

Solution for Smarter AI

똑똑한 AI를 위한 선택



1

셀렉트스타 소개



설립

2018 (8년차)

재직 인원

150명

누적 투자액

379억 원

누적 고객사

324개사

수행 프로젝트

500 +

셀렉트스타(주)는 Data-centric AI 기업으로
AI 서비스의 품질과 안전성을 보장하는
신뢰성 검증 분야를 선도하고 있습니다.

SAMSUNG

NAVER

kakao

LG AI Research

SK telecom

kt

LG U+

LG전자

신한은행

하나금융그룹

kakaobank

삼성생명

한화생명

KB증권

SAMSUNG SDS

LG CNS

HYUNDAI

HYUNDAI
MOBISHYUNDAI
AutoEver

LG 에너지솔루션

우리은행

NH농협은행

DaehanSteel

upstage

현재 라운드

Series B

총 투자 유치

379억

투자 현황

Series B 205억

Series A 170억

Seed 4억

kakaovertures

KDB산업은행



(주)코오롱인베스트먼트

kakaovertures

Company Partners



KB인베스트먼트

KB증권

신한벤처투자



삼성증권



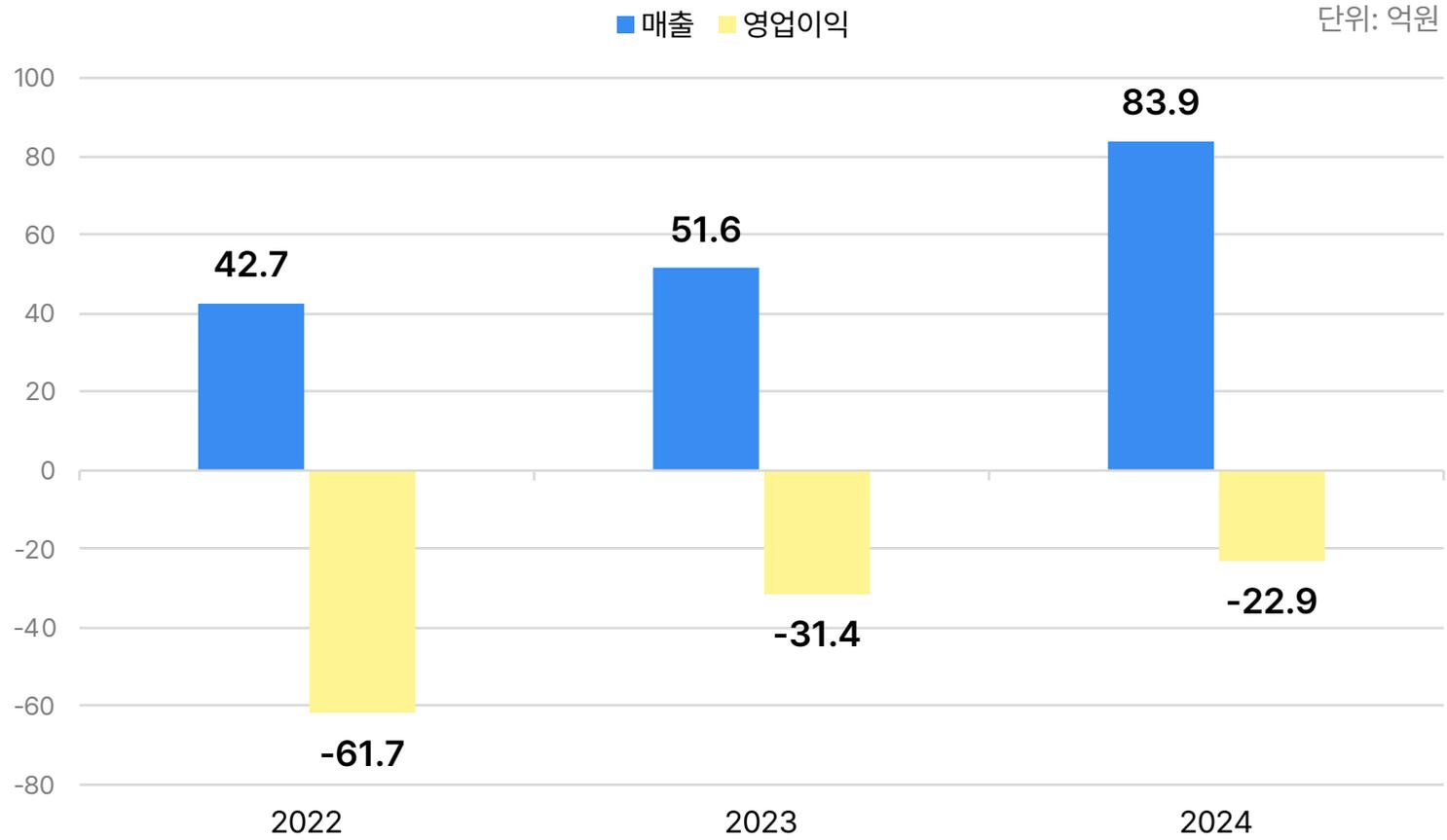
SBI인베스트먼트

MOORIM 무림캐피탈

infobank PARTNERS

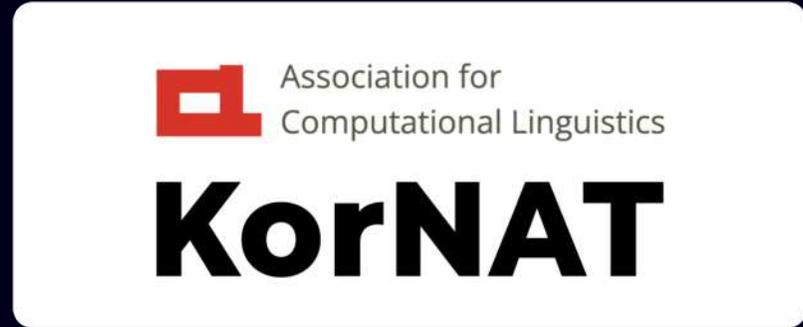
키움인베스트먼트

매출 현황



세계 최고 권위 AI 학회 등재

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



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



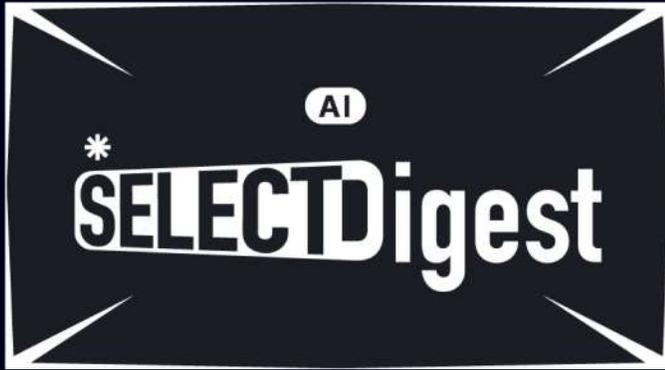
KPAS(Korea Promising AI Startup) 2년 연속 선정



