

# 금융 혁신을 위한 에이전틱 AI

- 신뢰와 자율성의 전략적 균형

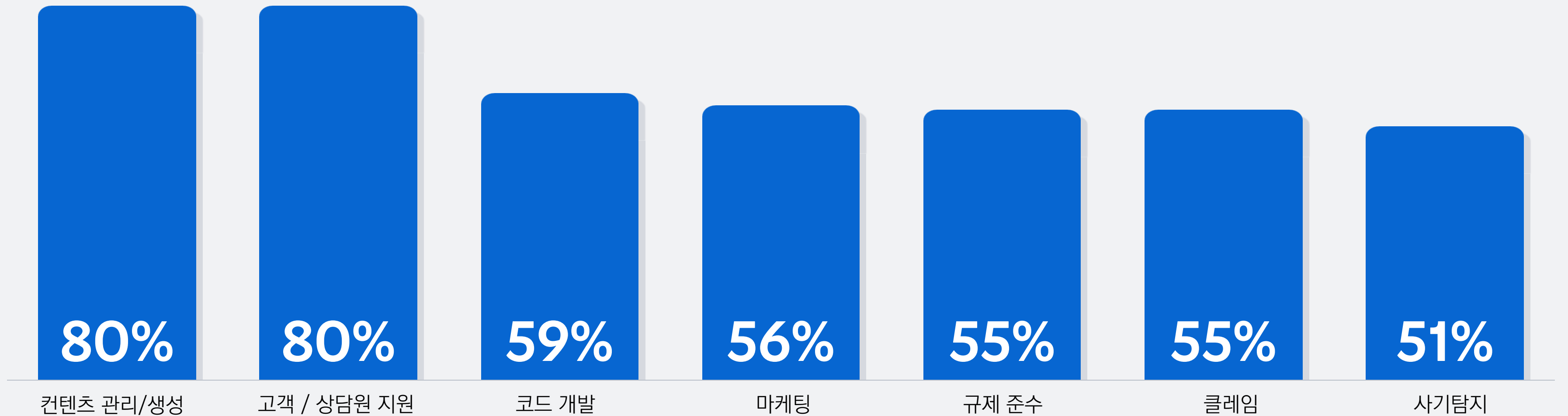
김근태 상무  
SAS Korea



# 확산되고 있는 생성형 AI 업무 적용

## 금융 서비스 분야

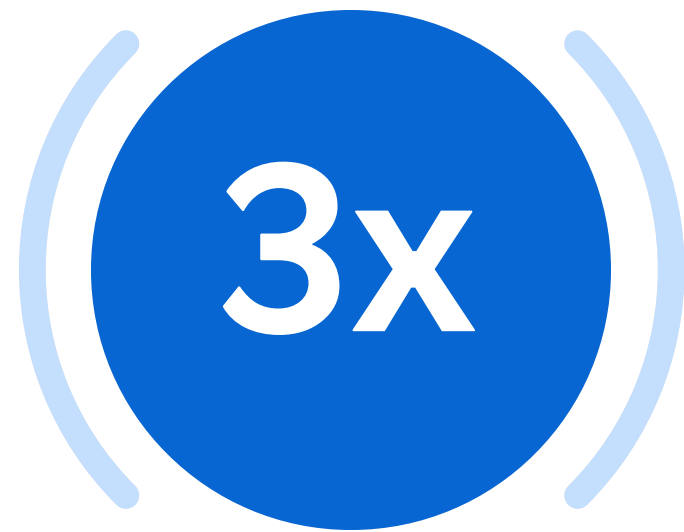
금융 조직이 적용을 검토하는 생성형 AI 솔루션의 주요 유형 (향후 12~24 개월):



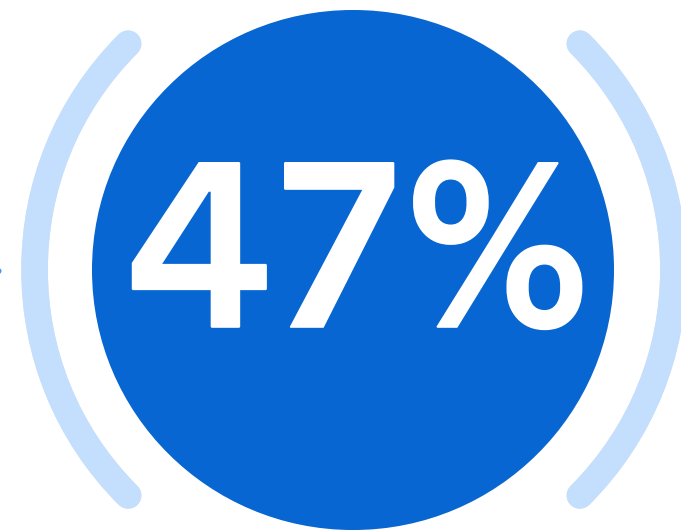
\* Source : Celent Gen AI: Lens on Use Cases – Retail Banking

# 생성형 AI 업무 적용 현황

글로벌 1600개 이상 기업, 조직에 대한 AI 설문조사 결과 (IDC, SAS)



