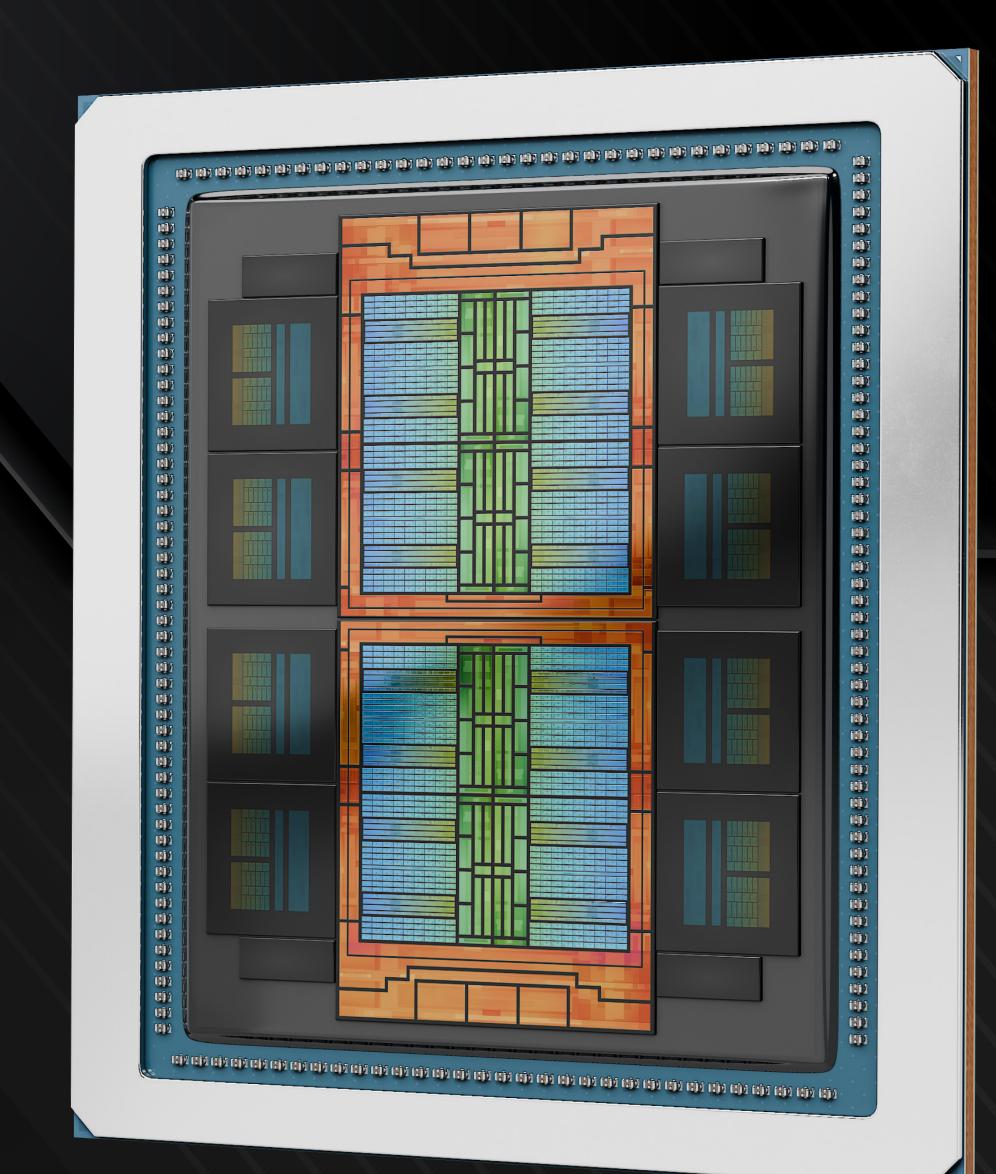
# 

WITH AMD INSTINCT<sup>™</sup> MI350 SERIES GPUs

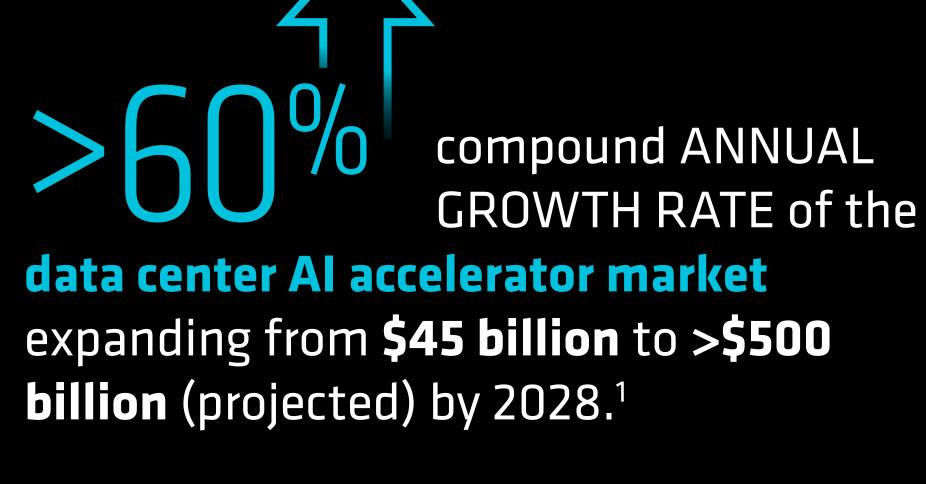




the optimal performance you need. AMD Instinct™ MI350 Series GPUs and the ROCm™ open-source software ecosystem help take you to the next level.

As you deploy your most demanding generative AI, training and HPC applications, get

AI EVERYWHERE



**AMD** MI350 Series

INSTINCT

C-SUITE WANTS MORE AI

of TECHNOLOGY EXECS say investments in data and AI are a top organizational priority in 2025, with 57.5% saying they are achieving exponential productivity gains, thanks to Al.<sup>2</sup>

> AMDA INSTINCT M1350



### LEADING HPC TRAIN AI FASTER

### offer up to 2x and 2.1x higher peak theoretical FP64 matrix performance,

respectively, with structured sparsity vs. the NVIDIA HGX B200 GPU. MI350-009

RESEARCH WITH EFFICIENCY

AMD Instinct MI350X and MI355X GPUs

### respectively, with structured sparsity

than the previous-generation AMD Instinct MI300X GPU. MI350-005

AMD Instinct MI350X and MI355X GPUs

offer up to 77% and 93% higher peak

theoretical FP8 matrix performance,

