

The logo for LIONBRIDGE, with the 'O' in LION underlined. The background of the top half of the page features a city skyline at night with various colored arrows (purple, blue, green, pink) pointing upwards, overlaid on a grid of dots that forms a wavy pattern across the bottom of the image.

LIONBRIDGE

AI 데이터 서비스

데이터 수집 및 주식 처리



라이온브리지(Lionbridge)는 고품질 데이터 수집과 주식 처리를 통한 AI 개발을 지원합니다. 당사의 글로벌 크라우드는 다양하고 문화적 맥락에 맞는 데이터세트를 제공함으로써 귀사의 모델이 다양한 언어, 지역, 실제 상황 전반에서 정확한 결과물을 생성할 수 있도록 합니다.



이유 및 대상

고품질 데이터 수집 및 주식 처리 서비스가 필요한 이유

AI 모델의 성능은 학습에 사용된 데이터의 품질에 따라 달라집니다. 많은 기업이 정확하고 신뢰할 수 있는 AI 시스템 개발에 필요한 대량의 다양한 고품질 데이터를 확보하는 데 어려움을 겪고 있습니다. 라이온브리지는 다음과 같은 지원을 제공합니다.

1

데이터 부족

대부분의 조직은 다양한 데이터셋을 필요한 만큼 대량으로 수집할 사내 자원이 부족합니다. 특히 특수 사례, 희소 지역, 희귀 언어와 같은 경우에 더욱 그렇습니다.

2

품질과 일관성

데이터 라벨링이 올바로 되어 있지 않거나 일관성이 없으면 모델이 편향되고 정확도가 떨어져 모델 도입에 실패하게 됩니다. 당사는 전문적인 주식 처리 워크플로와 엄격한 품질보증(QA)을 통해 첫날부터 정확성을 보장합니다.

3

다국어 및 멀티모달 수요

음성 보조장치부터 컴퓨터 비전까지, AI 시스템은 다양한 언어, 형식, 환경 전반에서 작동해야 합니다. 라이온브리지는 문화적 맥락에 맞는 대량의 다국어 및 멀티모달 데이터셋을 제공합니다.

4

속도 및 확장성

내부 데이터 팀을 구축하기 위해서는 시간과 비용이 많이 듭니다. 라이온브리지는 귀사의 모델이 생성 속도를 높일 수 있도록 빠른 확장, 글로벌 확대, 신속한 반복 주기 지원을 제공합니다.

5

휴먼인더루프(HITL) 통합

최고 수준의 모델조차 사람의 검증을 필요로 합니다. 당사는 사람이 프로세스에 직접 참여하여 학습 데이터를 검토함으로써 성능을 지속적으로 개선합니다.

데이터 수집 및 주석 처리 서비스의 주요 고객

당사의 서비스 지원팀은 다양한 업종에 걸쳐 AI를 구축하고 규모를 확장할 수 있도록 지원합니다. 일반적인 사용자는 다음과 같습니다.



기계학습 및 데이터 학습팀

AI 모델을 학습시키고 검증하고 미세조정하기 위해서는 대용량 고품질 데이터세트가 필요합니다.



AI 제품 및 엔지니어링 팀

음성 보조장치, 컴퓨터 비전 도구, 검색 알고리즘 등 기능을 개발하려면 정확히 라벨링된 데이터가 중요합니다.



NLP, CV, 멀티모달 AI 팀

개체 인식, 감정 분석, 대상 탐지, 전사 등 특정 작업에 맞춰 주석 처리된 전문 데이터(텍스트, 오디오, 이미지, 동영상)가 필요합니다.



의료, 금융, 자동차, 리테일 회사

AI를 진단, 사기 탐지, 자동화, 개인화 등에 활용하기 위해서는 각 분야에 특화된 정확한 데이터로 학습시켜야 합니다.



생성형 AI를 개발 중인 스타트업 및 기업

경쟁력 있는 AI 애플리케이션을 구축하기 위해서는 프롬프트 데이터, 합성 데이터, 미세 조정된 학습용 데이터세트, 사람의 평가를 거친 결과물이 필요합니다.

