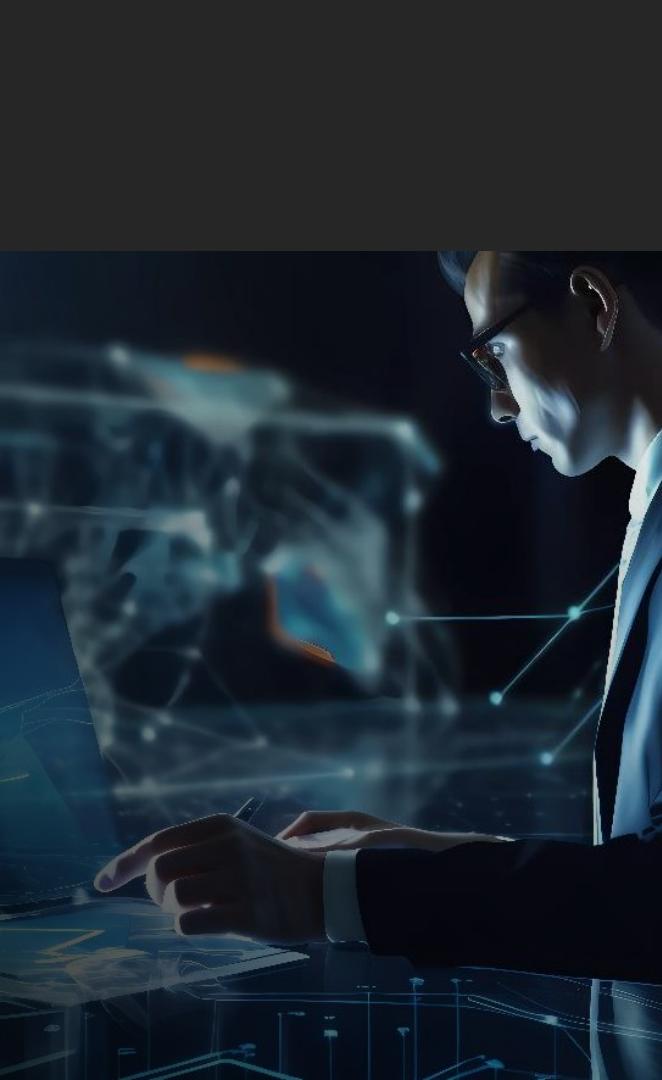


Session 1. AI 구현 E2E (전략에서 구축 및 운영까지)

# 기업의 성공적인 Gen-AI 도입 및 실행을 위한 Co-Innovation 활용전략

PwC컨설팅 조용민 파트너



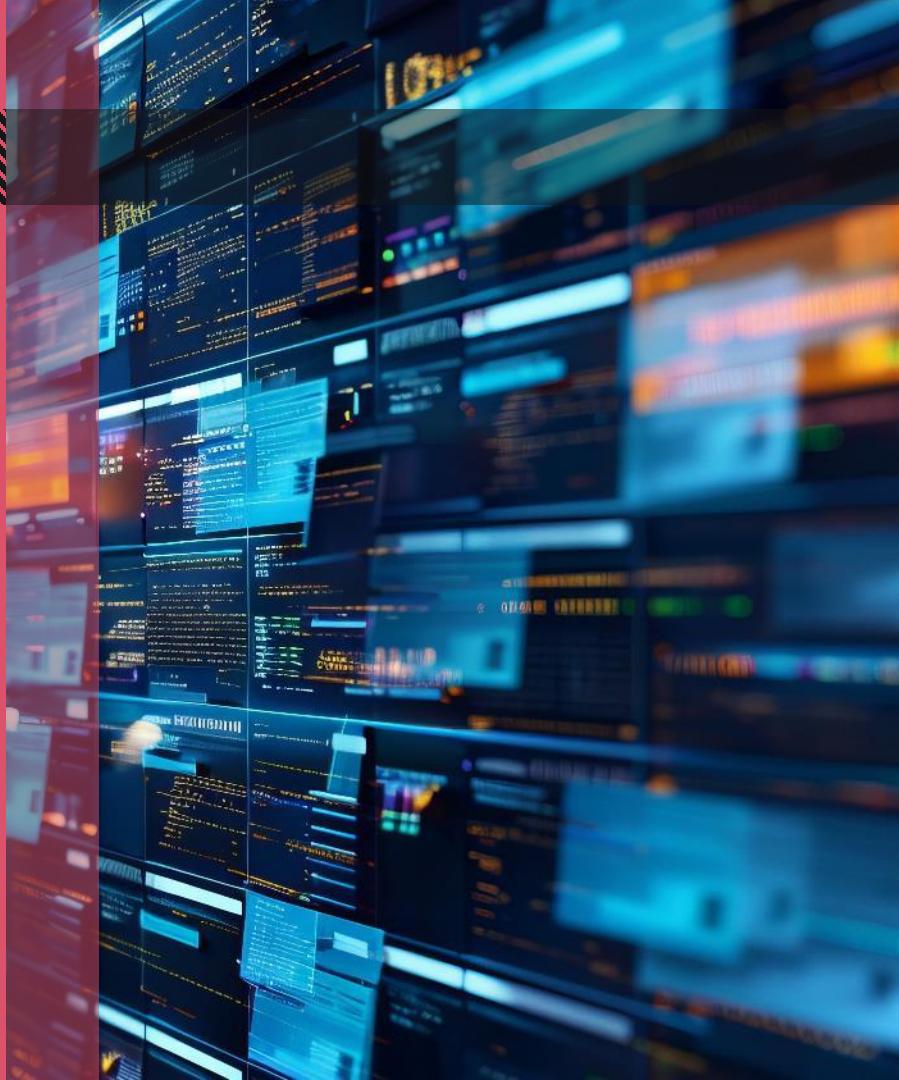


# Contents

<b>1. Generative AI 트렌드</b>	03
<b>2. 기업의 Generative AI 도입 전략</b>	09
<b>3 Key Use-Case</b>	15
<b>4. AI Co-Innovation 모델</b>	22
<b>5. Key Consideration</b>	25

# Generative AI 트렌드

1



# Generative AI Market Trend

국내외 기업의 Generative AI 서비스에 대한 투자는 연평균 59.2% 성장할 것으로 전망

405  
억\$

- 24년에 Generative AI 솔루션과 개발에 405억 달러를 투자할 것으로 예상
- 23년~28년까지 연평균 성장률(CAGR) 59.2% 수준으로 전망

\* Source : IDC FutureScape

13  
Unicorns

- 13개 Generative AI 기업이 Unicorn (기업가치 \$1B+) 기업으로 성장
- Open AI \$29B, Anthropic \$4.4B, Cohere \$2B, Hugging Face \$2B

