



Entendiendo la Era del dato: tecnologías y propuestas para gestionar la datificación

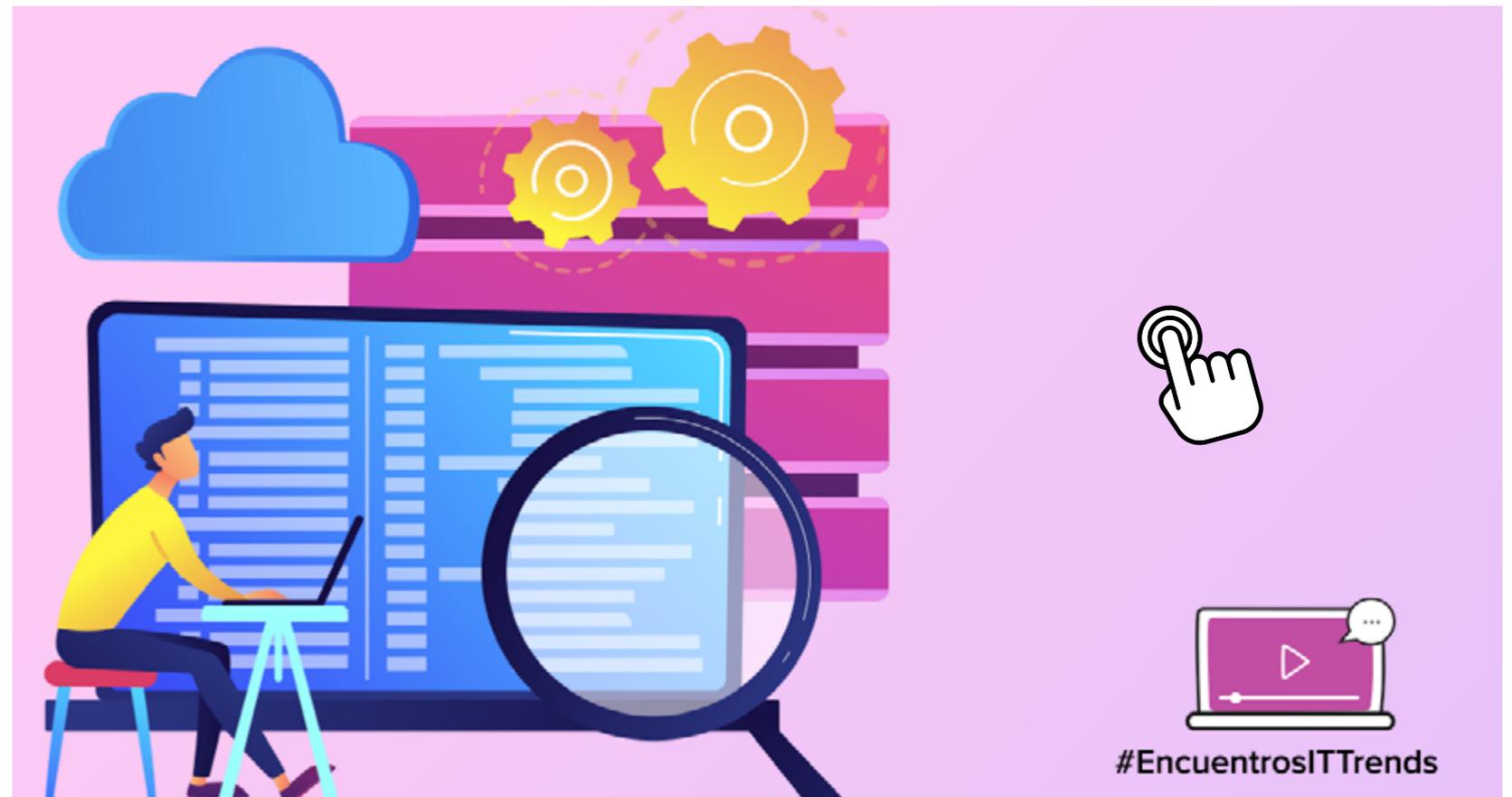
#ENCUENTROSITRENDS

Entendiendo la Era del dato: tecnologías y propuestas para gestionar la datificación

Actualmente se generan datos de todo lo que ocurre en el mundo digital y más allá. Las empresas, las personas, los sensores, las tecnologías... todo movimiento produce hoy en día datos que deben capturarse, almacenarse, analizarse, entenderse, hacer accesibles, garantizar su disponibilidad, protegerse... Y todo ello independientemente de su naturaleza, magnitud o uso.

