

“En Cloudera tenemos uno de los mejores productos para dar valor al dato”

Cloudera, compañía especializada en datos para la inteligencia artificial empresarial, fortalecía su presencia en España hace unos meses con la apertura de una nueva oficina en Madrid. Francia e Italia también se veían reforzadas en una apuesta de la multinacional por el sur de Europa. Con más de 3.000 empleados, más de 2.000 clientes, 100 en nuestro país, y 24 grandes oficinas, la apertura de su oficina madrileña significa para la compañía, desde el punto de vista de negocio, “un punto de encuentro con nuestros clientes y *partners*”, asegura Juan Carlos Sánchez, vicepresidente regional en Iberia de Cloudera. Esto se ha llevado a cabo en un momento en el que las expectativas de la multinacional son buenas, con la mirada puesta en la continuidad y el crecimiento.

Inma Elizalde



Juan Carlos Sánchez, vicepresidente regional Iberia Cloudera

Su apuesta por el *open source* y Hadoop están en el origen de una compañía que se creó en 2008 de la mano de cuatro socios que provenían de organizaciones como Google, Oracle, Yahoo y Facebook, lanzando su primera distribución basada en un concepto de *big data*. Hoy en día tan sólo queda una pequeña parte de aquella compañía, asegura Juan Carlos Sánchez, transformada en una plataforma de datos integral que aporta capacidades a todos los niveles.

Marcando la diferencia

Esta organización vivió un antes y un después con la compra de Hortonworks y la creación de Cloudera Data Platform. Dos competidores que en un momento dado decidieron unir sus caminos y fusionarse manteniendo la marca Cloudera. Esto también les proporcionó un diferencial tecnológico, admite. "Hortonworks era especialista en algunas capas y nosotros en otras. Con la fusión duplicamos capacidades y plantilla, vislumbrando lo que hoy es nues-



tra estrategia de producto. Creamos Cloudera Data Platform y lanzamos nuestra primera distribución en la nube. Llevamos a cabo un salto para tener una tecnología que cubre las necesidades de nuestros clientes", apunta.

"Nuestra diferencia estriba en que somos la única empresa que cuenta con una plataforma que puede dar servicios independientemente de dónde estén los datos. Una plataforma híbri-

da real", señala en directivo. "Nuestros clientes pueden trabajar con nosotros en cualquier tipo de nube, en contenedores o en entornos más tradicionales. Debajo de ese *framework* adoptamos toda la serie de componentes que van saliendo de la distribución *open source*", avanza. Uno de los últimos que han adoptado es Apache Iceberg, una tecnología nacida en 2018 a la que considera la más importante para desarro-

llar un concepto de *data lakehouse*, una de las tendencias actuales del mercado. Un formato de alto rendimiento para grandes tablas analíticas que cuando se aplica a nubes privadas permite aplicar IA en grandes cantidades de datos críticos tanto en la nube como en *on-premise*, permitiendo desarrollar nuevos casos de uso.

Cloudera se convierte, además, en la única plataforma en ofrecer Apache Iceberg en nube pública y privada, permitiendo a las organizaciones desarrollar la inteligencia artificial a nivel empresarial, con el fin de generar mayor valor de negocio. También es única al permitir ejecutar esta tecnología de manera independiente junto con modelos grandes de lenguaje, garantizando a los usuarios mejoras de seguridad y privacidad de datos, además de un rendimiento optimizado y una reducción en los costes operativos. Además son la única compañía que ha adoptado Apache NiFi, un componente utilizado para la gestión del movimiento de información en tiempo real.

A tener en cuenta

- **El *open source* es para Cloudera la clave de su ecosistema.**
- **Salió a Bolsa en 2021.**
- **Su aproximación en nube es un servicio de plataforma como servicio donde trabajan con los principales *vendors*: AWS, Google y Microsoft Azure.**
- **De la ampliación de la colaboración de Cloudera con NVIDIA ha nacido Cloudera Powered by NVIDIA que integrará los microservicios empresariales NVIDIA NIM en Cloudera Machine Learning, un servicio de Cloudera Data Platform para flujos de trabajo de IA y ML con el fin de que sean más rápidos, seguros y simples a la hora de llevarlos a producción.**
- **Con el acuerdo entre Cloudera y AWS ambas compañías acelerarán e impulsarán la gestión y el análisis de datos en la nube nativa de AWS, aprovechando Cloudera los servicios de AWS para ofrecer una continua innovación y una reducción de los costes para los clientes que usan Cloudera Open Data lakehouse para desarrollar IA generativa de confianza.**
- **El *chatbot* LMM de Cloudera permite a las organizaciones usar IA generativa basada en grandes modelos lingüísticos . Así, las empresas podrán construir su propia aplicación de IA eligiendo el LLM de código abierto que prefieran para utilizar sus datos.**

¿Qué valor da Cloudera a todo ello? Lo embobemos en nuestra plataforma, responde. “Llevamos a cabo la gestión y administración desde un punto de vista seguro y gobernado. Estos son los pilares que nos demandan los clientes. Necesitamos tener una visión de que siempre tengo el control del dato sin problemas de seguridad, contando con una trazabilidad y go-

bierno del dato”, argumenta. “Esto es lo que da la capa que desarrolla Cloudera por encima de todos los componentes *open source*”.

