

AI-DATATJÄNSTER

INSAMLING OCH UPPMÄRKNING AV DATA



Lionbridge bidrar till AI-utvecklingen genom insamling och uppmärkning av kvalitetsdata. Vår globala crowd levererar mångfasetterade och kulturellt relevanta datauppsättningar som hjälper AI-modeller att producera korrekta resultat oberoende av språk, region och världsläget.



VEM OCH VARFÖR?

Varför behöver du insamling och uppmärkning av kvalitetsdata?

AI-modeller blir aldrig bättre än de data som de tränas med. Många företag har svårt att få tillgång till data med den omfattning, kvalitet och variation som krävs för att utveckla korrekta och tillförlitliga AI-system. Lionbridge kan vara till hjälp med:

1

Databrist

De flesta organisationer har inga egna resurser som kan samla in de enorma och mångfasetterade datauppsättningar som krävs. Det gäller fram allt specialområden, underrepresenterade regioner och sällsynta språk.

2

Kvalitet och konsekvens

Dåligt uppmärkta eller inkonsekventa data leder till modeller som genererar fördomsfullt och inkorrekt innehåll eller till att implementeringen fallerar helt. Våra kvalificerade arbetsflöden för uppmärkning och strikta kvalitetskontroller säkerställer hög precision från första dagen.

3

Behov av flerspråkiga och multimodala datauppsättningar

AI-system måste fungera oberoende av språk, format och miljö när det gäller allt från röstassistenter till datorseende. Vi levererar kulturellt relevanta, flerspråkiga och multimodala datauppsättningar i stor skala.

4

Snabbhet och skalbarhet

Det är tidskrävande och kostsamt att bygga upp egna datateam. Vi snabbar på arbetet med dina modeller genom effektiva startfaser, global uppskalning och stöd för korta uppdateringscykler.

5

Integrering med en människa i processen

Även förstklassiga modeller behöver valideras av en människa. Våra tjänster sluter cirkeln med träningsdata som granskats av en människa och leder till ständigt bättre resultat.

Vilka behöver vanligtvis tjänster för insamling och uppmärkning av data?

Våra tjänster används av team i olika branscher som arbetar med att utveckla eller skala upp AI-tjänster. Exempel på vanliga användare:



Team inom maskininlärning och datavetenskap

Behöver stora datauppsättningar av hög kvalitet för att träna, validera och finkalibrera AI-modeller.



AI-produkt- och teknikteam

Förlitar sig på korrekt uppmärkta data för att införa funktioner som röstassistenter, datorseendeverktyg och sökalgoritmer.



Team som arbetar med NLP, CV och multimodal AI

Behöver specialiserade data (text, ljud, bild, video) som märks upp för specifika uppgifter, till exempel identifiering av entiteter, sentimentanalys, objektidentifiering och transkribering.



Vård-, finans-, bil- och detaljhandelsföretag

Använder AI för diagnostik, automatisering, anpassning, upptäckt av bedrägerier med mera – områden som alla kräver korrekta och ämnesspecifika träningsdata.



Nya och etablerade företag som utvecklar GenAI

Är beroende av promptdata, syntetiska källdata, finkalibrerade träningsuppsättningar och resultat utvärderade av en människa för att utveckla konkurrenskraftiga AI-appar.

● VÅRA TJÄNSTER ●



Uppmärkning
av data



Datainsamling



Utvärdering
av modell

CROWD

Vår crowd är din konkurrensfördel

Lionbridges globala crowd rymmer såväl vanliga användare som ämnesexperter – så att du får både djup och volym. Oavsett om du tränar en AI-modell eller ska märka upp data inom ett visst ämnesområde, väljer vi ut de mest lämpliga bidragsgivarna för uppgiften. Vår crowd backas upp av engagerade projektledare och robusta arbetsflöden för kvalitetskontroll och vi ser till att du får kulturellt relevanta kvalitetsdata som stämmer överens med den faktiska användningen och ger bättre resultat från din modell.

För att uppnå det krävs en global community som tillsammans har bred och högkvalificerad kompetens och kunskaper inom ett stort antal intresseområden.

