



ÉTUDE DE CAS SUR LA COLLECTE DE DONNÉES IA

# COLLECTE DE 28 000 POINTS DE DONNÉES EN UNE SEMAINE

1 SEMAINE  
DE DÉLAI

8 000  
ENSEMBLES DE DONNÉES  
SUPPLÉMENTAIRES LIVRÉS

28 000  
IMAGES LIVRÉES

## LE DÉFI

Ce client de Lionbridge avait besoin d'un volume important de données dans un délai serré, pour entraîner son modèle d'IA à identifier des objets. Il avait demandé environ 20 000 images en quelques jours. Le client avait besoin de données humaines de haute qualité, et pas d'images synthétiques (qui peuvent être de moins bonne qualité et donner lieu à des performances du modèle nettement inférieures). En particulier, le client avait besoin d'images d'objets dans des objets, pour aider le modèle d'IA à apprendre à identifier les cas où une image montre des objets dans d'autres objets, comme un panier, un sac, un bol, etc. Le client demandait quatre images dans un ensemble :

- Chaque objet séparément
- Un objet dans l'autre
- L'arrière-plan vide sans objets

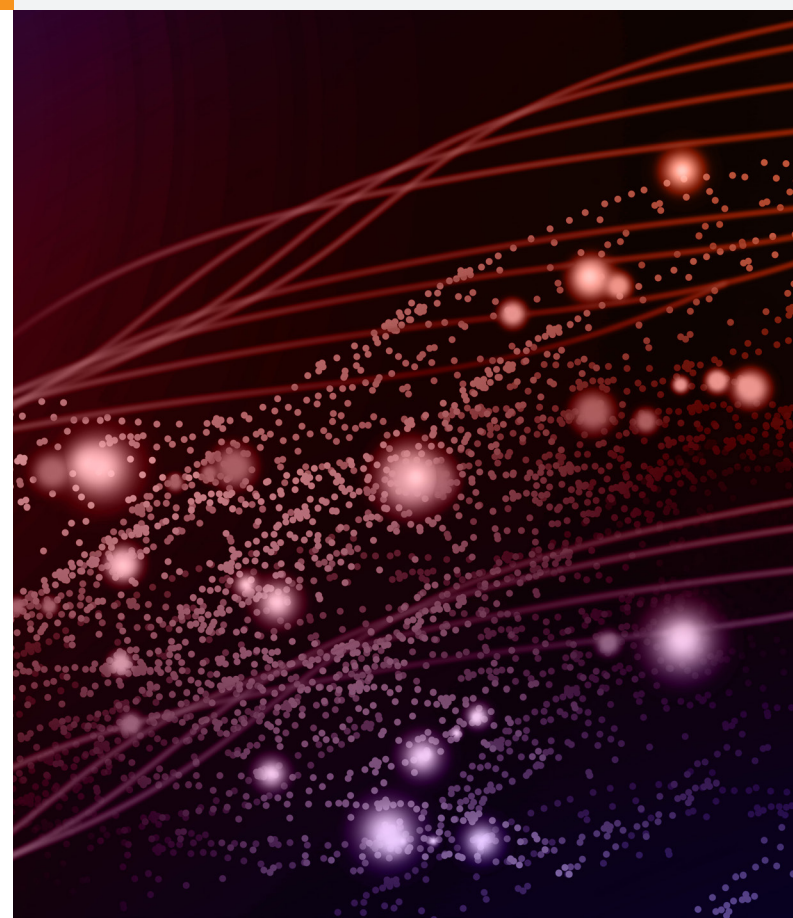
Il faut également noter que le client donnait la priorité aux données issues de différents points de vue. 20 000 images des mêmes objets n'auraient pas entraîné son modèle aussi efficacement que 20 000 images différentes. Des images de participants de nombreux groupes démographiques dans le monde entier allaient aider le modèle à comprendre les publics de tous les groupes démographiques et à livrer de bonnes performances.

Les personnes de différents pays, tranches d'âges, sexes, etc. sont susceptibles de fournir de nombreux types d'images différents sur le site Web du client. Un ensemble de données plus diversifié pour entraîner le modèle du site Web favoriserait de bonnes performances pour une base de clients plus étendue.

## À PROPOS DU CLIENT

Lionbridge AI™ a soutenu un client qui aide à mettre en relation les marques et les professionnels de la création dont elles ont besoin pour leurs projets.

Cette entreprise propose une plateforme innovante qui met en relation des talents créatifs et des marques qui ont besoin de faire appel à eux pour leurs projets. Le talent créatif peut publier son portfolio et activer la recherche selon différents facteurs. Les marques peuvent publier des projets ou des présentations. De manière générale, ce client de Lionbridge facilite les liens et le travail en réseau entre les marques et les talents créatifs. Ils peuvent prendre contact sur la plateforme et nouer des relations pour des projets actuels et futurs.



