

Transcript (edited):

Simon Robinson: We talk a lot about storage, but the important part is the data. That's the stuff we care about. Managing data according to policy in a cloud environment or multicloud environment is much more complex because it's much bigger and there are strict rules about what data can go away. So enabling that longer term I think is a really important problem that the industry is focused on addressing. One of the big challenges in the future is going to be: how do we get enough and the right sort of data into those models? So that has a couple of dimensions. One dimension is if I'm building massive GPU clusters, I want to be able to feed data very, very quickly. So I need highly parallelized capabilities on the storage side to keep those GPUs saturated. So that requires some very particular, very high performance storage capabilities.

[Related: [4 Trends Defining the Future of Enterprise AI](#)]

So that is definitely happening. More broadly, I think for the typical enterprise, they're now at the point where they're thinking, okay, so we're going to maybe have some models, build some models ourselves. We're going to be thinking about what these applications may be. They may be running at the edge. Where is the data going to be for those applications and how do we get data into those models? Storage today is very, very siloed and it's trapped into different locations. It's in different flavors of storage. You've got sand, you've got nasa, you've got object storage, you've got storage in the cloud, you've got storage at the edge, you've got storage in multiple data centers. What is the right data that I need? Right? Organizations struggle with visibility into their data. They always have, and there's always been kind of, well, what is the benefit of solving that problem? And I think AI could be a major enabler in persuading organizations to finally solve that problem. If they don't have visibility into their data, they're not able to optimize their AI environments. So I think that's where we see a real opportunity in some really interesting things going on in the industry in terms of innovation there.

[Related: [Enterprise IT Teams Jump Into AIOps](#)]

When we are talking about billions of parameters about generating insight from stuff that wasn't created by humans, that can only at scale, that's accessible globally, that can only be done in the cloud and very likely that will be the case for a great many years. Again, the question is how do we take some of those learnings and apply that to our on-premises environments? And it could be that actually we don't need a large language model. We actually might need many small language models. So how do we build an environment that is able to ingest elements of the public cloud, of the model that's in the public cloud and apply it for our own data? I think that's the key. The challenge for organizations, they can become sort of paralyzed by, I dunno where to start. The important point is just start. Just do something. Because the sooner you get going, the faster you're going to learn. The faster you're going to perhaps fail or make some mistakes and then learn again and do it again. And the faster you're going to get to a kind of maturity in terms of delivering some real capability to the business.