Con los datos, las empresas están construyendo hoy sus estrategias empresariales, pero sin una correcta gestión en cada una de sus fases, se pierde el valor del llamado "petróleo del siglo XXI".

En este Encuentro IT Trends, celebrado bajo el título Entendiendo la [Era del dato: tecnologías y propuestas para gestionar la datificación](#), hemos abordado los últimos enfoques tecnológicos para hacer que esta era del dato aporte valor a cada una de las estrategias empresariales, y, para ello, hemos reunido a un amplio grupo de expertos, tanto en el ámbito académico como en el empresarial. Así, hemos contado con las opiniones de Alfonso Castro,



director del Máster en Data Science & Big Data en el Centro Universitario U-tad; Santiago Moral Rubio, director de DCNC Sciences; Quique Sánchez Candorcio, director comercial para

EMEA de Ikusi; Ramsés Gallego, International Chief Technology Officer de CyberRes, a Micro Focus line of business; e Israel Serrano, director regional para Sur de Europa de Scality. ■

#ENCUENTROSITRENDS

El dato como activo de valor para las organizaciones

La datificación es un proceso por el que se plasma un fenómeno en un formato cuantificado para su tabulación y análisis en un dato. Es una carrera imparable y con numerosas aplicaciones. Para entenderlas y conocer el futuro de estas ciencias que se pueden aplicar a los datos, Alfonso Castro, director del Máster en Data Science & Big Data en el Centro Universitario U-tad, y Santiago Moral Rubio, director de DCNC Sciences, ahondaron en el primer debate del Encuentor IT Trends [Entendiendo la Era del dato: tecnologías y propuestas para gestionar la datificación](#), en la importancia de los datos para las estrategias corporativas y el comportamiento futuro de esos datos. Hoy, ya se suben cada minuto alrededor de 500 horas de vídeo en YouTube, mientras que la suma de las diferentes actividades digitales, como redes sociales, banca on-line... generan una cantidad de bits equivalente a un 1 seguido de 21 ceros, lo que nos deja sobrecargados de información.

itTRENDS

#EncuentrosITrends

(De arriba hacia abajo) Alfonso Castro (Centro Universitario U-tad) y Santiago Moral Rubio (DCNC Sciences), en un momento del debate. **Clica en la imagen para ver el vídeo**



“Ahora, esas arquitecturas que las empresas tenían que levantar on-premise para la gestión de datos, con un coste importante, pueden ser usadas en modo cloud, permitiendo una democratización del dato, siendo capaces de gestionar los datos para decidir y hacer predicciones sin grandes inversiones”

ALFONSO CASTRO, CENTRO UNIVERSITARIO U-TAD

La importancia del dato está generando un mayor interés por parte de los profesionales y un acercamiento estratégico de las empresas. En palabras de Alfonso Castro, “esta cantidad ingente de datos proporcionan un conocimiento más profundo de su organización, y lo que hacen es

incrementar el beneficio de las mismas. Estas empresas son capaces de orientar toda su estrategia al dato, y, a partir de estos, tomar decisiones de negocio. Un ejemplo es el sector de la Banca. En nuestro móvil tenemos información anticipada de la estimación de nuestros gastos o, si pedimos

un crédito, vamos a tener una respuesta basada en un análisis que se apoya en una información tan variada como, por ejemplo, cómo navegas por la propia aplicación. Las empresas financieras son un claro ejemplo de cómo se aprovecha el Big Data. Pero hay otros sectores, como la Fórmula 1, donde se analizan 400 canales de datos en tiempo real para saber cómo está el coche y ser capaces de ganar o no una carrera”.