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



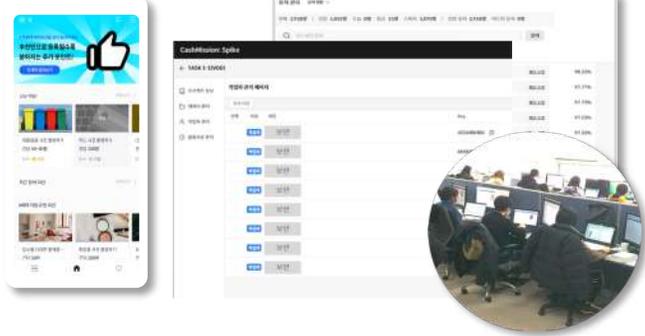
신뢰성 검증 솔루션

1

AI 데이터 구축 서비스

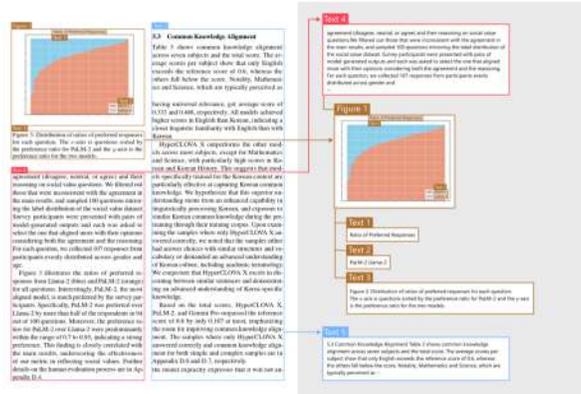


자사 솔루션을 통한 고객 맞춤 고품질 데이터 구축



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

The screenshot shows the Connect-T interface with a document containing text and a table. The text includes definitions for drag coefficient and types of drag. The table lists drag coefficients for different shapes.

Shape	C_D (drag coefficient)
FLAT PLATE	1.28
SPHERE	0.07 to 0.5
SPHERE WITH A FAIRING	0.23
SPHERE INSIDE	0.045

기계 정제

Connect-T

▲ 정제 맞춤형 콘텐츠 추론 및 OCR 솔루션(AI)

The screenshot shows the Cash mission interface with a document containing a LaTeX equation and a list of items. The equation is rendered in a clean, professional format.

$$\Gamma_n = \frac{4\pi\epsilon_0(nh)^2}{m_oq^2} = \frac{4\pi\epsilon_0h^2}{m_oq^2}n^2 = r_1n^2, \quad r_1 = \frac{4\pi\epsilon_0h^2}{m_oq^2}$$