\* Source : CB Insight

719,659  
LLMs

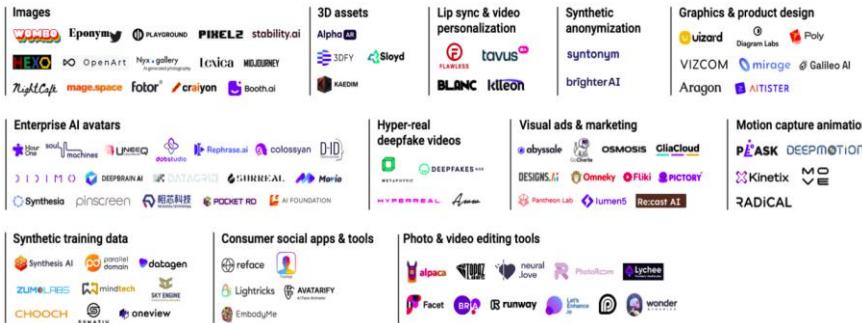
- 24.6월 기준 Hugging Face 등록된 LLM 모델의 수 : 약 70만개
- Open Source LLM 활용/투자 급격히 증가 (Llama 2 다운로드 : 3천만건+)

\* Source : Medium article

# Generative AI 서비스 사업자

Generative AI 응용 솔루션 영역도 Media, Text, Code, Speech/Audio 등으로 급속히 확장

## Media Generation



## Text Generation

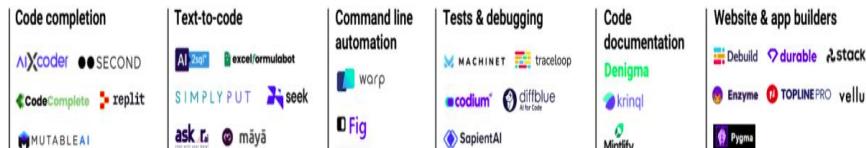


\* Source : CB Insights research : Generative AI Market Map

## Speech/Audio Generation



## Code Generation



## Generative Interface



# Generative AI 서비스 사업자 비중

...특히 Front-end 어플리케이션 영역에서 Text, Image, Audio 생성 서비스 사업자의 증가

→ Middle-end 서비스의 경우 기술 난이도 및 높은 초기 Capex로 진입 장벽이 존재

최근 5년간 투자 유치한 Gen-AI 기업 598개 대상 Research 결과

서비스모델구분		#of 연도별 투자유치 기업			회사 비율(%)
		Pre-2018	2018-2020	2021-2023	
Middle-end	Foundational models	2	2	2	6 (1%)
	MLOps/End-to-end Platforms	11	18	17	46 (8%)
	Data	10	12	7	29 (5%)
Front-end	Code	7	8	13	28 (5%)
	Image	10	27	40	77 (13%)
	Chatbot/Conversational AI	27	16	13	56 (9%)
	Text	31	53	60	144 (24%)
	Audio	21	30	13	64 (11%)
	Semantic Search	-	7	9	16 (3%)
	Video	10	29	16	55 (9%)
	Other Applications <sup>2</sup>	16	29	32	77 (13%)
	총계	145	231	222	598 (100%)

\* Source : PwC Analysis

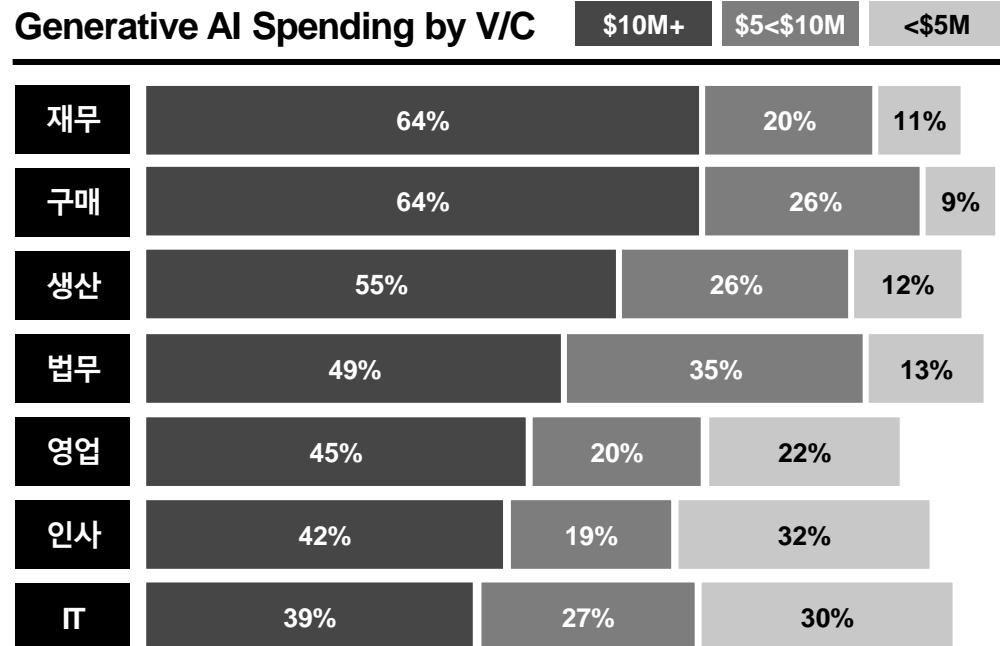
# 글로벌 기업의 Generative AI 도입 및 투자

글로벌 기업은 재무, 구매, 생산, 법무 영업 등 핵심 업무에 Gen-AI 대규모 투자 진행

## Top Generative AI Use-Cases



## Generative AI Spending by V/C



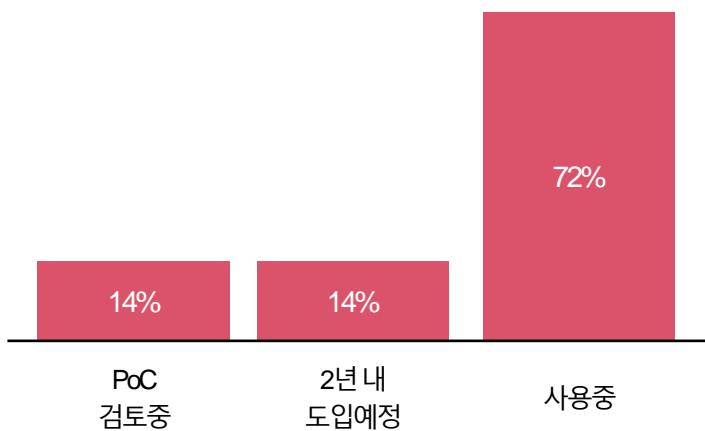
\* Source : Wharton, Navigating Generative AI's Early Years – AI Adoption Report

