NEBIUS

2025 Sustainability Report

Contents

Sustainability highlights

Disclosures referencing ESRS

Highlights 2025

About Nebius

Nebius is the AI cloud company. We deliver a unified platform that spans the complete AI journey, from data and model training and tuning to production runtime and deployment.

Founded on deep, in-house technological expertise, we design and operate high-performance platforms that run reliably at scale. We serve startups and enterprises across a range of industries, removing bottlenecks in the AI industry with a specialized AI cloud that meets the demand for sustainable, trustworthy, and resilient infrastructure.

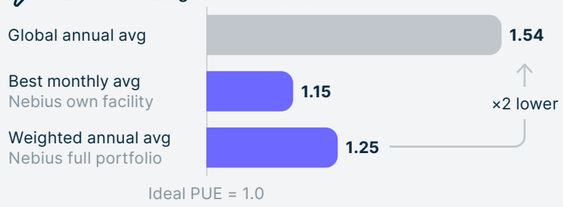
How we build

Energy efficiency and avoided consumption[1]

102 GWh
combined avoided electricity consumption = annual electricity use of 9,800 American households

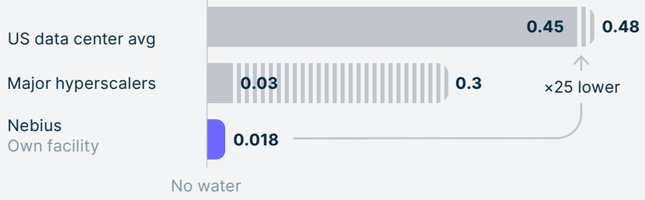
~20%
more AI compute per watt delivered by a Nebius-designed server

PUE — Power usage effectiveness



Water efficiency[2]

WUE — Water usage effectiveness, L/kWh



Water use avoidance when cooling high-density GPU racks, without sacrificing performance

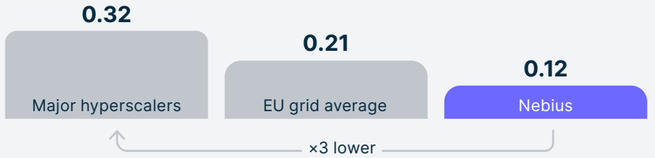
Free cooling
filtered outside air used

+

Closed-loop liquid cooling
water circulates in a sealed system. No intake from, or discharge to, external water sources

Carbon-efficiency

Carbon intensity[3], $tCO_2e/MWh



Low-carbon energy across portfolio

80%

2025 — EU / EEA

100% renewable energy sourcing for sites in France, Iceland, UK.

Finland-1: **100% renewable** from mid-2026

2026

328 MW behind-the-meter fuel-cell capacity bringing **on-site low-carbon** power generation, reducing grid dependency

Reliability

56.6 hrs
mean time between failures on a 3,000-GPU cluster vs. ~9.8 hrs industry-typical

168,900 GPU-hours on 3,000-GPU cluster converted to wall-clock hours, more details in the article

Enterprise-grade cybersecurity standards
6 ISO cybersecurity certifications, SOC 2 with HIPAA, NIS2, DORA, CSA STAR Level 1

Obtained in 2025 and 2026 (YDT)

How we create impact

Supported communities

19.5 GWh
exported to district heating +28% YoY

10%
heating cost reduction for residents

4K tCO_2e
heat production-related emissions avoided by the municipality

Free of charge
exports starting 2026

18,000+ people
people reached through AI education provided by Nebius Academy

NEBIUS ACADEMY

Opportunities for AI innovators

440 startups
with $2.1 million worth of compute credits

16 institutions
receiving research grants

Including UCL, Stanford, Caltech, UCSF, USC, LIMS

(1) Avoided electricity consumption and PUE calculation methodologies, and relevant references, are provided in the Efficiency by design chapter. Figure for Nebius's own facility reflects the Finland-1 data center, located in Mäntsälä, the only owned facility operational for the full year in 2025.
(2) Nebius's figure reflects water usage at the Finland-1 data center, the only owned facility operational for the full year in 2025, where liquid cooling was deployed. Peer figures are publicly reported global WUE values from major hyperscalers for the same period.
The US average is a projected 2025 figure from Lawrence Berkeley National Laboratory's 2024 United States Data Center Energy Usage Report.
(3) Figure reflects Scope 2 (location-based) over total electricity consumption, comparable figures for major hyperscalers are taken from the study The carbon and water footprints of data centers and what this could mean for artificial intelligence;
EU grid average from Ember Global Electricity Review 2025.

