



SU EMPLEO CADA VEZ ES MÁS POPULAR EN LOS HIPERESCALA

# ¿LA GRAN PROMESA DE LOS CPD EFICIENTES?

POR CRISTINA LÓPEZ ALBARRÁN

*Su incursión en los centros de datos cada vez es mayor, pues permite mejorar la infraestructura y facilitar una gestión de datos más inteligente y automatizada. Se presenta como el camino hacia el CPD autónomo y eficiente, pero son los grandes del cloud los que están innovando en este campo. A su vez, este avance genera muchos más datos que los que hay que procesar, lo que inicia un círculo vicioso a tener en cuenta.*

La inteligencia artificial ya no es una tecnología de ciencia ficción. En 1997 se hizo famoso el computador Deep Blue por derrotar al campeón mundial de ajedrez Gary Kasparov. Años después, en 2011, fue el superordenador Watson el que ganó jugando en el popular programa televisivo Jeopardy y en 2014, Eugene, otro ordenador, superó con éxito el test de Turing haciendo creer a un interrogador que se trataba de un chico de 13 años. Pero más allá de estos hechos históricos y tan conocidos, a día de hoy, la IA está muy presente en nuestras vidas cotidianas, aunque no nos damos cuenta. Robot aspiradora, juegos de ordenador, chatbot y asistentes de voz e incluso smartphones equipados con funciones de habla o reconocimiento facial basados en algoritmos de aprendizaje automático. Eso es IA. La robótica, el machine learning, el big data... y todas aquellas cosas o conocimientos que consiguen que las máquinas puedan percibir, comprender, actuar y aprender para ampliar las capacidades humanas.

Sin embargo, su incursión no sólo se ha producido a nivel doméstico: en el ámbito empresarial su empleo continúa creciendo. Ligado a la mejora de la analítica, la eficiencia operativa y la experiencia del cliente, las organizaciones cada vez apuestan más por esta innovación, definiendo soluciones que responden a las necesidades de su negocio. Su popularidad coincide precisamente con una tendencia mayor a utilizar la analítica para mejorar el uso de los datos disponibles y la velocidad y precisión de la toma de decisiones.

Desde una perspectiva más específica, su entrada en el negocio de los centros de datos se ha guiado por la optimización de la eficiencia energética, el gran reto de estas instalaciones. Los costes de energía se disparan alrededor de un 10% cada año -en España mucho más-, lo que supone un mayor coste por kWh. Sólo en Estados Unidos, los CPD consumen más de 90.000 millones de kWh de electricidad en un año. Este uso alcanza

un nivel superior a escala global, ya que los data center de todo el mundo utilizan alrededor de 416 Teravatios de electricidad. La IA y el aprendizaje automático pueden otorgar numerosos beneficios al uso energético en estos enclaves mediante el mantenimiento predictivo y preventivo, controlando la temperatura y ajustando los mecanismos de refrigeración. Baste recordar el caso de Google y su sistema Deepmind, que ha permitido al gigante reducir el consumo de energía en sus CPD en un 40%. Cada cinco minutos esta solución obtiene una instantánea del sistema de enfriamiento del data center de miles de sensores, y con esos datos se entrena una red neuronal que realiza la predicción sobre cómo las diferentes elecciones afectarán el consumo de energía en el futuro.

Visto lo visto, la inteligencia artificial puede emplearse, por ejemplo, como herramienta para evitar el sobrecalentamiento de los equipos, consiguiendo ahorrar energía y reducir los costes. Igualmente se puede aplicar a la monitorización y control de las salas técnicas. Sus posibilidades son extensas, pues también permite mejorar la distribución de cargas de trabajo, haciendo que estas sean más predecibles y fáciles de

gestionar. Los hiperescalares ya la están implementando para estas funciones y el resto de actores empiezan a verla con buenos ojos a medida que mejora la tecnología y disminuyen los costes. Mediante estas técnicas se consigue hacer un mejor seguimiento del rendimiento del servidor, la utilización del disco y la congestión de la red. Se pueden buscar posibles fallos en el sistema, optimizar los tiempos de procesamiento y conseguir que la reducción de los factores de riesgo sea más rápida, colaborando a detectar y prevenir interrupciones de datos.