# 국내 기업의 Generative AI 도입 현황

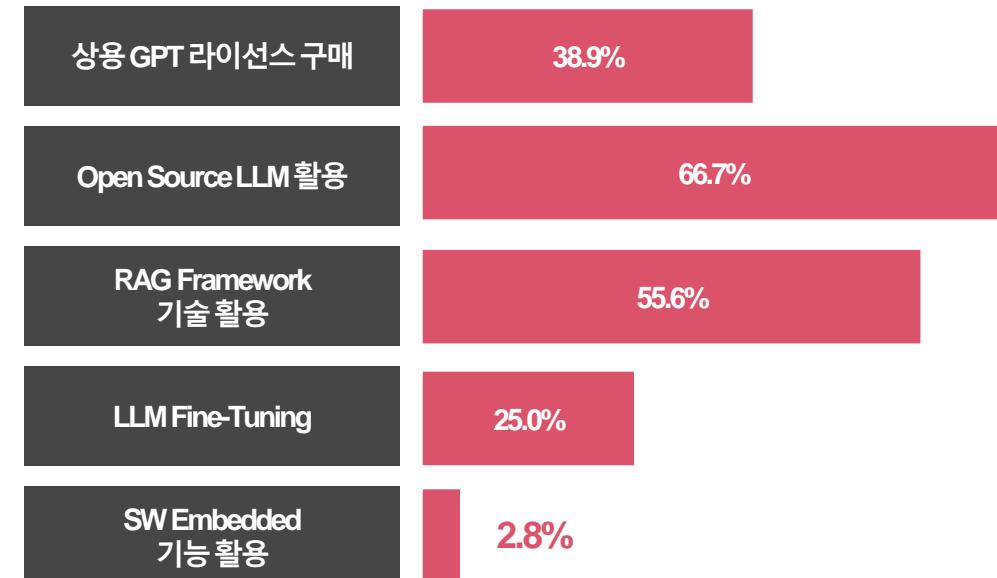
국내 기업이 약 72% Gen-AI를 활용하고 있으며, 특히 Open Source LLM 활용율 높게 나타남

## Gen-AI 활용 현황

“기업형 상용 LLM, Open Source LLM 등 활용 및 도입 활성화 단계로 진입”



## LLM 모델 활용 방식



\* Source : IDC, 국내 50개 기업 대상 Research 결과

# LLM 모델별 활용 방식

기업은 정보 보안, 보유 인프라, 서비스 목적 등에 따라 LLM 모델을 선택적으로 도입하여 활용

## LLM 모델 유형

## 활용 방식

## 고려 사항

상용 GPT 라이선스 구매	<ul style="list-style-type: none"><li>문서 (메일, 회의록, 보고서 초안 등) 작성, 번역</li><li>코드 생성 및 SW 개발 지원 (Co-Pilot) 등</li></ul>	<ul style="list-style-type: none"><li>클라우드 기반 → 민감정보 보안 관리 필요</li><li>Customizing 한계 (Domain 특화 기능 부족)</li></ul>
Open Source LLM 활용	<ul style="list-style-type: none"><li>Domain 특화 솔루션 개발 (Domain 지식 학습)</li><li>기업 내부 문서 요약/검색 등</li></ul>	<ul style="list-style-type: none"><li>On-Premise 형태로 보안 유출 Risk 低</li><li>LLM 실행을 위한 인프라 자원 필요</li></ul>
RAG Framework 기술 활용	<ul style="list-style-type: none"><li>기업 내/외부 문서에 대한 검색 기반 QA 시스템 구축</li></ul>	<ul style="list-style-type: none"><li>전처리 및 Vector DB Embedding 필요</li><li>검색 속도 및 정확도 최적화 필요</li></ul>
LLM Fine-Tuning	<ul style="list-style-type: none"><li>기업 내부 지식에 대한 학습 및 특정 기능 수행 목적 LLM - Domain, Function 특화 LLM</li></ul>	<ul style="list-style-type: none"><li>데이터 지속적 품질 관리 도구/조직 필요</li><li>Fine-Tuning을 위한 GPU 자원 필요</li></ul>
SW Embedded 기능 활용	<ul style="list-style-type: none"><li>상용 SW 포함된 LLM 모델 : SAP Joule, MS Copilot, Salesforce Einstein GPT 등</li></ul>	<ul style="list-style-type: none"><li>Customizing 한계 (Add-On 데이터 이해부족)</li><li>특정 SW 종속될 Risk 존재</li></ul>

\* Source : IDC, 국내 50개 기업 대상 Research 결과

# 기업의 Generative AI 도입 전략

2



# Generative AI 활용 전략

기업은 Generative AI 활용한 신규 비즈니스 모델 개발, 서비스 고도화, 운영 모델 최적화 관점에서의 대응

1

신규 사업 개발  
(New Business)

## ▪ Generative AI 기술 기반의 성장성이 높은 신규 Business 발굴/진출

- Domain 특화 솔루션 : 법률/제약/인사 등 전문 Knowledge 보유 사 중심
- Function 특화 솔루션 : 문서, 코딩, 영상, 번역, 음악생성 등
- 업무지원 솔루션 개발 : LLM Agent, RAG 솔루션, 업무 가이드 Chat-Bot 등

2

제품/서비스 지능화  
(Product/Service Intelligence)

## ▪ Product 지능화 및 Customer Service 혁신

- 제품/상품 지능화 : Product(로봇, 자동차, 전자기기 등)에 Gen-AI 탑재하여 서비스 지능화
- 고객 응대 Service 지능화 : 계약정보 및 관리, A/S 등의 고객 응대 서비스 고도화

3

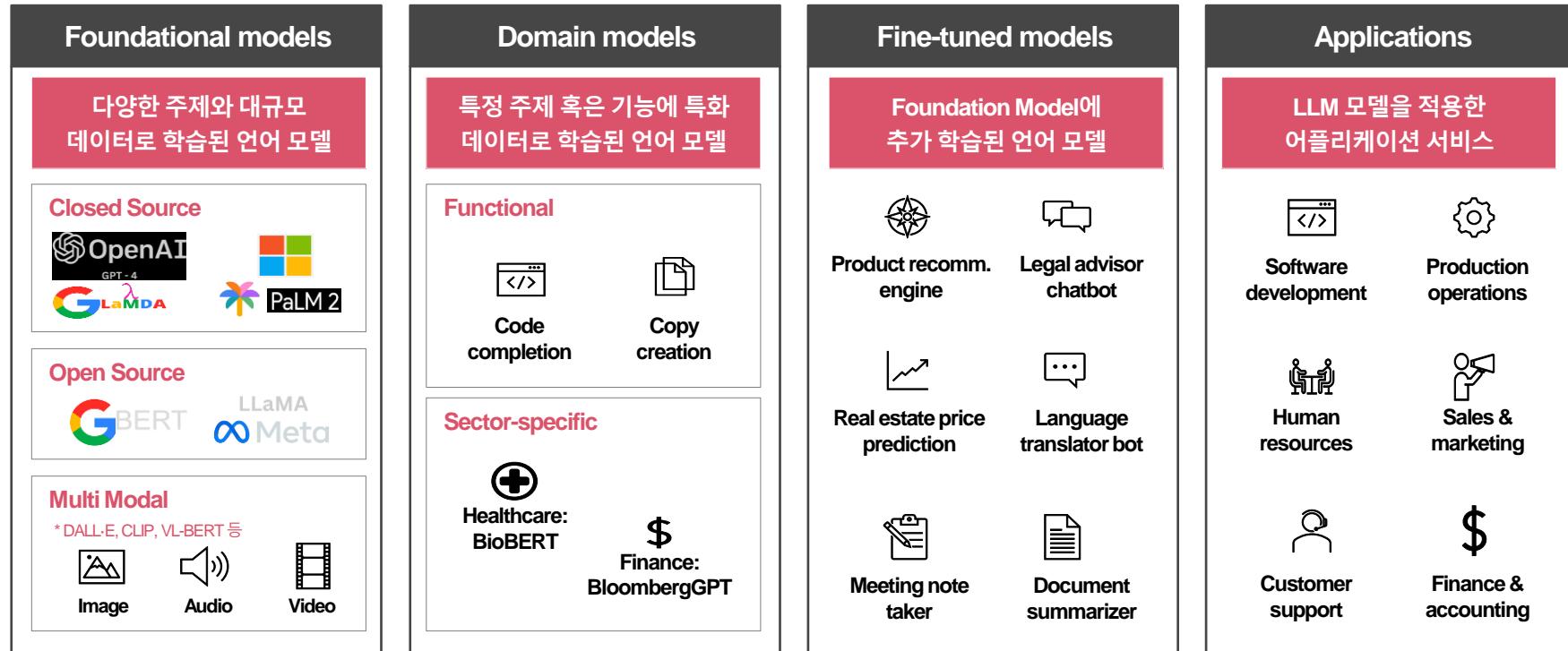
운영 모델 최적화  
(Operational Optimization)