의사 결정권자는  
AI 이니셔티브가  
최소 3배의 ROI를  
제공할 것으로 기대



생성형 AI를 개념  
단계에서 실제  
적용으로 전환하는데  
어려움이 있음



생성형 AI에 대한  
포괄적 신뢰성 및  
거버넌스 부족

# The GenAI Divide

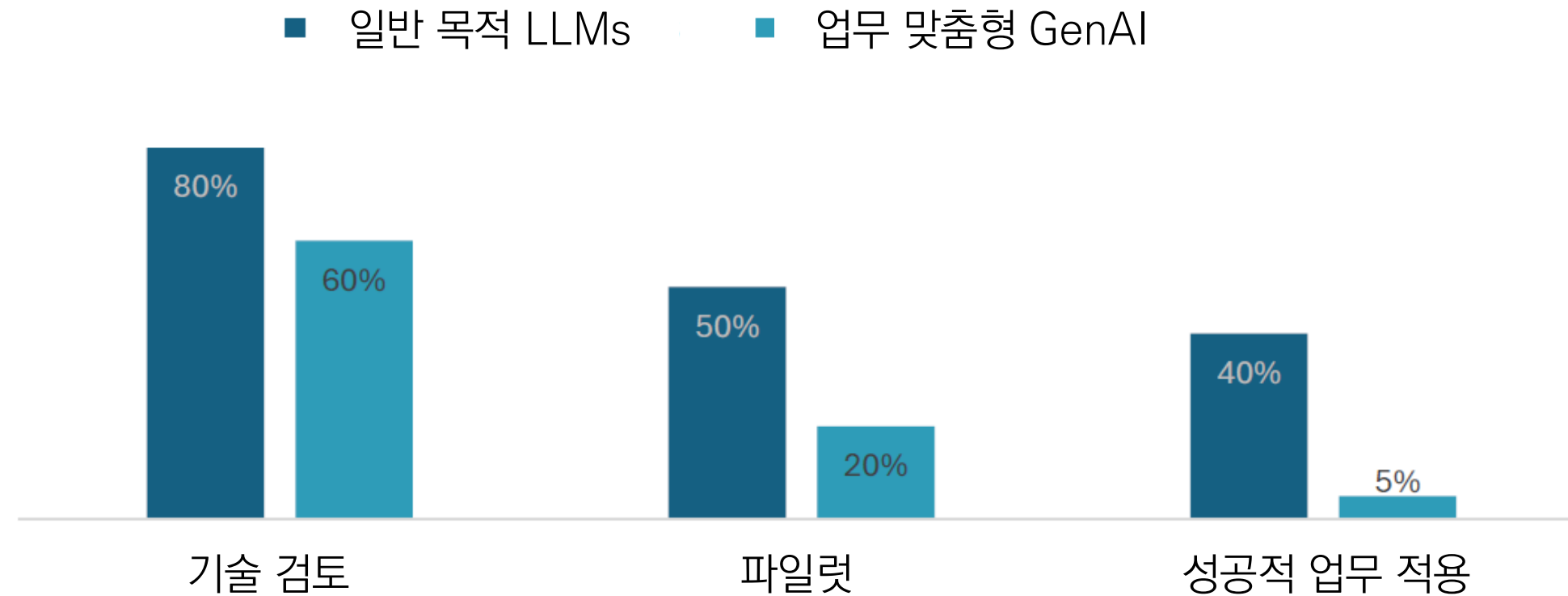
## STATE OF AI IN BUSINESS 2025 (MIT)

“ 지난 3년간 전 세계 기업들이 GenAI 프로젝트에 300억 ~ 400억 달러(약 43~57조원)를 투자했지만, 이 중 95%가 측정 가능한 수익을 얻지 못함 ”



Source: MIT NANDA – The GenAI Divide, STATE OF AI IN BUSINESS 2025

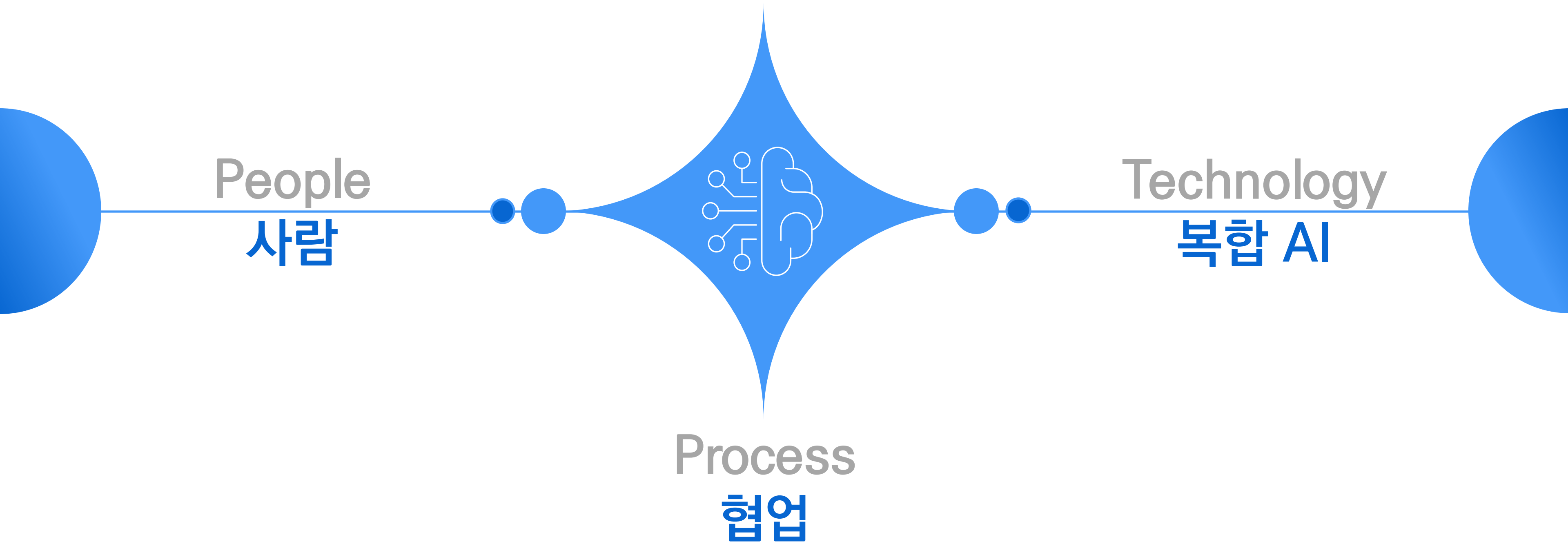
“ 일반 목적 LLMs”와 달리 “업무 특화 GenAI”는 파일럿 단계에서 성공적 운영으로의 전환 비율이 급격히 감소하는 GenAI 격차가 나타남



