

KI-DATENSERVICES

DATENZUSAMMENSTELLUNG
UND -ANNOTATION



Lionbridge treibt die KI-Entwicklung durch die Zusammenstellung und Annotation hochwertiger Daten voran. Unsere globale Community schafft diverse, kulturell relevante Datenmengen, damit KI-Modelle über Sprachen, Regionen und reale Szenarien hinweg wunschgemäß funktionieren.



DAS WER UND DAS WARUM

Warum Sie auf Zusammenstellung und Annotation hochwertiger Daten angewiesen sind

KI-Modelle sind nur so gut wie die Daten, mit denen sie trainiert werden. Viele Unternehmen haben Schwierigkeiten, Daten in der für die Entwicklung akkurater und zuverlässiger KI-Systeme erforderlichen Menge, Qualität und Diversität zusammenzustellen. Lionbridge bietet:

1

Datenvielfalt

Den meisten Unternehmen fehlen die internen Ressourcen, um die erforderlichen umfangreichen und diversen Datenmengen zusammenzustellen – insbesondere für sehr spezielle Einsatzszenarien, unterrepräsentierte Regionen oder Nischensprachen.

2

Qualität und Konsistenz

Schlecht gekennzeichnete oder inkonsistente Daten führen zu Bias und unzureichender Genauigkeit von Modellen sowie zum Fehlschlagen ihrer Bereitstellung. Unsere fachkundigen Annotationsworkflows und eine strenge Qualitätssicherung sorgen vom ersten Tag an für die erforderliche Akkuratessse.

3

Mehrsprachige und multimodale Anforderungen

Von Sprachassistenten bis hin zu Computer Vision müssen KI-Systeme sprach-, format- und umgebungsübergreifend funktionieren. Wir liefern kulturell relevante, mehrsprachige und multimodale Datenmengen in jeder benötigten Größe.

4

Schnelligkeit und Skalierbarkeit

Der Aufbau interner Datenteams dauert lang und ist kostspielig. Wir beschleunigen die Einsatzbereitschaft Ihrer Modelle durch schnelle Bereitstellung, globale Skalierung und Unterstützung besonders kurzer Iterationszyklen.

5

HITL(Human-in-the-Loop)-Integration

Selbst die besten Modelle müssen von Menschen validiert werden. Unsere Services umfassen von Menschen korrigierte Trainingsdaten, die der kontinuierlichen Verbesserung der Performance dienen.

Wer benötigt typischerweise Services für Datenzusammenstellung und -annotation?

Unsere Services unterstützen Teams unterschiedlicher Branchen, die KI-Lösungen entwickeln oder skalieren:



Teams für maschinelles Lernen und Data Science

Benötigen umfangreiche, hochwertige Datenmengen zum Trainieren, Validieren und Optimieren von Modellen.



Teams für KI-Produkte und -Entwicklung

Verlassen sich auf akkurat gelabelte Daten, um Features wie Sprachassistenten, Computer-Vision-Tools und Suchalgorithmen bereitzustellen.



Teams für NLP, CV und multimodale KI

Benötigen spezielle Daten (Text, Audio, Bild, Video), die für spezifische Aufgaben wie Eigennamenerkennung, Stimmungsanalyse, Objekterkennung oder Transkription annotiert wurden.



Unternehmen der Sektoren Gesundheit, Finanzen, Automobil und Einzelhandel

Nutzen KI für Diagnosen, Betrugserkennung, Automatisierung, Personalisierung und vieles mehr. Für all diese Aufgaben werden spezifische, akkurate Trainingsdaten benötigt.



Start-ups und Konzerne, die generative KI entwickeln

Sind auf Prompt-Daten, synthetische Eingabedaten, optimierte Trainingsdatenmengen und die Evaluierung der Ausgaben durch Menschen angewiesen, um wettbewerbsfähige KI-Anwendungen zu entwickeln.

