



# SERVICES DE DONNÉES D'IA

## COLLECTE ET ANNOTATION DE DONNÉES



Lionbridge soutient le développement de l'IA grâce à des services de collecte et d'annotation de données de haute qualité. Notre équipe internationale fournit des corpus de données diversifiés et pertinents culturellement, ce qui aide les modèles à livrer de bonnes performances dans les différents scénarios concrets, langues et régions.



## QUI ET POURQUOI

### Pourquoi avez-vous besoin de services de collecte et d'annotation de données de haute qualité ?

La qualité des modèles d'IA dépend des données avec lesquelles ils sont entraînés. La plupart des entreprises ont des difficultés à accéder à l'échelle, à la qualité et à la diversité des données nécessaires pour développer des systèmes d'IA fiables et précis. Lionbridge peut apporter son aide dans les domaines suivants :

1

#### Rareté des données

La plupart des organisations ne disposent pas des ressources internes nécessaires pour collecter les corpus de données vastes et diversifiés, en particulier pour les cas spécifiques, les régions sous-représentées ou les langues rares.

2

#### Qualité et cohérence

Les données mal étiquetées ou incohérentes donnent lieu à la présence de biais dans le modèle, à une mauvaise précision et à un échec des déploiements. Nos processus experts d'annotation et notre AQ rigoureuse garantissent la précision dès le premier jour.

3

#### Exigences multilingues et multimodales

Les systèmes d'IA, depuis les assistants vocaux jusqu'à la vision par ordinateur, doivent fonctionner dans toutes les langues, tous les formats et tous les environnements. Nous fournissons des corpus de données culturellement pertinents, multilingues et multimodaux à grande échelle.

4

#### Vitesse et évolutivité

La création d'équipes de données internes est lente et coûteuse. Nous accélérons la mise en production de vos modèles grâce à une montée en puissance agile, une expansion mondiale et la prise en charge de cycles d'itération rapides.

5

#### Intégration d'une intervention humaine

Même les meilleurs modèles requièrent une validation humaine. Nos services bouclent la boucle avec des données d'entraînement révisées par des humains, pour améliorer les performances en continu.

## Qui a besoin de services de collecte et d'annotation des données en général ?

Nos services aident des équipes de différents secteurs qui créent l'IA ou la font évoluer. Nos principaux utilisateurs sont les suivants :



### Équipes dédiées à l'apprentissage automatique et à la science des données

Elles ont besoin de corpus de données volumineux et de haute qualité pour entraîner, valider et affiner les modèles.



### Équipes de produit et d'ingénierie de l'IA

Elles doivent s'appuyer sur des données étiquetées avec précision pour lancer des fonctionnalités, comme des assistants vocaux, des outils de vision par ordinateur et des algorithmes de recherche.



### Équipes de programmation neurolinguistique (PNL), de CV et d'IA multimodale

Elles ont besoin de données spécialisées (texte, audio, image, vidéo) annotées pour des tâches spécifiques, comme la reconnaissance d'entité, l'analyse des sentiments, la détection d'objets ou la transcription.



### Entreprises de soins de santé, financières, automobiles et de vente au détail

Elles utilisent l'IA pour les diagnostics, la détection des fraudes, l'automatisation, la personnalisation et plus encore. Chaque entreprise a besoin de données d'entraînement précises et spécifiques à leur domaine.



### Start-ups et entreprises qui développent une IA générative

Elles ont besoin de données d'invite, d'entrées synthétiques, de corpus d'entraînement affinés et de résultats évalués par des humains pour créer des applications d'IA concurrentielles.

