



CASO PRÁCTICO SOBRE RECOPIACIÓN DE DATOS DE IA

RECOPIACIÓN DE 28 000 PUNTOS DE DATOS EN UNA SEMANA

1 SEMANA
DE PLAZO

8000
CONJUNTOS DE DATOS
ADICIONALES ENTREGADOS

28 000
IMÁGENES ENTREGADAS

EL RETO

El cliente de Lionbridge necesitaba recopilar un elevado volumen de datos en un plazo muy ajustado para entrenar su modelo de IA para la identificación de objetos. En concreto, solicitó alrededor de 20 000 imágenes en cuestión de días. Era imprescindible que se tratase de datos humanos de alta calidad, y no de imágenes sintéticas, que suelen ser de calidad inferior y pueden provocar un rendimiento notablemente peor del modelo. El cliente necesitaba, específicamente, imágenes que mostraran objetos contenidos dentro de otros objetos, ya que estas ayudarían al modelo de IA a aprender a identificar correctamente situaciones en las que un objeto aparece dentro de otro, como en el caso de cestas, bolsas, cuencos, etc. Para cada conjunto, se solicitaron cuatro imágenes:

- Cada objeto de manera individual
- Un objeto dentro del otro
- El fondo vacío sin ningún objeto

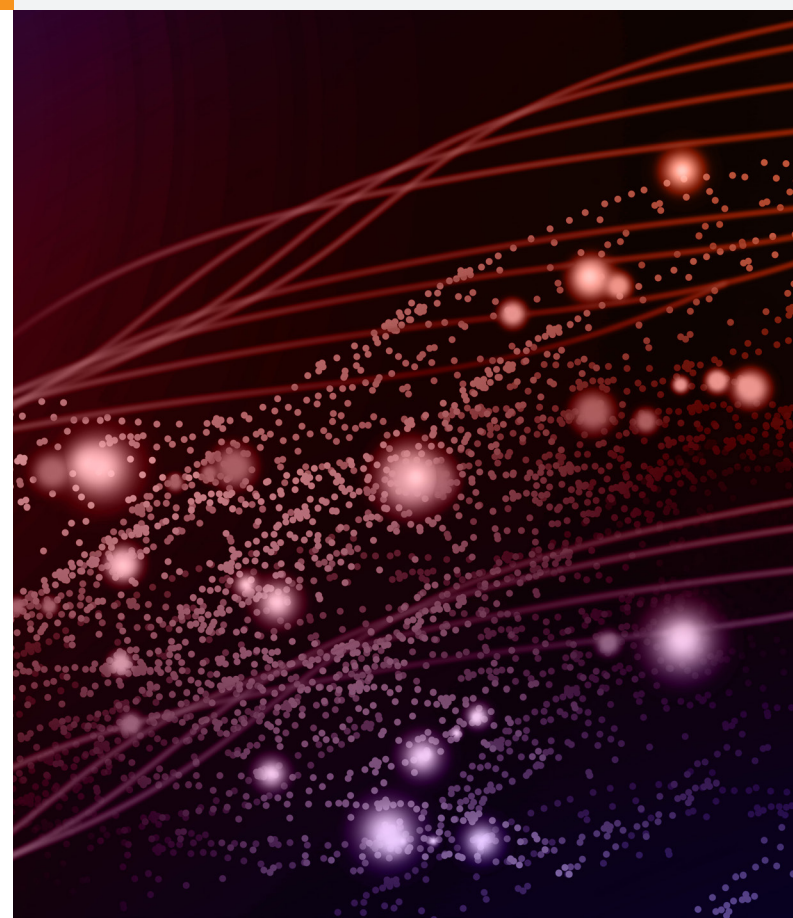
Cabe subrayar asimismo que el cliente otorgó especial relevancia a la diversidad de perspectivas en la recopilación de datos. Contar con 20 000 imágenes de los mismos objetos no habría resultado tan eficaz para el entrenamiento del modelo como disponer de 20 000 imágenes distintas. Obtener imágenes de participantes de todo el mundo con distintos perfiles demográficos permitiría al modelo captar mejor la diversidad de sus usuarios y ofrecer un rendimiento óptimo para un público tan diverso.

Usuarios de diversas nacionalidades, edades, géneros y contextos contribuirán probablemente con una amplia variedad de imágenes al sitio web del cliente. Contar con un conjunto de datos de entrenamiento más diverso permitiría que el modelo funcionase de manera eficaz para una base de clientes mucho más amplia.

ACERCA DEL CLIENTE

Lionbridge AI™ prestó sus servicios a un cliente cuya misión es conectar a las marcas con los profesionales creativos que necesitan para sus proyectos.

Esta empresa gestiona una innovadora plataforma que pone en contacto a talentos creativos con marcas que requieren este tipo de perfiles para sus proyectos. Los profesionales creativos pueden publicar sus carteras y hacerlas visibles mediante distintos criterios de búsqueda, mientras que las marcas pueden publicar ofertas de trabajo o informes. En definitiva, este cliente de Lionbridge facilita la creación de conexiones y redes de colaboración entre marcas y talento creativo. A través de la plataforma, pueden conectarse y establecer relaciones para colaborar en proyectos actuales y futuros.



