

EXECUTIVE BRIEFING

HACIA ARQUITECTURAS HÍBRIDAS DE DATOS: RESOLVIENDO LA COMPLEJIDAD DE LOS ENTORNOS ON-PREMISE Y CLOUD

Hacia arquitecturas híbridas de datos: resolviendo la complejidad de los entornos on-premise y cloud

En un contexto económico cada vez más incierto, los datos se han convertido en un activo imprescindible para la toma de decisiones. En otras palabras, los negocios se han convertido en organizaciones orientadas al dato. Cualquier actividad de la empresa necesita apoyarse en datos, desde el marketing para conocer al cliente y mejorar su experiencia, hasta las operaciones para optimizar los costes o la cadena de suministro.

En consecuencia, están proliferando los casos de uso del dato a todos los niveles de la organización. Por otro lado, la base de usuarios del dato se está ampliando día a día, cada uno con un perfil distinto de capacidad. Finalmente, poner el dato en el centro obliga a asegurar su calidad, gobierno y accesibilidad, en especial en relación con su compartición. Por tanto, se hace necesario triangular casos de uso, usuarios y datos y no existe una fórmula universal; cada empresa tiene que encontrar su equilibrio.

Todo esto tiene que hacerse sobre una arquitectura en evolución que parte de una base de tecnología heredada. Al mismo tiempo, el modelo cloud permite el acceso a tecnologías avanzadas con una flexibilidad e inmediatez de acuerdo con los nuevos casos de uso. Las empresas están incorporando cada vez más cloud tanto en las iniciativas en marcha como en las nuevas, pero necesitan soluciones que soporten un modelo híbrido que va a seguir operativo.

En este documento se desarrollan estos aspectos tal y como se debatieron en una sesión que tuvo lugar en diciembre de 2022 y en la que participaron responsables de datos y tecnología de distintos sectores de actividad



Los casos de uso de datos proliferan

Actualmente se observa un aumento en los casos de uso de las empresas para mantenerse competitivas y aprovechar oportunidades de innovación. A continuación, se presentan los que se destacaron en el debate:

Casos de uso internos, hacia la organización. El dato ha abierto nuevas posibilidades para mejorar la eficiencia y productividad, reduciendo la necesidad de personal en ciertas actividades (FTE). Algunos ejemplos son los procesos logísticos, la planificación de recursos, las decisiones operativas y la gestión del riesgo.

- **Planificación eficiente de recursos.** Consiste en asignar adecuadamente los recursos, tanto humanos como de equipamiento. En palabras de un asistente “utilizamos analítica de datos para realizar una asignación dinámica de personal que da servicio en ventanas de tiempo concretas”. A medida que el tamaño de las empresas aumenta, la planificación de personal a proyectos o servicios se vuelve más compleja porque necesita ajustar las cargas de trabajo con la disponibilidad de los empleados (ej. periodos vacacionales).

- **Gestión del riesgo.** La incertidumbre existente en el mercado está llevando a un mayor número de empresas a analizar y gestionar su riesgo. En palabras de un asistente “necesitamos ser capaces de establecer patrones y detectar anomalías para reaccionar con rapidez”. Este aspecto es especialmente relevante en empresas catalogadas como infraestructuras críticas, que no pueden permitirse interrupciones en el servicio.

- **Optimización de procesos logísticos.** Al combinar distintas fuentes de información, las empresas son capaces de optimizar las rutas de entrega de producto al consumidor. En palabras de un asistente, “además de reducir costes conseguimos ser más rápidos apoyándonos en modelos geoespaciales”. Adicionalmente, se consigue reducir el impacto medioambiental.





Casos de uso externos, hacia el cliente. Las empresas necesitan conocer a sus clientes y caracterizar su comportamiento para explotar oportunidades de negocio y prever la demanda futura.

- **Ofrece valor añadido al cliente.** Las organizaciones están recopilando datos de uso de los servicios para encontrar oportunidades de innovación. En palabras de un asistente “si somos capaces de caracterizar el comportamiento del cliente podemos ofrecerles servicios de valor añadido”. Por ejemplo, si detectan patrones anómalos o erróneos de uso, se puede enviar una alarma de forma proactiva para que el cliente realice un uso correcto.
- **Previsión de la demanda.** Para los negocios en los que hay desfase temporal entre la compra de insumos y su utilización, la precisión en la predicción de la demanda es clave para su rentabilidad. El problema es que hay una gran volatilidad de precios que afecta a todo el aprovisionamiento, incluyendo materias primas, energía o productos perecederos. En palabras de un asistente “te juegas el negocio”. Además del precio hay que acertar con la cantidad adecuada para evitar extremos como exceso de stock o ruptura.

Otros casos particulares. Algunos negocios están desarrollando casos de uso ad-hoc que responden a preguntas concretas. Por ejemplo, las administraciones públicas cuentan con un grandes históricos de información, muchas veces desestructurados, susceptible de dar respuesta a preguntas asociadas a políticas públicas. En muchas organizaciones de servicios la información propia puede reconvertirse en productos de datos que pueden ser monetizados.