### B200 GPU. MI350-009 Get up to 2x the peak theoretical FP6 matrix PFLOPS performance with structured sparsity

when comparing the AMD Instinct MI355X vs. the NVIDIA GB200 GPU. MI350-008

Get up to 2.1x and 2.2x the peak theoretical FP6 matrix PFLOPS performance, respectively, with

structured sparsity when comparing AMD Instinct MI350X and MI355X vs. the NVIDIA HGX

SPEED UP INFERENCING GET SUPERPOWERS AMD Instinct MI350X and MI355X Embrace scale and efficiency. AMD

### FP4 matrix performance, respectively, with structured sparsity. MI350-010

DESIGNED FOR AI

A massive 288GB of HBM3E memory

and up to 8TB/s bandwidth deliver a

huge leap in performance over previous

8xGPU platforms offer up to 36.8 and

respectively, with structured sparsity

and up to 147.6 and 161.1 PFLOPs FP6/

40.3 PFLOPs FP16 matrix performance,

### than half of the top 20 supercomputer systems on the Green500 list.4

OPTIMAL ARCHITECTURE

Enhanced processing combined with

added next-gen FP6 and FP4 capabilities

positions the AMD Instinct MI350 Series

GPUs to deliver exceptional performance

for generative AI models, pushing the

powers five of the top 10 supercomputer

systems on the Top500 list<sup>3</sup> and more

of AI and HPC innovation.

