

Solution for Smarter AI

# 똑똑한 AI를 위한 선택



# 1

## 셀렉트스타 소개



설립

2018

재직 인원

200+

누적 투자액

434억 원

누적 고객사

330+

수행 프로젝트

500+

셀렉트스타(주)는

AI 서비스의 품질과 안전성을 보장하는  
신뢰성 검증 분야를 선도하고 있습니다.

SAMSUNG

NAVER

kakao

LG AI Research

SK telecom

SK 주식회사  
AX

kt

LG U+

LG전자

신한은행

하나금융그룹

kakaobank

삼성생명

경기도교육청

한화생명

KB 증권

SAMSUNG SDS

LG CNS

HYUNDAI

HYUNDAI  
MOBIS

HYUNDAI  
AutoEver

LG 에너지솔루션

우리은행

NH농협은행

DaehanSteel

upstage

현재 라운드

Series B

총 투자 유치

434억

투자 현황

Series B 260억

KB 인베스트먼트

KB 증권

신한벤처투자

salesforce ventures

NCUC  
ACVC PARTNERS

MOORIM 무림캐피탈

infobank  
PARTNERS

키움인베스트먼트

MIRAE ASSET  
미래에셋벤처투자

PLUX  
VENTURES

Series A 170억

KDB산업은행

CJ 인베스트먼트

삼성증권

(주)코오롱인베스트먼트

kakaoventures

SBI 인베스트먼트

삼성벤처투자