USA och Kanada: 26 %

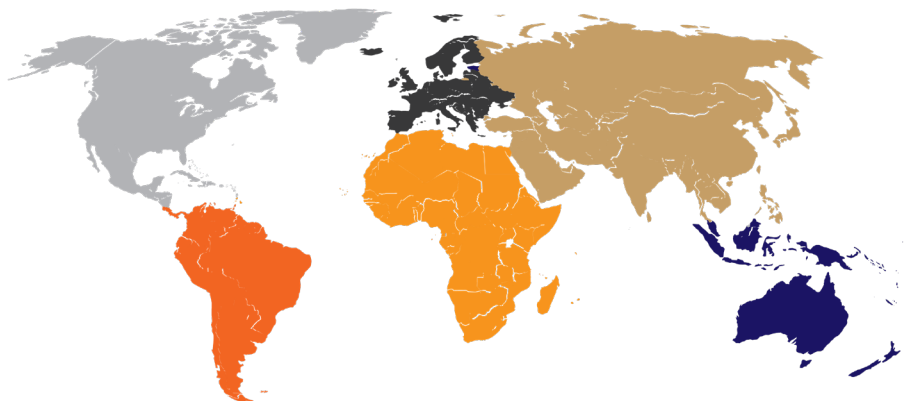
Europa: 29,5 %

Asien: 27 %

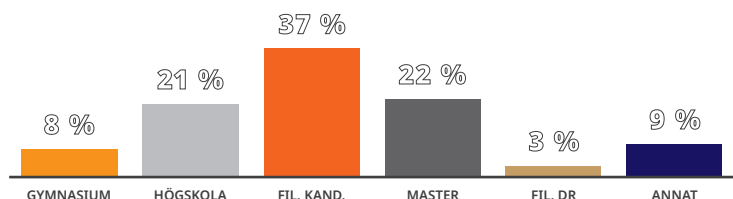
Afrika: 6,5 %

Sydamerika: 8,5 %

Södra Stillahavsområdet/
Australien: 2,5 %



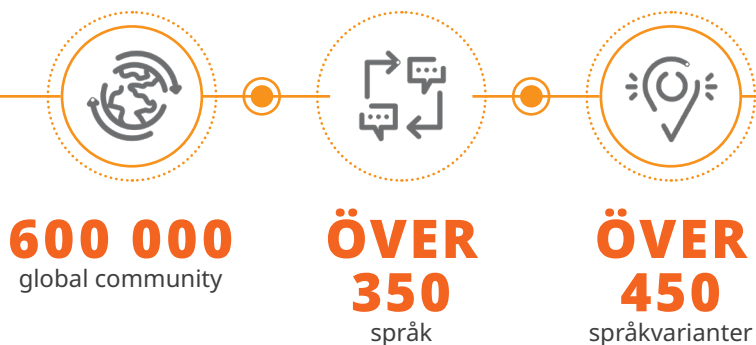
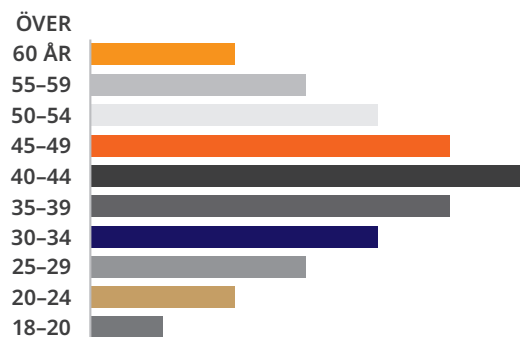
Högsta avslutade utbildning



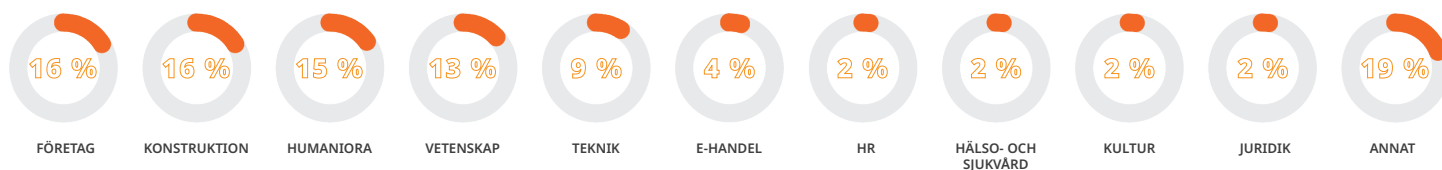
Kön



Ålder



Branscher



Bygga eller köpa? Göra själv eller göra rätt. Verkligt användningscase om avkastning.

Att bygga upp ett eget datateam kan kosta tiotals miljoner kronor och kräva enorma insatser i form av tid, verktyg och arbetstimmar – utan någon garanti för att vare sig kvalitet eller omfattning blir som du har tänkt dig. Ett globalt personalanalysföretag behövde märka upp arbetsmarknadsdata på 30 språk – ett projekt som beräknades ta 1 200 arbetstimmar per språk, dvs. totalt 36 000 timmars uppmärkning.

Beräkning av internt projekt:

- **Kostnad för språkexperter:** ca 1,8 miljon dollar
- **Rekryteringstid:** Rekrytera och introducera 30 språkexperter under 3–4 månader
- **Projekttid:** ca 6 månader på grund av uppstart och delfleveranser
- **Färdiga träningsdata:** ca 8 månader totalt innan data kan användas för att träna modeller

Lionbridges metod:

- Global crowd på begäran med kvalificerade språkexperter och tvåspråkiga uppmärkare
- Hela projektet hanteras av Lionbridges projektledare med kvalitetskontroll i flera nivåer
- Projektet lanseras inom några veckor i stället för månader
- Träningsdata klara på mindre än 12 veckor, vilket gör att modellutvecklingen kan gå flera månader snabbare
- **40 % snabbare leverans och besparing på över 1 miljon dollar jämfört med internt projekt**

VIKTIGA LÄRDOMAR

- **30–50 % snabbare** driftsättning av modell genom att lägga mindre tid på rensning av data och omarbeten
- **50–60 % lägre kostnad** jämfört med att bygga upp ett eget datateam, tack vare vårt globala crowdnätverk och en effektiv infrastruktur
- **3 gånger kortare** tid till produktion genom att undvika tidskrävande rekryterings-, verktygs- och utbildningscykler
- **Mindre intern arbetsbelastning**, så att maskininlärningsteam kan fokusera på att utveckla modellen i stället för att hantera data
- **Vi gör AI-dataproduktion skalbar**, prisvärd och resultatinriktad. Du får resultat i stället för kostnader.

