



CASE STUDY SULLA RACCOLTA DATI AI

RACCOLTA DI 28.000 PUNTI DATI IN UNA SETTIMANA

1 SETTIMANA
PER LA CONSEGNA

8.000
SET DI DATI AGGIUNTIVI
CONSEGNATI

28.000
IMMAGINI CONSEGNATE

INFORMAZIONI SUL CLIENTE

Lionbridge AI™ ha supportato un cliente che aiuta a mettere in contatto i brand con i professionisti creativi di cui hanno bisogno per i loro progetti.

Questa azienda gestisce una piattaforma innovativa che mette in contatto talenti creativi e brand alla ricerca di competenze creative per i propri progetti. I professionisti creativi possono caricare i loro portfolio e consentirne la ricerca in base a diversi criteri. I brand possono pubblicare offerte di lavoro o brief. Nel complesso, questo cliente di Lionbridge agevola le connessioni e la creazione di reti lavorative tra brand e talenti creativi. Le parti possono entrare in contatto sulla piattaforma e instaurare relazioni per progetti presenti e futuri.

LA SFIDA

Il cliente di Lionbridge aveva bisogno di un volume elevato di dati in tempi molto stretti per addestrare il proprio modello AI per l'identificazione degli oggetti. Ha richiesto circa 20.000 immagini nel giro di pochi giorni. Il cliente necessitava di dati umani di alta qualità, non di immagini sintetiche (che possono essere di qualità molto inferiore e portare a prestazioni del modello sensibilmente peggiori). In particolare, il cliente aveva bisogno di immagini di oggetti contenuti all'interno di altri oggetti. Queste immagini servivano per aiutare il modello AI a imparare a identificare quando un'immagine mostra oggetti inseriti in altri oggetti, come un cestino, una borsa, una ciotola e così via. È stato richiesto un set di quattro immagini:

- Ogni oggetto separatamente
- Un oggetto dentro l'altro
- Lo sfondo vuoto senza oggetti

È inoltre importante sottolineare che per il cliente era importante ricevere dati con prospettive diverse. Ottenere 20.000 immagini degli stessi oggetti non sarebbe stato altrettanto efficace per addestrare il modello quanto ottenere 20.000 immagini differenti. Coinvolgere partecipanti provenienti da contesti demografici differenti di tutto il mondo aiuta infatti il modello a comprendere meglio i destinatari e a offrire loro prestazioni ottimali.

Persone di paesi, età e generi diversi tendono a caricare sul sito del cliente molte tipologie differenti di immagini. Addestrando il modello con un set di dati più diversificato, è possibile ottenere prestazioni migliori per un pubblico più vasto.



DATI STATISTICI SUI NOSTRI PROFESSIONISTI IN CROWDSOURCING



Lingue parlate: **oltre 350**



Provenienza da **oltre 450** paesi in tutto il mondo:

- America del Nord: **26%**
- Europa: **29,5%**
- Asia: **27%**
- Africa: **6,5%**
- America del Sud: **8,5%**
- Oceania/Australia: **2,5%**



- Donne: **52%**
- Uomini: **48%**



Fascia di età da **18 a oltre 60**

Settore:

- **16%** economia
- **16%** ingegneria
- **15%** discipline umanistiche
- **13%** scienza
- **9%** tecnico
- **4%** eCommerce
- **2%** risorse umane
- **2%** sanità
- **2%** arti figurative
- **2%** legale
- **10%** altro

LA SOLUZIONE

Per supportare questo cliente, abbiamo usato la piattaforma Lionbridge Aurora AI Studio™. La piattaforma ci mette in contatto con una community di oltre mezzo milione di esperti in diverse materie in tutto il mondo. Queste persone assicurano diversità per quanto riguarda il luogo, le lingue parlate, l'età, il genere e i settori di specializzazione. Coinvolgiamo questa community per svolgere numerose attività, inclusa la raccolta dati, con tempi di esecuzione estremamente rapidi.