## NOS SERVICES



Annotation  
de données



Collecte  
de données



Évaluation  
de modèle

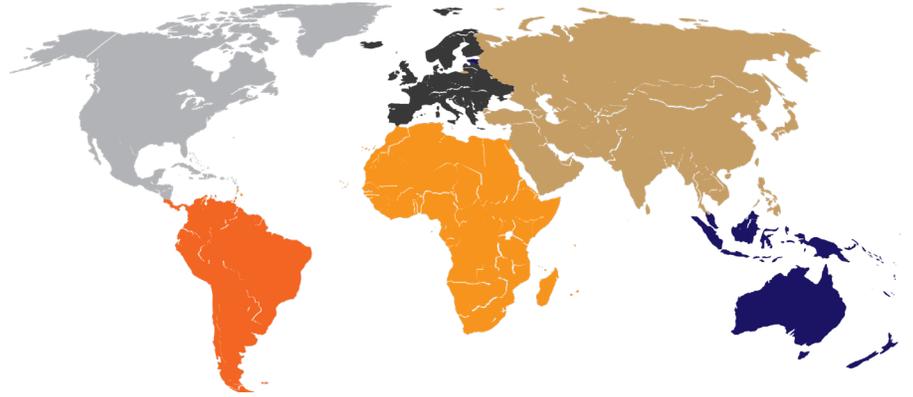
# ÉQUIPE

## Notre équipe constitue votre avantage concurrentiel

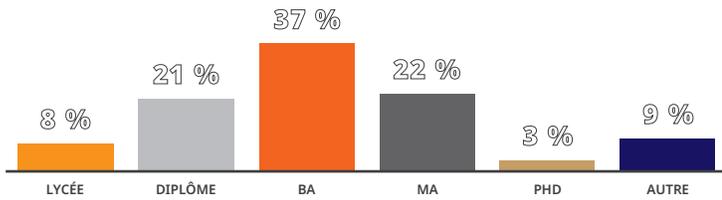
L'équipe internationale de Lionbridge regroupe des utilisateurs du quotidien et des experts métiers, pour vous offrir à la fois un vaste champ d'action et une profondeur d'analyse. Que vous entraîniez un modèle d'IA ou annotiez des données complexes spécifiques à un domaine, nous trouvons les meilleurs contributeurs pour chaque tâche. Soutenus par des chefs de projet dédiés et par des processus d'AQ solides, nous fournissons des données de haute qualité et culturellement pertinentes, qui correspondent aux usages concrets et améliorent les performances des modèles.

**La communauté internationale doit être hautement spécialisée et posséder un savoir-faire et des centres d'intérêt diversifiés.**

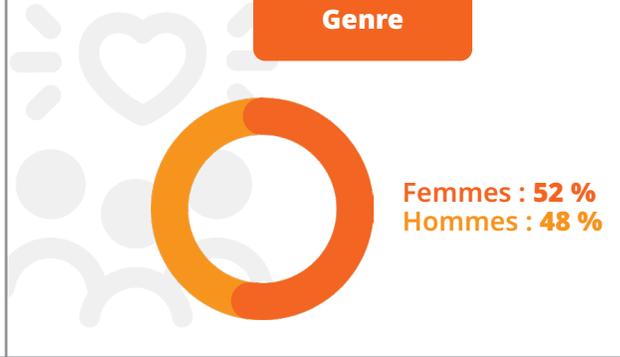
- Amérique du Nord : 26 %
- Europe : 29,5 %
- Asie : 27 %
- Afrique : 6,5 %
- Amérique du Sud : 8,5 %
- Océanie/Australie : 2,5 %



### Niveau d'études le plus élevé



### Genre



### Âge



**600 000**  
Communauté internationale



**350+**  
Langues



**450+**  
Sites

### Secteur d'activité



### Créer ou acheter ? Le coût d'un projet interne comparé au meilleur choix. Cas d'utilisation ROI concret.

Mettre sur pied une équipe d'opérations de données interne peut coûter des millions d'euros en effectifs, en outils et en temps, sans garantie de qualité ni d'échelle. Une société internationale d'analyses sur les effectifs avait besoin d'annoter les données du marché de la main-d'œuvre pour 30 langues, avec 1 200 heures de travail par langue, soit un total de 36 000 heures d'annotation.