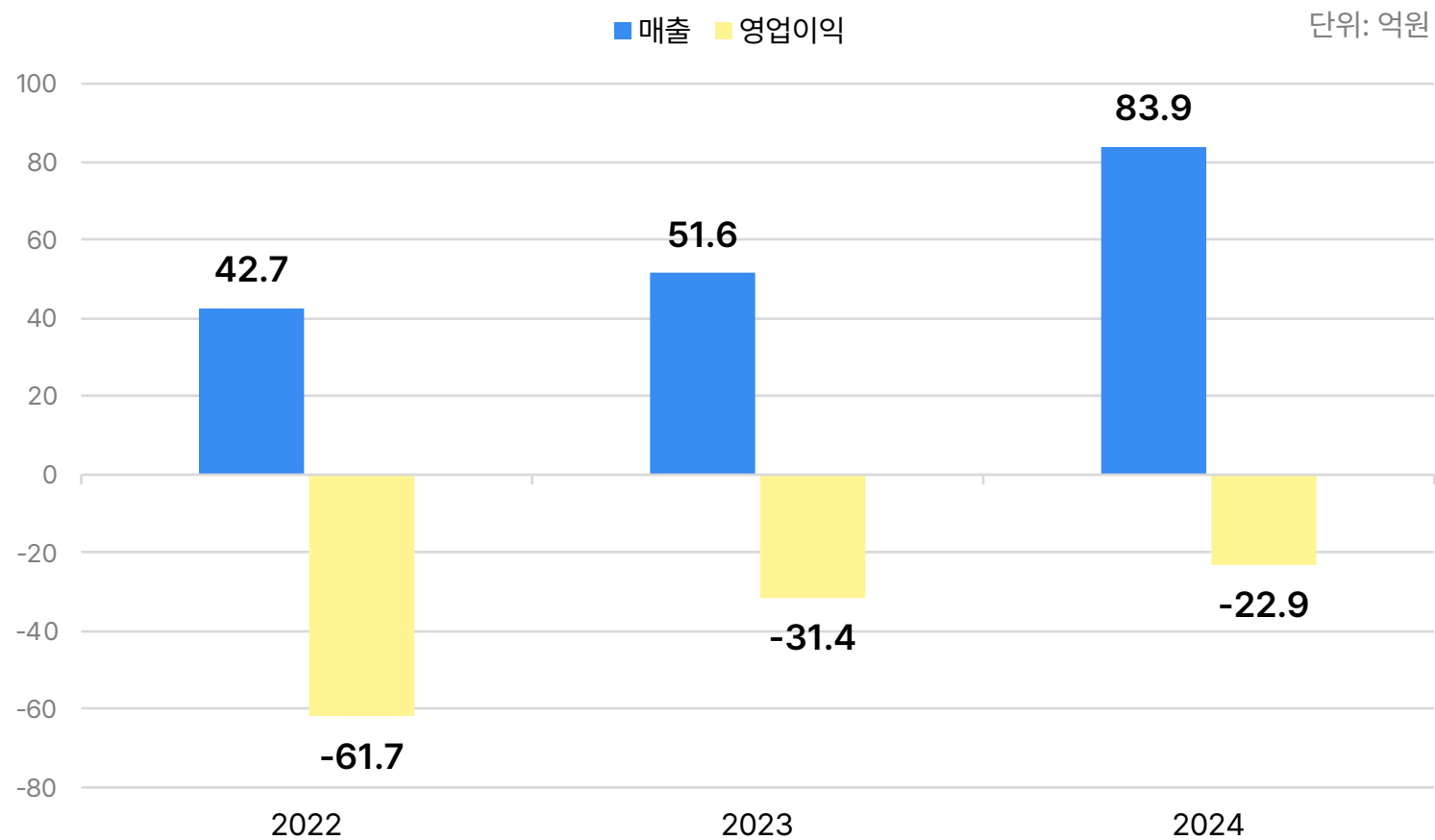
Company partners

NAU

Seed 4억

kakaoventures

# 매출 현황



# 세계 최고 권위 AI 학회 등재

NeurIPS EMNLP CVPR 등 글로벌 Top AI 학회 등재 데이터셋 국내 최다 구축



## KorNAT

국내 최초 한국형 LLM 평가 데이터셋



인공지능 윤리정책 포럼



## KorQuAD 2.0

한국지능정보사회진흥원 인공지능 학습용 데이터 구축 사업

LLM 신뢰성 벤치마크 데이터: 국내 최초 한국어 언어 모델 신뢰성 기준 제작

# AI 국가대표 셀렉트스타

주요 성과



민간 AI 신뢰성 인증 AI-마스터 시험기관 참여



독자 AI 파운데이션 모델 SKT 컨소시엄 합류



포브스코리아 '2025 대한민국 AI 50' 선정



AI + X Top 100 5년 연속 선정



KPAS(Korea Promising AI Startup) 2년 연속 선정

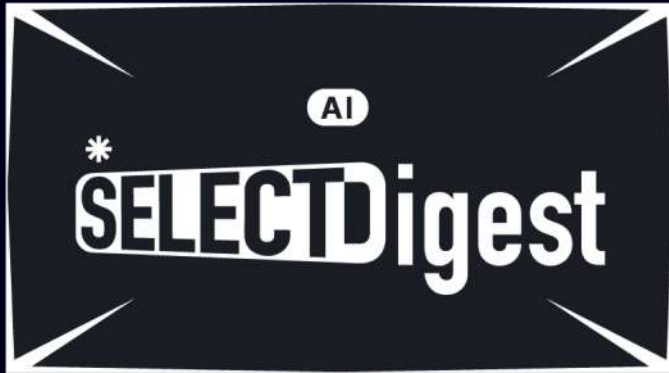


포브스지 선정 '2021 아시아 30세 이하 리더' 선정



# AI 국가대표 셀렉트스타

주최 행사



국내 동종업계 1위 뉴스레터 운영



AI Talk with 앤드류 응 기획 / 운영  
진행 및 키노트 연사



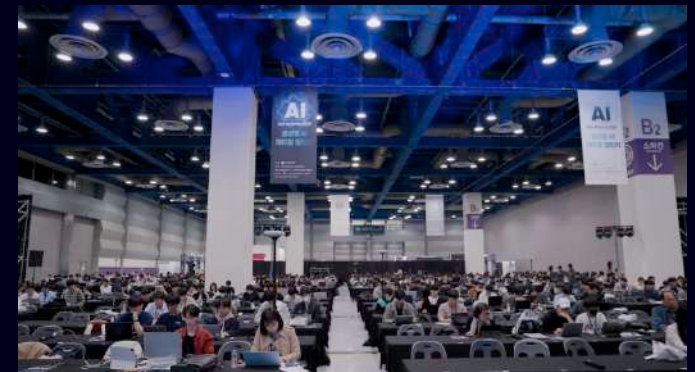
분기별 AI IGNITE 기획 및 진행  
AI 업계 네트워킹 및 인사이트 공유 행사