Sustainability vision

In conversation with our leadership team

A note from Daria Mukhortova, Head of Sustainability

The month we published this report, Nebius turned two. It is also our second sustainability report. For a company this young, there is already a great deal to say about what we've built and about how we've built it.

We invited the people leading this work, who also have been with the company from the very beginning and have seen the journey, to discuss what often stays unseen about sustainability — the choices, the tradeoffs and a lot of figuring out that shape the company we are building.

What follows is an edited version of those conversations. The commentaries have been condensed for clarity and length while remaining true to their original content and speakers' intent.

You have built infrastructure for decades and witnessed major technology waves throughout your career. How does it feel that what you do is suddenly at the center of every AI debate?

Andrey Korolenko
Chief Infrastructure & Product Officer

"There were a few technological shifts that changed everything: the steam engine, the internet, and now AI. Every one of them needed infrastructure. Factories. Railroads. Supply chains. People. The difference is time. The steam engine took decades. The internet took decades. AI will probably make the same shift in five to ten years. **The challenge is not building the infrastructure — it's the pace and the scale.**"

Danila Shtan
Chief Technology Officer

"Humanity hasn't seen anything on this scale for at least a hundred years. If AI is going to change every industry and every household, and we all work under the assumption that it will, then the infrastructure underneath it has to be extremely reliable. **We're talking about systems that will support major business decisions at a scale we've never seen before.**"

Building AI takes more than building technology, it is also about building trust. From your experience, how does a company earn it?

John Boynton
Chairman of the Board

"**By doing what it says it'll do, even when no one's watching yet.** Capital and velocity are table stakes at this point. What differentiates over a decade is whether the people inside the company actually believe in what they're building, and whether they can prove it with their choices."

Tom Blackwell
Chief Communications Officer

"By doing a better job of explaining and doing, and it's a process. None of this happens overnight. In a very new industry, a level of skepticism is not just natural but valid. Even substance alone isn't enough. We have the substance: Andrey and his team are extremely efficient in how they build, and our energy and heat metrics give us a good story to tell. But every new region, country or state brings new questions. And **building trust is only possible by answering them openly and consistently.**"

When forced to choose between discipline and ambition, you choose ambition. Then you make efficiency the core focus of your teams — which may be the most disciplined choice of all. How do you reconcile the two?

Danila Shtan
Chief Technology Officer

You don't want to waste GPUs running at peak power while doing nothing. **Our goal is to get every last compute cycle out of the hardware** — to fit more workloads, more user sessions and more model responses onto the same GPUs. Andrey and his team build the capacity. Then the software layer takes over. Our job is to fit as much as we possibly can into the infrastructure without wasting compute cycles.
In a sense, sustainability and efficiency go hand in hand. **At the scale AI infrastructure operates, efficiency immediately becomes business efficiency.**"

Andrey Korolenko
Chief Infrastructure & Product Officer

"The more we optimize through the stack, the more value we create.
When you're building the stack, you're building a solution for a wide variety of customers and businesses. You have to build reliably, with the ability to scale, while ensuring alignment with the whole environment and community, everything around you.
The scale of these projects is huge.
The impact is huge. Almost every effort to make the platform better [more efficient] pays back immediately. We have a very high pragmatic motivation to do better."

John Boynton
Chairman of the Board

"For us, efficiency and responsibility were always pointing in the same direction. A server that runs without interruption and draws less power is also a better product. Energy discipline that reduces your footprint also frees capital. Those aren't tradeoffs - they're the same goal."

One thing we hope readers understand about the company that the numbers alone cannot show

John Boynton
Chairman of the Board

"That the **people inside this company were asking these questions before anyone required them to.** That's not a small thing. It tells you something about character - about what a team values when there's no external pressure shaping the answer."