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 추울텐데 놀러 가게 했다.
- ③ 날씨가 추워서 한강에 놀러 가기는 불편하다.
- 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 증입자처속	• 투기과열지구 및 청약과열지역 : 가입 후 2년이 경과한 분	납입인정금액이 지역별 예치금액 이상인 분
	청약예금	• 위촉지역 : 가입 후 1개월이 경과한 분 • 투기과열지구 및 청약과열지역, 위촉지역 외	이상인 분
	청약부금 (85㎡ 이하만 청약 가능)	- 수도권 지역 : 가입 후 1년이 경과한 분(다만, 필요 한 경우 시·도지사가 2개월까지 연장 가능) - 수도권 외 지역 : 가입 후 6개월이 경과한 분(다만, 필요한 경우 시·도지사가 12개월까지 연장 가능)	매월 약정납입액에 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 청약통장의 가입 기간이 3개월인 사람
- ④ 투기과열지구, 청약과열지구, 위촉지역이 아닌 수도권에서 청약통장의 가입 기간이 10개월인 사람

```
{
  "data_id": 1349,
  "data_file": "2025 귀환시험 사회통합프로그램 종합평가 모의고사.pdf",
  "data_title": "2025 귀환시험 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 춥다 / 한강에 놀러간다\n보기 : ㉠ 날씨가 추웠기 때문에 놀러 가라고 했다.\n㉡ 날씨가 추울텐데 놀러 가게 했다.\n㉢ 날씨가 추워서 한강에 놀러 가기는 불편하다.\n㉣ 날씨가 추웠던 한강에 놀러 갈지도 모른다.\n해설 : 08\''-기는 불편하다 \''는 바라거나 하려는 일이 순조롭게 되지 못함을 나타내는 말이다. 날씨가 너무 추운 탓에 한강에 놀러 갈 수 없게 되었음을 이야기하고 있으므로\''놀러 가기는 불편하다 \''가 들어가는 것이 적절하다.\n* 순조롭다 : 일 따위가 아무 말썽 없이 예정대로 잘되어 가는 상태에 있다.\n* 많은 사람의 걱정과는 다르게 두 사람의 결혼생활은 순조로웠다.\n[오답해설]\n㉠ -(으)라고 하다 : 다른 곳에서 좋은 명령을 상대방에게 전달할 때 쓰는 말\n* 과장스럽게서 이 말을 오늘날까지 끝내라고 했어요.\n* -게 하다 : 앞말의 행동을 시키거나 앞말이 뜻하는 상태가 되도록 힘을 나타내는 말\n* 그 사람은 내게 이쪽 복도를 청소하게 했다.\n* -a지도 모르다 : 어떤 일이 일어날 것을 추측할 때 사용하는 말\n* 내에 이제 정마찰이 시작되었으니 내일부터 비가 올지도 몰라.\n정답 : ㉠"}
  "add_info": {
    "page_num": 98,
    "문제": "[07-08] 다음을 한 문장으로 알맞게 연결한 것을 고르시오. 08 날씨가 춥다 / 한강에 놀러간다.",
    "보기": "㉠ 날씨가 추웠기 때문에 놀러 가라고 했다.\n㉡ 날씨가 추울텐데 놀러 가게 했다.\n㉢ 날씨가 추워서 한강에 놀러 가기는 불편하다.\n㉣ 날씨가 추웠던 한강에 놀러 갈지도 모른다.\n",
    "해설": "08\''-기는 불편하다 \''는 바라거나 하려는 일이 순조롭게 되지 못함을 나타내는 말이다. 날씨가 너무 추운 탓에 한강에 놀러 갈 수 없게 되었음을 이야기하고 있으므로\''놀러 가기는 불편하다 \''가 들어가는 것이 적절하다.\n* 순조롭다 : 일 따위가 아무 말썽 없이 예정대로 잘되어 가는 상태에 있다.\n* 많은 사람의 걱정과는 다르게 두 사람의 결혼생활은 순조로웠다.\n[오답해설]\n㉠ -(으)라고 하다 : 다른 곳에서 좋은 명령을 상대방에게 전달할 때 쓰는 말\n* 과장스럽게서 이 말을 오늘날까지 끝내라고 했어요.\n* -게 하다 : 앞말의 행동을 시키거나 앞말이 뜻하는 상태가 되도록 힘을 나타내는 말\n* 그 사람은 내게 이쪽 복도를 청소하게 했다.\n* -a지도 모르다 : 어떤 일이 일어날 것을 추측할 때 사용하는 말\n* 내에 이제 정마찰이 시작되었으니 내일부터 비가 올지도 몰라.",
    "정답": "㉠"
  }
}
```

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

일반 → 금융 도메인
GPT4 성능 하락

Dataset	Metrics	Chat GPT	GPT 4
NER	EntityF1	0.77*	0.83*
FINER-ORD	EntityF1	0.28	0.77
FinRED	F1	0.00	0.02
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*
FOMC	F1	0.64	0.71
	Acc	0.6	0.69
FinArg-ACC	MicroF1	0.50	0.60
FinArg-ARC	MicroF1	0.39	0.40
MultiFin	MicroF1	0.59	0.65
MA	MicroF1	0.85	0.79
MLESG	MicroF1	0.25	0.35
FinQA	EmAcc	0.58*	0.63*
TATQA	EmAcc	0.00*	0.13*
Regulations	Rouge-1	0.12	0.11
	BertScore	0.64	0.62
ConvFinQA	EmAcc	0.60*	0.76*