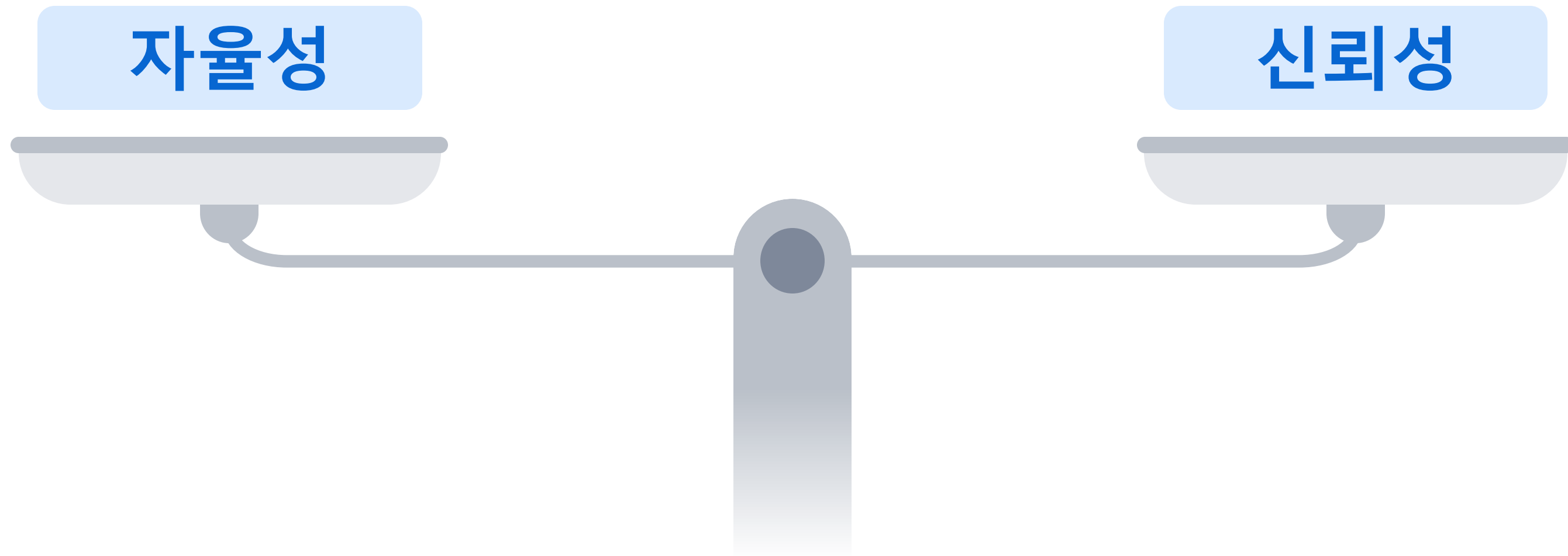
- ChatGPT와 같은 범용 GenAI (General-Purpose LLMs)는 80% 기업이 검토 후, 40%가 정식 적용
- 다만, 주요 Use Case는 개인 업무 효율성 향상의 목적(요약, 검색, 코딩)이며, 회사 전체의 수익 개선 기여는 미미
- 업무 맞춤형 GenAI(Embedded or Task-Specific GenAI)의 경우, 60% 기업이 검토 후, 5%만이 정식 운영

GenAI를 개인 작업에 활용했던 사용자들은 기업 시스템 내 사용 목적에서는 GenAI의 신뢰성 부족을 지적함. 이 패러독스는 사용자 관점에서의 GenAI 격차를 시사함.

# 에이전틱 AI의 신뢰성 향상



# AI 적용을 고려하는 조직의 고민



## AI 딜레마

인공지능이 더욱 강력해지고 자율성을 가지게 됨에 따라,  
AI의 신뢰성을 확보하는 것이 가장 중요한 과제 중 하나

# 인간의 개입 없이 AI가 작업을 수행하고 결정을 내리는 능력

사람이 프로세스에 참여  
(Human *in* the loop)

사람이 감독  
(Human on the loop)

사람이 개입하지 않음  
(Human out of the loop)

AI :  
수동적 지원

AI :  
감독하의  
자율성

AI :  
완전 자율성

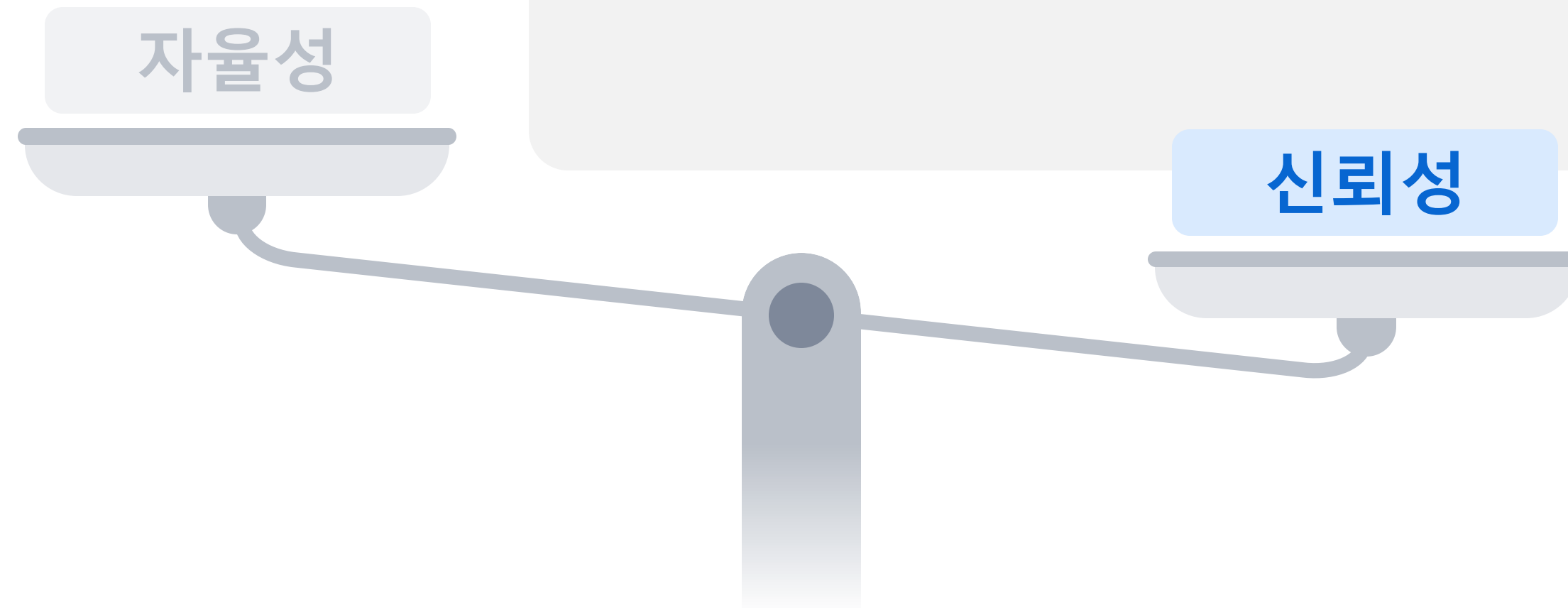
자율성

신뢰성

## AI 딜레마

인공지능의 신뢰성 향상을 통해,  
시스템이 합법적이고 윤리적이며 견고함을 보장

- 공정성** 유해한 편견을 지속하거나 만들지 않음
- 투명성** 이해 가능한 의사 결정 과정
- 신뢰성** 다양한 조건에서 일관되게 작동
- 책임성** 인공지능의 결과에 대한 명확한 책임