## ▪ 기업 내부 운영 프로세스 효율화/자동화

- 정보 검색 : 기업 내부정보(ERP, MES 등), 일반지식, 전문지식 등의 정보 검색
- R&D : 신규 제품 개발을 위한 Recipe 생성, 생산공정 Simulation, 공정설계 생성 등
- S/W : 기업 시스템 개발/유지보수 목적의 Code Generation
- 기타 : 이메일 전송, 회의록/보고서 등의 문서 작성 서비스

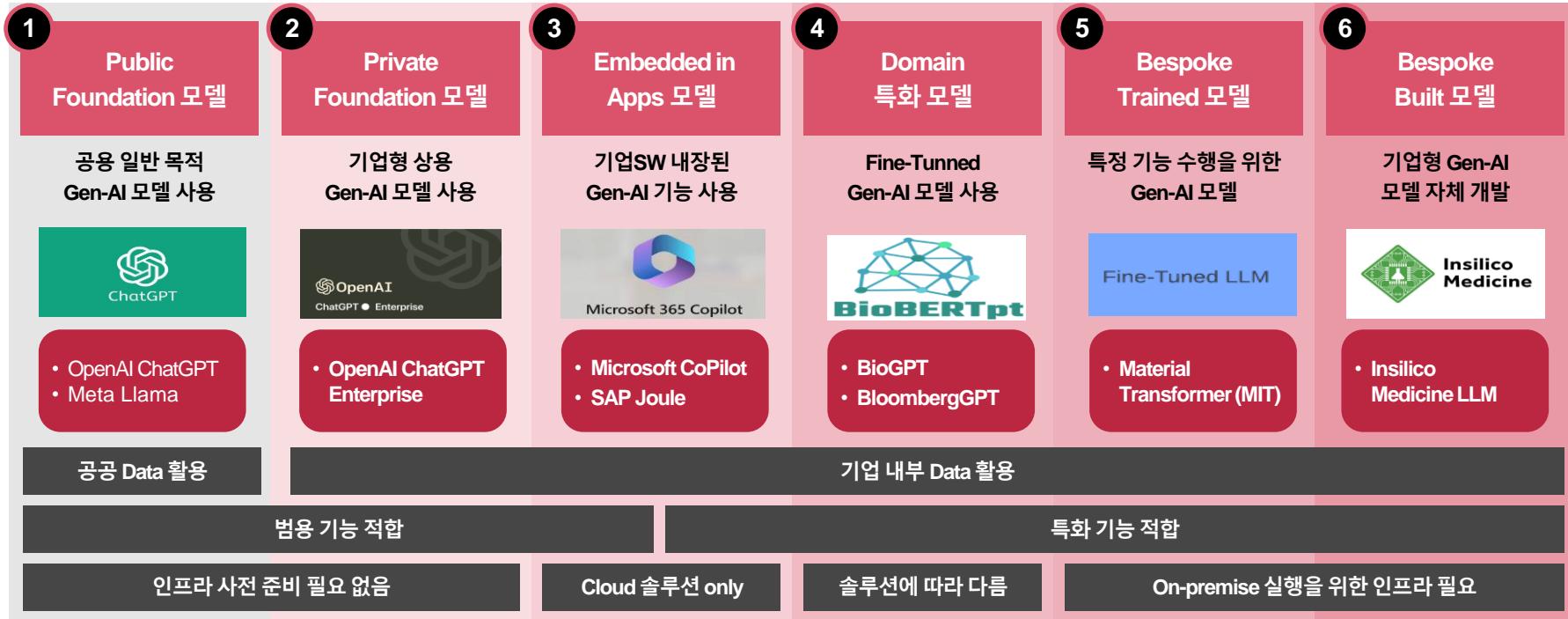
# Generative AI 생태계

Generative AI 기술 도입 및 활용 체계를 구축하기 위해서는 Generative AI 생태계에 대한 이해 필요



# 도입 가능한 Generative AI 모델 유형

기업은 활용 목적 및 인프라 현황에 따라 다양한 Generative AI 모델을 선택적으로 도입 가능



# 기업 단계별 Gen-AI 도입 방향성

국내 다수 기업은 초기 상용 LLM 도입 후 → Open Source sLLM 병행 활용하는 방식 채택

## 활용 단계

### 1단계

상용 LLM 도입  
(License/API)

- 기업형 상용 LLM 라이선스/API 구매
  - OpenAI GPT, Anthropic Claude 등

### 2단계

sLLM 모델 활용  
(Open Source)

- 오픈소스 Small LLM On-premise 활용
  - Llama, Mistral, Yi, Falcon, Solar 등

### 3단계

자체 LLM 개발  
(Built)