Andrey Korolenko
Chief Infrastructure & Product Officer

"I'm proud of what we've built. **That's practical engineering.** Years ago in Finland, we designed hardware that could operate at much higher temperatures, so we didn't need to rely on energy-intensive cooling. We recover much of the heat generated by the compute and feed it into the city's district heating system. In other words, we use the energy once for AI and then again to heat people's homes."

Tom Blackwell
Chief Communications Officer

"There are many companies involved in the AI cloud buildout, and not all are building the same way. It's **important to have a more nuanced view** of who does what and how."

Danila Shtan
Chief Technology Officer

"The numbers show the results, but **not the mindset**. We don't treat sustainability as separate from the business, it's **essential part of what we do**."

Daria Mukhortova
Head of Sustainability

"**We're a new industry, and we don't have everything figured out** — and that's quite alright. We are a new industry, and we are in this remarkable moment where everyone is learning, and **we are confident we are learning well**."

Sustainability agenda

Pillar 1

Sustainable AI starts with the efficiency of the underlying infrastructure. To deliver more compute per unit of energy consumed, we invest in vertical integration across hardware, data center operations, orchestration, and software, maximizing efficiency gains at and across different levels of our technology stack.

- Power-efficient server and rack design
- Higher operational temperatures
- Extended hardware lifecycles
- No water-intake cooling
- Low-carbon energy sourcing
- Integrated heat recovery systems
- Software-driven workload orchestration

Pillar 2

Breakthroughs with AI require more equitable access to the systems that power it. We use our AI expertise to serve local communities, expand access to cutting-edge tools and education, power disruptive research, and help the next wave of innovators develop solutions to real-world challenges.

- Community benefit initiatives
- Research and academic partnerships
- Applied AI training and education
- Infrastructure for critical sectors
- Local economic development
- Start-up and builder support

Pillar 3

Reliable computing keeps AI workloads running without failure, reruns, or data exposure. Customers depend on us to deliver stable performance, protect critical workloads, and support continuity when issues occur. We combine robust governance with software features that improve recovery, visibility, and user control over access, activity, and security.

- Robust cybersecurity framework
- Confidential computing controls
- Data sovereignty measures
- Resilient 24/7 infrastructure
- Performance observability
- Global standards alignment
- Risk management and response

Key outcomes

For customers: Greater cost efficiency, and more reliable compute, lower workload emissions.

For local communities: Less pressure on local energy and water resources, added value from recovered heat supplied into local energy systems.

For our business: Lower TCO, effective use of capital, reduces exposure to energy constraints and climate risks, faster scaling.

For innovators and researchers: Faster progress to testing and deployment, with easier access to compute and technical support.

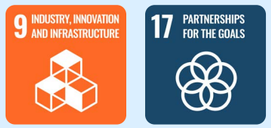
For local communities: Economic participation, workforce readiness for the AI economy.

For individuals: Broader access to AI skills, tools, education, and workforce development opportunities.

For customers: Trusted infrastructure, stronger security, and fewer service disruptions across workloads.

For partners and regulators: Confidence through strong governance, compliance, and responsible AI practices.

For our business: Stronger foundation for scaling across customer segments, with less operational friction, deeper customer trust.



Efficiency by design

Efficiency by design

Demand for AI is accelerating. The infrastructure requirements behind it are becoming more demanding. As a leading AI cloud, we see this as both an engineering and a business challenge. Our focus is on scaling our capacity to deliver high-performance computing power, while avoiding unnecessary overheads in resource consumption, or system complexity.

For us, efficiency is a starting point. Every component of our AI cloud, from hardware to the data center systems and software, is designed so that more energy goes into valuable compute. That means designing a full-stack system where everything works together. The aim is to enable the positive potential of AI without losing sight of the resources required to power it: delivering more compute with less energy and managing the footprint behind every workload.

Our advantage lies in how we operate across the full stack, enabling us to optimize how components interact rather than treating them in isolation. Small gains at the level of individual component compound across the system, turning incremental improvements into cost and energy optimizations.

The result is infrastructure that delivers consistent, high-performance compute while using fewer resources. Efficient systems are more reliable, more accessible, and better suited to integrate with their environments, whether that’s local energy grids or the local communities they serve.

