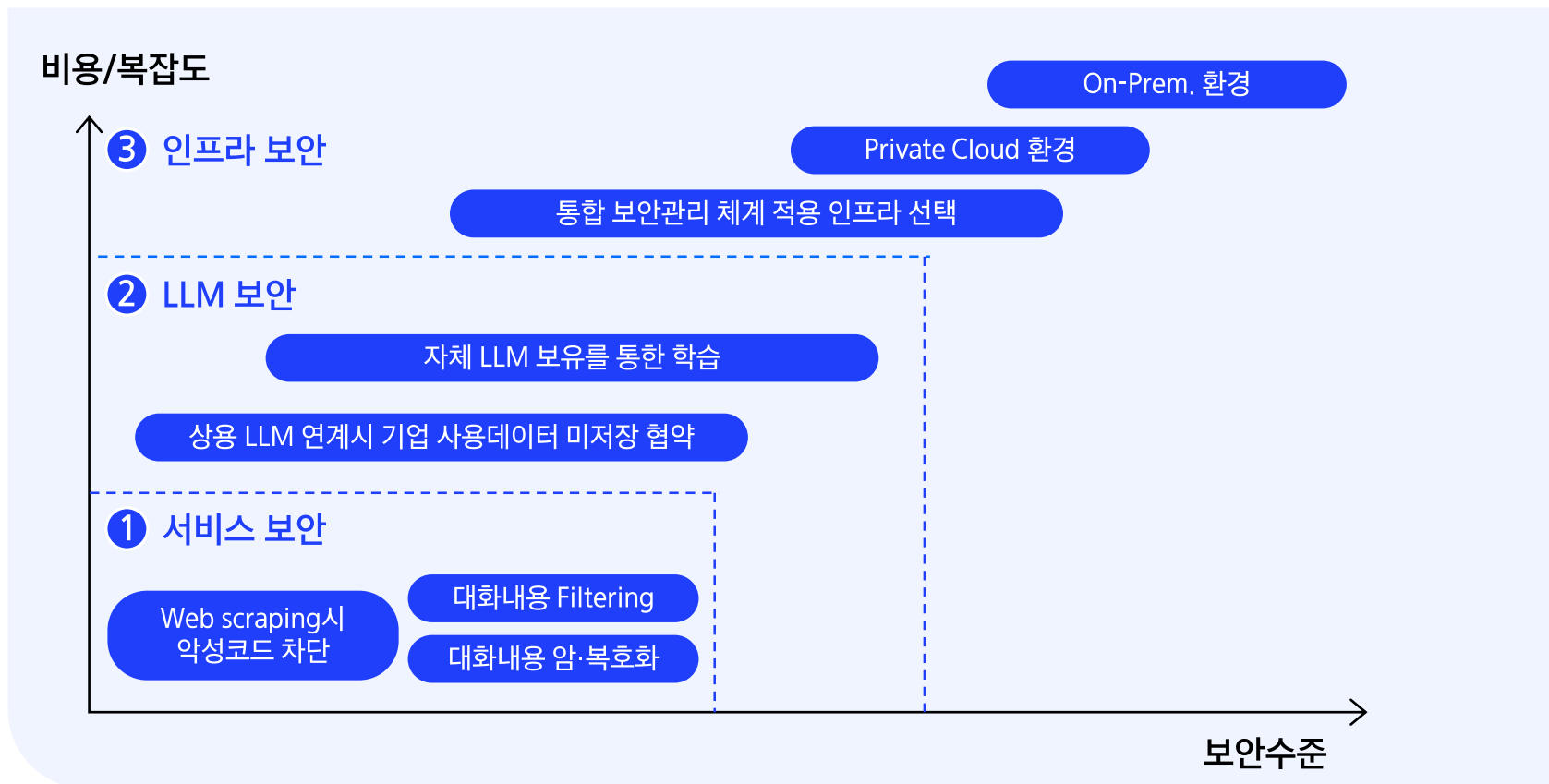
# 기업데이터 유출위험은 없을까?

비용 및 복잡도를 감안하여 Service, LLM, Infra Layer 상 기업이 우려하는 보안위험에 맞는 보안장치를 적절하게 결합하여 사용하면 안전

## Challenge



## Solution



# Why SDS?

## 국내 최초/최고의 Gen AI 솔루션

- Gen AI 솔루션과 확장 가능한 플랫폼, 업무와 Gen AI 결합 가속화
- Real 2023, CES 2024 출품, Live Demo



## LLM 최적화 기술 및 노하우

- Own LLM + 멀티 LLM의 서빙/최적화 경험
  - Samsung LLM
  - Llama2 Core, Mixtral 8x7b, Sola
- Hallucination 저감 노하우 (정확도 56% → 98%<sup>2)</sup>)
  - Multi-Turn, Adaptive Chunking, 임베딩/리랭크 모델, Tool Aware LLM, Prompt Augmentation, Fine Tuning 등

## 국내 No.1 규모 AI특화 인프라<sup>1)</sup>

- H100 및 A100 수만장으로 구성된 대규모 GPU Cluster 구축/최적화/운영 역량
  - 쉽고 빠른 모델 학습/평가 LLMOps 환경 구성
  - GPU As a Service 제공



국내 최초 HPC 전용 데이터센터 (동탄)



클라우드 특화 춘천 데이터 센터

국내 최대규모(현재 20만명, 연내 50만명) 서비스 제공경험 기반 Co-Optimization 역량 (Gen AI 솔루션 + 자체 LLM + GPUaaS)

1) B2B 서비스용 인프라 기준 2) 고객사 HR규정 질의응답 Case 기준

— Thank You —

문의 : 김공환 그룹장 (kh915.kim@samsung.com, 010-9915-7336)