MWC25 주최  
GSMA 공동 세계최초 '글로벌 AI 레드팀 챌린지' 주최



AI 리더스 포럼  
사상 최초 NeurIPS 한국인 조직 위원들과 포럼 진행



2024 생성형 AI 레드팀 챌린지 기획 / 운영  
국내 최초, 세계 최대 규모



All-in-one Data Service

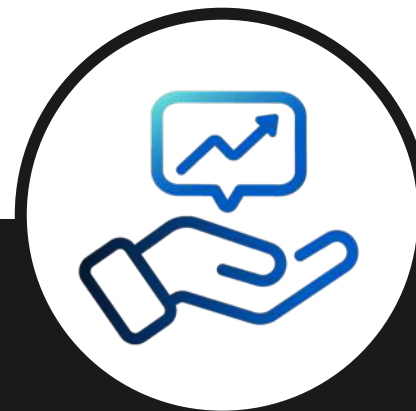
# 더 나은 AI를 위한 데이터 설계, 구축, 신뢰성 검증



AI 데이터 구축



AI 데이터 판매



신뢰성 검증 컨설팅



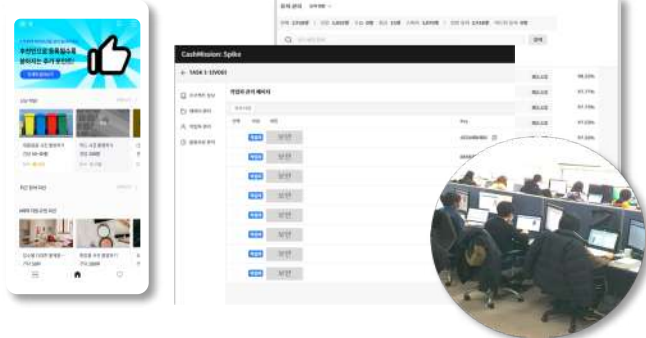
신뢰성 검증 솔루션

# 1

## AI 데이터 구축 서비스

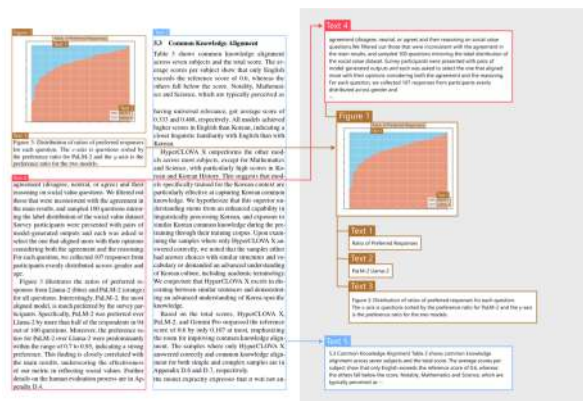


# 자사 솔루션을 통한 고객 맞춤 고품질 데이터 구축



## 캐시미션: 데이터 구축 플랫폼

- 클라우드소싱 / 보안망 데이터 구축 환경
- 이미지, 텍스트, 오디오 등 다양한 데이터 구축
- 고객 요구사항에 맞춘 최적화 세팅



## Connect-T: 문서 파싱 및 구조화

- 문서의 구성 요소를 분석해 디지털 형태로 위치 추론
- 문서 텍스트(폰트, 언어)를 인식해 데이터로 변환
- 복잡한 수식이나 표를 LaTeX 형식으로 변환