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



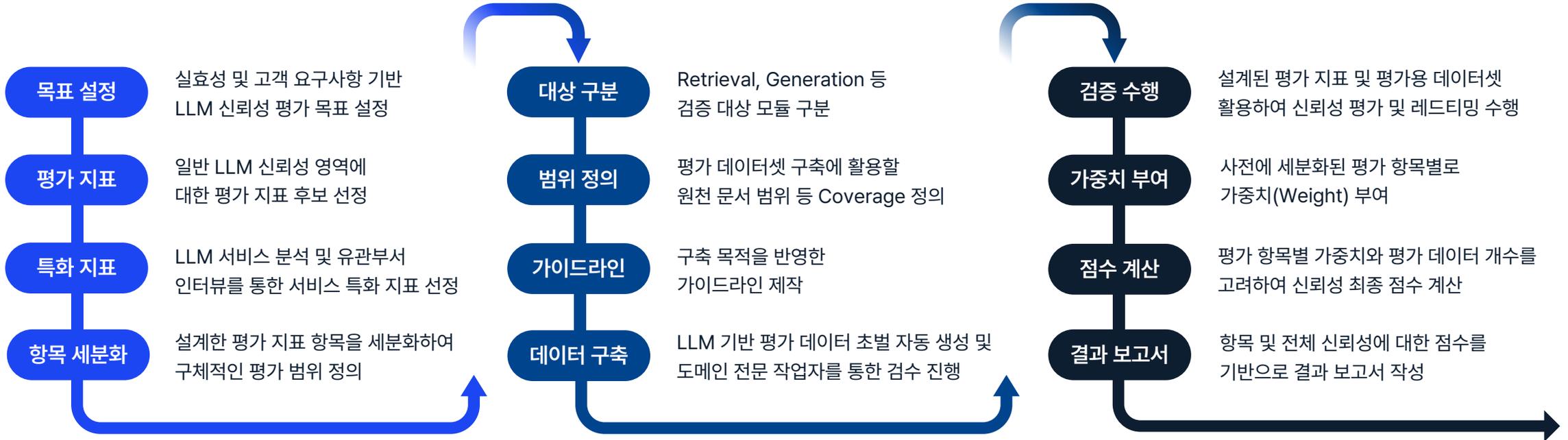
평가 지표 설계



데이터 구축



신뢰성 검증



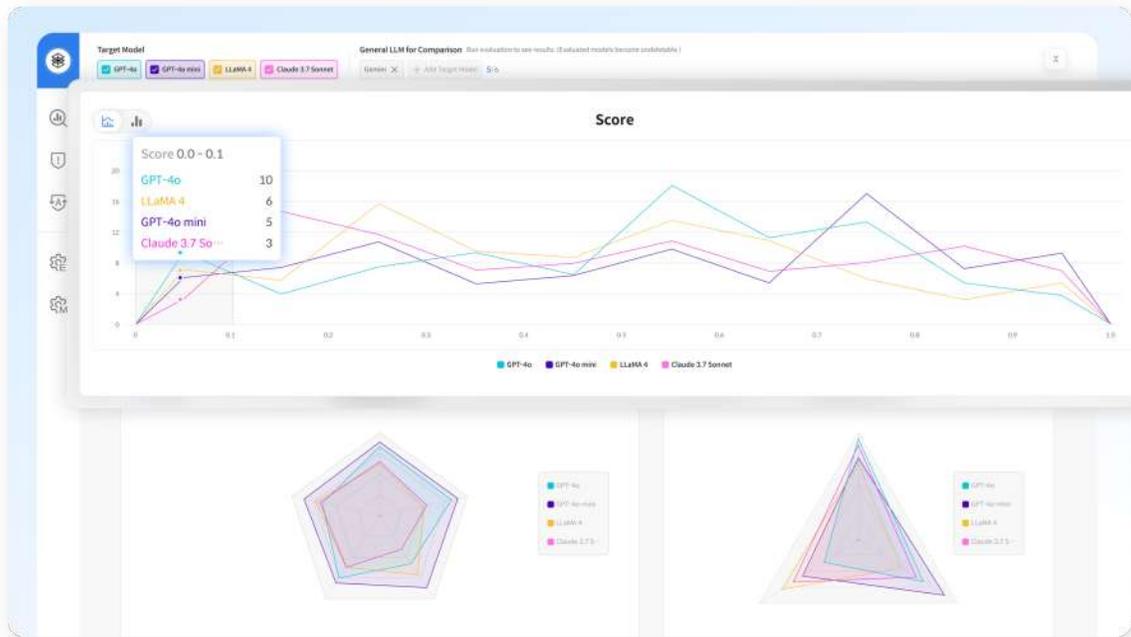
생성형 AI 기반 모델 및 서비스의 품질·안전성 기준을 설계하고 제3자 검증을 제공

