



# SERVICIOS DE DATOS DE IA

## RECOPIACIÓN Y ANOTACIÓN DE DATOS



Lionbridge impulsa el desarrollo de la IA con servicios de recopilación y anotación de datos de alta calidad. Nuestro equipo global proporciona conjuntos de datos diversos y culturalmente relevantes, lo que permite a los modelos ofrecer resultados precisos en múltiples idiomas, regiones y contextos del mundo real.



## **POR QUÉ Y QUIÉNES**

### **¿Por qué necesita una recopilación y anotación de datos de alta calidad?**

La calidad de los modelos de IA depende directamente de la calidad de los datos con los que se entrenan. Muchas empresas enfrentan grandes dificultades para obtener el volumen, la calidad y la diversidad de datos necesarias para desarrollar sistemas de IA precisos y fiables. Lionbridge puede ayudar con aspectos como los siguientes:

1

#### **Escasez de datos**

La mayoría de las organizaciones carecen de recursos internos que se ocupen de recopilar los conjuntos de datos necesarios, con la variedad y el volumen adecuados, especialmente en casos poco comunes, regiones subrepresentadas o idiomas poco habituales.

2

#### **Calidad y coherencia**

Los datos incoherentes o mal etiquetados generan sesgos en los modelos, baja precisión y problemas de implementación. Nuestros flujos de trabajo de anotación de datos especializados y rigurosos procesos de control de calidad garantizan la precisión desde el primer día.

3

#### **Demandas multilingües y multimodales**

Desde asistentes de voz hasta herramientas de visión artificial, los sistemas de IA deben ser capaces de trabajar en diversos idiomas, formatos y entornos. Lionbridge ofrece conjuntos de datos culturalmente relevantes, multilingües y multimodales a escala.

4

#### **Velocidad y escalabilidad**

Formar equipos de datos internos es una tarea lenta y costosa. Nosotros podemos acelerar el paso de sus modelos a producción con una rápida puesta en marcha, soluciones de ampliación globales y rápidos ciclos de iteración.

5

#### **Integración de la intervención humana**

Incluso los modelos más avanzados necesitan la intervención humana. Nuestros servicios cierran el ciclo con datos de entrenamiento revisados por humanos, que permiten mejorar continuamente el rendimiento.

## ¿Quiénes suelen necesitar servicios de recopilación y anotación de datos?

Nuestros servicios prestan su ayuda a equipos de diversos sectores que desarrollan o amplían soluciones de IA. Entre nuestros usuarios más habituales se incluyen:



### Equipos de aprendizaje automático y análisis científico de datos

Necesitan conjuntos de datos de gran tamaño y alta calidad para entrenar, validar y perfeccionar los modelos.



### Equipos de ingeniería y productos de IA

Dependen de datos correctamente etiquetados para el lanzamiento de funciones, como asistentes de voz, herramientas de visión artificial y algoritmos de búsqueda.



### Equipos de NLP, visión artificial e IA multimodal

Requieren datos especializados (texto, audio, imagen y vídeo) con anotaciones para tareas específicas, entre ellas, reconocimiento de entidades, análisis de sentimientos, detección de objetos o transcripciones.



### Empresas del sector sanitario, financiero, automovilístico y comercial

Utilizan la IA para tareas de diagnóstico, detección de fraudes, automatización y personalización, entre otras muchas, donde se requieren datos de entrenamiento precisos y específicos de cada dominio.



### Empresas emergentes y empresas que desarrollan IA generativa

Dependen de datos de prompts, entradas sintéticas, conjuntos de datos de entrenamiento optimizados y resultados evaluados por humanos para crear aplicaciones de IA competitivas.