“Pronto, los centros de datos podrán automatizar casi todas las operaciones, desde la ciberseguridad hasta el mantenimiento y el monitoreo”, predice Jeff Kavanaugh, director del Infosys Knowledge Institute, una organización centrada en el análisis de tendencias comerciales y tecnológicas. “Sin embargo, el volumen y la gestión de las cargas de trabajo continuarán evolucionando a medida que la cantidad de datos aumente exponencialmente y, también, a medida que encontremos nuevos usos para la IA en la empresa”.

Otra de las grandes aportaciones en los centros de datos tiene que ver con la disminución y resolución de incidentes.

## 10 TENDENCIAS EN TECNOLOGÍA DE DATOS Y ANÁLISIS PARA 2021

Cuando estalló la Covid-19, las empresas que usaban técnicas de análisis tradicionales (que dependen en gran medida de grandes cantidades de datos históricos), se dieron cuenta de algo importante: muchos de estos modelos ya no eran útiles. Básicamente, la pandemia lo cambió todo, y muchos datos se volvieron improductivos. A su vez, los equipos de análisis y datos con visión de futuro están cambiando de las técnicas de inteligencia artificial tradicionales que se basan en “grandes” datos a una clase de análisis que requiere menos datos, o “pequeños” y más variados. La transición de grandes datos a datos pequeños y amplios es una de las principales tendencias de análisis y datos de Gartner para 2021:

1. IA MÁS INTELIGENTE, MÁS RESPONSABLE Y MÁS ESCALABLE
2. DATOS Y ANÁLISIS COMPONIBLES
3. EL TEJIDO DE DATOS COMO BASE
4. DE GRANDES DATOS A DATOS PEQUEÑOS Y AMPLIOS
5. XOPS (DATOS, MACHINE LEARNING, MODELO, PLATAFORMA)
6. INTELIGENCIA DE DECISIONES DISEÑADA
7. LOS DATOS Y EL ANÁLISIS COMO FUNCIÓN EMPRESARIAL
8. LOS GRAFOS LO RELACIONAN TODO
9. EL AUJE DEL CONSUMIDOR POTENCIADOR
10. DATOS Y ANÁLISIS EN EL BORDE

Fuente: Gartner 2021.

Aquellos data center que se autogestionan, hacen uso de aplicaciones de deep learning para predecir fallos antes de tiempo. Esto unido a sistemas de recomendación basados en machine learning posibilitan localizar y atajar cualquier problema con antelación.

La monitorización inteligente y el almacenamiento de datos son otros usos de la IA, que tiene el potencial de tomar decisiones basadas en datos.

Finalmente, también estaría la seguridad lógica, en tanto en cuanto se introducen sondas recolectoras de datos capaces de correlacionar logs que hacen posible detectar y bloquear automáticamente ataques difíciles de detectar, como las Amenazas Persistentes Avanzadas (APT). Esta tecnología identifica el comportamiento normal en las redes y detecta los riesgos cibernéticos en función de las anomalías y desviaciones en las infraestructuras.

## Almacenamiento inteligente

Detrás de la aplicación de deep learning e IA están los datos, que van más allá de la información limitada que se puede extraer de los registros y las métricas de las plataformas de hardware tradicionales. Actualmente, las empresas no almacenan su información únicamente en los CPD: también lo hacen en la nube híbrida. Además, con el edge computing, el consumo de esta se realiza en el mismo lugar en el que se produce. Los datos deben estar disponibles dónde y cuándo se precise, si no fuera así los costes subirían y la eficiencia bajaría.