Pero, ¿cómo ha evolucionado este interés por el dato en las empresas? Para Santiago Moral, “estamos viendo que, cuanto más jóvenes son las empresas, antes crean un modelo competitivo frente a las ya asentadas que es paradigmático, porque todas las nuevas empresas se basan en datos. Todas las compañías, tecnológicas o no, se están basando en datos. No hay ninguna que esté naciendo ahora que no piense en una estrategia comercial, de negocio o de posicionamiento que no pase por estar analizando los datos de su actividad comercial, de negocio y operativa. Sin embargo, las empresas muy establecidas tienen dificultades para asumir esta tendencia y esta velocidad. Lo que genera es un mercado importante de transformación digital poniendo siempre en el eje de todos los proyectos el dato. Es el elemento sobre el que se construye todo lo demás. Por tanto, tener buenos profesionales que puedan trabajar esa información y puedan proponer ideas de cambios operativos transformacionales, es el quid de la siguiente revolución

industrial que tenemos en marcha: la revolución industrial del dato”.

EL IMPACTO DEL VOLUMEN

La generación exponencial e imparable de datos está impactando en el mundo tecnológico. Según Santiago Moral, “se está generando una situación curiosa e inesperada. Hace unos años pensábamos que iba a ser inmanejable por el volumen, pero están surgiendo tecnologías que, con menor coste y consumo, son capaces de manejar muchísimos más datos. El ejemplo más claro son los modelos de ciencia de datos. Ahora mismo, un modelo analítico bien basado en ciencia de datos es capaz de analizar volúmenes exponencialmente superiores a los de hace diez años, y de almacenarlos. Pensar en modelos basados en estructuras de datos no clásicas es complicado desde un punto de vista tradicional, pero existe tanto la tecnología base como los algoritmos para manejar volúmenes de datos que hace un lustro eran impensables”.

En opinión de Alfonso Castro, “Big Data no es solo un gran volumen de datos, sino que es información procedente de muy diversas fuentes, de diferentes tipos, estructurados y no estructurados o semiestructurados, y que cambian con mucha rapidez. El tratamiento de esta información resulta muy complejo. Por eso, hace algunos años comenzaron a afianzarse nuevas arquitecturas y tecnologías de programación, procesamiento y almacenamiento de datos distribuidos, que nos permitió abordar la



Santiago Moral Rubio, DCNC Sciences

“Tener buenos profesionales que puedan trabajar la información y proponer ideas de cambios operativos transformacionales, es el quid de la siguiente revolución industrial que tenemos en marcha: la revolución industrial del dato”

SANTIAGO MORAL RUBIO, DCNC SCIENCES

realidad de Big Data. Ser capaces de generar valor a partir de los datos y su análisis. Hemos contado con algoritmos que han podido ejecutar estos modelos basados en estas arquitecturas. Pero, incluso estas arquitecturas que las empresas tenían que levantar on-premise, con un coste importante, pueden ser usadas en modo cloud, permitiendo

una democratización del dato, siendo capaces de gestionar los datos para decidir y hacer sus predicciones, sin grandes inversiones”.

EL FUTURO DE LOS DATOS

Los datos han evolucionado, tanto por volumen como por tipología. Pero, ¿qué podemos esperar

a partir de ahora? En palabras de Alfonso Castro, “el análisis de datos es una técnica que se ha usado desde hace mucho tiempo. Todos recordamos conceptos como Business Intelligence, que lo que nos aporta es la primera parte de la analítica de datos. Pero, a partir de un conjunto de métricas se obtienen cuadros que permiten a las empresas tomar decisiones estratégicas. Pero esto ha evolucionado a una analítica de datos, que facilita, a partir de estas arquitecturas distribuidas de procesamiento y almacenamiento de datos, ser capaces de manejar algoritmos que no eran capaces de ser ejecutados en los plazos de tiempo necesarios. Ahora conseguimos ejecutarlos y generar modelos predictivos y prescriptivos que nos hacen anticipar qué es lo más probable que vaya a ocurrir, o nos capacitan para predecir las situaciones que podremos abordar en la empresa”.

En opinión de Santiago Moral, “uno de los cambios que más está costando en las empresas, es pasar de modelos puramente exactos a modelos de comportamiento aleatorio. Cuando acabas el programa sólo has empezado a trabajar, y esto es complicado para las unidades de negocio y de tecnología. Ahora, el científico de datos tiene que estar pegado al algoritmo para ayudar a este algoritmo a seguir aprendiendo y evolucionando. Pasamos de una informática determinista, con una cadena de valor en la que cuando uno acaba empieza otro, a una informática no determinista, donde la efectividad siempre es relativa, las me-

diciones tienen que ser permanentes. Otro gran cambio que nos viene, y que estamos a ver cómo lo desarrollamos, es el de la ética de los algoritmos. Como cada vez ponemos más decisiones de forma automática en algoritmos, y estos aprenden de los datos, si los datos tienen sesgo, el algoritmo va a tener un comportamiento sesgado, aunque el propio algoritmo no lo tenga. La calidad del resultado tiene que ver con la calidad de la entrada. Pero son dos modelos que nos cambian mucho la forma de pensar”.