Per questo progetto, abbiamo configurato l'attività per il cliente in Aurora AI Studio. In poche ore, abbiamo coinvolto oltre 2.000 partecipanti. Ai partecipanti sono state concesse otto ore per completare l'attività di raccolta dati. Parallelamente, abbiamo creato un team interno di controllo qualità più ampio, per rivedere l'elevato volume di dati nel rispetto delle tempistiche particolarmente ristrette del progetto. Questo team ha esaminato i dati raccolti

per garantirne l'accuratezza e la conformità alle esigenze del cliente. Al termine della revisione qualitativa, abbiamo assegnato i metadati a ciascuna immagine, facendo risparmiare al cliente molte ore di lavoro.

Lionbridge ha consegnato al cliente i dati raccolti in tre batch. Il primo è stato consegnato in anticipo, per consentire al cliente di iniziare l'addestramento del modello il prima possibile. Questo primo batch comprendeva 12.000 immagini, consegnate in tre giorni. Il secondo batch, consegnato in cinque giorni, comprendeva le restanti 8.000 immagini.

Il cliente è rimasto così soddisfatto di questi set di dati di addestramento dell'AI da richiedere ulteriori 8.000 immagini in altri due giorni. Abbiamo accolto volentieri la richiesta e consegnato complessivamente 28.000 immagini al cliente, il tutto nell'arco di una sola settimana. Il batch inizialmente richiesto di 20.000 immagini è stato interamente consegnato in soli quattro giorni.

METODOLOGIA

Per supportare il cliente, Lionbridge ha usato un processo articolato in più fasi:

- Fase 1:** creazione dell'attività in Aurora AI Studio
- Fase 2:** coinvolgimento dei partecipanti
- Fase 3:** ricezione dei dati raccolti
- Fase 4:** revisione dei dati e assegnazione dei metadati
- Fase 5:** consegna dei dati raccolti al cliente
- Fase 6:** raccolta e consegna di ulteriori dati, su richiesta del cliente





TEAM MULTIDISCIPLINARE

Program Director | Technical AI Engineer | QA Lead | Project Manager | 10 QA Specialist interni



— I RISULTATI

Lionbridge AI ha aiutato il cliente in tre modi principali. In primo luogo, grazie alla nostra piattaforma Aurora AI Studio e a una community estremamente diversificata, abbiamo raccolto un volume enorme di dati di alta qualità, acquisiti da persone reali e in grado di riflettere prospettive e immagini provenienti da tutto il mondo. Abbiamo permesso al cliente di evitare le conseguenze legate all'uso esclusivo di dati sintetici, che possono portare a un addestramento non adeguato del modello. Lionbridge ha inoltre coinvolto un team di controllo qualità ampliato per verificare i dati prima della consegna. Questi dati di altissima qualità aiuteranno ad addestrare il modello affinché supporti efficacemente il brand e gli utenti creativi, indipendentemente dalla loro provenienza, dalle caratteristiche demografiche o dal contenuto dei loro portfolio.

Il secondo modo in cui abbiamo supportato il cliente è stato consegnare questo volume estremamente elevato di dati (addirittura superiore a quanto inizialmente richiesto) nell'arco di una sola settimana. Tempi di consegna rapidi sono fondamentali per l'addestramento dei modelli: ogni giorno perso può risultare estremamente costoso. Grazie all'uso della piattaforma Aurora AI Studio per completare il progetto in tempi così rapidi, Lionbridge ha consentito al cliente di risparmiare migliaia di dollari in costi di manodopera e opportunità di business perse. In questo modo, il cliente ha potuto addestrare in modo adeguato il modello della sua piattaforma in appena una settimana.

Infine, abbiamo superato le aspettative del cliente andando oltre la richiesta iniziale. Sebbene il cliente avesse inizialmente richiesto 20.000 immagini, ne abbiamo raccolte 28.000. Lionbridge è orgogliosa di aver fornito non solo i dati di cui il cliente aveva bisogno, ma anche quelli che non sapeva ancora di volere. Riuscire a offrire al cliente tutto ciò che desiderava, e anche di più, in una sola settimana è stato un notevole successo.



PER SAPERNE DI PIÙ:

LIONBRIDGE.COM