FULLY CONNECTED

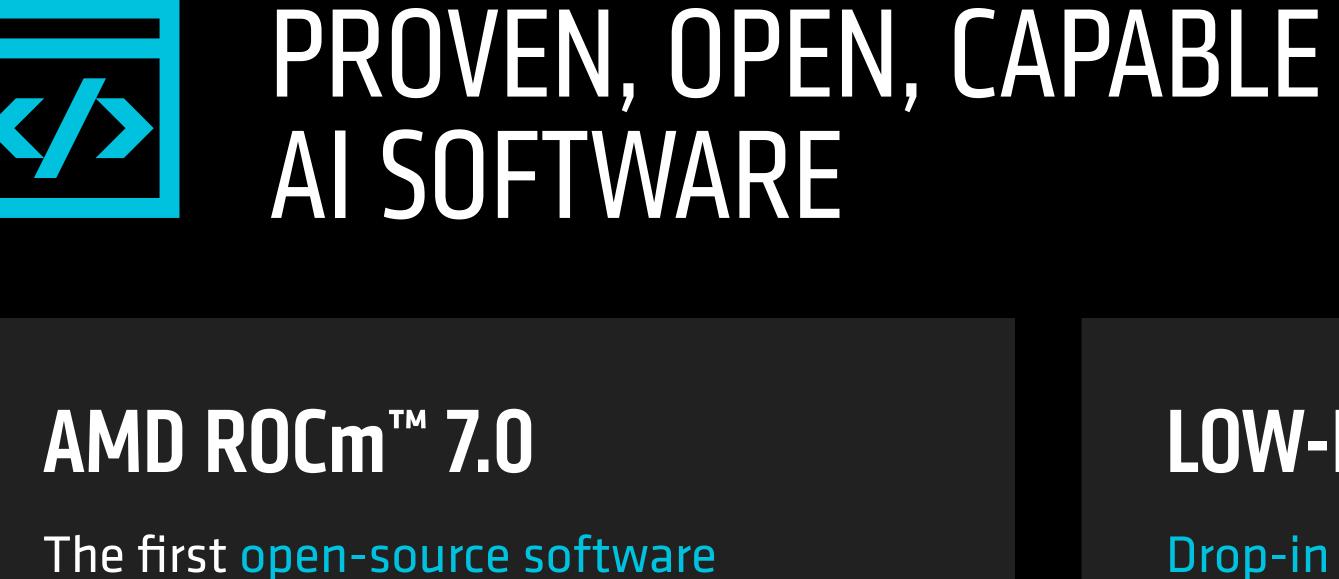
AMD Instinct generations.

REVVED-UP PERFORMANCE PER WATT Built on the cutting-edge 4th Gen AMD CDNA™ architecture, the AMD Instinct MI350 Series GPUs feature more powerful and energy-efficient cores, enhancing

boundaries of AI acceleration.

SPARSITY SUPPORT Native hardware support for sparse matrices on AMD Instinct GPUs helps save power, lower compute cycles and reduce memory use during Al and ML training.

performance per watt to drive the next era



ecosystem and programming toolset

for AI and HPC/hyperscale-class GPU

POWERFUL PARTNERSHIPS

development and deployment.