- 기업형 Foundation 모델 자체 개발
  - Alibaba Tongyi Qianwen LLM

## 활용 방식 (summary)

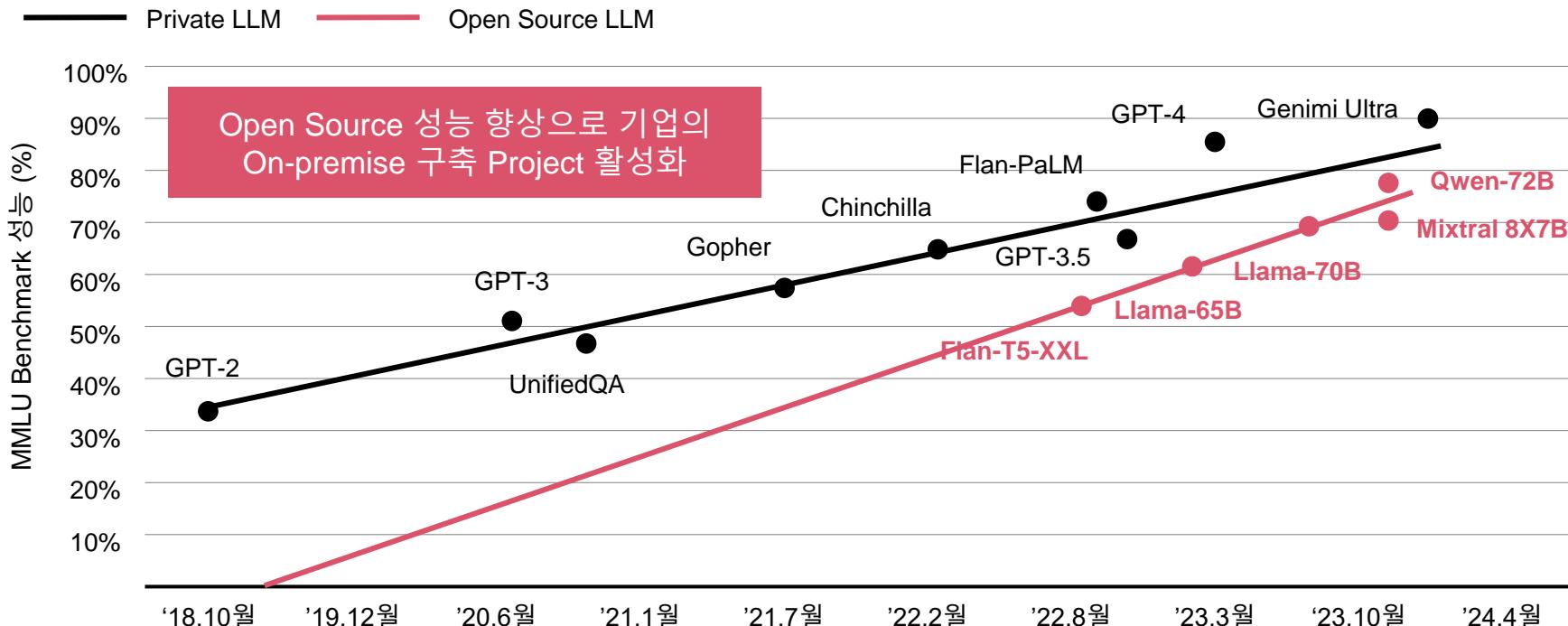
- 지식검색 – 외부정보, 내부문서 (RAG 방식 적용)
- 업무지원 – 요약, 번역, 초안작성 (Email, 보고서 등)

- 보안 Risk 예방 → On-premise LLM 환경 구축
- 기업 내부 정보 검색 및 특정 기능 수행 (Fine-Tuning)
  - 기준정보 (Code) 생성, SoP 문서 작성 등

- 기업 핵심 정보의 자산화 및 검색/활용 목적
  - 반도체, 배터리, 제약 등 산업별 생산/공정 정보 학습

# Open Source sLLM 성능 향상

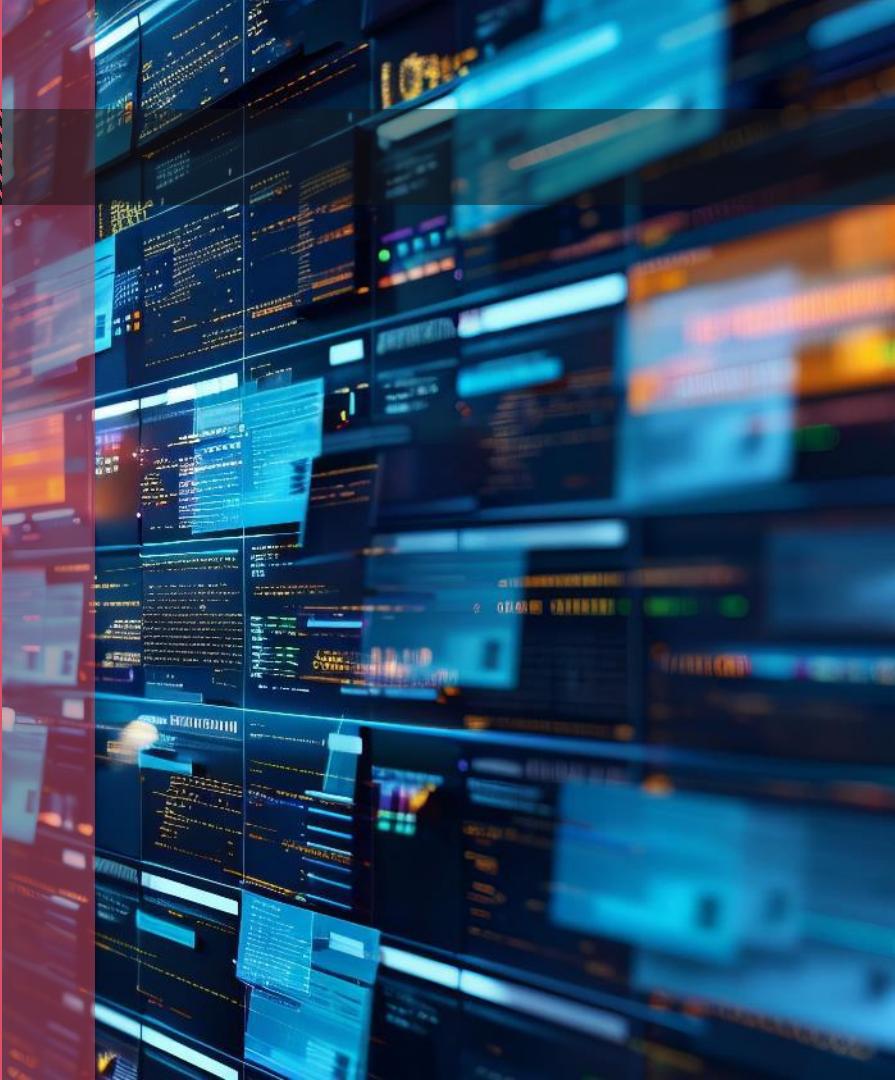
Llama 2/3, Mixtral, Falcon 등의 Open Source LLM의 성능이 상용 LLM과 유사 수준으로 향상



\* Source : sLLM 연구논문, OpenCompass 발표 자료 참고

# Key Use-Case

3



# Generative AI 주요 기능

Generative AI 기술을 활용하여 요약, 검색, 번역, 증강, Q&A, 코드/이미지 생성 등에 적용



Summarization



Deep retrieval



Transformation



Augmentation



Q&A (Dialogue)



Net-new creation

문서, 코드 프로그램  
또는 기타 텍스트의  
요약 생성

문서내에서 특정  
정보를 검색

이미지 스타일 변환,  
텍스트 번역 또는  
텍스트 생성

자동 완성 또는 합성  
데이터 생성과 같이  
기존 콘텐츠를 확장

챗봇, 서비스 봇, 가상  
비서와 같이 질문에  
응답(Q&A)

사용자가 제공한  
프롬프트(지시) 활용  
콘텐츠 생성

## 활용 예시 :

- 정형 데이터 (Structured Data) 세트를 조회하여 재무/손익정보 요약
- 설문 조사 또는 소셜 데이터를 통해 고객 Feedback 분석