We design the full stack in-house to optimize every component of AI cloud

This means we can choose exactly where to innovate or how components should interact.
We can also balance trade-offs to deliver maximum efficiency, turning individual improvements into compound gains.
For example, our custom hardware operates at higher temperatures, increasing system reliability and drawing less power for cooling.

Layer	Efficiency gains
AI Operations	Lower power draw from right-sized inference
Machine learning operations	Avoided energy use through faster training and inference cycles fewer failed job
Infrastructure-as-a-Service for AI-native workloads	Reduced idling through multi-tenancy and virtualization
Data center systems	Lower energy overheads; Heat reuse potential; Close-to-zero water use; Capex and opex savings
Servers & Racks	Hardware longevity; Higher operational temperatures; Ease of maintenance

Calculating efficiency gains from AI cloud operations

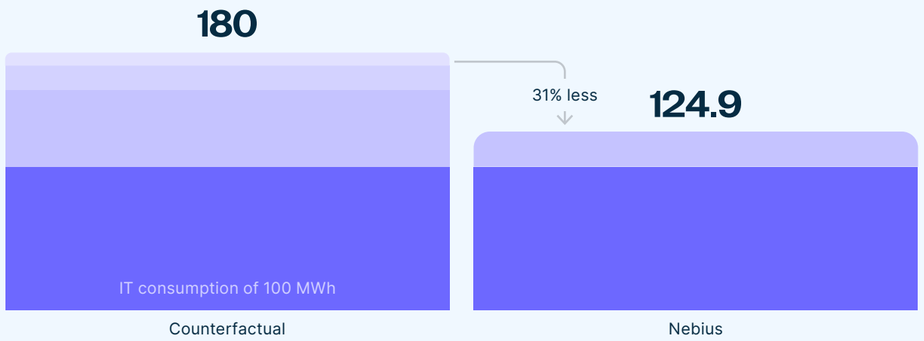
Absolute electricity consumption grows as Nebius scales. Another useful measure is relative efficiency: how much energy we consume compared to what an equivalent AI cloud would have consumed to deliver the same compute output.

- Compound effect
- Hardware gap
- Facility overhead
- IT consumption of 100MWh
- Nebius portfolio total
- Heat recovered

55 MWh

Avoided per 100 MWh IT consumption in 2025[1]

Per 100 MWh of IT consumption, MWh



Two layers, one compound effect

Hardware: Nebius-designed servers consume ~20% less electricity per unit of compute than the most popular off-the-shelf alternatives. This reduces the IT load drawn from the grid across both owned and colocation facilities.

Facility: For every MWh of IT load, a data center adds overhead energy for cooling, lighting, and power conversion. PUE measures this ratio: a PUE of 1.54 means 54 MWh of overhead for every 100 MWh of IT load. Nebius's data center portfolio-wide weighted average PUE was 1.25 in 2025[2]

Compound effect: Because Nebius servers draw less IT load to begin with, the PUE advantage applies to a smaller base. The overhead an average-PUE operator would add on top of the additional IT load they would need — but Nebius does not — is a third, distinct source of avoidance.

102 GWh

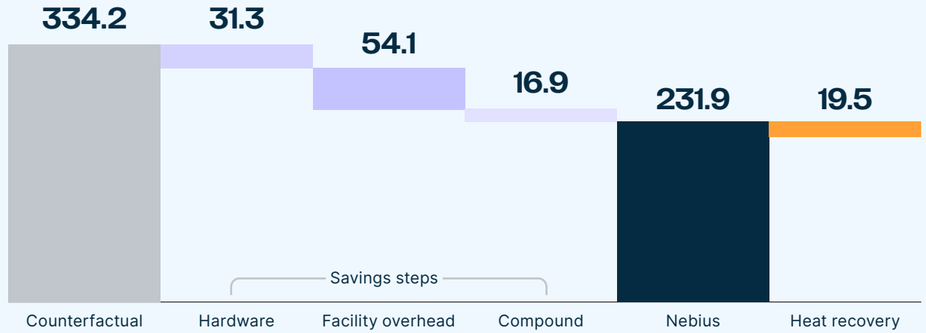
Total electricity avoided by Nebius portfolio[1] in 2025