## 전문 데이터 작업자 보유

- 금융, 보험 등 전문 분야 작업자 소싱채널 보유
- 2,000명 규모 우수 작업자 인하우스 관리
- 25만명 이상 클라우드소싱 작업자 보유

# 2

## 데이터셋 스토어

- 판매 데이터 목록
- 데이터 정제 툴
- 데이터 정제 구조 예시



# 판매 데이터

라이선스 이슈가 해결된, 다양한 형태 및 도메인의 데이터셋을 고객의 요구사항(LaTeX, 이미지, 추출/정제 등)에 맞춰 제공



## 전문가 질의응답 데이터셋

- 법률, 의료, 금융 등 분야별 질의에 대한 전문가 답변 Set 230만 건 이상



## 도서 데이터셋

- 200여 개 카테고리로 구분된 종이책/전자책
- 국내 유통되는 대부분의 도서 제공



## 언론 데이터셋

- 경제, 정치, 문화 등 다양한 분야에 대한 언론 기사 및 영상



## LLM 무해성 평가 데이터셋

- LLM의 무해성을 평가할 수 있는 질문 데이터
- 편향성, 혐오, 위법성, 민감성 및 시의성 반영



## 다국어 번역 데이터셋

- 외국어-외국어, 외국어-한국어 번역 코퍼스
- 52개 이상 언어



## 사진, 일러스트, 영상 데이터

- 글로벌 이미지 플랫폼 제휴
- 인물, 교통, 동물, 음식 등 약 5억 건



## 코딩 테스트 데이터셋

- Python3, Java, C++ 등 다양한 언어에 대해 코딩 테스트 문제-풀이 데이터셋



## 카드 데이터셋

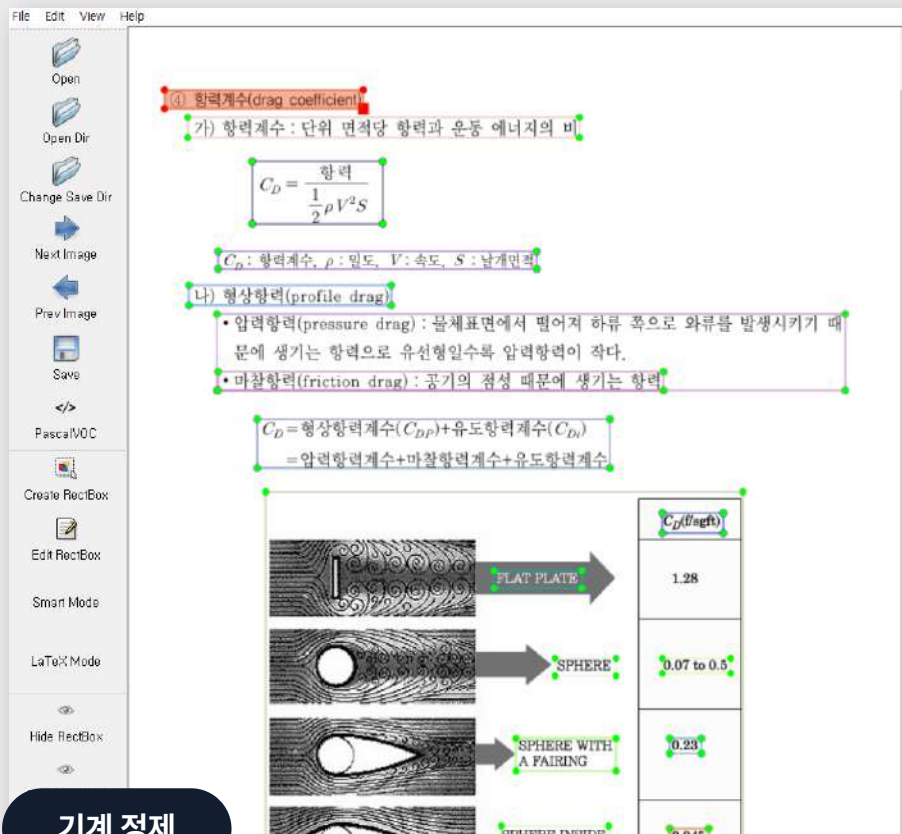
- 가맹점 소재지별 카드 사용(소비) 데이터셋



