

AI Hardware Atlas: Talk about it

Two or three questions for every exhibit. There are no answers printed here, on purpose.

Pick two or three exhibits and one question from each. Ten minutes is plenty.

For grown-ups: there are no right answers here. Ask a question, then ask "why do you think that?" The reasons matter more than the answer.

CHIP

The Transistor

hibot.space/hardware/transistor

A switch with no moving parts. AI is billions of them flipping.

- Q1** A transistor is just an on/off switch. What else in your house is an on/off switch?
- Q2** Early computers filled whole rooms with clicking switches. What would your day be like if computers were still that big?
- Q3** Switches keep getting smaller, but they can't shrink forever. What do you think people will try when they hit the limit?

The GPU Die

hibot.space/hardware/gpu-die

Thousands of small cores doing the same math at the same time.

- Q1** A GPU does thousands of small jobs at once. What chore at home goes faster when lots of people help at once?
- Q2** Is there a job where lots of helpers would actually make things slower? Why?
- Q3** Most big AI chips come from just a few factories. Is that risky, or is it a smart way to do it?

HBM: Stacked Memory

hibot.space/hardware/hbm

Memory chips stacked like pancakes and drilled through with copper.

- Q1** Stacked memory works because the data's trip is short. Where in your life does a shorter trip save lots of time?
- Q2** Often the chip sits waiting for memory, not doing math. Is a fast worker useful if their supplies arrive slowly?

The Accelerator Module

hibot.space/hardware/accelerator

GPU, memory, power delivery and a cold plate on one board.

- Q1** People say a lab trained on 10,000 GPUs. How much space do you think that many boards would take?
- Q2** One of these modules uses about as much power as a microwave running all the time. Where should that electricity come from?

The CPU

hibot.space/hardware/cpu

The general-purpose brain that runs everything around the GPUs.

- Q1** The CPU is the coordinator that keeps the GPUs busy. Who plays that role on a sports team or in a kitchen?
- Q2** A slow coordinator can leave expensive helpers waiting around. Have you ever been stuck waiting on one person? What fixed it?

Systolic Array (TPU-style)

hibot.space/hardware/tpu

A grid where numbers pulse through like a heartbeat.

- Q1** In this grid, numbers pass from cell to cell like a bucket line. What else works like a bucket line?
- Q2** A chip built for one job is faster at it, but bad at new jobs. When is being a specialist worth it?

BOARD

DRAM: System Memory

hibot.space/hardware/dram

The server's big, fast scratchpad. Forgets everything when the power goes off.

- Q1** This memory forgets everything when the power goes off. Why might forgetting be fine for a scratchpad?
- Q2** Fast memory forgets and slow storage remembers. If a computer could only have one, which would you pick, and why?

NVMe SSD

hibot.space/hardware/ssd

Flash storage: no moving parts, keeps data when the power's off.

- Q1** An SSD has no moving parts. Why might that matter for something that runs all day and all night?
- Q2** A training computer saves its progress every so often. What do you save as you go, and what happens if you forget?

Hard Disk Drive

hibot.space/hardware/hdd

Spinning mirrors and a flying needle: the cheapest way to keep a lot of data.

- Q1** A hard drive is a spinning disk with a tiny arm floating above it. What other spinning things store sound or pictures?
- Q2** Big storage systems keep extra copies so one broken drive loses nothing. How many copies of your favorite photo would you keep?

Network Card + Optics

hibot.space/hardware/nic

Turns electricity into flashes of light and back again.

- Q1** This card turns electricity into flashes of light. Why do you think light is used for long trips instead of plain wires?
- Q2** Thousands of chips must keep sharing results while they learn. What happens in a group project when people stop sharing?

Network Switch

hibot.space/hardware/network-switch

The traffic cop: dozens of ports, one giant routing chip.

- Q1** The switch is a traffic cop for data. What makes a good traffic cop on a busy road?
- Q2** If every chip talks at the same time, the network jams. What rules would you make so everyone gets a turn?

Power Supply

hibot.space/hardware/psu

Turns wall power into the steady DC the chips want.

- Q1** One AI server can use as much power as several homes. What uses the most power in your home?
- Q2** Some power always turns into heat instead of useful work. Why would wasting a little less power save a lot of money?

Liquid Cooling Loop

hibot.space/hardware/cooling

Water carries heat away far better than air can.

- Q1** Water carries heat away far better than air. Where have you seen water cool something down faster than air?
- Q2** One way to cool uses lots of water, another uses more electricity. Which would you pick for your town, and why?

SERVER + RACK

The AI Server

hibot.space/hardware/server

Eight GPUs, two CPUs, eight network cards, one very loud box.

- Q1** This box holds eight GPUs and costs about as much as several houses. Why do you think it costs so much?
- Q2** A small computer with one GPU is great for learning. What would you build if you had one at home?

The Rack

hibot.space/hardware/rack

A steel cabinet: servers stacked like drawers, with power, water and network threaded through.

- Q1** An AI rack can need its own water pipes. What other things in a building need their own plumbing?
- Q2** Older buildings weren't made for racks this hot. Is it better to fix up old buildings or build new ones? Why?

FACILITY

The Data Center

hibot.space/hardware/datacenter

A building that is mostly power plant and plumbing, with computers inside.

- Q1** Most of this building is power and cooling gear. Why do the computers inside need so much help?
- Q2** If a data center came to your town, what would you want to know before deciding how you feel about it?
- Q3** A data center can bring tax money and some jobs, but needs lots of power and water. How would you weigh those?

DATA

The Data Lake

hibot.space/hardware/datalake

Where raw data pours in and comes out clean enough to train on.

- Q1** Cleaning data means throwing out junk and private details. What would you throw out before teaching a robot about your school?
- Q2** AI learns from huge piles of writing and pictures people made. Should people be asked first? Why or why not?

HOW IT ALL WORKS

A Training Run

hibot.space/hardware/training-run

Thousands of GPUs learning one model together for weeks.

- Q1** Thousands of GPUs work for weeks to teach one model. What have you learned that took weeks of practice?
- Q2** Starting from a model someone else trained takes hours instead of months. When is it fair to build on someone else's work?

Your Prompt's Round Trip

hibot.space/hardware/inference-path

What happens in the second after you hit send.

- Q1** Your question travels to a far-off building and back in about a second. What else do you know that moves that fast?
- Q2** A small AI can run on your phone without sending anything away. When would you want that instead of a big one?

The Software Stack

hibot.space/hardware/software-stack

Eight layers between 'hi' and a transistor flipping.

- Q1** Better software can make the same chip about twice as fast. When has a better method made you faster at something?
- Q2** You can use someone else's AI over the internet, or run your own. What are good reasons to pick each one?