● UNSERE SERVICES ●



Datenannotation



Datenzusammenstellung



Modellevaluierung

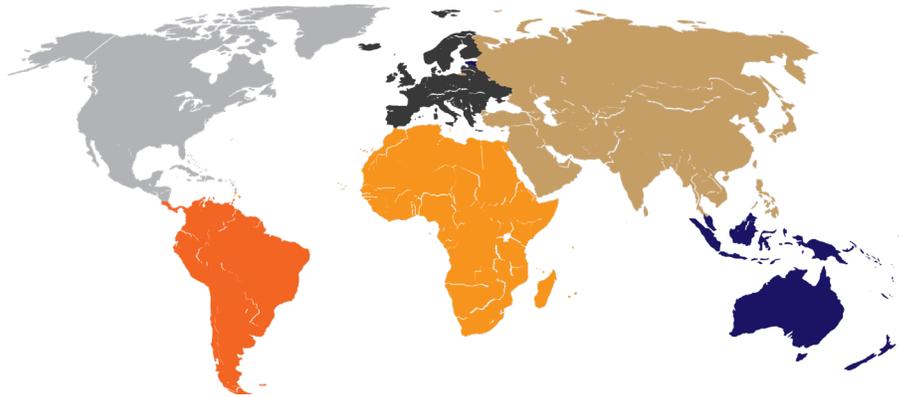
COMMUNITY

Unsere Community ist Ihr Wettbewerbsvorteil

Die globale Community von Lionbridge vereint alltägliche Benutzer mit Experten, also Breite und Tiefe. Wir finden die besten Kräfte für die jeweilige Aufgabe, vom Training eines KI-Modells bis zur Annotation komplexer themenspezifischer Daten. Mit engagierten Projektmanagern und robusten QA-Workflows stellen wir hochwertige, kulturell relevante Daten bereit, die auf den Praxiseinsatz abgestimmt sind und zu einer besseren Modellleistung beitragen.

Benötigt wird zudem eine globale Community mit spezialisierter und diverser Expertise und vielfältigen Interessen.

- Nordamerika: 26 %
- Europa: 29,5 %
- Asien: 27 %
- Afrika: 6,5 %
- Südamerika: 8,5 %
- Ozeanien/Australien: 2,5 %



Höchstes Bildungsniveau



Geschlecht



Alter



600 000
Mitglieder der globalen Community



350+
Sprachen



450+
Regionen

Branche



Entwickeln oder kaufen? Kosten im Überblick. Wertschöpfung in der Praxis.

Der Aufbau eines internen Datenteams kann Kosten in Millionenhöhe für Personal, Tools und Zeitaufwand verursachen – ohne Garantie für Qualität und Skalierbarkeit. Ein globales Unternehmen für Workforce Analytics musste Arbeitsmarktdaten in dreißig Sprachen annotieren. Der Aufwand betrug 1200 Arbeitsstunden pro Sprache und insgesamt 36 000 Stunden.

Schätzung für die interne Abwicklung:

- **Kosten für Sprachexperten:** ca. 1,8 Millionen USD
- **Einstellungsaufwand:** Einstellung von 30 Sprachexperten zuzüglich Onboarding, 3 bis 4 Monate erforderlich
- **Projektdauer:** voraussichtlich über 6 Monate mit Anlaufphase und Teillieferungen
- **Bereitstellung der Trainingsdaten:** voraussichtlich über 8 Monate, bevor die Daten zum Trainieren der Modelle verwendet werden können

Lionbridge als Partner:

- Weltweite On-Demand-Community qualifizierter Linguisten und zweisprachiger Annotationsexperten
- Umfassende Betreuung durch Lionbridge-Projektmanager, dazu mehrstufige Qualitätssicherung
- Projektstart innerhalb von Wochen statt Monaten
- Trainingsdaten in weniger als 12 Wochen verfügbar, Modellentwicklung dauert einige Monate weniger
- **40 % schnellere Lieferung und Einsparungen von über 1 Million USD im Vergleich zur internen Abwicklung**

WICHTIGE ERKENNTNISSE

- **30–50 % schnellere** Modellbereitstellung durch Reduzierung des Zeitaufwands für Datenbereinigung und -nachbearbeitung.
- **50–60 % geringere Kosten** im Vergleich zum Aufbau interner Datenteams, dank unserer globalen Community und unserer effizienten Infrastruktur.
- **Dreimal schneller** erreichte Einsatzbereitschaft durch Vermeidung langwieriger Einstellungsverfahren, schnelle Toolbereitstellung und kurze Trainingszyklen.
- **Weniger interner Aufwand**, sodass die ML-Teams sich auf die Modellentwicklung anstelle der Datenaufbereitung konzentrieren können.
- **Mit uns wird die Produktion von KI-Daten skalierbar**, erschwinglich und ergebnisorientiert. Sie erhalten Ergebnisse, keinen Mehraufwand.



ERFAHREN SIE MEHR AUF
LIONBRIDGE.COM