



Getty Images

[STORAGE](#) > [HIGH SPEED STORAGE](#)

## AI-Based Storage Helping Companies Get More Out of Their Data

AI-based storage enables organizations to analyze data quickly and intelligently, delivering insights almost instantaneously.

[Karen D. Schwartz](#) | Mar 10, 2020



What if you could think of storage more like a [self-driving car](#) and less like a hands-on, labor-intensive necessity? What if, like a self-driving car, your storage infrastructure could predict what you need, navigating roadblocks along the way?

---

that information to predict and prevent issues.

The technology at the heart of these charged-up capabilities is artificial intelligence. In fact, AI is having a moment, and according to experts, this isn't likely to change anytime soon. A recent survey from McKinsey found that nearly half of companies are using AI in some capacity today, but the vast majority expect their investments in AI to increase over the next few years.

When it comes to meeting today's storage requirements, [AI is fast becoming critical](#). It's what allows so much data to be analyzed so quickly and intelligently and helps avoid bottlenecks, availability issues and security concerns. AI-based storage allows IT staff to spend less time fighting fires and drives higher availability and greater productivity out of the infrastructure.

The goal, said David Wang, director of product marketing for HPE Storage, is creating an autonomous, AI-driven infrastructure that can deliver insights almost instantaneously.

"We want to get to a place where insights can drive immediate change," he said. "That's an argument for having an end-to-end AI pipeline from on-premises edge all the way out to the cloud."

## Looking at Things Differently

AI changes the mission of storage, and that means that organizations should be looking at storage and data differently, said Doug O'Flaherty, IBM's director of storage marketing.

"You have to stop thinking about storage as something you need for a database or a particular use case, and think of how you can use the access you have to data from different departments in different ways," he said. "If you can make that data

---

Along with thinking more broadly, it's important to take a more application-centric approach to storage instead of the traditional data-centric approach.

"In version 1.0 of this big data AI world, companies thought they had to be data-driven. So they focused on getting all of their data in a repository and all of their AI people in that group," explained Monte Zweben, CEO of Splice Machine, which focuses on distributed NoSQL database technology.

The result, he said, is that too quickly, these data lakes often turned into "data swamps," mostly because the people responsible for the business processes that could best use that data and the applications that are used in those business processes were left out of the equation.

By focusing on the application instead of the data, you'll be better able to match the storage and its capabilities to the business. For example, an insurance company that processes a lot of claims (the business process) would identify an application along with claims experts and application developers responsible for the claims system. By putting them all together, they can best decide how to use the data to create an intelligent claims processing system.

"It's about making the applications smart with data as opposed to trying to collect all the data in the world and feed it to people who may be interested in claims," he said. "It's a simple idea, but it can have a profound impact on how companies can operationalize artificial intelligence."

Building a successful AI-based storage infrastructure also means addressing each of three distinct phases in the [AI storage](#) pipeline: data ingestion (ingesting and normalizing data from different environments so you can look across it as a whole), training (using [machine learning](#) to examine the data to understand what's really inside of it) and inference (the stage of delivering insights).

---

processing. Put another way, AI at scale requires capacity at scale, retention and performance.

The ability to support very high storage capacity is critical, said George Crump, principal analyst of Storage Switzerland. Organizations rarely delete the data points they use to train their [AI workloads](#) because of the initial cost in acquiring it, he said. In addition, these data sets do not follow the typical data access model of decreased chance of use as they age. "The chances of the AI workload needing to reprocess old training data are almost 100%, so the entire data set needs to remain readily accessible," he added.

Long-term retention is equally important, especially as the amount of storage scales upward.

"We're implying that decisions will be made by machines based on the data that has been fed into them. That means the data can't be deleted. It continues to grow," O'Flaherty said. "And the more data you have, the better your accuracy and efficacy of applying AI grows."

In addition to simply storing more data, more *types* of data must be stored. That includes data about the data (metadata), which many believe is becoming one of the most valued commodities, especially when it comes to data governance.

The third requirement is high performance processing.

"Training an AI application is an iterative process, [and] improving accuracy is a process of repeated training, tweaking the AI algorithm and then training again," Crump said. "The faster the iteration occurs, the more accurate the developer can make the model, which increases the pressure on the storage infrastructure."

The key in most AI workloads, Crump said, is to ensure that the graphics

---

workloads tend to be very parallel, and a parallel, scale-out storage cluster may meet the challenge even with hard disk drives," he said.

## Taking Advantage of AI

For companies just starting out, it's possible to simply augment what you have by extracting data out of your systems and applying AI methodologies to chosen data sets to look for correlations. Eventually, though, you'll want to go deeper. Once you have tied your critical applications and systems together, getting to the real benefits of AI probably will require deploying new infrastructure and ways of approaching data.

The AI-based storage system you choose should have the intelligence to manage metadata quickly and store the right type of data on the right type of storage. If you choose to run your infrastructure on premises, Crump said you might be able to start with an all-flash storage system, but eventually, it makes sense to move to a mixed flash and hard disk environment. Often, the environment also will include software-defined storage that can automate the data movement between environments.

While some organizations prefer to keep everything on premises—especially those with sensitive workloads and compliance/data governance issues—others could benefit from a [cloud-based](#) AI/storage environment.

"A lot of this is happening in the cloud because of the shared compute power and data it needs," said Rochna Dhand, a senior director of product management at Nimble Storage, an HPE company. "The quality of the results you get from any AI model depends on the diversity and quantity of data available to train those models, so using a system that collects global data from the cloud can make a lot of sense."

apply them as updates locally behind the firewall.

Over time, Dhand believes that the technology will evolve and infrastructure management will eventually become completely hands-off. "You'll be able to predict and then prevent even more issues than you can today, and you will become more confident in those predictions and preventions," she said. "You'll not only be able to predict issues and identify what will solve the problem, but it will go a step further and actually determine the right solution and act on it."

TAGS: [ARTIFICIAL INTELLIGENCE](#)

## 1 COMMENT

## RELATED



### Why Computational Storage May Be the Next Big Thing

JAN 17, 2020



### Storage Briefings Roundup: Computational Storage, Micro Edge

JAN 13, 2020



### 2019 Data Storage Trends: Cloud-Based Storage Shone

DEC 19, 2019



### Lightbits Brings Clustered Storage to On-Premises Data Centers

DEC 14, 2019

ITProToday

[About](#)

[Advertise](#)

[Contact Us](#)

[CCPA: Do not sell my personal info](#)

[Privacy Policy](#)

[Terms of Service](#)

MENU

**ITPro**Today™

NEWSLETTER SIGN UP

Q SEARCH

LOG IN

REGISTER

f t in

.

tech

© 2020 Informa USA, Inc., All rights reserved

[Privacy Policy](#) | [Cookie Policy](#) | [Terms of Use](#)