En este variopinto escenario entra en escena un concepto relacionado con la incorporación de la IA en los CPD: la hiperconvergencia. Esta combinación de componentes virtuales y físicos de una infraestructura gestionados por software hace posible que sistemas locales, centros de datos externos y servicios en la nube se integren perfectamente y de forma segura. Permiten escalar el almacenamiento y la potencia de cálculo de forma flexible para adaptarse a las necesidades cambiantes de cada momento. Asimismo, conceden la oportunidad de mover grandes cantidades de datos a la nube y, cuando es preciso, transferirlos de un lado a otro entre sistemas locales y remotos.

## ACLARANDO CONCEPTOS

### ALGORITMO

Es una secuencia de pasos matemáticos o instrucciones computacionales. Es un conjunto de instrucciones automatizado. Puede ser una sola instrucción o una secuencia de instrucciones: su complejidad depende de cuán simple o compleja sea cada instrucción individual y / o la gran cantidad de instrucciones que el algoritmo necesita ejecutar.

### MODELO

En IA, un modelo se refiere a un modelo matemático que es capaz de procesar datos y proporcionar la respuesta esperada o el resultado de esos datos. Por ejemplo, si se aplica un algoritmo a un conjunto de datos, el resultado sería el modelo. Por lo tanto, el modelo es el resultado de uno o muchos algoritmos. Un modelo cambia si los datos introducidos en el algoritmo cambian, o si los mismos datos se alimentan a través de un algoritmo diferente. Otra distinción muy importante es entre los dos tipos principales de técnicas de IA que se utilizan en los CPD hoy en día: el aprendizaje automático y el aprendizaje profundo:

### APRENDIZAJE AUTOMÁTICO

Hay tres tipos principales de técnicas:

- **APRENDIZAJE SUPERVISADO:** Los humanos proporcionan un modelo y datos de entrenamiento. Los algoritmos toman los datos de entrenamiento y ajustan el modelo para que las entradas y salidas / respuestas estén más alineadas. A medida que se agregan más datos con el tiempo, los algoritmos mejoran aún más el modelo y pueden hacer predicciones razonables para las respuestas a los nuevos datos. El aprendizaje automático supervisado se utiliza con mayor frecuencia en centros de datos y otras industrias.

- **APRENDIZAJE NO SUPERVISADO:** Los algoritmos encuentran patrones o estructuras intrínsecas en datos no etiquetados. En algunos escenarios, las técnicas de aprendizaje automático no supervisadas se combinan con las supervisadas. En efecto, la salida del aprendizaje automático no supervisado puede

convertirse en los datos de entrenamiento para el aprendizaje automático supervisado.

- **APRENDIZAJE POR REFUERZO:**

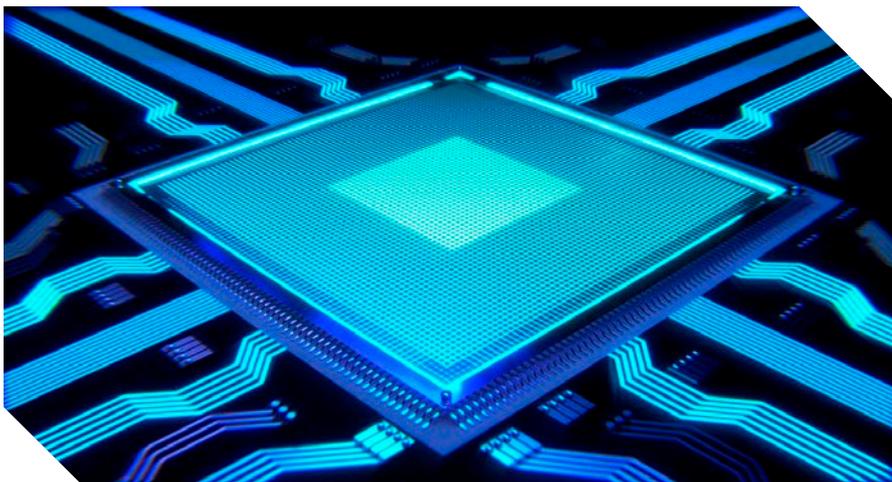
Los humanos proporcionan un modelo y datos no etiquetados. Cuando un algoritmo determina un resultado óptimo para los datos, se ve reforzado por una "recompensa" matemática positiva. (Un modelo de aprendizaje por refuerzo de código abierto de Google se llama apropiadamente Dopamina). Al proporcionar retroalimentación, aprende a través de diferentes variaciones. De estos, el aprendizaje por refuerzo es la técnica de aprendizaje automático más nueva.