# AI 딜레마

# 자율성과 신뢰성 측면의 AI 기술에 대한 고찰

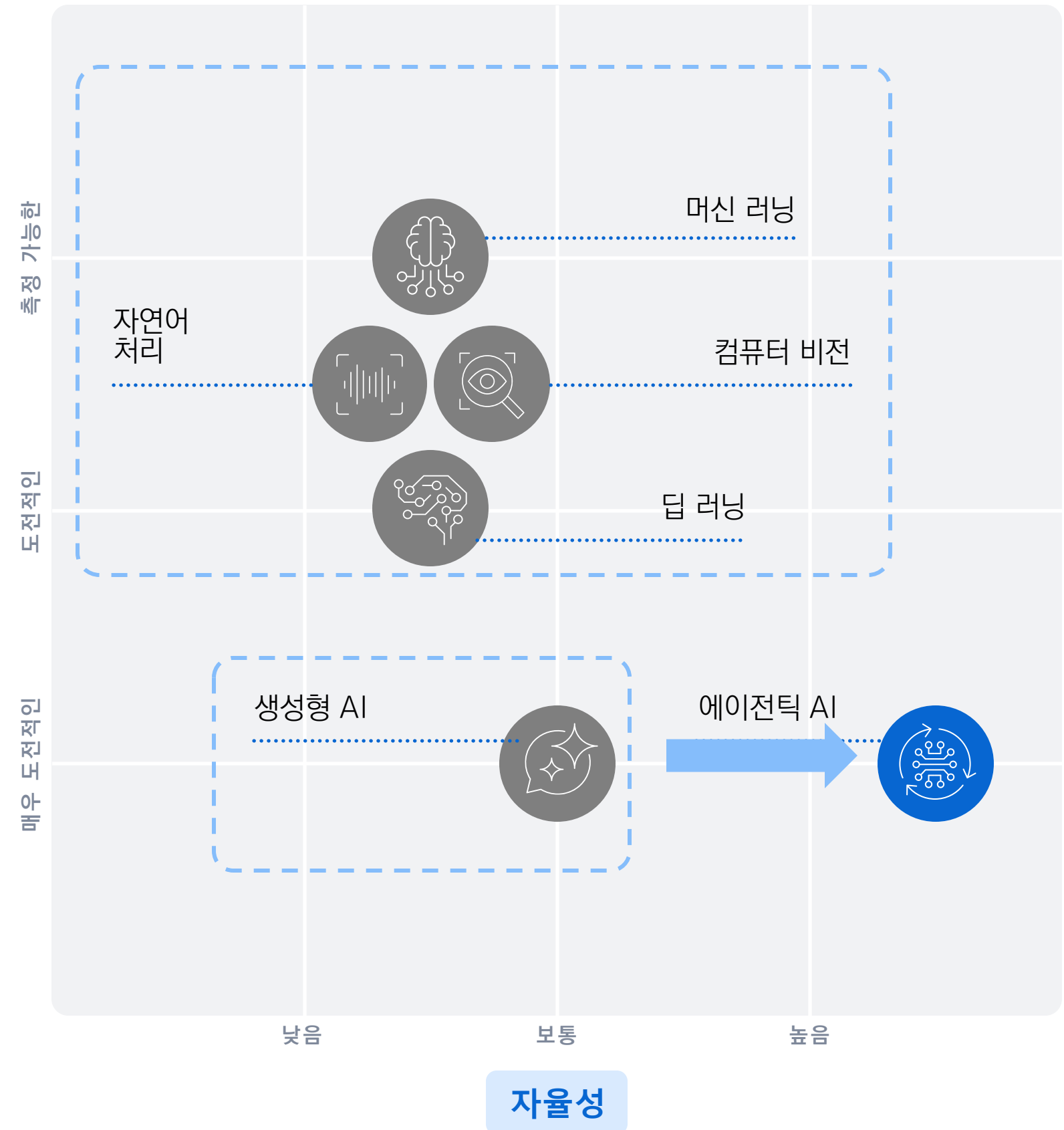
## 에이전틱 AI

주변 환경을 인지하고, 결정을 내리며, 특정하고 복잡한 목표를 달성하기 위해 행동을 취할 수 있는 자율 시스템

### 해결 과제

사람의 통제성 보장, 모든 행동 예측, 자율적 행동에 대한 책임 소재.