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 이견한 기자] 셀렉트스타가 국내 최초로 생성형 AI 서비스 '신뢰성 검증 자동화 솔루션'을 출시했다고 20일 밝혔다. 다투모 이벨(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

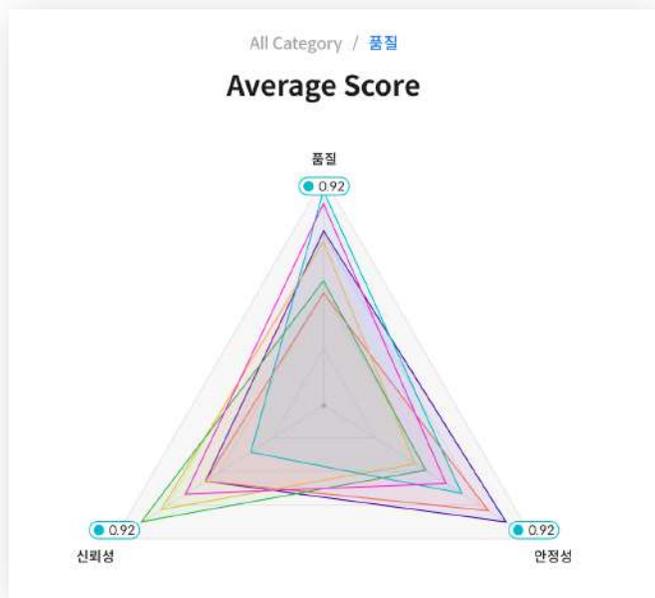
2. 고품질 평가 데이터셋 자동 생성 및 관리

- ✓ 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

Judgement



1. LLM-as-a-Judge (AI 자동 평가)

Model name 3	Claude	Model name 1	Model name 2	Model name 2	Model name 2
데이터를 완전히 명문화하지 말...	0.40	데이터 수명을 위해 주기적으로...	0.00		
여행 필수품: 여권, 티켓, 현금/카드...	0.20	여행 준비 시, 여권과 항공권은...	0.10		
최근 인기 있는 지역 유리는 간단...	0.50	간단한 저녁 요리로는 샐러드, 빵...	0.20		
효과적인 방법으로는 매일 열매...	0.90	열매 직송은 꾸준한 판매 학습과...	0.30		
스마트폰 추천 제품: 구글 네스트...	0.90	스마트폰을 위한 추천 제품으로...	0.60		
반려동물과 피 신비 관계를 밝히...	0.80	반려동물과 잘 지내기 위한 신...	0.30		
운동 루틴을 시작할 때는 골목 우...	0.90	운동 루틴은 현재 체력 상태에 맞...	0.20		
중고 가구를 살 때는 사기를 방지...	1.00	판매자는 물품의 상태를 정확히...	1.00		
비 오는 날에는 출퇴근 환승기...	0.40	비 오는 날에는 모자, DIY 프로젝트...	0.70		
면접에서 자주 묻는 질문: 자기소개...	0.40	면접에서 자주 묻는 질문은 본인...	0.70		
직업 공간을 바로 마련하고, 후속...	0.50	재택근무 시 작업 환경을 청결하...	0.90		
단어의 문법 기본기를 다진 후...	0.30	새 언어를 배우려면 먼저 알아듣...	0.80		
자원의 경쟁 거리의 사고 아역들...	0.50	중고차를 살 때는 사고 이력, 장...	0.70		
오트밀, 실론 쌀, 무지방 요거...	0.20	건강한 간식으로서는 곡물, 계소...	0.50		
우선순위를 설정하고 일정 관리...	1.00	효율적인 시간 관리를 위해 작업...	0.30		