### APRENDIZAJE PROFUNDO

Es un subconjunto del aprendizaje automático y utiliza múltiples capas de redes neuronales artificiales para construir algoritmos, basados en amplios datos, que encuentran una forma óptima de tomar decisiones o realizar tareas por su cuenta. Los humanos suministran datos y algoritmos de entrenamiento, y la computadora descompone estas entradas en una jerarquía de conceptos muy simples. Cada concepto se convierte en un nodo matemático en la red neural. En lugar de utilizar modelos de aprendizaje automático de humanos, el aprendizaje profundo utiliza los datos de entrenamiento como una red neuronal, que funciona como un árbol de decisión. Construye nuevos modelos a partir de su propio análisis de los datos de entrenamiento.

### ¿QUÉ TÉCNICA ES LA MEJOR?

Depende de la calidad y sofisticación del algoritmo, así como del modelo y los datos que se utilicen. Sin embargo, si todos estos parámetros son iguales, hay ciertas técnicas que son particularmente adecuadas para ciertos casos de uso. Algunos dicen que el aprendizaje profundo puede encontrar mayores niveles de ineficiencias porque no está sujeto a modelos conocidos. Por otro lado, el aprendizaje automático supervisado es más transparente (lo que facilita que los humanos expertos en el dominio validen los resultados) y, posiblemente, más rápido de automatizar. (Ver gráfico)



## IA EN EL CPD: 5 CLAVES

• EL AUGE DE LOS CHIPS DE IA PERSONALIZADOS

• EL MOVIMIENTO HACIA "AUTO" TODO

• TODOS NECESITAMOS ESTÁNDARES

• EL CIENTÍFICO DE DATOS SE VUELVE VIRTUAL

• LA IA IMPREGNA LAS OPERACIONES DEL CPD

Fuente: Western Digital.