신뢰성



# 자율성과 신뢰성 측면의 AI 기술에 대한 고찰

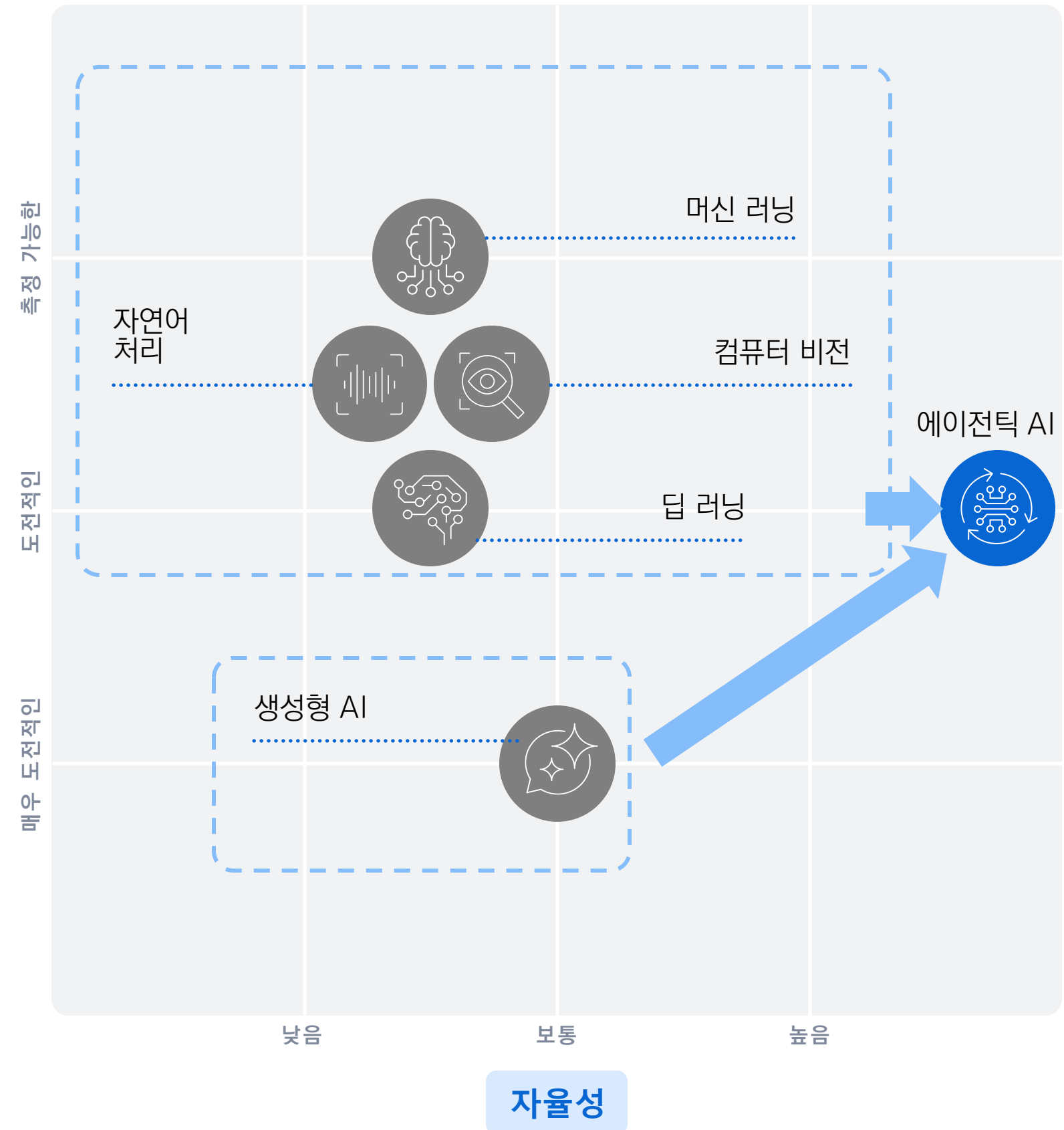
## 에이전틱 AI

주변 환경을 인지하고, 결정을 내리며, 특정하고 복잡한 목표를 달성하기 위해 행동을 취할 수 있는 자율 시스템

### 해결 과제

사람의 통제성 보장, 모든 행동 예측, 자율적 행동에 대한 책임 소재.

신뢰성



자율성

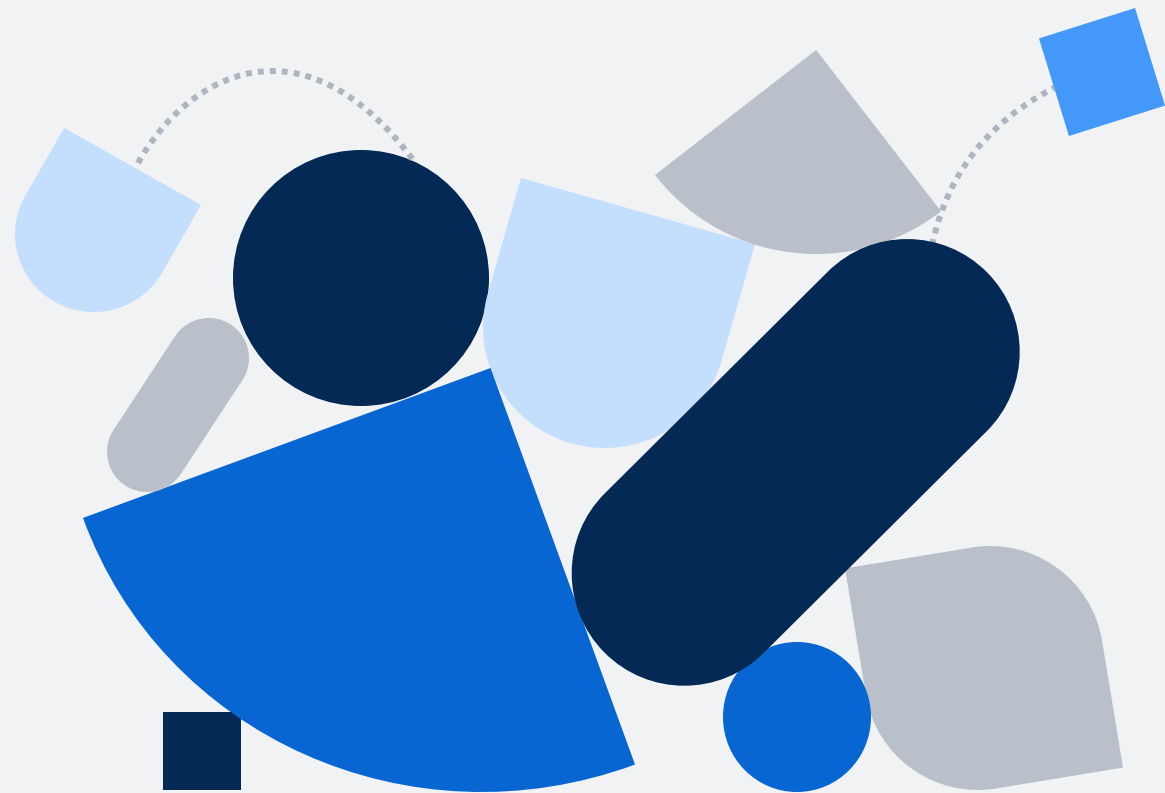
# 복합 AI (Composite AI)



“ 여러 AI 기술을 활용하면 AI의 영향력과 신뢰성이 향상됩니다.  
생성형 AI, LLM을 넘어 머신 러닝, 최적화 등을 포함한 종합적 AI 솔루션을 구축해야 합니다.”

– Gartner Identifies Top Trends in Data and Analytics for 2025

# 조직의 주요 업무에 완전 자율 AI 에이전트를 적용하려면?



## 엔터프라이즈 세계에서는 완전한 자율성 이상의 것이 필요 :

### 의사 결정 추적성 및 설명 가능성

결정은 어떻게 이루어졌으며 그 이유는? LLM으로 구동되는 에이전트는 종종 환각을 일으키고, 일관성이 부족하며, 비즈니스 로직을 무시함.

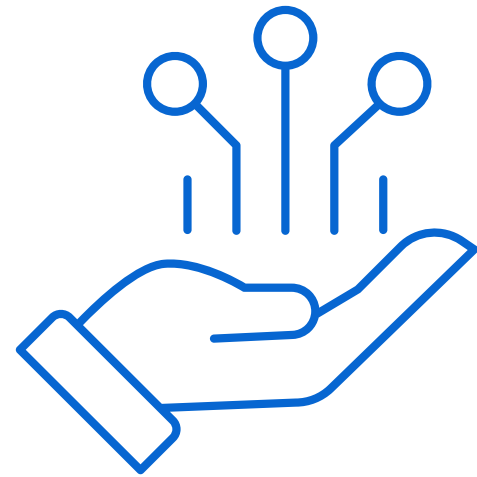
### 위험 완화 및 제어

에이전트가 잘못되면 어떻게 할 것인가? 기업 환경에서 감독이나 통제 없이 행동하는 것은 순식간에 골칫거리가 될 수 있음.

### 거버넌스 및 감사 가능성

에이전트가 윤리적이고 규제 및 조직 표준에 부합하는가? AI 적용이 가속화됨에 따라 오용의 위험도 커지고 있음.

오늘날 거론되는 완전 자율 에이전트는  
엔터프라이즈 환경에서 한계를 가짐



## 이것이 바로 **엔터프라이즈용 AI 에이전트**는 **하이브리드 ( AI + 사람 )** 여야 하는 이유

기업의 특정 비즈니스 조건하에서 설명 가능하고  
통제된 액션으로 인간과 함께 작업하는 협업 에이전트

# 신뢰성 있는 에이전틱 AI

## 통제된 액션으로 인간과 AI가 함께 작업하는 하이브리드 방식의 협업 에이전트

### 인간과의 협업 수준이 낮은 AI 에이전트



Agent 행동 이유?  
추적? 제어?