## 합성 데이터셋

- 주형, 제조, 안전 등 희귀 사례에 대한 언리얼 엔진 시뮬레이션 데이터

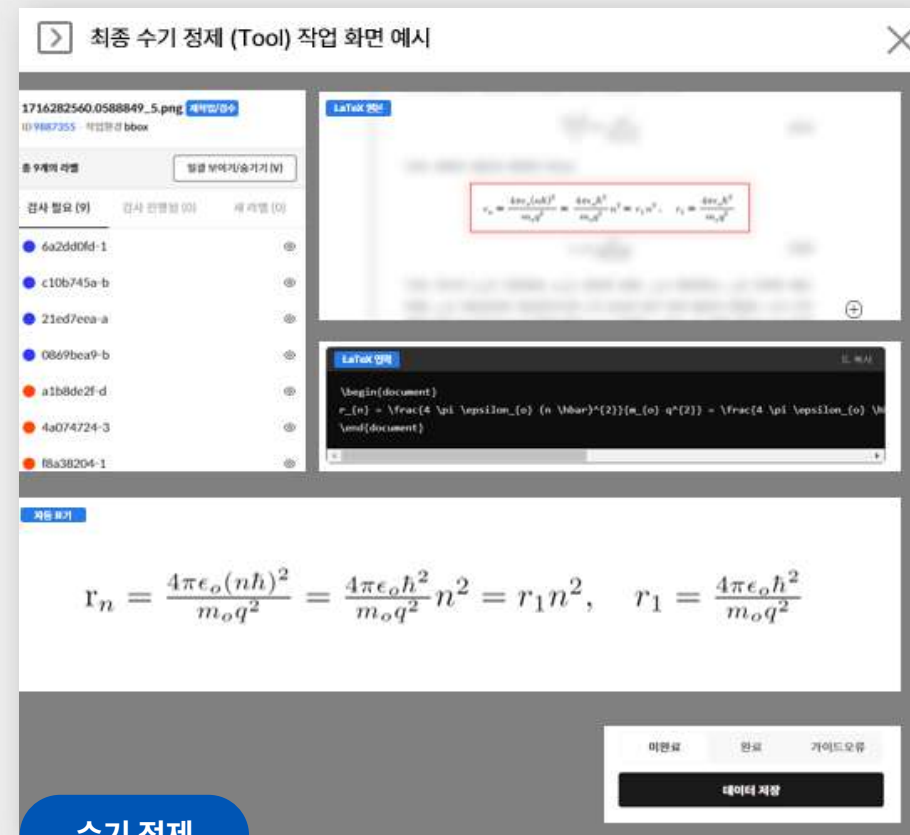
# 데이터 정제 툴



기계 정제

## Connect-T

▲ 정제 맞춤형 콘텐츠 추론 및 OCR 솔루션(AI)



수기 정제

## Cash mission(보안용)

▲ 정제 맞춤형 작업 환경(자사 플랫폼)



예시: Lv. 4(표 LaTeX)

데이터  
정제 구조

08 날씨가 춥다 / 한강에 놀러간다

- ① 날씨가 추웠기 때문에 놀러 가라고 했다.
- ② 날씨가 추웠는데 놀러 가게 했다.
- ③ 날씨가 추워서 한강에 놀러 가기는 틀렸다.
- ④ 날씨가 추웠던 한강에 놀러 갈지도 모른다.

09 이 글에서 설명한 민영주택의 주택청약 1순위에 해당하는 사람은?

민영주택을 청약받기 위한 자격요건은 다음과 같다.		
청약순위	청약종장 (입주자처속)	순위별 조건
		청약통장 가입기간      납입금
1순위	주택청약 종합저축	• 투기과열지구 및 청약과열지역 : 가입 후 2년이 경과한 분      납입인정금액이 지역별 예치금액 이상인 분
	청약예금	• 위촉지역 : 가입 후 1개월이 경과한 분      납입인정금액이 지역별 예치금액 이상인 분
	청약부금 (85㎡ 이하만 청약 가능)	• 투기과열지구 및 청약과열지역, 위촉지역 외 매월 약정납입일에 납입한 납입인정 금액이 지역별 예치금액 이상인 분
2순위		1순위에 해당하지 않는 분(청약통장 가입자만 청약 가능)