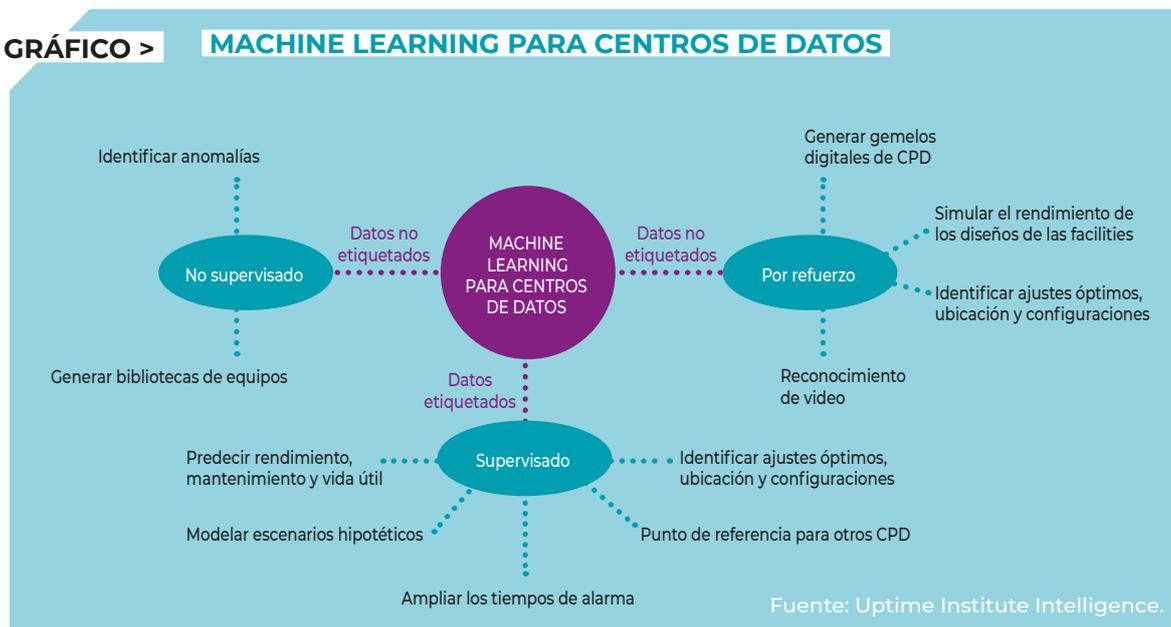
No obstante, desde el punto de vista del almacenamiento, muchas arquitecturas no fueron diseñadas para operar con soltura en estos contextos híbridos y es aquí donde la inteligencia artificial se presenta como una perfecta aliada. Su entrada en el mundo del 'storage' concede la posibilidad de que los sistemas se autogestionen, se optimicen e, incluso, en caso de error, se recuperen de forma autónoma. Pero la cuestión no sólo tiene que ver con el almacenamiento de los datos en sí, sino también con el hecho de poder garantizar la disponibilidad de la información en el momento y lugar que se demande.

La gran aportación de esta innovación a los data center es que les concede autonomía, pues gracias a la información que se obtiene de sensores (medicio-

nes, estados de los sistemas, fallos...) y al análisis predictivo de la misma, estas instalaciones se pueden anticipar a los problemas y solventarlos sin la necesidad de intervención humana. A través de simulaciones autorregresivas, estudio de usos históricos y patrones de comportamiento, los equipos aprenden de forma continua de la experiencia. A ello se une los dinámicos motores de recomendaciones que sugieren alojar los datos donde resulta más eficiente y aportan mayor accesibilidad. Se minimizan los cuellos de botella o las caídas de rendimiento, preparándolos para afrontar mejor los imprevistos en materia de consumo de almacenamiento. Puede predecir las incidencias de una aplicación y gestionarlas sin que afecten al resto. Con su conocimiento constante