- 자연어 질의를 통한 문서 검색 및 Business Insight 탐색
- 경영진의 의사결정을 위한 정보 추출 생성

- 회사 표준 양식에 맞춘 데이터 변환
- 디자인 개인화
- 번역 등

- 결측값을 합성 데이터로 대체
- 텍스트 자동 완성 (이메일, 코드 등)
- 고객 경험의 개인화

- 고객 문의 응답 (지능형 Retail Bot)
- 제품내 대화형 기능 탑재 (Robot, Video Game 등)

- 마케팅을 위한 비디오, 이미지 생성
- 뉴스, 기사, 블로그 제목과 내용 생성
- 기능 명세서를 바탕으로 코드 생성

# Generative AI 적용 영역

Response의 유형과 Business 활용 방식에 따라 지식검색, 거래생성, 분석/예측, 운영/개발 영역에 적용 가능

Information Retrieving 정보 검색	Data Analytics 분석/예측	Automated Operation 거래 생성	System Management 운영/개발
<ul style="list-style-type: none"><li>자연어 기반으로 방대한 Document 내 <b>Fact 정보 발췌, 번역, 답변</b> 등 Answer</li><li>자연어를 기반 주어진 정보를 <b>요약, 분류, 해석</b> 등 Intuitive Document Search 지원</li><li>경영진의 <b>의사결정 지원 및 Risk 모니터링</b>을 위한 내부/외부 정보 추출 및 가공</li></ul>	<ul style="list-style-type: none"><li>자연어 기반의 DB Interaction (SQL, Python 등)을 통한 데이터 추출, 분석, 시각화</li><li>현업 부서에서 개발/운영 중인 <b>분석 모델을 Gen AI 서비스 연계</b>하여 제공</li><li>분석 모델링에 필요한 <b>학습 데이터</b> 보간, 보정, 생성</li></ul>	<ul style="list-style-type: none"><li>자연어를 기반으로 Text, DB 정보를 <b>해석하고, 보고서의 형태로 자동 생성</b></li><li>Transaction 및 Master 데이터 등의 정보 생성</li><li>자연어 기반의 System Command 생성을 통한 <b>Code 실행, RPA Activation, 정보 입력</b></li></ul>	<ul style="list-style-type: none"><li>문의 대응, 시스템 모니터링, 품질 관리(기준정보 등), 운영 보고서 작성 등의 <b>사용자 지원 (Help Desk, Agent)</b></li><li>ABAP Programming, Interface/API 개발, Data Migration 등의 <b>프로그램 개발 지원</b></li></ul>

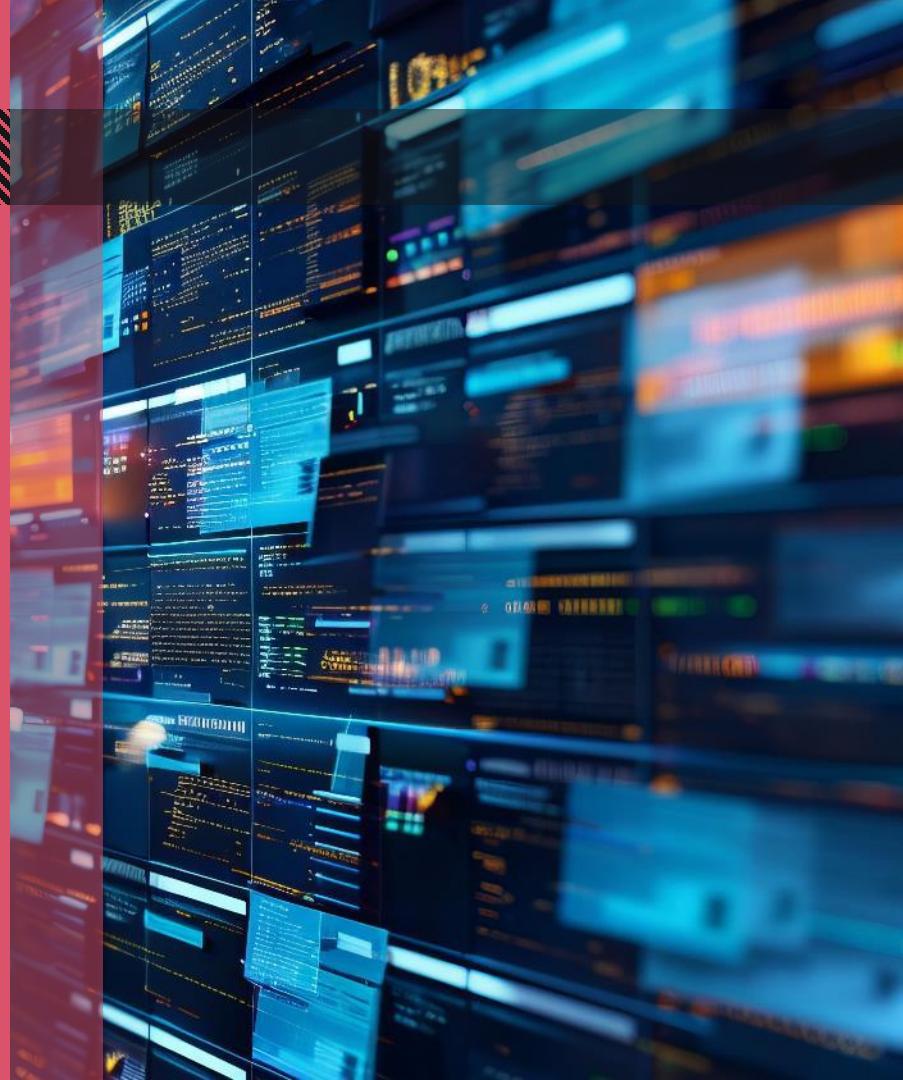
# Generative AI Use-Case

기업의 R&D, 제조(Manufacturing), 물류(SCM), 마케팅(CRM), 경영관리, IT 등의 전 분야에 적용

R&D	Manufacturing/SCM	Marketing/Sales	
新 제품/제약/소재 탐색 및 발굴	제조/공정 가이드 제공	고객 문의 대응 (Q&A) 및 상품 추천	
新 제품 및 공정 설계/시뮬레이션	업무표준, SoP, 지침서 등 문서 자동작성	시장/시황 조사 및 소비자 트렌드 분석	
On-Demand 지식 제공	제품 품질 검사 문서 자동 작성	마케팅 콘텐츠 생성/번역 (캠페인 문구 등)	
R&D 문석 작성 (연구 제안서, 실험 계획서 등)	엔지니어 교육 자료 생성 (Video, AR/VR)	상품/서비스 추천 서비스	
Technology/IT			
UX 디자인 및 코드생성/검증	DevOps 프로세스 최적화	디지털 프로젝트 개발 지원	사용자 Self Service (Q&A)
G&A (General & Administrative)			
BS/PL 등 Report 검색/작성	기업 Risk 션싱 및 분석	계약서 검토 및 자동 작성	업무 지식 제공(Q&A)

# AI Co-Innovation

4



# Generative AI 실행 가속화를 위한 Co-Innovation Lab

기업의 AI Agility 확보를 위해 전문가 그룹과 Co-Innovation Lab을 통한 혁신 가속화 가능

## Co-Innovation Lab 역할

## 수행 내용

기술 센싱 및 전략 수립  
(Research & Strategy)