## QUELQUES STATISTIQUES SUR NOTRE COMMUNAUTÉ



Parlent **plus de 350** langues



Habitent dans **plus de 450** régions  
différentes dans le monde :

- Amérique du Nord : **26 %**
- Europe : **29,5 %**
- Asie : **27 %**
- Afrique : **6,5 %**
- Amérique du Sud : **8,5 %**
- Océanie/Australie : **2,5 %**



- Femmes : **52 %**
- Hommes : **48 %**



Âgés de **18 à plus de 60 ans**

Ils travaillent dans les domaines  
suivants :

- **16 %** Commerce
- **16 %** Ingénierie
- **15 %** Lettres
- **13 %** Science
- **9 %** Technique
- **4 %** Commerce électronique
- **2 %** RH
- **2 %** Santé
- **2 %** Beaux-arts
- **2 %** Juridique
- **10 %** Autres



## LA SOLUTION

Nous avons utilisé notre plateforme **Lionbridge Aurora AI Studio™** pour servir ce client. Cette plateforme nous relie à une communauté d'un demi-million d'experts du monde entier, diversifiés de par leur zone géographique, leur langue, leur âge, leur sexe et le secteur dans lequel ils travaillent. Nous faisons appel à eux pour effectuer de nombreuses tâches, y compris la collecte de données, à des vitesses record.

Nous avons configuré la tâche dans Aurora AI Studio pour ce projet. En quelques heures seulement, nous avons déjà engagé plus de 2 000 participants. Les participants disposaient de huit heures pour effectuer la tâche de collecte de données. En parallèle, nous avons mis sur pied une équipe d'assurance qualité interne chargée de réviser les données. Cette équipe, de grande envergure en raison du volume et des délais serrés du projet, a vérifié si les données collectées étaient exactes et respectaient

les besoins du client. Après cette révision de la qualité, nous avons attribué un nom aux métadonnées de chaque image, économisant ainsi plusieurs heures de travail au client.

Lionbridge a livré les données collectées au client en trois lots. Nous avons livré le premier ensemble rapidement, afin que le client puisse commencer à entraîner son modèle aussi rapidement que possible. Ce lot se composait de 12 000 images sur trois jours. Le deuxième lot, livré en cinq jours, comportait les 8 000 images restantes.

Le client a été tellement satisfait de ces lots de données collectées pour l'entraînement de l'IA qu'il a demandé 8 000 images supplémentaires en deux jours. Nous avons été ravis de livrer 28 000 images au client en seulement une semaine. Le lot de 20 000 images initialement demandé a été livré en seulement quatre jours.



## MÉTHODOLOGIE

Lionbridge a utilisé un processus en plusieurs étapes pour aider le client :

Étape 1 : Créer la tâche dans Aurora AI Studio

Étape 2 : Engager les participants

Étape 3 : Accepter les données collectées

Étape 4 : Réviser les données et attribuer un nom aux métadonnées

Étape 5 : Livrer les données collectées au client

Étape 6 : Collecter et livrer plus de données, à la demande du client







## ÉQUIPE PLURIDISCIPLINAIRE

Directeur de programme | Ingénieur en IA technique | Responsable de l'AQ | Chef de projet |  
10 spécialistes en AQ internes



## LES RÉSULTATS

**Lionbridge AI a aidé ce client de trois façons essentielles.** Tout d'abord, grâce à notre plateforme Aurora AI Studio, nous avons amassé une quantité colossale de données de haute qualité, collectées par des humains, qui reflétaient des points de vue et images du monde entier. Nous avons ainsi fait en sorte que le client évite les problèmes liés à l'utilisation de données synthétiques uniquement, pouvant mener à un entraînement incorrect du modèle. Lionbridge a même fait appel à une grande équipe d'assurance qualité pour vérifier les données avant la livraison. Ces données irréprochables aideront à entraîner le modèle du client afin de soutenir sa marque et les talents créatifs, quels que soient le lieu, la tranche démographique ou le portfolio de ces talents.

Nous avons également aidé notre client en livrant ce volume de données titanesque (et même supérieur à ce que le client avait initialement demandé) en une semaine. Des délais rapides sont essentiels lors de l'entraînement de modèles : chaque journée perdue peut coûter très cher. Lionbridge a permis au client d'économiser des milliers de dollars en main-d'œuvre et en perte d'activité, et a utilisé sa plateforme Aurora AI Studio pour terminer le projet extrêmement rapidement. Le client a ainsi pu s'assurer que le modèle de sa plateforme était correctement entraîné en seulement une semaine.

Enfin, nous avons ébloui notre client en dépassant ses exigences initiales. Il avait demandé 20 000 images à l'origine, mais nous en avons livré 28 000. Nous avons été fiers de fournir non seulement les données dont le client avait besoin, mais aussi celles dont il ignorait avoir besoin. Nous considérons que le fait de livrer au client tout ce qu'il attendait, et plus encore, en seulement une semaine, fut une réussite considérable.

POUR EN SAVOIR PLUS, RENDEZ-VOUS SUR  
**LIONBRIDGE.COM/FR**