2. Human Evaluation (인력을 통한 정성평가)

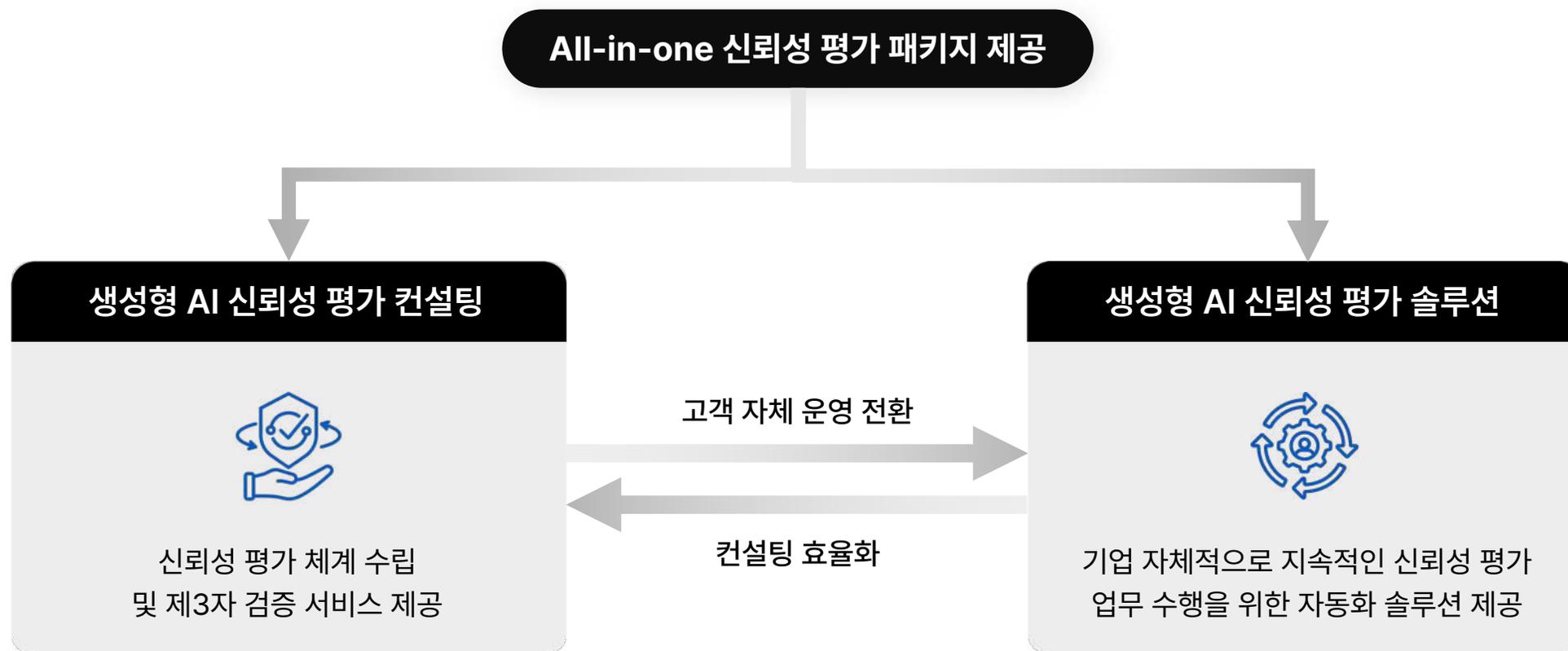
Financial QA Retrieval Benchmark

Leader Board

Rank	Target Model	NDCG@1	NDCG@3	NDCG@5
1	Word2Vec	1.0	0.9	0.8
2	Word2Vec	1.0	0.8	0.8
3	Word2Vec	0.9	0.7	0.6
4	Word2Vec	0.8	0.7	0.5
5	Word2Vec	0.8	0.4	0.4
6	Word2Vec	0.6	0.4	0.1
7	Word2Vec	0.5	0.3	0.3

3. Benchmark (정량평가)

국내 최초 1위 생성형 AI 신뢰성 전문 기업





SELECTSTAR

THE DATA-CENTRIC AI COMPANY

contact@selectstar.ai