- 글로벌 네트워크, VC 및 내부 R&D 기반 주요 기술 트렌드 및 마켓 센싱 활동
- 글로벌 선도사 및 주요 경쟁사 신기술 적용 및 혁신 사례 B/M

과제 발굴 및 구체화  
(Discovery)

- 글로벌 혁신 사례 및 기술 동향 기반 당사 적용 AI 과제 Use-Case 구성
- AI 모델링 Concept 및 아키텍처 설계

기술 검증 및 평가  
(PoC & Prototype)

- AI 과제 업무 요건 정의, 알고리즘/모델링 효과 입증 및 사용자 효과성 검증
- 초기 제품 Prototype 설계/제작 및 기능 검증

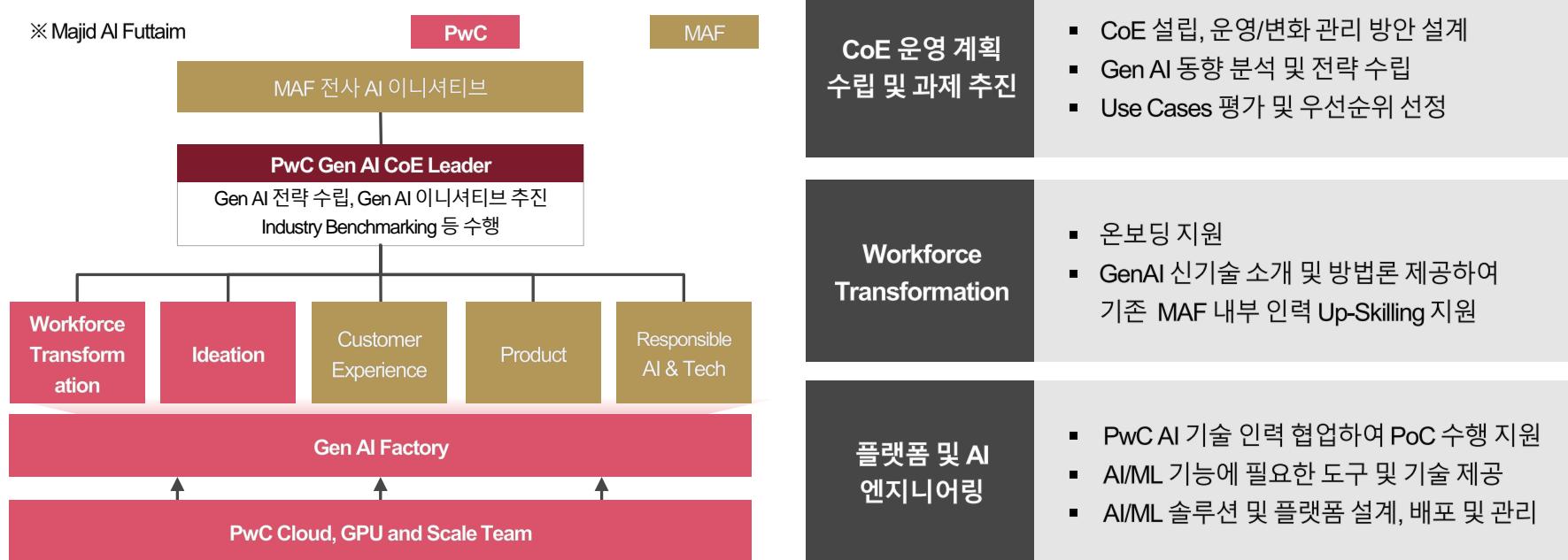
기술 개발 및 적용  
(Implement. & Operate)

- 과제 적용을 위한 Value Chain별 업무 영향도 평가 및 프로세스 재설계 수행
- 사용자 및 운영자를 위한 UI/UX 개발 및 모델 서빙 인프라 기능 설계 및 개발

# Co-Innovation Lab 운영 사례

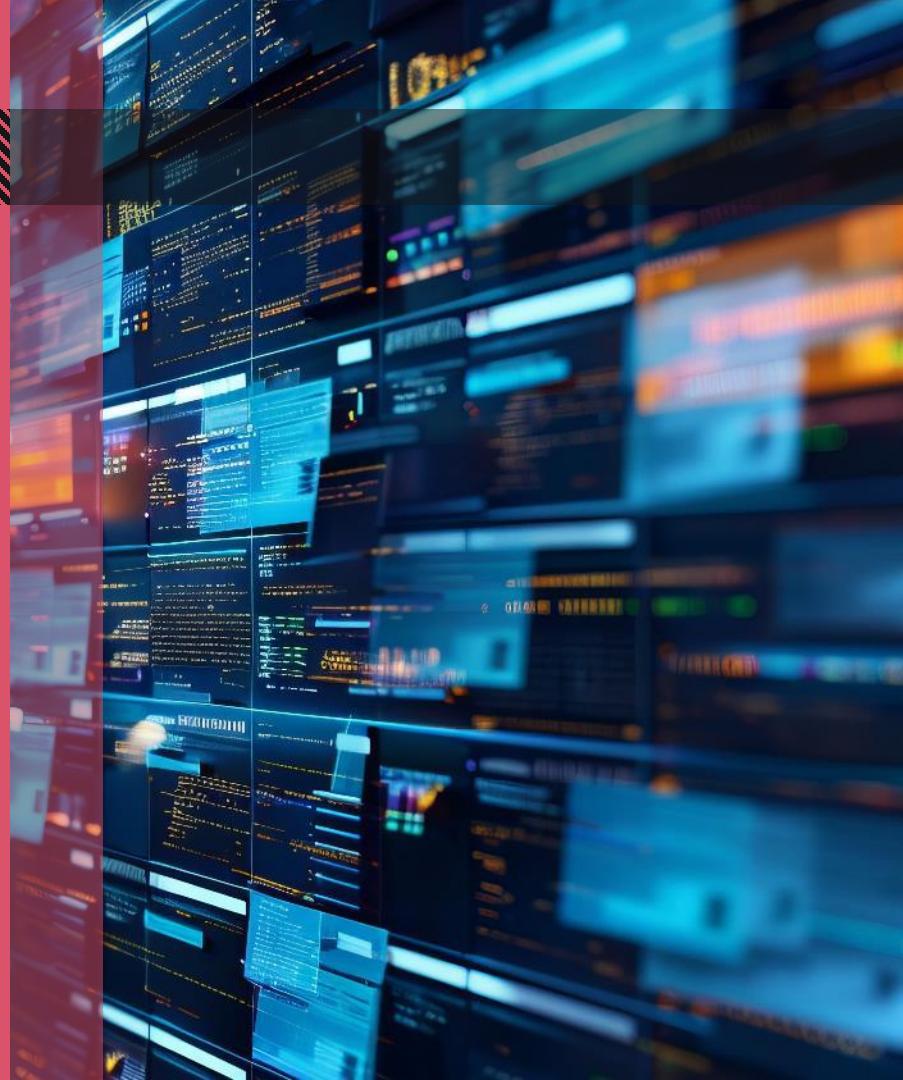
MAF 그룹은 PwC와 Global Gen AI CoE를 통한 Gen AI 전 영역의 혁신 추진