## NUESTROS SERVICIOS



Anotación de datos



Recopilación de datos



Evaluación de modelos

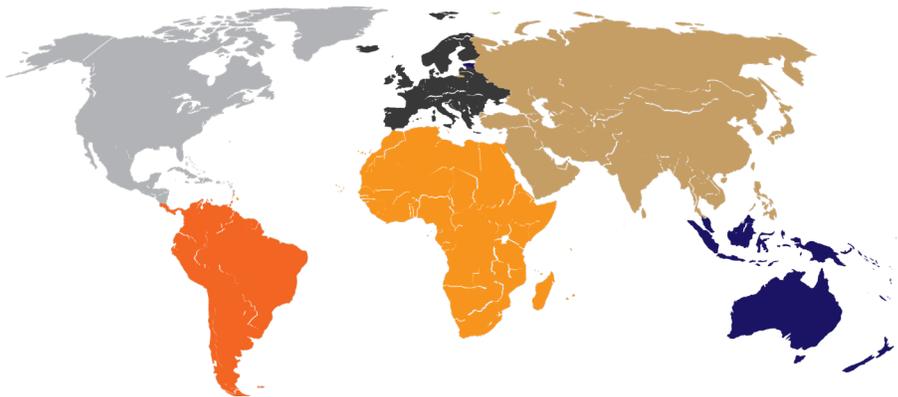
## EQUIPO

### Nuestro equipo es su ventaja competitiva

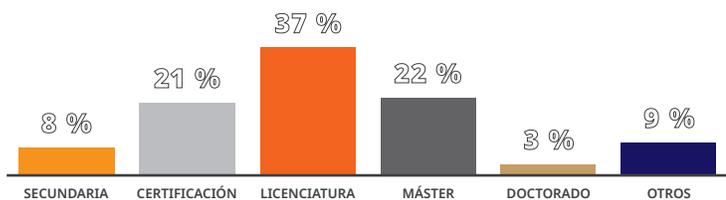
El equipo internacional de Lionbridge combina usuarios comunes con expertos en la materia, lo que nos permite ofrecerle escalabilidad y personalización a partes iguales. Tanto si necesita entrenar un modelo de IA como anotar datos complejos y específicos de un sector concreto, contamos con los colaboradores más adecuados para cada tarea. Con el respaldo de gestores de proyectos especializados y rigurosos procesos de control de calidad, podemos garantizarle datos culturalmente relevantes aplicables a la vida real y optimizar el rendimiento de su modelo.

Es necesaria una comunidad con conocimientos e intereses diversos y muy especializados.

- América del Norte: 26 %
- Europa: 29,5 %
- Asia: 27 %
- África: 6,5 %
- América del Sur: 8,5 %
- Oceanía/Australia: 2,5 %



### Máximo nivel educativo



### Género



### Edad



Comunidad de  
**600 000**  
personas de todo el mundo



**+350**  
idiomas



**+450**  
variedades regionales

### Sector



### ¿Crear o comprar? El coste de trabajar solo frente a la elección correcta. Caso práctico real de ROI.

Poner en marcha un equipo interno de operaciones de datos puede suponer millones en personal, herramientas y tiempo, sin ninguna garantía de calidad o escalabilidad. Una empresa internacional de análisis de personal necesitaba anotar datos del mercado laboral en 30 idiomas, lo que requería 1200 horas de trabajo por idioma: un total de 36 000 horas de anotación.

#### Estimación de un enfoque interno:

- **Costes en lingüistas:** aprox. 1,8 millones de dólares
- **Tiempo de contratación:** selección de 30 lingüistas + proceso de incorporación = 3 a 4 meses
- **Duración del proyecto:** aprox. 6 meses, debido al periodo de puesta en marcha y al proceso de entrega escalonada
- **Disponibilidad de los datos de entrenamiento:** aprox. 8 meses, o más, para que los datos se puedan utilizar para entrenar modelos

#### El enfoque de Lionbridge:

- Equipo global de anotadores bilingües y lingüistas cualificados bajo demanda
- Proceso completamente gestionado por gestores de proyectos de Lionbridge, con controles de calidad en cada etapa
- Lanzamiento del proyecto en semanas, no en meses
- Datos de entrenamiento listos en menos de 12 semanas, lo que acelera el desarrollo del modelo en varios meses
- **Entrega un 40 % más rápida, con ahorros de más de 1 millón de dólares, en comparación con la ejecución interna**

### CONCLUSIONES PRINCIPALES

- **Implementación del modelo entre un 30 % y un 50 % más rápida** al reducir el tiempo dedicado a la limpieza y la corrección de los datos
- **Costes entre un 50 % y un 60 % más bajos** en comparación con la creación de un equipo interno de datos, gracias a nuestra red global de colaboradores y a nuestra eficiente infraestructura
- **Tiempo de producción 3 veces más rápido**, lo que evita largos ciclos de contratación, equipamiento y entrenamiento
- **Reducción de la carga interna**, lo que libera a los equipos de aprendizaje automático para que puedan centrarse en el desarrollo del modelo y no en la gestión de datos
- **Convertimos la producción de datos en un proceso escalable**, rentable y orientado a los resultados. Ofrecemos un impacto real y acabamos con la sobrecarga.