- 에이전트 행동의 이해가 어렵고
- 통제 및 관리가 어려운 AI



- LLM이 에이전트의 두뇌
- LLM 학습을 위한 비용 및 리스크
- LLM 환각현상, 블랙박스 특성



- 복잡한 코드상에서 Agent의 행동을 추적, 설명, 수정하기 어려움

### 인간과의 협업 수준이 높은 하이브리드 AI 에이전트



Agent 행동에 대한  
인간의 원활한  
추적, 설명, 제어!

- AI 추적성 및 설명 가능성
- AI 리스크 완화 및 통제



- LLM의 주요 역할은 사람과의 커뮤니케이션



- 비즈니스 룰을 준수하는 에이전트 행동을 가시적이고 직관적인 프로세스 플로우로 정의
- 생성형 AI와 기존 AI의 손쉬운 연계

# LLM과 신뢰성 있는 SAS의 AI 분석을 통합한 하이브리드 에이전틱 AI 예시

## LLM을 통한 사용자 인터페이스

고객 ID 1275에 대한 이탈 가능성을 분석하기 위해 churn\_agent를 호출하겠습니다.

고객 ID 1275가 신용카드를 취소할 가능성에 대한 분석 결과는 다음과 같습니다:

- 소득:** 이 고객은 평균 비이탈 고객보다 소득이 상당히 높습니다. 높은 소득은 재정적 안정성을 나타낼 수 있지만, 동시에 고객이 기대하는 서비스 수준이 높을 수 있음을 의미합니다. 만약 고객의 경험이 이러한 기대에 미치지 못한다면, 다른 옵션을 찾고 싶어할 수 있습니다.
- 거부된 거래 수:** 다행히도 이 고객은 거부된 거래가 없습니다. 이는 고객이 재정을 잘 관리하고 적극적인 신호입니다. 그러나 서비스와의 상호작용이 적을 경우, 고객이 소외감을 느끼거나 충성도가 낮아질 수 있습니다.
- 신용 점수:** 이 고객의 신용 점수는 평균 이상으로 매우 우수합니다. 높은 신용 점수는 신뢰성을 나타내며, 고객이 기대하는 프리미엄 서비스와의 격차를 느낄 경우 이탈을 고려할 수 있습니다.

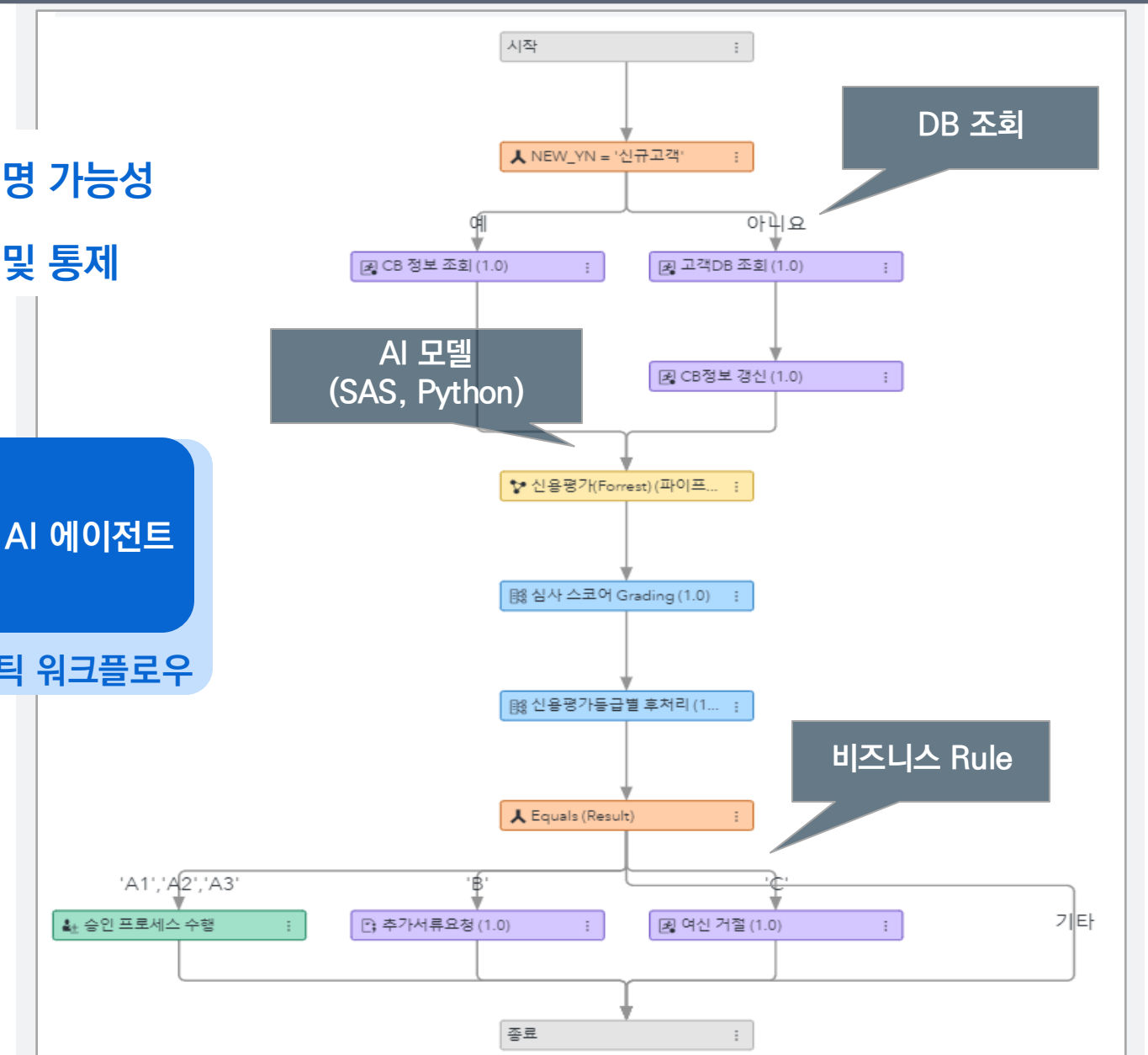
이러한 요소들을 이해함으로써 고객의 우려를 보다 효과적으로 해결하고, 고객을 유지할 수 있는 방법을 모색할 수 있습니다. 추가적인 지원이 필요하시면 말씀해 주세요.

이탈을 방지할 수 있는 제공할 혜택은 무엇이 있나요?