POTENCIANDO LAS CAPACIDADES PARA ANALIZAR LOS DATOS

A nivel de tecnologías, según Santiago Moral, “vamos a arquitecturas basadas en que jamás se borra. Esto hace que podamos tener sistemas de procesamiento casi infinitos. Estamos mezclando tecnologías, on-premise o en la nube, con posibilidades casi infinitas de almacenamiento, capacidades enormes de procesamiento, y con modelos donde la aplicación de la ciencia de datos exige menos conocimientos de la propia ciencia de datos. Son tecnologías que se van mucho de los modelos clásicos de bases de datos, muy basadas en tecnologías No-SQL, integrando datos no estructurados con estructurados, modelos en los que nunca se borra... tecnologías como Elastic en Amazon o BigQuery en Google. Y todo esto con programación en diferentes lenguajes”. Para Alfonso Castro, “debe-

mos hacer una división clara entre la gestión de datos y el análisis de datos. Dentro del análisis de datos, las tecnologías son las relacionadas con algoritmos de IA que permiten esos modelos para predecir y abordar el análisis de datos. Dentro de estos modelos, hay dos, fundamentalmente: los asociados a Machine Learning y los asociados al Deep Learning. En el primer caso, entrenamos el modelo con algunos datos y usamos otros para realizar las predicciones. En el segundo, se basan en algoritmos de redes neuronales que toman como referencia el funcionamiento del sistema nervioso. A nivel de gestión del dato, tenemos que ser capaces de manejar todos estos elementos para que el entorno funcione correctamente. Y abordar diferentes elementos, como la calidad del dato, la seguridad del dato, la integración de datos de diferentes fuentes, cómo se almacenan, cómo se operan... En este punto, estamos trabajando en la creación de frameworks que generen procesos claramente definidos para que cada uno de los participantes sepan lo que tienen que hacer en cada momento. No son solo avances tecnológicos, sino también en los procesos”. ■

**Si te ha gustado este artículo,
compártelo**



#ENCUENTROSITRENDS

Maximizando el valor del dato

Tenemos una huella digital enorme. De hecho, el recuento de los datos del año pasado alcanza los 59 Zetabytes. Pero este número, de por sí muy grande, está creciendo con una tasa anual compuesta del 61%. Si esto lo trasladamos a las empresas, supone un gran volumen de información.

A sí las cosas, los retos que planean los datos a las empresas son muchos, y para saber cómo responde la tecnología a ellos se abordaron las posibilidades tecnológicas junto a Quique Sánchez Candorcio, director comercial para EMEA de Ikusi; Ramsés Gallego, International Chief Technology Officer de CyberRes, a Micro Focus line of business; e Israel Serrano, director regional para Sur de Europa de Scality, en la segunda mesa de debate de este Encuentro IT Trends, [Entendiendo la Era del dato: tecnologías y propuestas para gestionar la datificación](#).

RETOS CON LOS QUE ENFRENTARSE

¿Cuáles son las dificultades para almacenar, proteger, analizar o aprovechar toda la información que circula por las empresas? En opinión de Quique Sánchez, "son muchas, pero cuestiones como quién es el dueño del dato, la gestión distribuida, múltiples accesos, diferentes fuen-

Arancha Asenjo, IT Televisión

Quique Sánchez Candorcio, Ikusi

Ramsés Gallego, CyberRes, a Micro Focus line of business

Israel Serrano, Scality

itTRENDS #EncuentrosITrends

(De arriba hacia abajo) Quique Sánchez Candorcio (Ikusi), Ramsés Gallego (CyberRes, a Micro Focus line of business) e Israel Serrano (Scality). Clica en la imagen para ver el vídeo.