- ① 투기과열지구에서 청약통장의 가입 기간이 12개월인 사람
- ② 위촉지역에서 청약통장의 가입 기간이 6개월인 사람
- ③ 투기과열지구, 청약과열지구, 위촉지역이 아닌 수도권 외 지역에서 청약통장  
의 가입 기간이 3개월인 사람
- ④ 투기과열지구, 청약과열지구, 위촉지역이 아닌 수도권에서 청약통장의 가입  
기간이 10개월인 사람

```
{
  "data_id": 1349,
  "data_file": "2025 귀화시험 사회통합프로그램 종합평가 모의고사.pdf",
  "data_title": "2025 귀화시험 사회통합프로그램 종합평가 모의고사",
  "data_source": "도서",
  "category_main": "기타",
  "category_sub": "기타",
  "data_type": "Text",
  "collected_date": "2024.10.11",
  "content": "문제 : [07-08] 다음을 한 문장으로 알맞게 연결한 것을 고르시오. 08 날씨가 추  
다 / 한강에 놀러가다\n보기 : ㉠ 날씨가 추웠기 때문에 놀러 가라고 했다.\n㉡ 날씨가 추웠는데 놀  
러 가게 했다.\n㉢ 날씨가 추워서 한강에 놀러 가기는 틀렸다.\n㉣ 날씨가 추웠던 한강에 놀러 갈지도  
모른다.\n해설 : 08\n㉠-기는 틀렸다\n㉡는 바라거나 하려는 일이 순조롭게 되지 못함을 나타내는 말  
이다. 날씨가 너무 추운 탓에 한강에 놀러 갈 수 없게 되었음을 이야기하고 있으므로\n㉡'놀러 가기는  
틀렸다\n㉢'가 들어가는 것이 적절하다.\n㉣ 순조롭다 : 일 따위가 아무 말썽 없이 예정대로 잘되  
어 가는 상태에 있다.\n㉤ 많은 사람의 걱정과는 다르게 두 사람의 결혼생활은 순조로웠다.\n[오답해설]\n㉠- (으)라고 하다 : 다른 곳에서 들은 명령을 상대방에게 전달할 때 쓰는 말\n㉤에 과장남겨서 이 일  
을 오늘까지 끝내라고 했어요.\n㉢- 게 하다 : 앞말의 행동을 시키거나 앞말이 뜻하는 상태가 되  
도록 힘을 나타내는 말\n㉤에 그 사람은 내게 이쪽 복도를 청소하게 했다.\n㉢- =지도 모른다 : 어떤 일이  
일어날 것을 추측할 때 사용하는 말\n㉤에 이제 정마찰이 시작되었으니 내일부터 비가 올지도 몰  
라.\n정답 : ㉢
}
{
  "data_id": 1350,
  "data_file": "2025 귀화시험 사회통합프로그램 종합평가 모의고사.pdf",
  "data_title": "2025 귀화시험 사회통합프로그램 종합평가 모의고사",
  "data_source": "도서",
  "category_main": "기타",
  "category_sub": "기타",
  "data_type": "Text",
  "collected_date": "2024.10.11",
  "content": "문제 : [07-08] 다음을 한 문장으로 알맞게 연결한 것을 고르시오. 08 날씨가 추다 / 한  
강에 놀러가다.\n보기 : ㉠ 날씨가 추웠기 때문에 놀러 가라고 했다.\n㉡ 날씨가 추웠는데 놀러 가게 했  
다.\n㉢ 날씨가 추워서 한강에 놀러 가기는 틀렸다.\n㉣ 날씨가 추웠던 한강에 놀러 갈지도 모른  
다.\n해설 : 08\n㉠-기는 틀렸다\n㉡는 바라거나 하려는 일이 순조롭게 되지 못함을 나타내는  
말이다. 날씨가 너무 추운 탓에 한강에 놀러 갈 수 없게 되었음을 이야기하고 있으므로\n㉡'놀러 가기는  
틀렸다\n㉢'가 들어가는 것이 적절하다.\n㉣ 순조롭다 : 일 따위가 아무 말썽 없이 예정대로 잘되  
어 가는 상태에 있다.\n㉤ 많은 사람의 걱정과는 다르게 두 사람의 결혼생활은 순조로웠다.\n[오답해  
설]\n㉠- (으)라고 하다 : 다른 곳에서 들은 명령을 상대방에게 전달할 때 쓰는 말\n㉤에 과장남겨서 이 일  
을 오늘까지 끝내라고 했어요.\n㉢- 게 하다 : 앞말의 행동을 시키거나 앞말이 뜻하는 상태가 되  
도록 힘을 나타내는 말\n㉤에 그 사람은 내게 이쪽 복도를 청소하게 했다.\n㉢- =지도 모른다 : 어떤  
일이 일어날 것을 추측할 때 사용하는 말\n㉤에 이제 정마찰이 시작되었으니 내일부터 비가 올지도 몰  
라.\n정답 : ㉢
}
```

# 3

## 신뢰성 검증 컨설팅

- 신뢰성 검증 프로세스



# 도메인 및 서비스 특화 평가 중요성

상용 및 오픈소스 생성형 AI는 도메인 및 기업 내부 데이터를 통한 학습·검증이 어렵기에  
도메인/서비스 특화로 갈수록 성능이 하락하며, 이에 대한 맞춤 평가가 필수

## 도메인에 따른 생성형 AI 성능 저하

	GPT-4 Evaluated few-shot
MMLU [49] Multiple-choice questions in 57 subjects (professional & academic)	<b>86.4%</b> 5-shot
HellaSwag [52] Commonsense reasoning around everyday events	<b>95.3%</b> 10-shot
AI2 Reasoning Challenge (ARC) [54] Grade-school multiple choice science questions, Challenge-set.	<b>96.3%</b> 25-shot
WinoGrande [56] Commonsense reasoning around pronoun resolution	<b>87.5%</b> 5-shot
HumanEval [43] Python coding tasks	<b>67.0%</b> 0-shot
DROP [58] (F1 score) Reading comprehension & arithmetic.	80.9 3-shot
GSM-8K [60] Grade-school mathematics questions	<b>92.0%*</b> 5-shot chain-of-thought