## evolving software needs.

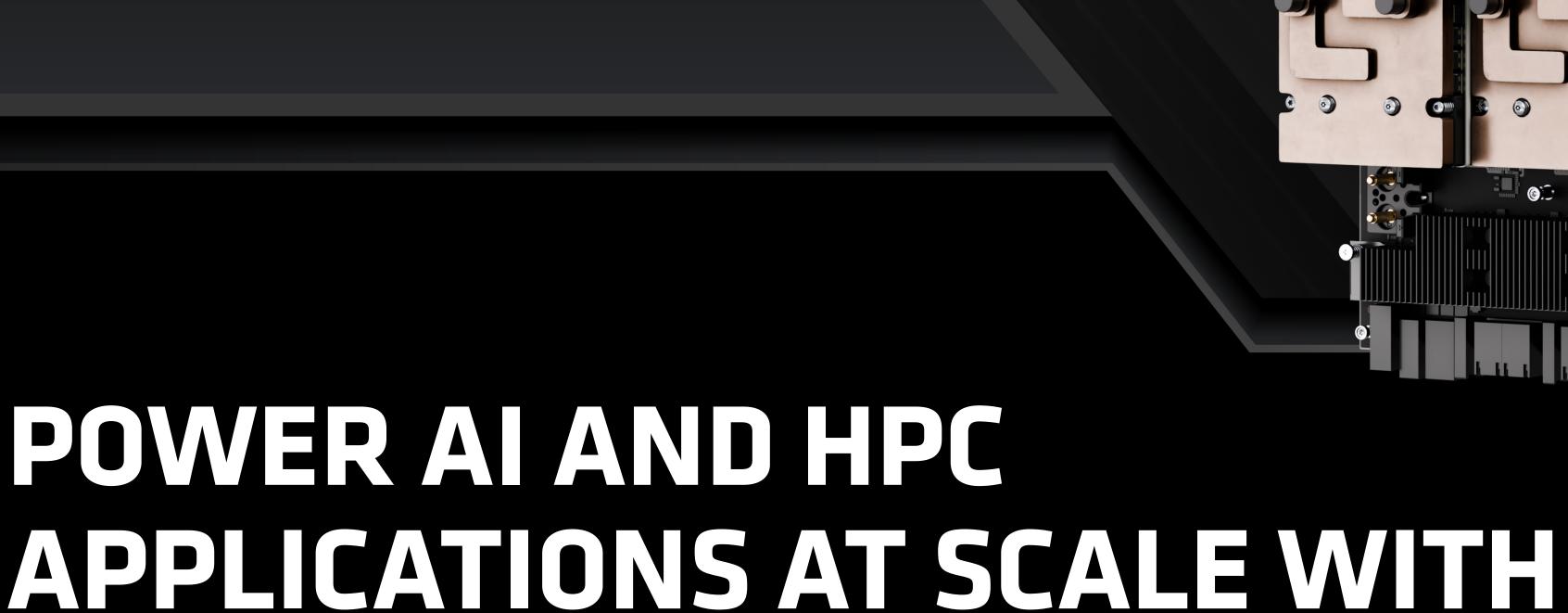
LOW-FRICTION ECOSYSTEM Drop-in support for major Al frameworks and HPC programming models, a key advantage for rapidly