ALGUNAS ESTADÍSTICAS SOBRE NUESTRA COMUNIDAD



Hablan **más de 350** idiomas



Residen en **más de 450** ubicaciones en todo el mundo:

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



- Mujeres: **52 %**
- Hombres: **48 %**



Edades de entre **18 y más de 60 años**

Ámbitos profesionales:

- **16 %** Empresa
- **16 %** Ingeniería
- **15 %** Humanidades
- **13 %** Ciencia
- **9 %** Técnico
- **4 %** Comercio electrónico
- **2 %** RR. HH.
- **2 %** Sanidad
- **2 %** Bellas artes
- **2 %** Jurídico
- **10 %** Otros



LA SOLUCIÓN

Para dar respuesta a las necesidades de este cliente, utilizamos nuestra plataforma **Lionbridge Aurora AI Studio™**.

Esta plataforma nos conecta con una comunidad de más de medio millón de expertos de todo el mundo, que destacan por su diversidad en cuanto a ubicación, idiomas, edad, género y sectores profesionales. Gracias a ellos, podemos llevar a cabo múltiples tareas, incluida la recopilación de datos, a una velocidad excepcional.

Para este proyecto, configuramos la tarea del cliente en Aurora AI Studio. En cuestión de horas, habíamos incorporado a más de 2000 participantes, que dispusieron de ocho horas para completar la tarea de recopilación de datos. De forma paralela, creamos un equipo interno de control de calidad de mayor tamaño para revisar los datos obtenidos, dada la magnitud del proyecto y la exigencia de los plazos. Este equipo se encargó de verificar la precisión de los datos recopilados y su adecuación

a los requisitos del cliente. Tras esta revisión de calidad, asignamos los metadatos correspondientes a cada imagen, lo que permitió ahorrar al cliente numerosas horas de trabajo manual.

Lionbridge entregó los datos recopilados al cliente en tres lotes. El primero se entregó antes de lo previsto para que el cliente pudiera comenzar a entrenar su modelo lo antes posible. Este primer lote constaba de 12 000 imágenes y se entregó en un plazo de tres días. El segundo lote, entregado en cinco días, contenía las 8000 imágenes restantes.

El cliente quedó tan satisfecho con estos lotes de datos recopilados para el entrenamiento de IA que solicitó 8000 imágenes adicionales con un plazo de entrega de tan solo dos días más. Attendimos encantados su petición y entregamos un total de 28 000 imágenes, todo ello en una sola semana. El lote inicial de 20 000 imágenes solicitado se entregó íntegramente en apenas cuatro días.



METODOLOGÍA

Lionbridge utilizó un proceso de varios pasos para ayudar al cliente:

Paso 1: Creación de la tarea en Aurora AI Studio

Paso 2: Incorporación de los participantes

Paso 3: Integración de los datos recopilados

Paso 4: Revisión de los datos y asignación de metadatos

Paso 5: Entrega de los datos recopilados al cliente

Paso 6: Recopilación y entrega de datos adicionales, según la petición del cliente





EQUIPO MULTIDISCIPLINAR

Director de programas | Ingeniero técnico de IA | Responsable de control de calidad | Gestor de proyectos | 10 especialistas internos en control de calidad



LOS RESULTADOS

Lionbridge AI ayudó a este cliente de tres maneras clave. En primer lugar, gracias a nuestra plataforma Aurora AI Studio y a nuestra comunidad extremadamente diversa, logramos recopilar un inmenso volumen de datos humanos de alta calidad, que reflejaban perspectivas e imágenes procedentes de todo el mundo. De este modo, garantizamos que el cliente evitara los riesgos asociados a depender exclusivamente de datos sintéticos, que pueden dar lugar a modelos mal entrenados. Además, Lionbridge recurrió a un equipo ampliado de control de calidad para verificar los datos antes de su entrega. Este conjunto de datos, de una calidad impecable, permitirá entrenar el modelo para dar un soporte eficaz tanto a la marca como a los usuarios de talento creativo, independientemente de su procedencia, perfil demográfico o del contenido de sus carteras.

La segunda forma en la que ayudamos fue entregando este volumen tan elevado de datos — de hecho, superior al inicialmente solicitado— en el plazo de una semana. La rapidez en la entrega es fundamental en el entrenamiento de modelos, ya que cada día de retraso puede resultar extremadamente costoso. Gracias al uso de Aurora AI Studio, Lionbridge permitió al cliente ahorrar miles de euros en costes de mano de obra y oportunidades de negocio perdidas al completar el proyecto en un tiempo récord. De este modo, el cliente pudo asegurarse de que el modelo de su plataforma pudiera entrenarse adecuadamente en tan solo una semana.

La última forma en que impresionamos a este cliente fue superando sus expectativas iniciales. Aunque en un principio se habían solicitado 20 000 imágenes, conseguimos recopilar un total de 28 000. Lionbridge se enorgullece de haber proporcionado no solo los datos que el cliente necesitaba, sino también aquellos que ni siquiera sabía que podía necesitar. Lograr ofrecer todo lo solicitado —y aún más— en el transcurso de una sola semana fue un logro verdaderamente notable.



MÁS INFORMACIÓN:
LIONBRIDGE.COM