

Nombre del Cliente

DefinedCrowd

Sector

Big data

Reto

Proporcionar una conexión de alta velocidad, segura y fiable para acceder a un gran volumen de datos

Negocio

Proveedor de datos de entrenamiento para inteligencia artificial

Productos y servicios

IP Access de 1 Gbps

La startup DefinedCrowd cuenta con una plataforma que permite a los científicos de datos recoger, estructurar y enriquecer información de alta calidad para entrenar algoritmos de inteligencia artificial. La calidad de los datos que maneja y su velocidad de procesamiento son claves para su negocio.

Colt proporciona la conectividad a la empresa para acceder a todos sus datos de forma rápida y eficiente. Ricardo Gusmão, IT & Information Security Manager de DefinedCrowd, destaca la presencia internacional de Colt como un valor decisivo y la contempla no solo como un proveedor, sino como un partner para acompañar a la startup en su fase de crecimiento.

DefinedCrowd recurre a Colt para conectar con su smart data platform

Las posibilidades de la inteligencia artificial están en boca de todos. Sin embargo, se presta menos atención a los cimientos que sustentan esta tecnología: los datos. En concreto a la calidad de estos, imprescindible para entrenar los algoritmos de IA. Aquí ha avistado una oportunidad de negocio DefinedCrowd. La empresa se fundó en 2015 con la intención de crear un sistema eficiente para transformar datos de entrenamiento no estructurados en datasets de alta calidad.

DefinedCrowd trabaja con grandes cantidades de datos en la nube que requieren de una conectividad fiable y de alta velocidad, para lo que ha confiado en Colt. Este servicio le permite acceder a su smart data platform y gestionarla. En ella se recogen datos, se estructuran y se enriquecen para entrenar modelos de machine learning. Los datos pueden tener dos orígenes diferentes. Los clientes pueden aportarlos, para que DefinedCrowd los estructure, o bien, solicitar su recogida.

Se estima que los científicos de datos dedican un 80% de su tiempo a estructurar la información. El objetivo es librarlos de esta pesada tarea para que se centren en crear los algoritmos de IA de alto rendimiento. Catarina Salteiro, Global PR & Communication Manager en DefinedCrowd, explica el funcionamiento del sistema:

“ Los clientes van a la plataforma y pueden customizar el tipo de datos que quieren. Pongamos que trabajan en un sistema de reconocimiento de voz y necesitan datos de entrenamiento. Deciden cuántos prompts quieren que la gente lea, cuántas veces y en qué idiomas. Después, estos requerimientos se transforman en microtareas que se canalizan hacia nuestra comunidad HITL (Human-in-the-Loop). ”

La comunidad de DefinedCrowd es una parte esencial de lo que ofrece la compañía, pues cuenta con más de 130.000 usuarios que cubren más de 70 lenguas y dialectos. Todos ellos etiquetan imágenes, leen textos y hacen anotaciones, así como otras microtareas específicas que contribuyen a los proyectos de los clientes. Esto, junto a la especialización de su equipo de ingenieros, permite a la empresa destacarse en tecnologías de voz y procesamiento del lenguaje natural (NLP). Para complementar sus servicios, también trabajan con otros ecosistemas de machine learning, como visión artificial.

“Nuestra comunidad es la que hace las microtareas que nos permiten obtener datasets de alta calidad. En este punto tenemos miembros altamente cualificados que proporcionan los datos que nuestros clientes necesitan, que nos permiten garantizar entre un 95% y un 98% de calidad”, apunta Salteiro.



Para acceder a toda esta infraestructura de datos, DefinedCrowd utiliza un Acceso IP de Colt, de 1 Gbps simétricos. Con esta conexión maneja su smart data platform y gestiona todo el volumen de datos con el que trabaja. Para que la productividad sea óptima es necesaria una alta velocidad de acceso a los datos, pero también son decisivos otros factores.

Ricardo Gusmão, IT & Information Ssecurity Manager, destaca del servicio de Colt la seguridad de la red, que protege de posibles ataques al sistema, así como las condiciones del SLA y la fiabilidad de la conexión. Aunque hay un aspecto sobre el que el responsable de IT se detiene. “La disponibilidad es muy importante, especialmente en compañías como la nuestra, que dependemos de las comunicaciones. Esta disponibilidad, y la fiabilidad, son los aspectos más importantes y Colt puede proporcionárnoslos”, puntualiza, para después añadir: “Realmente veo a Colt como un partner, no solo como un proveedor”. DefinedCrowd tiene un largo recorrido por delante y Gusmão está convencido de que Colt puede ser un acompañante en su proceso de crecimiento.

La transversalidad de su smart data platform le permite trabajar con clientes de distintas industrias, desde la atención al cliente al sector salud, pasando por la automoción, el energético, las fintech, el retail o los medios de comunicación. Su producto busca que las empresas mejoren la calidad y escalabilidad de sus aplicaciones de machine learning y aceleren su lanzamiento al mercado.

Los clientes de DefinedCrowd pertenecen mayoritariamente a compañías del ranking Fortune 500, entre las que destacan nombres como BMW, MasterCard o Yahoo Japan. Además, la empresa tiene lazos sólidos con compañías pujantes en inteligencia artificial. Amazon está entre sus inversores principales y recomienda sus servicios tras su papel en el

desarrollo de Alexa. La compañía es co-selling partner de Microsoft y su solución también está integrada en la plataforma IBM Watson.

El rápido crecimiento de DefinedCrowd se ha acelerado desde el anuncio en verano de 2018 de una ronda de inversión de 11,8 millones de dólares.

Empezó 2019 con menos de 80 empleados y tiene previsto terminar el año con 150. Ahora mismo tienen oficinas en Lisboa, Oporto, Seattle y Tokio. Colt tiene comunicaciones de fibra óptica en estas ciudades y ofrece la posibilidad de contar con un solo proveedor en todas sus localizaciones globales.

Para sostener su futura expansión, DefinedCrowd se plantea la posibilidad de un acceso cloud dedicado, para Microsoft Azure, plataforma que utiliza para sus servicios. “Por eso es tan importante la presencia de Colt en los países donde estamos, porque para nosotros es muy imprescindible la velocidad de

acceso a la información”, afirma Gusmão. “El estar ubicado en los lugares donde nosotros tenemos las oficinas nos puede beneficiar a la hora de manejar las conexiones”. Estos son los cimientos para el objetivo que se ha propuesto la startup: ser el proveedor número uno de datos de entrenamiento para iniciativas de inteligencia artificial.

“ La disponibilidad es muy importante, especialmente en compañías como la nuestra, que dependemos de las comunicaciones. Esta disponibilidad y fiabilidad son los aspectos más importantes y Colt puede proporcionárnoslos”, puntualiza el responsable de IT y añade: “Realmente veo a Colt como un partner, no solo como un proveedor”

Ricardo Gusmão,
IT & Information Security Manager

DefinedCrowd[®]

Para más información, contacte con nosotros:

935 502 211
sales@colt.net
colt.net/es