La base de usuarios se expande dentro y fuera de la empresa

Los casos de uso vistos en el apartado anterior tienen que encajar en la base de usuarios para que se conviertan en valor. En estos años la proporción de empleados que utilizan datos en las organizaciones ha ido aumentando de forma sostenida.

Para los responsables del dato, está cambiando la forma de dar servicio a la organización. En palabras de uno de ellos, “antes era suficiente con entregar a los usuarios el dato en bruto y ellos mismos lo analizaban. Ahora, con más usuarios y aplicaciones más complejas, hay que darles también herramientas, soluciones analíticas y de visualización de forma que puedan trabajar con los datos.”

Los responsables del dato buscan empoderar al usuario para que sea autónomo, y no necesite recurrir a ellos. Esto es particularmente importante en el caso de las empresas con pocos empleados, que necesitan ser muy productivos. De ahí que estas áreas estén realizando actividades de formación, así como grabando y desarrollando materiales.

No obstante, hay un equilibrio entre empoderar al empleado y gobernar el uso. Es importante que el usuario pueda trabajar y analizar el dato, pero al mismo tiempo mantener una única fuente de la verdad y su integridad sea cual sea el uso.

Además, cada organización tiene su propia casuística de usuarios, relacionada con los contratos laborales y la permanencia en la empresa. Por ejemplo, en la Administración, el funcionario tiene una edad promedio superior a los 50 años y se producen jubilaciones, pero no hay reposición. Por otro lado, el sector privado tienen que hacer frente a una elevada rotación del personal. Los contratos de 4 años son habituales, y ese es el horizonte temporal con el que tienen que trabajar en relación con el talento. Todos estos factores hacen necesario “retener y aprovechar el know-how analítico dentro de la organización”.





Otro aspecto a tener en cuenta es la cultura, tanto del usuario como de la organización. Un ejemplo es el hábito del uso de Excel para el *reporting* o para ciertas analíticas, que varias empresas han empezado a prohibir. A nivel de organización, es importante gestionar las expectativas temporales de los proyectos analíticos. En palabras de un participante, “se necesita un plazo hasta que se obtienen resultados, y ante la gran cantidad de oportunidades que se han abierto se puede morir de éxito”. Otro reto es la falta de una visión transversal, con una percepción muy vinculada a las necesidades y oportunidades propias. Tal y como comentaba un participante “es difícil ver las oportunidades más amplias en la organización”.

Por último, para dar un servicio adecuado a una base heterogénea de usuarios se hace necesario segmentarlos a ellos y a las propias aplicaciones. Por ejemplo, hay una diferencia entre el usuario final (empleado o incluso cliente) y el talento digital (científico o analista del dato). Asimismo, los requisitos cambian según se trate de un usuario interno o de un usuario externo. En este último caso, puede tratarse de colaboradores que tienen que acceder a la información, introducirla en los sistemas corporativos, o bien realizar analíticas como contratistas externos.

La calidad y gobierno del dato son prioritarios

Ante la multiplicación de casos de uso y usuarios, gestionar y gobernar el dato es una prioridad. Se trata en primer lugar de que se pueda compartir en la organización, y por tanto pueda llegar a los usuarios. En segundo lugar, su calidad no es negociable porque va a determinar el valor de los resultados. Finalmente, los participantes destacaron la necesidad de conectar los datos para tener una visión unificada y de este modo explotar todas las posibilidades.

La compartición del dato no es fácil, porque persiste un sentimiento de propiedad en los diferentes departamentos. Uno de los participantes renombró la soberanía del dato como “soberbia del dato”, lo que fue refrendado de forma unánime en la mesa.

En el caso de las administraciones públicas, entra en juego la compartición de datos entre diferentes organismos, lo que supone un cuello de botella. En primer lugar, no existe una normativa común que facilite el intercambio de datos entre entidades. A esto hay que sumar condicionantes culturales y políticos. En la práctica, se hace necesario realizar convenios entre organizaciones uno a uno, que pueden llevar muchos meses e incluso alcanzar el año. Ante esta situación, se están realizando esfuerzos para desarrollar una normativa que agilice los intercambios de información.

Por otro lado, compartir datos sensibles requiere anonimizarlos para preservar la privacidad. A esto se añade la dificultad técnica de las bases de datos legacy. Los participantes hicieron referencia al movimiento de open data, que tuvo un fuerte impulso inicial hace algo más de una década, pero no llegó a cuajar y quedó relegado a iniciativas puntuales de diferentes administraciones.





En relación con la calidad del dato, varios participantes destacaban la importancia que tiene la calidad para la precisión en los análisis predictivos de los que depende la rentabilidad del negocio. La calidad aplica al “iceberg de datos”, de forma que estén limpios, ordenados y completos. De hecho, “sin un dato de calidad que funcione para las analíticas descriptivas, es imposible que funcione en las predictivas”.