~full year of continuous operation of a smaller facility running 24/7[2]

~full year of consumption of over 9,800 American households[4]

over 2/3 of New York City's daily electricity demand[3]

Portfolio total, GWh



Finland-1: a fourth dimension

At our Finland-1 data center, 19.5 GWh of heat — a byproduct of server operation — was exported to local district heating networks in 2025, equivalent to 20% of the site's annual electricity consumption. Unlike the avoided-consumption figures above, heat recovery does not reduce electricity drawn from the grid. However, it converts a consumption byproduct into a useful output. Both metrics are disclosed separately for this reason. More details on data center heat recovery as part of data center systems design is provided in the Heat recovery section.

19.5 GWh

heat exported to district heating network in 2025[6]

(1) Avoided consumption quantifies the electricity Nebius did not consume in 2025 relative to a counterfactual, or the same IT workloads delivered on industry-average infrastructure. The counterfactual is a constructed benchmark built from two main assumptions: Nebius-designed servers, which make up a significant part of the fleet in 2025, draw ~20% less electricity per unit of compute than the most popular off-the-shelf alternatives (empirical benchmarking), and the counterfactual operator runs at the global average PUE of 1.54 (Uptime Institute 2025). We use the global average because it is the most widely cited reference. A peer-specific benchmark based on AI cloud operators would be more precise comparison, but consistent and publicly available efficiency data for this segment does not yet exist at sufficient quality to support a defensible calculation; we intend to update the methodology when it does. The avoided total decomposes into three components: hardware savings, facility savings, and a compound effect. The model holds server utilization of the counterfactual operator consistent with Nebius utilization. If Nebius achieves higher average utilization per server, delivering more compute output per MWh of IT load, the counterfactual operator would require proportionally more servers and more energy to match the same workload. This effect is not currently quantified. Heat recovered at the Finland-1 site is disclosed separately and excluded from avoided consumption.
(2) Small site comparison assumes a typical 10-15 MW IT capacity data center at portfolio average PUE and 90% utilization
(3) NYC reference: New York City consumed ~49,700 GWh of electricity in 2024 (~5.7 GWh per hour on average). 99 GWh avoided is equivalent to roughly 17.5 hours of that demand. Source: NYISO 2025 Gold Book, Table I-2.
(4) American household consumption: according to EIA data, the average residential electric-utility customer in the US purchased 0.0108 GWh of electricity in 2022. Applying EIA's recently projected 2% annual growth rate (likely above actual growth for the period) gives an estimated 2025 consumption of 0.0114 GWh.
(5) Data center portfolio weighted average PUE is calculated using the methodology described in the PUE performance section.
(6) Heat recovery is metered at the point of handoff and reported gross — electricity generating the heat is fully counted in consumption. Not included in avoided-consumption figures.

Hardware

Our in-house R&D teams designs custom servers to make them simple, reliable, and energy-efficient. By rethinking every design feature, such as thermal and power designs as well as maintenance simplicity, we optimize their overall performance.

~20% more AI output

Per watt consumed by a Nebius server[1]

Up to 45°C

operating temperature of liquid and up to 40°C — of air used by cooling systems. Elevated temperatures lower cooling demand

Cable-less board connections

Fewer physical connections reduce failure points, lower material use, and improve signal integrity

Tool-less design

for fast maintenance. Components can be replaced without tools, enabling easier serviceability and reduced downtime

Server design

Our custom servers are designed for efficient heat removal. Different workloads, such as training, inference or storage, place different demands on the hardware and generate different heat profiles. Because we design across the full stack, server architecture and data center cooling are developed together, matching workload, GPU type, thermal profile and cooling approach into one integrated system.

Training servers

Powering the most energy-intensive workloads

Training workloads are the most energy-intensive, so this is where cooling design comes into play. We have developed training servers that operate in both air- and liquid-cooling environments, giving us the flexibility to choose the best fit depending on a combination of GPU specifications, workload requirements, and the cooling system available at a given data center.