일반 → 금융 도메인  
GPT4 성능 하락

Dataset	Metrics	Chat GPT	GPT 4
NER	EntityF1	0.77*	<b>0.83*</b>
FINER-ORD	EntityF1	0.28	<b>0.77</b>
FinRED	F1	0.00	<b>0.02</b>
SC	F1	0.80	0.81
CD	F1	0.00	0.01
FNXL	EntityF1	0.00	0.00
FSRL	EntityF1	0.00	0.01
FPB	F1	0.78*	0.78*
	Acc	0.78*	0.76*
FiQA-SA	F1	0.60	0.80
TSA	RMSE↓	0.53	0.50
Headlines	AvgF1	0.77*	0.86*
	F1	0.64	<b>0.71</b>
FOMC	Acc	0.6	<b>0.69</b>
FinArg-ACC	MicroF1	0.50	<b>0.60</b>
FinArg-ARC	MicroF1	0.39	0.40
MultiFin	MicroF1	0.59	<b>0.65</b>
MA	MicroF1	0.85	0.79
MLESG	MicroF1	0.25	<b>0.35</b>
FinQA	EmAcc	0.58*	<b>0.63*</b>
TATQA	EmAcc	0.00*	0.13*
Regulations	Rouge-1	0.12	0.11
	BertScore	0.64	0.62
ConvFinQA	EmAcc	0.60*	<b>0.76*</b>

## 서비스 개발 현장의 목소리

**S사)**

범용적인 벤치마크 데이터로 성능을 평가한 LLM을  
금융 도메인에 적용했을 때의 성능 간극에 질렸다

**L사)**

매우 다양한 사용자 발화 케이스를 자체  
수작업 검증으로는 커버가 불가능하고,  
어떤 사항을 놓쳤을 지 알 수 없어 불안하다

# 인공지능 서비스 / 모델의 신뢰성 검증 프로세스



## 평가 지표 설계



## 데이터 구축



## 신뢰성 검증



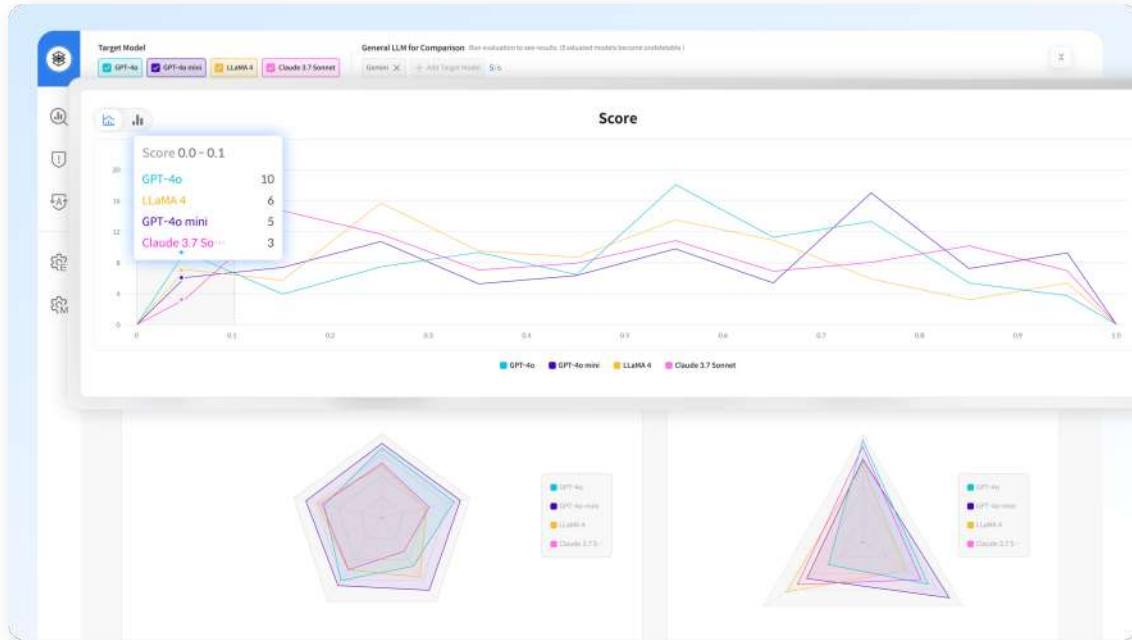
생성형 AI 기반 모델 및 서비스의 품질·안전성 기준을 설계하고 제3자 검증을 제공

# 4

## 신뢰성 검증 솔루션



# 생성형 AI 신뢰성 검증 솔루션



## 셀렉트스타, 국내 최초 'AI 신뢰성 검증 자동화 솔루션' 출시

디지털데일리 | 발행일 2025-02-20 11:04:45

[디지털데일리 이건한 기자] 셀렉트스타가 국내 최초로 생성형 AI 서비스 '신뢰성 검증 자동화 솔루션'을 출시했다고 20일 밝혔다. 다투모 이밸(DATUMO Eval)로 명명된 본 서비스는 AI 서비스 운영 중 발생할 가능한 오류와 위험성을 사전에 식별 및 대응할 수 있도록 지원하는 AI 신뢰성 검증 프로세스 자동화 솔루션이다.