“Almacenar datos que no tienen utilidad tiene un coste, y hay que saber cuándo y cómo eliminarlos”

**QUIQUE SÁNCHEZ CANDORCIO,
IKUSI**

tes, un contexto hiperconectado, un volumen ingente... convierten en cruciales temas como el almacenamiento, la seguridad, la migración de los datos... y la capacidad de establecer cuál es el dato útil, porque no todos los datos son útiles. Almacenar datos que no tienen utilidad tiene un coste, y hay que saber cuándo y cómo eliminarlos. Estos puntos se convierten en cruciales, sobre todo para la seguridad”.

En palabras de Ramsés Gallego, “estas dificultades pasan por entender el contexto glo-

bal en el que ya no solo la pandemia, sino el mundo en general nos ha llevado. Pero, sin duda, la pandemia acelera la necesidad de un trabajo remoto. La dificultad está en entender que, si todos los datos no son creados de la misma manera o con la misma finalidad, por qué debemos protegerlos, gestionarlos, maximizarlos... de la misma manera. Si los datos no tienen el mismo valor, deberíamos acercarnos a ellos de forma desacoplada, de manera distribuida. Ciertamente es que con políticas centralizadas y simplificadas, pero con una voluntad de maximización no solo de un dato, sino de los datos en su conjunto, para que se convierta en información. Lo contrario del gobierno no es el desgobierno del dato, que de por sí es malo, sino la complacencia, asumir que no tiene el valor suficiente el dato que estamos creando y protegiendo”.

Desde el punto de vista de Israel Serrano, “hay que analizar si los datos que se están almacenando tienen valor o no, y la dificultad para las empresas es poder hacer ese análisis a priori. Muchas veces no sabes si tiene valor hasta que no lanzas el proceso de análisis. Lo más importante para las empresas es que sean capaces de hacer económicamente factibles este tipo de iniciativas. Este componente económico es muy importante, así como la decisión de dónde quieren guardar estos datos, con elementos económicos, técnicos y de legislación”.

ENTENDIENDO LA ERA DEL DATO

En un entorno empresarial cada vez más complejo, con realidades híbridas, una computación que se reparte tanto en la nube como en el extremo, con millones de dispositivos conectados, y con un acceso por parte de los usuarios desde cualquier lugar y en cualquier momento, la gestión de los datos es un elemento fundamental. A partir de una adecuada y correcta gestión de los datos, es posible tomar las mejores decisiones para el negocio, apoyándose en las tecnologías más adecuadas en cada caso y modelo, pero analizando las operaciones de cada cliente como punto de partida para transformar el negocio de la mejor manera posible.

INCREMENTO DE VOLUMEN Y DE COMPLEJIDAD DE LOS ENTORNOS

¿Qué herramientas serían básicas para poner cierto orden en la situación que enfrentan las empresas? Según Ramsés Gallego, “hay que aplicar principios básicos de efectividad y eficiencia. No se trata solo de hacer las cosas bien, sino de hacerlas



“La dificultad está en entender que si todos los datos no son creados de la misma manera o con la misma finalidad, por qué debemos protegerlos o gestionarlos de la misma manera”

**RAMSÉS GALLEGO, CYBERRES,
A MICRO FOCUS LINE OF BUSINESS**

correctamente. Otro axioma es que el trabajo es lo que se hace, no donde se está, de ahí que sean fundamentales herramientas de archivado, de resiliencia, de gestión y gobierno del contenido, de protección, de analítica avanzada...”.

Para Israel Serrano, “tenemos que hacer que los proyectos sean viables. Da igual que el dato tenga valor o no, si el proyecto no es viable, no lo vamos a saber nunca. Hay que ser capaces de establecer entornos de analítica de pruebas de concepto. Hay que facilitar que las empresas tengan resultados rápidos para poder evaluar si el proyecto tiene sentido. Además, tenemos que ser capaces de facilitar nuevos modelos de consumo de nuestros productos, movernos hacia modelos de pago por uso o de suscripción. Asimismo, como no todos los datos son iguales, no podemos tratarlos igual. Hay que analizar el coste por terabyte. No puedes pedir el mismo coste a todos los datos, porque para obtener el mismo valor necesitas más datos, y el coste debe estar ajustado al valor que van a proporcionar. Por último, si vamos a trabajar con datos no estructurados, no podemos estar trabajando con las mismas herramientas de las últimas décadas. Y aquí es importante el cambio de paradigma y mirar al almacenamiento de objetos”.