라이온브리지 서비스



데이터 주석 처리



데이터 수집



모델 평가

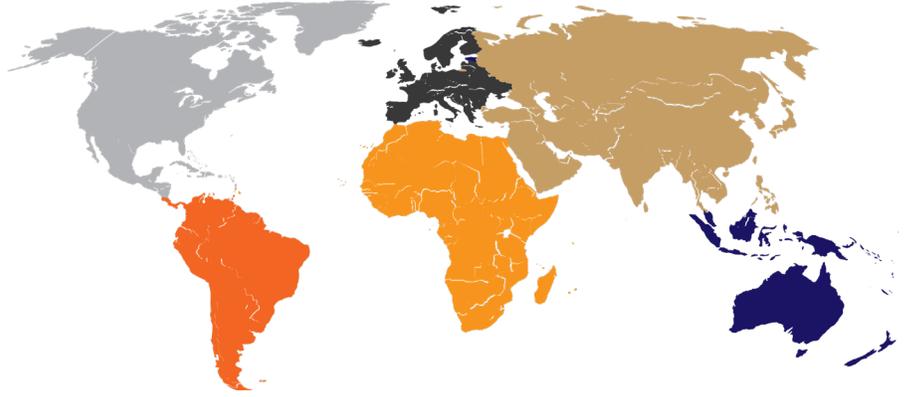
클라우드

라이온브리지 클라우드는 귀사의 경쟁력입니다

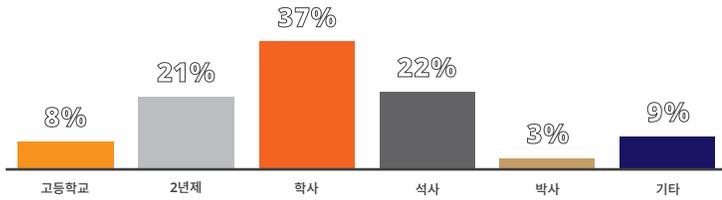
일반 사용자와 분야별 전문가 모두가 참여하는 라이온브리지의 글로벌 클라우드는 심층적인 데이터를 대량으로 제공합니다. 이 클라우드에서는 AI 모델 훈련이든 복잡한 분야별 데이터 주석 처리든, 해당 작업에 가장 적합한 기여자를 찾을 수 있습니다. 당사는 전담 프로젝트 매니저와 강력한 QA 워크플로를 통해 실제로 사용할 수 있고 모델 성능을 개선할 수 있는 문화적으로 적합한 고품질의 데이터를 제공합니다.

고도로 전문화되고 다양한 전문지식과 관심사를 갖춘 글로벌 커뮤니티의 필요성

- 북미: 26%
- 유럽: 29.5%
- 아시아: 27%
- 아프리카: 6.5%
- 남미: 8.5%
- 오세아니아/호주: 2.5%



최고 학력



성별



연령



60만 명
글로벌 커뮤니티



350+
언어 수



450+
지역

업종



직접 구축할 것인가, 전문가에게 맡길 것인가? 실제 사용 사례를 통해 직접 구축과 제대로 된 구축 시의 비용과 ROI 비교
사내 데이터 운영 팀을 갖추기 위해서는 인력과 도구를 확보하는 데 수백만 달러의 비용이 들며 시간도 많이 걸리지만 품질이나 확장성을 보장할 수 없습니다. 한 글로벌 인력 분석 기업은 30개 언어로 된 노동 시장 데이터를 주석 처리하는 데 각 언어당 1,200 시간씩 총 36,000시간이 걸렸습니다.

자체 처리 방식의 추정치:

- 언어전문가 비용: 약 180만 달러
- 채용 시간: 언어전문가 30명 고용 및 온보딩에 3~4개월 소요
- 프로젝트 기간: 순차적 진행 및 시차를 둔 납품으로 인해 6개월 이상 소요
- 학습 데이터 준비 기간: 모델 학습에 사용할 수 있는 데이터를 준비하는 데 총 8개월 이상 소요

라이온브리지 접근방식:

- 검증된 언어전문가 및 이중 언어 주석 처리자로 구성된 글로벌 온디맨드 인력 풀
- 라이온브리지 프로젝트 매니저(PM)가 다층적 품질보증(QA)을 포함해 전 과정 관리
- 수개월이 아니라 수주 내에 프로젝트 시작
- 12주 이내 훈련 데이터를 준비하여 모델 개발 기간을 수개월 단축
- 납품 속도를 40% 개선하여 자체적으로 처리했을 때에 비해 백만 달러 이상 비용 절감

주요 인사이트

- 데이터 정리 및 재작업 시간 단축으로 모델 배포 속도 30~50% 단축
- 글로벌 기여자 네트워크와 효율적인 인프라를 통해 내부 데이터 팀 구축 대비 50~60% 비용 절감
- 시간이 많이 걸리는 채용, 도구 구축 및 학습 단계를 제거함으로써 제작까지 걸리는 속도 3배 단축
- 내부 부담 완화로 사내 기계학습 팀이 데이터 정리 대신 모델 개발에 집중할 수 있음
- 결과물에 중점을 두고 비용 효율적인 방식으로 확장 가능한 AI 데이터 생성 가능. 추가 비용 없이 결과를 확인할 수 있음.