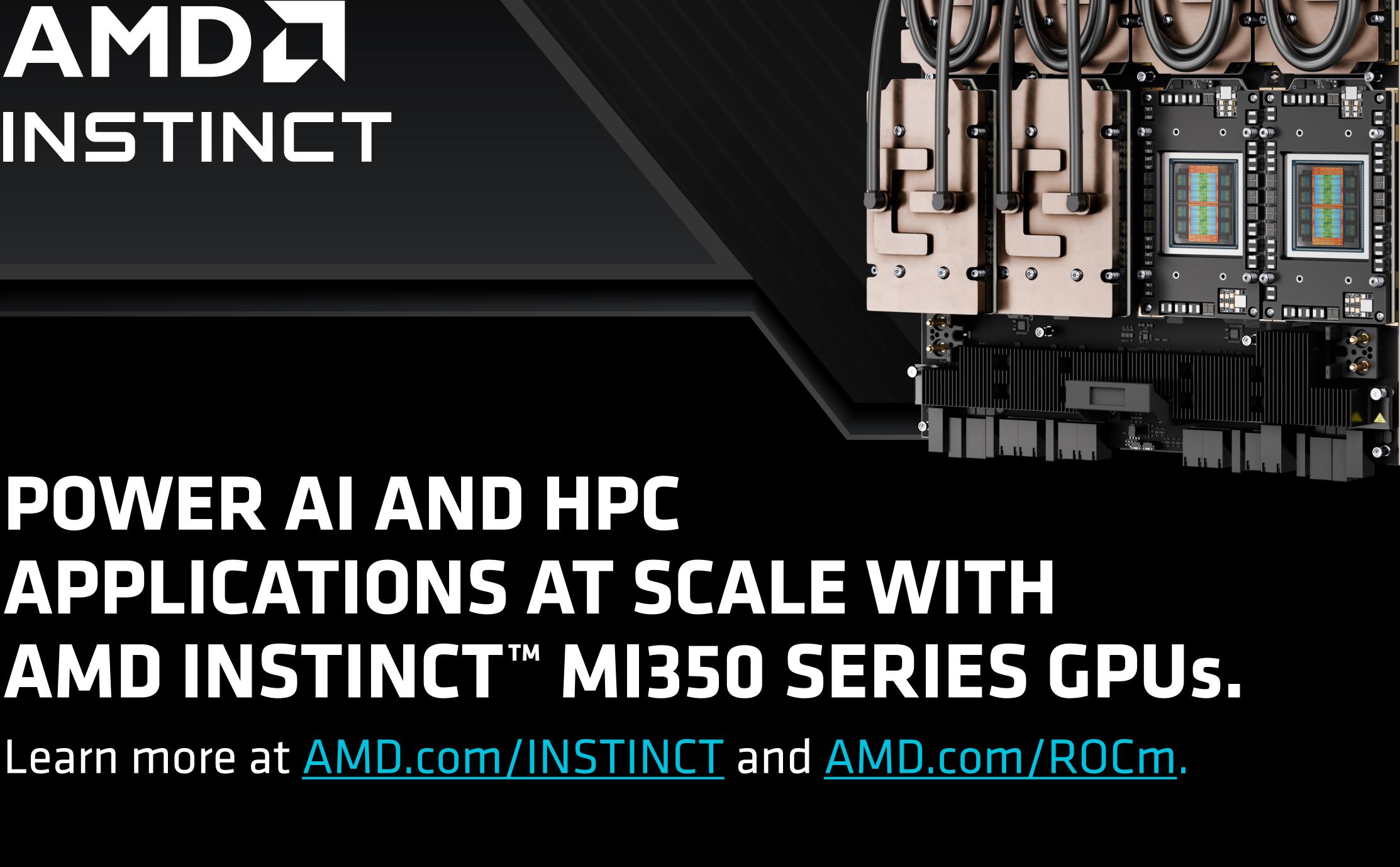
### participates in open-source cross-platform initiatives such as JAX, Kokkos, ONNX, PyTorch, Raja, Runtime, TensorFlow, Triton and vLLM to enable effortless AI and HPC deployment with minimal code changes.

AMD works with AI industry leaders, including Hugging Face, Meta and OpenAI, and



INSTINCT





# AMDA

1 Source: AMD Internal Data.

together we advance\_Al

25933/2025+AI+%26+Data+Leadership+Executive+Benchmark+Survey+120624.pdf. 3 Top500 list, November 2024, <a href="https://top500.org/lists/top500/2024/11/">https://top500.org/lists/top500/2024/11/</a>. 4 Green500 list, November 2024, <a href="https://top500.org/lists/green500/list/2024/11/">https://top500.org/lists/green500/list/2024/11/</a>.

2 Data & Al Leadership Exchange and datalQ, "2025 Al & Data Leadership Executive Benchmark Survey," 2025, https://static1.squarespace.com/static/62adf3ca029a6808a6c5be30/t/67642c0d40b42a7d7e684f49/17346181