### 1. 평가 지표 설정 및 관리

- ✓ RAG Quality & Safety 자체 평가 지표 보유
- ✓ Use case 기반 기업 맞춤형 지표 설계

### 2. 고품질 평가 데이터셋 자동 생성 및 관리

- ✓ 목적(도메인&사용자)별 맞춤 고품질 평가 데이터 생성
- ✓ 벤치마크 및 Safety 평가 데이터셋을 탑재

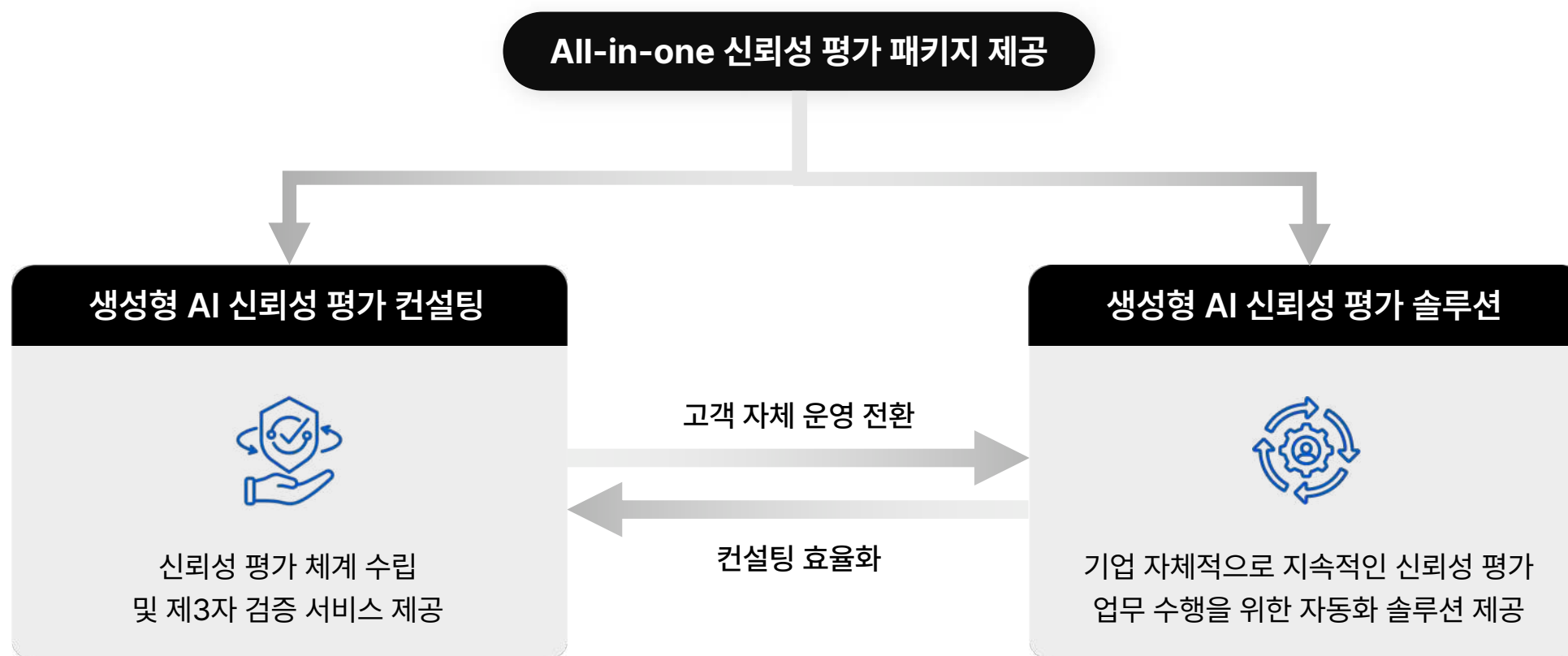
### 3. 평가 결과 분석 및 레드티밍 수행

- ✓ 모델 평가 모니터링 (복수의 모델 및 버전 별 성능 관리)
- ✓ 레드티밍 전략 및 자동 레드티밍 수행





# 국내 최초 1위 생성형 AI 신뢰성 전문 기업





**SELECTSTAR**  
THE DATA-CENTRIC AI COMPANY

[contact@selectstar.ai](mailto:contact@selectstar.ai)