Otro aspecto que se destacó en el debate es la necesidad de conectar los datos para tener una visión de 360°. Las empresas admiten que se trata todavía de un trabajo en construcción. En el caso de las administraciones, la orientación a procedimiento hace necesario un cambio cultural, para orientarse alrededor del ciudadano. Prueba de ello es que persisten múltiples visiones de la misma persona en función de cada organismo, y de cada proceso en marcha. Algo similar sucede en el caso de IoT, donde se necesita conectar todos los datos para poderlos explotar (ej. flotas de vehículos). En estos casos, el reto se encuentra más en la complejidad tecnológica y la interoperabilidad.

Finalmente, los nuevos canales de interacción están provocando la emergencia de datos no estructurados a gran escala. Esto supone un reto para muchas herramientas, orientadas a tipologías de dato estructurado. Por ejemplo, el caso de la videollamada supone un cambio en el tipo de dato, que además lleva asociado el reto de su volumetría.

Las arquitecturas de datos se hibridan

Actualmente la mayoría de las empresas cuentan con una arquitectura de datos híbrida, que combina infraestructura en la nube y on-premise. Cloud se ha consolidado como una tendencia dentro del mapa tecnológico de las empresas y está ganando peso en la hibridación. Gran parte de las organizaciones están adoptando estrategias de “cloud first” para todo lo nuevo y contemplan migrar su core, aunque el horizonte temporal nos está determinado.

En conjunto existió un consenso sobre los beneficios de cloud al no tener que invertir recursos en su gestión y mantenimiento. En este sentido uno de los participantes sentenció que “trabajar bajo *serverless* conllevaba fuertes beneficios a la organización”. Otro asistente destacó que “cloud ha permitido detectar inconsistencias mucho antes que en on-premise”.

No obstante, la aproximación es pragmática y se realiza una evaluación caso a caso teniendo en cuenta el coste, la sensibilidad de los datos y los requisitos de rendimiento. Uno de los participantes mencionó que “en el caso de las aplicaciones de diseño, dado que requieren baja latencia, se mantienen on-premise”. Otro participante comentó que el OPEX no justificaba el caso de negocio. Cada organización tiene sus propios condicionantes, como por ejemplo el legacy o la existencia de “un alto volumen de desarrollos personalizados”.

El resultado es que la adopción de cloud en las arquitecturas de datos varía entre empresas y sectores. Por ejemplo, en el sector público, cloud está entrando más despacio debido a las idiosincrasias del sector. Por un lado, cuentan con una fuerte herencia tecnológica; y por otro, deben tener en cuenta aspectos específicos como la ley de contratación y sus implicaciones financieras al distinguir entre capítulos de inversión y gastos. Por tanto, los entornos tecnológicos tiene por delante un largo recorrido de hibridación.



Conclusiones

En el debate se destacó como están surgiendo múltiples oportunidades para explotar los datos que antes ni siquiera se habían contemplado, desde analizar el histórico de datos propios hasta lanzar nuevos productos de datos al mercado.

Para hacer realidad estas oportunidades, cloud hace posible acceder a múltiples herramientas de análisis y, de hecho, se está adoptando cloud-first en las iniciativas de innovación.

Por otro lado, las empresas constatan que los perfiles de usuario de datos se están ampliando, abarcando empleados que tienen escasa destreza digital. Además, los colaboradores externos gana peso y muchos de ellos están activamente involucrados en el procesamiento de los datos.

El modelo actual de infraestructura es de hibridación, donde usuarios, datos y herramientas residen en diferentes entornos, incluyendo cloud. Ante esta heterogeneidad ganan peso las plataformas que permiten gobernar y explotar los datos dondequiera que estén y hacerlos disponibles a los diferentes usuarios y aplicaciones.



Hacia arquitecturas híbridas de datos:
resolviendo la complejidad de los
entornos on-premise y cloud



CLOUDERA

Analistas principales de IDG Research

 @IDGResearch_ES

**Alberto Belle
Fernando Maldonado**

 @FmaldonadoF

© Todos los contenidos, textos e imágenes son propiedad de IDG COMMUNICATIONS, S.A.U. o de terceros a los que se han adquirido sus derechos de explotación, y están protegidos por los derechos de Propiedad Intelectual e Industrial. El usuario únicamente tiene derecho a un uso privado de los mismos, sin ánimo de lucro, y necesita autorización expresa para modificarlos, reproducirlos, explotarlos, distribuirlos o ejercer cualquier derecho perteneciente a su titular.



RESEARCH SERVICES

Ctra. de la Coruña Km 18.200

Edif. C - Bajo Izq. - 28231

Las Rozas de Madrid

Teléfono +3491 349 6600 - jbenedetti@idg.es