Desde la perspectiva de Quique Sánchez, “las herramientas son múltiples y variadas, pero es esencial la correcta organización de los datos, dónde tiene que estar para tener un acceso en tiempo real, y luego la correcta identificación y almacenamiento. Quizá es más sencillo con los datos estructurados, pero también hay que hacerlos con otros da-

it whitepapers **CONVERTIR LOS DATOS EN ACCIONES**

Turning Data into Action
The Evolution of Data and the Case for Information Management and Governance

RESEARCH BY
Moulayetou Khou
Security & Trust Products, LLC

APR 2017 | IDC Doc #A4470701

La información es el elemento que mueve las empresas en la Economía Digital. Son los datos la materia prima para poder tomar las mejores decisiones y aplicarlas al negocio como parte de la ejecución de la estrategia. No contar con una adecuada política o las correctas herramientas para la gestión y gobierno de los datos puede provocar desde decisiones erróneas para el negocio a brechas de seguridad con consecuencias catastróficas para la empresa.

tos, que tienen que confluir con los datos estructurados. Si bien esto es básico, también lo es la trazabilidad que hemos de tener de estos datos. Hablamos también de correlación y coherencia de los mismos, para empezar a hablar de información y no solo de datos”.



“Hay que analizar si los datos que se están almacenando tienen valor o no, y la dificultad para las empresas es poder hacer ese análisis a priori”

ISRAEL SERRANO, SCALITY

INNOVACIÓN EN TORNO AL DATO

En palabras de Israel Serrano, “hay mucha innovación en muchas áreas. En nuestro caso, en la infraestructura que soporta estos datos. La innovación es el uso de protocolos y estructuras de datos que se adaptan mejor al tipo de datos que se están generando. Para 2024, según IDC, hablamos de 138 ZB de datos generados, de los que el 80% van a ser no estructurados. Tenemos que ser capaces de dotar a las empresas de las

herramientas necesarias para analizar toda esta ingente información para extraer el verdadero valor. En términos de innovación, la pieza fundamental es el metadato, que me permite dotar de valor, de contexto de información adicional a un dato según va pasando por los diferentes procesos de negocio. Y esto permite a las organizaciones identificar si tiene valor. Por otra parte, cuando piensas dónde colocar el dato, no es algo que sea inamovible. Necesitamos que sea ubicuo y no tengamos la necesidad, a priori, de definir dónde está, salvo por criterios económicos, de negocio o legales, y tenemos que proporcionarles las herramientas para que puedan tomar la decisión adecuada”.

En opinión de Quique Sánchez, “el mayor impacto dentro de la gestión del dato lo está teniendo la Inteligencia Artificial, y es curioso, porque la IA tiene como punto de partida el dato, pero la mejora de la limpieza de datos para la propia calidad del dato se apoya en la propia Inteligencia Artificial. Por otra parte, podemos analizar dónde estamos obteniendo mayor beneficio del dato en base a la analítica. Y, pensando en la protección, encontramos la detección de amenazas analizando riesgos en tiempo real”.

En palabras de Ramsés Gallego, “en una era compleja y cambiante, todos hablamos de transformación y disrupción, pero hay que apostar por la innovación radical, haciendo las cosas de forma diferente. Hay que enriquecer el dato empleando

ALMACENAMIENTO DE DATOS EN CLOUD HÍBRIDA A ESCALA

La solución de almacenamiento escalable Scality RING, definida por software, permite a las empresas y proveedores de servicios cloud ejecutar servicios de datos enriquecidos, a escala petabyte, como aplicaciones web, vídeo bajo demanda, archivos activos, archivos de cumplimiento y nubes de almacenamiento privado.

IA y Machine Learning, y aprovechar los beneficios que aportan los algoritmos enriqueciendo los datos a cada paso. Cuando tienes inteligencia de datos para poder correlacionar los datos, es algo tremendamente innovador, un salto cualitativo y cuantitativo en la gestión del dato”. ■

Si te ha gustado este artículo, compártelo



Entendiendo la Era del dato: tecnologías y propuestas para gestionar la datificación



“El dato es un valor que hay que convertir en información para tomar decisiones”, Ikusi



“Es fundamental proteger a las personas y la información”, CyberRes, a Micro Focus line of business



“La tecnología debe ser capaz de manejar la ubicuidad del dato”, Scality