- AI 추적성 및 설명 가능성
- AI 리스크 완화 및 통제

## SAS 플랫폼 기반 에이전틱 워크플로우



- SAS : 고객 데이터 및 AI 기반 분석으로 고객의 이탈 가능성 분석
- LLM : SAS가 제공한 이탈 가능성 분석 결과를 문장으로 표현

# 금융기관의 AI 신뢰성 확보를 위한 고려사항의 확장



## 데이터 품질 및 보안

- AI 활용 데이터의 출처, 품질, 편향 여부를 철저히 검증
- 개인정보 사전 동의, 비식별화 등 법적 요건 충족 및 재식별 가능성 최소화



## AI 투명성 및 설명 가능성

- 금융 서비스에 AI 적용시, 고객에게 의사결정 과정과 결과 설명(XAI) 필요
- 특히 신용평가, 여신심사 등 고위험 서비스에 설명 가능한 AI 기술 적용



## 규제 및 거버넌스

- AI 규제 대응을 위해 투명성, 공정성, 소비자 보호 원칙 준수 필요.
- AI 적용의 전 과정(기획, 개발, 운영)에 걸쳐 책임 소재, 역할, 권한 명확화



## 복합 AI ( GenAI + 기존 AI 조합 )

- 여러 AI 기술을 활용하여 AI의 영향력과 신뢰성 향상
- 생성형 AI, LLM을 넘어 머신 러닝, 최적화 등을 포함한 종합적 AI 솔루션 구축



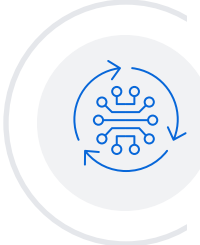
## 하이브리드 AI 에이전트 ( AI + 사람 협업 )

- 통제된 액션으로 인간과 함께 작업하는 협업 에이전트 환경
- 사람이 에이전트의 행동을 이해하고 추적하며, 필요시 제어 가능

AI / ML

생성형 AI  
에이전트 AI

# 에이전틱 AI의 성공 지표 및 적용 분야



## 에이전틱 AI

### 에이전틱 AI의 성공 지표

- 보안 및 감독 가능
- 감사 및 설명 가능
- 성공 여부 정의 및 측정 가능성
- 위험 감소 및 비용 효율성



### 고객 응대

- 아웃바운드 코멘트 지원
- 고객 불만 관리
- 콜센터 운영자 지원



### 리스크 & 컴플라이언스

- 영향 평가 초안 작성
- 모델 위험 문서화
- 컬렉션 최적화
- 대출 문서 처리/데이터 추출



### 사기방지 & 금융 범죄

- 감독관 또는 STR을 위한 초안 문서화
- 검토를 위한 규정 준수 테스트/감사 플래그 지정



### 은행 전략 & 운영

- 내부 정책 및 절차 질의
- 시장 정보
- 트렌드 감시
- 소프트웨어 규정 준수 테스트 및 위반사항 식별



# 에이전틱 AI 예시 : SAS Viya로 자금세탁방지 STR 프로세스 최적화

(Suspicious Transaction Report)

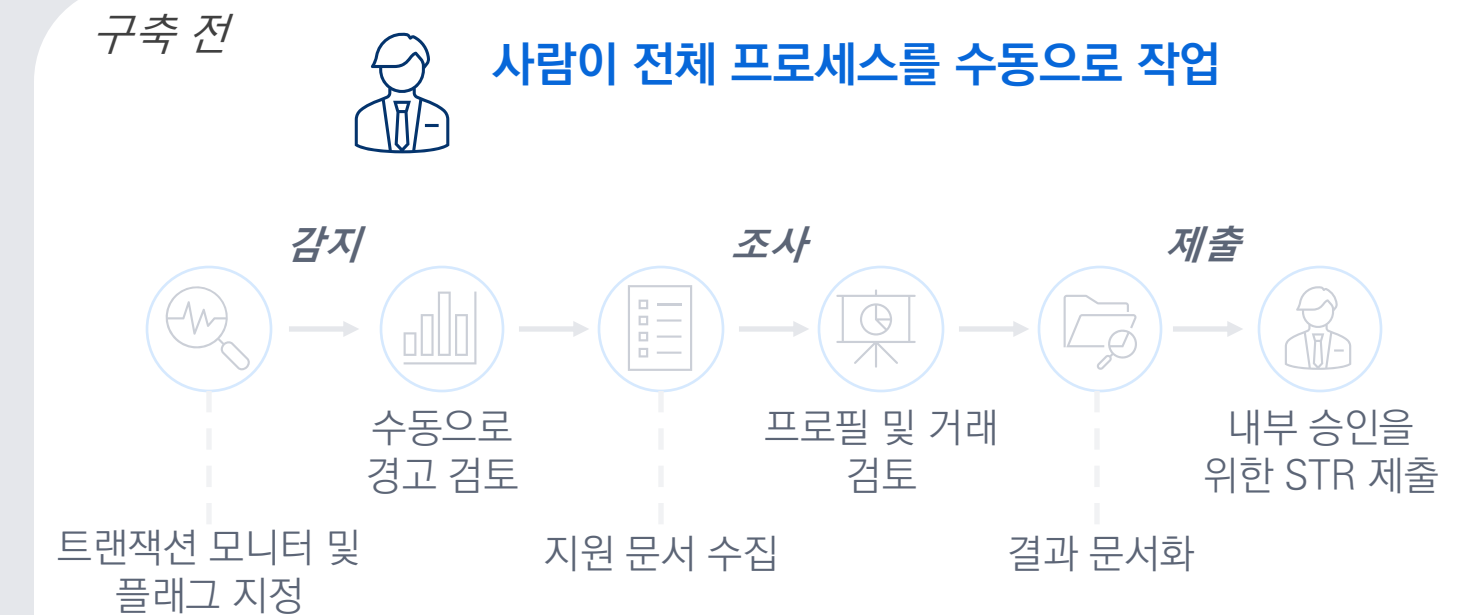
## 이슈

- **경고 피로현상:** AML(자금세탁방지) 팀은 전체 시간의 70~80%를 거짓 양성(false positive) 검토에 소비하며, 그 과정에서 잡음 속에 숨겨진 정교한 사기 수법을 놓치고 있음
- **내러티브 병목 현상:** STR 작성은 사건당 오랜 시간이 걸리므로 제출이 지연되고 품질이 일관되지 않음
- **30일 제출 압력:** 성급한 조사는 취약한 STR에 대한 방어적 과잉 제출로 이어져 법 집행에 대한 정보 가치를 희석시킴

## 솔루션

- **지능형 자동 조사:** AI는 대부분의 경고를 자동으로 처리하므로 분석가는 사람의 판단이 필요한 고위험 사례에만 집중할 수 있음
- **자동화된 내러티브:** AI는 30분 이내에 포괄적인 STR 초안을 작성하고 최종 검토를 위해 분석가에게 전달할 수 있음
- **가속화된 파이프라인:** 엔드 투 엔드 자동화는 STR 작업을 며칠에서 몇 시간으로 단축하는 동시에 품질을 향상시킴