### GRÁFICO >

## MACHINE LEARNING PARA CENTROS DE DATOS

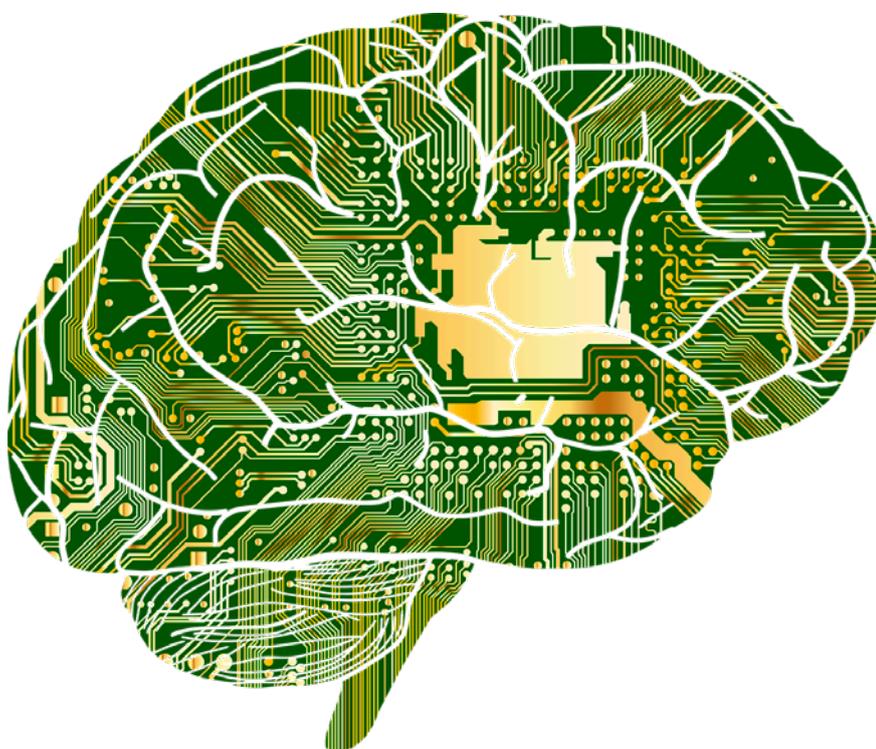


Fuente: Uptime Institute Intelligence.

## Eficiencia energética, seguridad, monitorización y automatización marcan la entrada de la IA en el CPD

optimizan los recursos, pudiendo mover una máquina virtual a otro host cuando el primero ya ha superado el número de licencias establecido o cambiando ajustes de red para evitar problemas de conmutación, por ejemplo.

Dotar de inteligencia estos activos se traduce en un ahorro de costes de infraestructuras del 30% hasta una reducción de los costes operacionales de almacenamiento de casi el 80%. Los hace capaces de adaptarse y ajustar la carga de trabajo en tiempo real, identificar y eliminar riesgos de seguridad y optimizar los datos a lo largo de todo su ciclo de vida.



### El pez que se muerde la cola

Si bien la IA se ha integrado en el centro de datos, el propio avance de la Inteligencia Artificial requerirá grandes infraestructuras de almacenamiento de datos. Es decir, esta tecnología se nutre de datos y esos datos hay que guardarlos, una razón que supondrá un claro impulsor del sector.

Pongamos números. Andrew Ng, científico jefe del Laboratorio de Silicon Valley de Baidu, dijo que entrenar uno de los modelos chinos de reconocimiento de voz de Baidu requiere no sólo

cuatro Terabytes de datos de entrenamiento, sino también 20 exaflops de computación, o 20.000 millones, 1.000 millones de operaciones matemáticas en todo el ciclo de entrenamiento.

Así pues, la aplicación de esta tecnología crea cada vez volúmenes mayores de datos, procesos pesados que requieren muchas horas de computación, grandes cantidades de ejecuciones repetitivas o experimentos exhaustivos, y esto repercute en el servidor, el almacenamiento y los dispositivos de red alojados en estos entornos críticos. Fundamentalmente en aquellos que no están diseñados con la flexibilidad re-

querida para amoldarse a los cambios. La HPC (High Performance Computing) es la solución común para los problemas inducidos por la IA, y su alta demanda de kW de potencia y refrigeración no parece estar disminuyendo en el futuro cercano. En lugar de HPC, algunos operadores continuarán confiando en servidores convencionales basados en CPU probados. Pero a medida que los sensores y las redes se preparan, la ley fundamental de la física que establece que para cada acción hay una reacción igual y opuesta, la realidad para los centros de datos es clara. Todos estos servidores basados en CPU necesitan energía, y se necesitarán más ser-

vidores para cumplir con los requisitos de procesamiento ante la proliferación prevista de datos.

La IA llega para resolver los problemas de energía y refrigeración que su propio aumento ha creado. Como el caso de Google con Deepmind, lo más probable es que veamos que más data center están reduciendo sus PUE a través de una mejor gestión centrada en los datos de sus infraestructuras de enfriamiento. Incluso para aquellas operaciones que están a pocos años de distancia de las bombas y torres de enfriamiento administradas por IA, exper-

tos del sector sostienen que continuaremos viendo una mayor dependencia de la energía a nivel de bastidor en forma de PDU inteligentes con capacidades remotas de administración de energía y monitoreo ambiental del CPD. Es donde los sensores de enfriamiento se comunicarán en tiempo real con los sistemas de gestión de enfriamiento para producir la mayor eficiencia posible.