## MAF Group AI Co-Innovation 운영 모델



# Generative AI 도입을 위한 Key Consideration

5



# Generative AI 도입을 위해 준비 사항

기업은 Generative AI를 성공적으로 도입·활용하기 위해 5 Focus Area에 대한 준비가 필요



# 1) 개발 옵션

국내외 다수의 기업은 하이브리드 방식 (Gen-AI 솔루션 도입+자체 개발)으로 개발 진행

	솔루션 도입	자체 개발	하이브리드 방식
방식	GPT-series, Google Cloud AI, Microsoft Azure AI 등 클라우드 기반 솔루션 도입	On-Premise 형태로 기업 내부에 AI 시스템을 개발하고 구축	외부 플랫폼을 구매하고 내부 개발팀을 통해 맞춤형 모듈 추가 진행
주요 목적	빠르고 효율적인 Gen.AI 솔루션	완벽하게 맞춤화된 Gen.AI 서비스 개발	외부 솔루션의 편리함 + 내부 개발의 유연성
장점	<ul style="list-style-type: none"><li>빠르고 간편한 구현</li><li>다양한 기능과 모듈 제공</li></ul>	<ul style="list-style-type: none"><li>맞춤형 솔루션 개발 가능</li><li>직접적인 품질 통제, 보안 통제</li></ul>	<ul style="list-style-type: none"><li>맞춤화와 신속한 구현의 조화</li><li>리소스 활용 최적화</li></ul>
단점	<ul style="list-style-type: none"><li>제한된 맞춤화로 특정요구 충족 한계</li><li>솔루션 범위 내로 유연성 제한</li></ul>	<ul style="list-style-type: none"><li>초기 투자비용과 시간 소요</li><li>기술 및 전문 지식 요구</li></ul>	<ul style="list-style-type: none"><li>관리 및 통합 어려움</li></ul>
사례	<p>[생산라인 품질관리] Microsoft Azure AI 도입하여 이미지 분석 및 머신 러닝 알고리즘 구현하여 제품 불량률 자동으로 감지</p>	<p>[고객 서비스 개선] 고객 데이터 분석으로 개인화 추천 시스템을 구현하고, 자연어처리 기술로 고객 실시간 응답 챗봇 개발</p>	<p>[실시간 다국어 채팅 개발] 내부 AI팀과 Google Cloud AI 플랫폼(NLP 및 기계 번역 기능)을 결합하여 스마트 채팅 어플리케이션 개발</p>

## 2) 아키텍처 (Architecture)

Generative AI 아키텍처 설계 시에는 호환성, 유연성, 확장성, 보완성을 고려한 설계가 필요



## 2) 아키텍처(Architecture) : 관리도구

LLM 업데이트 및 관리 등 LLM 운영을 위해 데이터 관리, 모델 관리, 상태 관리 도구가 필요

데이터 연계 및 데이터 관리			모델 관리 및 모니터링		서비스 상태 관리 및 모니터링	
문서 관리	<ul style="list-style-type: none"><li>문서 Raw Data 연계 및 관리 지원</li><li>데이터 버전 관리 및 신규 문서 관리 지원</li></ul>	LLM 관리	<ul style="list-style-type: none"><li>상용/오픈소스 <u>베이스 모델 연동</u></li><li>모델 산출물 Asset 등록 및 관리</li></ul>	운영 관리	<ul style="list-style-type: none"><li>모델 버전, 상태, 활용 비용, 배포 상태, 데이터 접근 등 관리 지원</li><li><b>무중단 배포 지원</b></li></ul>	
데이터 전처리	<ul style="list-style-type: none"><li>문서 및 자료 형태 별 <u>LLM이 인식 가능한 형태로 처리</u></li></ul>	성능 모니터링	<ul style="list-style-type: none"><li>평가할 <u>모델 선택 및 평가 방식</u> Click 기반 선택</li><li>다중 모델에 역량 비교 및 대시보드</li></ul>	서비스 상태 관리	<ul style="list-style-type: none"><li>사용자별 <u>CPU, 메모리 사용량</u> 모니터링 제공</li><li>서비스 Endpoint 상태 관리</li></ul>	
Vector DB 관리	<ul style="list-style-type: none"><li>상용/오픈소스 <u>Embedding 모델 연동</u></li><li>Chunking 사이즈 및 오버랩 설정</li></ul>	RAG 구성 지원	<ul style="list-style-type: none"><li>모델, 데이터 선택 등 Click 기반 RAG 모듈 설정</li><li><u>RAG 프롬프트 및 결과값 테스트</u></li></ul>	사용자 관리	<ul style="list-style-type: none"><li>사용자 조직, 권한, 민감 정보 등 서비스 관리 도구 제공</li></ul>	

### 3) 거버넌스 (Governance)

AI기술 도입·활용 시 준수해야 할 국내·외 가이드, 지침, 규정, 법률이 구체화되는 추세

#### 국내외 AI 관련 입법/감독기관의 가이드 및 규제



#### 1. 현행 법안 및 지침 총괄

- 데이터 3법 및 금융소비자법에 대한 지침 충족 필요
- 민감 AI 서비스로 인한 민형사상 책임 요소 도출

#### 2. 국제적 Trend에 따른 잠재적 입법 사항에 대한 리스크

- 국내 적용 가능성이 있는 규제사항 도출, 이에 대한 선제적 대응 필요

#### 3. 책무구조도 및 PSMOR 관련 검토

- 감독기관에서 제시하고 있는 AI 서비스와 관련된 책임 범위와 관리 의무조치 반영 필요
- ICT 리스크 관리 및 AI BCP 체계 필요

# Key Takeaways

## 활용 목적에 따른 Hybrid 모델 Approach

- 데이터 보안성, 중요도 등을 고려하여 상용 LLM 및 Open Source sLLM 모델 도입을 병행하는 방안 고려 → 초기 Capex, 아키텍처 유연성 확보

## LLM 재학습, Risk 관리 가능한 관리도구 구축

- LLM 버전관리, sLLM 재학습, Prompt 관리 등 Gen-AI 특화 관리도구 구축
- Hallucination Risk, Prompt 탈취, 학습 데이터 오염 예방을 위한 관리 도구 구축

## 지속적 LLM 모델 운영을 위한 조직/역량 확보

- LLM 관리도구 운영 및 Fine-Tuning, Prompt 최적화를 위한 전문 운영 인력 확보

## 외부 전문가 그룹 협업 모델 (Co-Innovation Lab) 활용

- 新기술 대응 전략 수립, 기술 검증 (PoC), 모델 개발 등 신속한 수행을 위한 외부 전문가 그룹과의 협업 모델 적극 활용 (Gen AI Co-Innovation)

# Thank you

[www.pwcconsulting.co.kr](http://www.pwcconsulting.co.kr)

© 2024 프라이스워터하우스쿠퍼스컨설팅. All rights reserved.