## SAS 에이전틱 AI를 사용한 STR 제출



구축 후

AI 에이전트가 수행한 작업을 사람이 검토하고 신뢰성을 모니터링

# 성공적인 에이전틱 AI 적용을 위한 SAS의 핵심 영역

## 에이전틱 AI 기술 스택



## SAS 에이전틱 AI



### 비 종속적 기술 :

특정 언어, 프레임워크, 데이터 원본 또는 모델에 종속되지 않음



### 사람과 Agent 협업 :

사람이 참여 가능한 에이전트 워크플로우



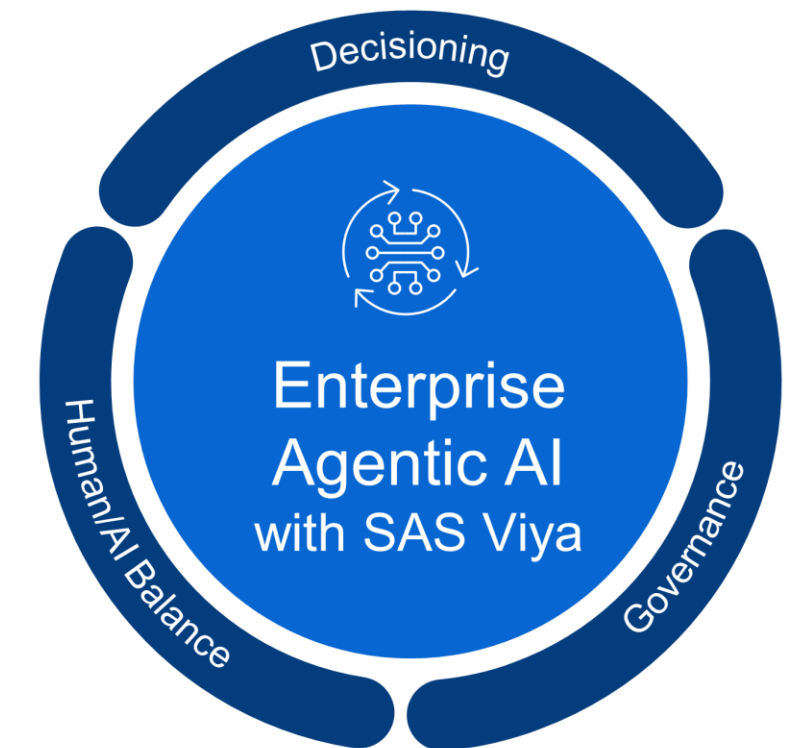
### 복합 AI :

기존 ML 기술과 에이전트 결합

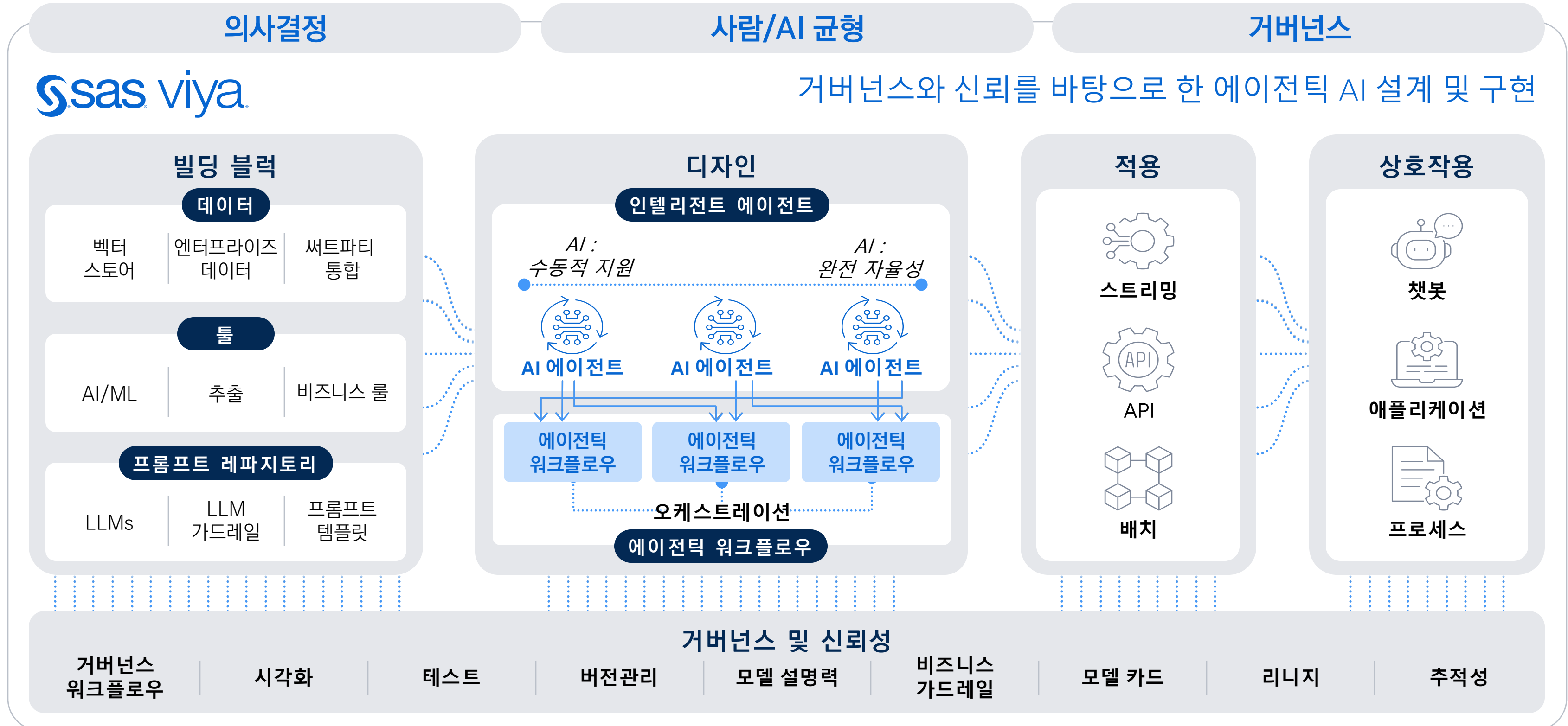


### 신뢰성 :

에이전트 시스템에 필요한 신뢰성 제공



# Agentic AI의 기업 운영환경 적용을 위한 프레임워크





가장 빠르고, 가장 생산적이며,  
신뢰할 수 있는 AI 플랫폼