En este sentido, el reciente anuncio de Google de su último hardware de IA, un ASIC personalizado llamado Tensor Processing Unit (TPU 3.0) que, en su diseño gigante, será ocho veces más potente que las TPU del año pasado con más de 100 petaflops obliga a replantear todo de nuevo. Empaquetar aún más potencia de cálculo en el silicio también aumentará la cantidad de energía para impulsarlo y, por lo tanto, la cantidad de calor, por lo que en el mismo anuncio la multinacional dijo que están cambiando a la refrigeración líquida al chip porque el calor generado por TPU 3.0 ha excedido los límites de sus soluciones de enfriamiento de centros de datos anteriores.

Pero es mucho más que eso. El machine learning se puede ejecutar en una sola máquina, pero gracias a la increíble cantidad de rendimiento generalmente se ejecuta en múltiples máquinas, todas interconectadas para garantizar una comunicación continua durante las fases de entrenamiento y procesamiento de datos, con baja latencia y absolutamente sin interrupción del servicio a nuestro alcance, pantallas o dispositivos de audio. Este ancho de banda debe distribuirse dentro y a través de múltiples instalaciones, utilizando diseños de arquitectura más complejos donde las redes creen una autopista para que toda la computación algorítmica fluya entre diferentes dispositivos y, en última instancia, regrese a nuestros receptores. Estas redes suelen ser de fibra y fuertemente malladas para garantizar que ningún servidor esté a más de dos saltos de red entre sí. Pero la demanda de ancho de banda es tal alta que se siguen innovando en eficiencia y latencias.

Así pues, a medida que la inteligencia artificial continúa remodelando el mundo de los negocios, las empresas dependerán de la misma para poder

remodelar la forma en la que operan sus centros de datos. Según IDC, para 2023, el 50% de activos TI en los CPD de las organizaciones podrán ejecutarse de forma autónoma gracias a funciones de IA integradas.

### Definir una estrategia

Gartner ya advirtió hace unos años de que más del 30% de los centros de datos que no implementan IA y aprendizaje automático no serán operativos ni económicamente rentables. Y es que, en definitiva, con IA en la gestión de la infraestructura de un CPD se anticipa en la predicción, prevención y resolución de cualquier contratiempo en las instalaciones, consiguiendo un aumento significativo de la eficiencia, mejoras operativas y una infraestructura más inteligente y fiable. Sin embargo, desarrollar soluciones basadas en ella sin una estrategia más amplia hacen que muchas iniciativas no consigan el potencial de rentabilidad que se esperaba de ellas. Como reconocen desde Uptime Institute, para satisfacer las necesidades del negocio hay que seleccionar casos de uso específicos y comprender los tipos de datos que influyen en los resultados de la IA, y luego validar tales resultados.

Al centrarse en estos casos, los éxitos que se pudieran lograr se pueden escalar y se puede extraer más valor de forma incremental. Los directores generales no necesitan ser expertos en IA, pero Uptime Institute

recomienda que comprendan la profundidad y amplitud fundamentales de la IA que se está aplicando. Hacerlo significa que pueden determinar mejor la cantidad de datos que se requieren y cómo la IA utilizará los datos, lo que será fundamental al validar los resultados y las recomendaciones.

Si la IA tiene un talón de Aquiles es la reacción de la tecnología, incluso a cambios relativamente sutiles en los sistemas y prácticas de los centros de datos. La mayoría de las técnicas de IA tratan de encontrar patrones estables, asumiendo un entorno fijo.

Si cambia el entorno de una manera que el modelo utilizado no lo puede ver, dará una respuesta incorrecta.

Una planificación cuidadosa antes de implementar modificaciones ayudará a mitigar esta preocupación.

Actualmente, los gigantes de la nube son los que están aplicando este avance en sus