#### Estimation de l'approche interne :

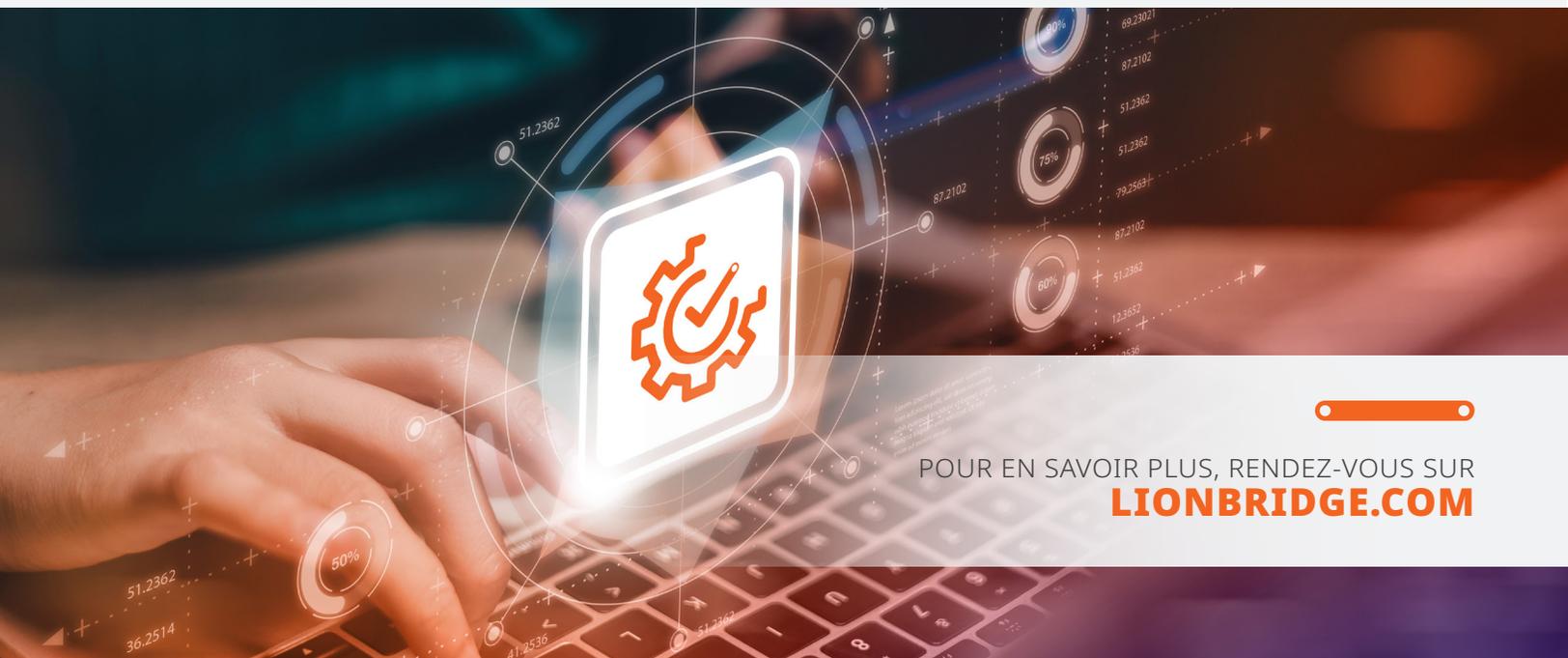
- **Coûts liés aux linguistes** : ~ 1,8 million \$
- **Temps dédié au recrutement** : Recrutement de 30 linguistes + intégration, 3 à 4 mois
- **Durée du projet** : ~ 6 mois et plus en raison du temps d'apprentissage et des livraisons échelonnées
- **Préparation des données d'entraînement** : ~ 8 mois et plus au total avant que les données ne puissent être utilisées pour entraîner des modèles

#### L'approche de Lionbridge :

- Équipe internationale à la demande de linguistes qualifiés et d'annotateurs bilingues
- Projet entièrement géré par les chefs de projet de Lionbridge avec une AQ en plusieurs étapes
- Lancement du projet en quelques semaines au lieu de plusieurs mois
- Données d'entraînement prêtes en moins de 12 semaines, diminuant de plusieurs mois le délai de développement du modèle
- **Livraison 40 % plus rapide, avec plus de 1 M \$ d'économies par rapport à l'exécution du projet en interne**

### INFORMATIONS CLÉS

- **30 à 50 % de réduction** du délai de déploiement du modèle, en réduisant le temps consacré au nettoyage des données et aux retouches
- **50 à 60 % de réduction des coûts** par rapport à la création d'une équipe de données en interne, grâce à notre réseau international de contributeurs et à notre infrastructure efficace
- **Réduction par 3** du délai de production, en évitant les longues démarches d'embauche, l'acquisition d'outils et les cycles d'entraînement
- **Réduction de la charge interne**, ce qui laisse à vos équipes de ML le temps nécessaire pour se concentrer sur le développement de modèles et non sur le traitement des données
- **Nous rendons la production de données d'IA évolutive**, abordable et axée sur les résultats. Vous obtiendrez des résultats, et non des frais supplémentaires.



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