Aportando valor al dato

Dar mayor valor a los datos es algo que no todas las organizaciones consiguen en un momento en el que el 30 % de los presupuestos

están dedicados a proyectos de datos. Según un estudio elaborado por la compañía, las organizaciones no son capaces de entender un tercio de sus datos, por lo que están perdiendo esa información, advierte Juan Carlos Sánchez. Y un 68 % muestra una gran complejidad con los silos de información. “Hay una gran dispersión y no tienen un gobierno central en una plataforma única que les ayude a llevar un seguimiento, encontrándose con muchos problemas a la hora de aportar valor”, asegura. A pesar de todo, considera que las organizaciones



“Cloudera ayuda a las empresas a monetizar todo el ciclo del dato, dando soporte a entidades mediana-grandes o muy grandes”

están operando bien porque funcionan pero, en muchos casos, se hace de manera tradicional y no de la manera más eficiente. “Extraer valor a los datos es la capacidad de ahorrar costes, ser más productivo y eficiente. Y nosotros tenemos si no el mejor, uno de los mejores productos que hay en el mercado para aportar valor al dato”, especifica.

Cloudera ayuda a las empresas a monetizar todo el ciclo del dato, dando soporte a entidades mediana-grandes o muy grandes que cuentan con zettabytes de información, convirtiéndose en una de las principales compañías, a nivel mundial, que manejan mayor cantidad de información. Apache Iceberg ayuda a tratar datos masivos de manera eficiente, extrayendo todas las ventajas del procesamiento natural en los repositorios de datos en los que las organizaciones guardan su información (*data warehouse*). Y montarlos en los lagos de datos analíticos (*data lakehouse*). “Juntamos ese procesamiento para obtener información en tiempo real”, recuerda.



El papel del CIO

¿Cómo lo afronta el CIO? Juan Carlos Sánchez estima que el CIO actual, en muchas ocasiones, pertenece al consejo de dirección, aportando valor a sus compañías. Entiende qué hay que hacer para mejorar, sabe de tecnología y arquitectura. Cloudera les ayuda en gran medida a través de experiencias con clientes, compar-

tiendo referencias e ideas con el fin de hacerles ver que hay cosas que pueden llevarse a cabo de manera mucho más exigente.

Teniendo en cuenta que los retos del CIO este año pasan por aportar valor al negocio a través de la IA, la automatización, la reducción de costes y no apostar por proyectos equivocados, ¿qué papel juega Cloudera en este sentido?

Cloudera Data Platform para la nube privada

La compañía lanzó Cloudera Data Platform para la nube privada con el fin de aumentar la velocidad, simplicidad y rentabilidad del ciclo de vida del dato conectado al centro de datos, permitiendo a los departamentos TI responder a las necesidades empresariales de manera más ágil, ofreciendo una mayor productividad en la gestión de los datos.

Cuenta con potentes herramientas de gestión basadas en contenedores de datos que reducen de semanas a minutos los tiempos de entrega del análisis y el aprendizaje automático. Optimiza la rentabilidad de los centros de datos con analítica basada en contenedores y *machine learning* al aumentar la capacidad de uso de los servidores hasta un 70 % y reduce la sobrecarga de almacenamiento del centro de datos.

¿Qué define a su plataforma en estos ámbitos? “Hacemos que el cliente no tenga que sacar sus datos y ponerlos al servicio de las diferentes soluciones ya que les ayudamos a crear esas soluciones para las mismas vengan a los datos”, responde. “Cloudera aporta una plataforma integral con el fin de ayudar en el ciclo de vida completo del dato. Podemos hacer que existan componentes que se conecten a esos dispositivos para capturar datos en tiempo real, procesarlos, analizarlos en un tiempo real absoluto para que pueda tener respuestas en ese momento”, enfatiza. Es decir, “llevar a cabo la transformación del dato, su ingeniería

y analítica, terminando en iniciativas de *machine learning*. Cloudera aporta su plataforma integral híbrida que les ayude a gestionar sus datos para que no exista complejidad a la hora de trabajar”.

IA generativa

El siguiente paso de la IA generativa pasará a crear imágenes, textos, videos y audio originales. El directivo admite que Cloudera lleva tiempo trabajando en este ámbito para llevar la IA a los datos, aportando un ecosistema de piezas para montar la IA generativa.

IA generativa

Seguridad

No podía faltar la seguridad, uno de sus diferenciadores principales, tal y como enfatiza Juan Carlos Sánchez, “y uno de los motivos por el que los clientes trabajan con Cloudera. Tenemos una capa transversal en la que aseguramos ciertos componentes para dar el valor de la seguridad al dato”, finaliza.

Seguridad