General-purpose servers

Flexible by design

For inference workloads, general compute, and flash storage, we have built flexible general-purpose servers. Their configuration can be adjusted to match specific use cases, and key components can be upgraded without redesigning or replacing the entire system. This extends the hardware's lifespan and improves resource efficiency by reducing the frequency of full server replacements.

Air-cooled servers

Thermal control with lower energy waste

- Less cooling waste from separated CPU and GPU airflow paths.
- Reduced electricity draw via dynamic fan speed control.
- Lower conversion losses from high-efficiency power components.

These servers improve thermal efficiency via more precise airflow management. They use smart fan systems that isolate CPU and GPU cooling paths, preventing thermal cross-contamination and directing airflow to where it's actually needed.

They also use dynamic airflow control, with fan speeds adjusted continuously to the minimum level needed for effective cooling. This reduces unnecessary power draw while maintaining stable operation.

We also select high-efficiency power components to reduce losses as electricity is delivered to and distributed inside the server.

Liquid-cooled servers

Cooling only what gets hot

- Targeted cooling for high-heat components.
- Lower material use through fewer cold plates.

Rather than liquid-cooling the entire server, this design applies liquid cooling only to components with the highest thermal load. This has both environmental and operational advantages. Using fewer cold plates reduces demand for copper-intensive components, lowering material use in the value chain. It also simplifies maintenance. In a fully liquid-cooled server, servicing may require disconnecting pipes, making maintenance slower and more expensive. In our hybrid approach, air-cooled components such as CPUs can be accessed and replaced more easily. This is why, in some cases where GPUs are not supplied as part of pre-configured reference-architecture racks, we deliberately move away from fully liquid-cooled designs, even where they are the default architecture proposed by chip providers.

Modular servers

Multiple workloads with one hardware investment

- Lower replacement needs through upgradable components
- One technician can support thousands of servers
- Fewer components in the supply chain per unit of compute delivered

(1) Custom Nebius server recorded an average power draw of 8.2 kilowatts, compared with 10.0 kilowatts for a comparable Gigabyte G593-SD0 GPU server, representing an approximate 20% reduction in server electricity consumption. Similar tests were run against comparable servers from other manufacturers. Measurements were taken under controlled test conditions across inlet temperatures of 25°C, 30°C, 35°C, and 40°C

Designing for circularity

Our approach to circularity encompasses the design, use, and end-of-life phases of the hardware in our stack.

We aim to extend the useful life of our components by building for durability and repurposing, and to retain as much value as possible at the end of life by recovering raw materials.

This helps reduce supply chain risks, lower material demand, limit landfill volumes and indirect emissions, and extract greater residual value from retired equipment.

Design phase

Reducing the strain on natural resources

Component-light design: By designing tool-less, cable-free architecture, we reduce the number of (replacement) parts needed and minimize the quantity of raw materials and rare metals required in the first place.

Chip partnerships: Through regular R&D-to-R&D engagement with chip manufacturers, we provide feedback on reference architectures before they are finalized, helping them fit efficiently across the physical stack.

Use phase

Prolonging the lifespan of GPUs

Thermal-aware controls: Our servers use a baseboard management controller, a built-in hardware monitor, to adjust fan speeds based on real-time thermal load. This delivers the airflow needed, reducing thermal stress so components last longer.

Shifting GPUs to serve inference workloads: By operating as a full-stack provider with our own inference platform, we can shift less new GPUs from more intensive training workloads to inference tasks, extending the service lifetimes of the chips.

End-of-life phase

Handling used products responsibly

Recycling: As we grow, we will be building a more structured approach to e-waste, focusing on recycling retired equipment to recover its valuable materials.

Recycling rate of nearly 100%

for e-waste collected at Finland-1 data center in 2025[1]

Energy-efficient supercomputing

Our two supercomputers, ISEG in Finland-1 and ISEG2 in Iceland, are among the most powerful commercially available systems in the world.

They are also built for optimal energy and thermal efficiency. As of June 2026 ISEG2 ranks #15 among the top 500 most powerful supercomputers globally. Within that elite group, it ranks 10th in energy efficiency.

Among the commercially-operated supercomputers powerful enough to crack the global top 15, ISEG2 is the third most energy-efficient in the world.

Nebius's ISEG2

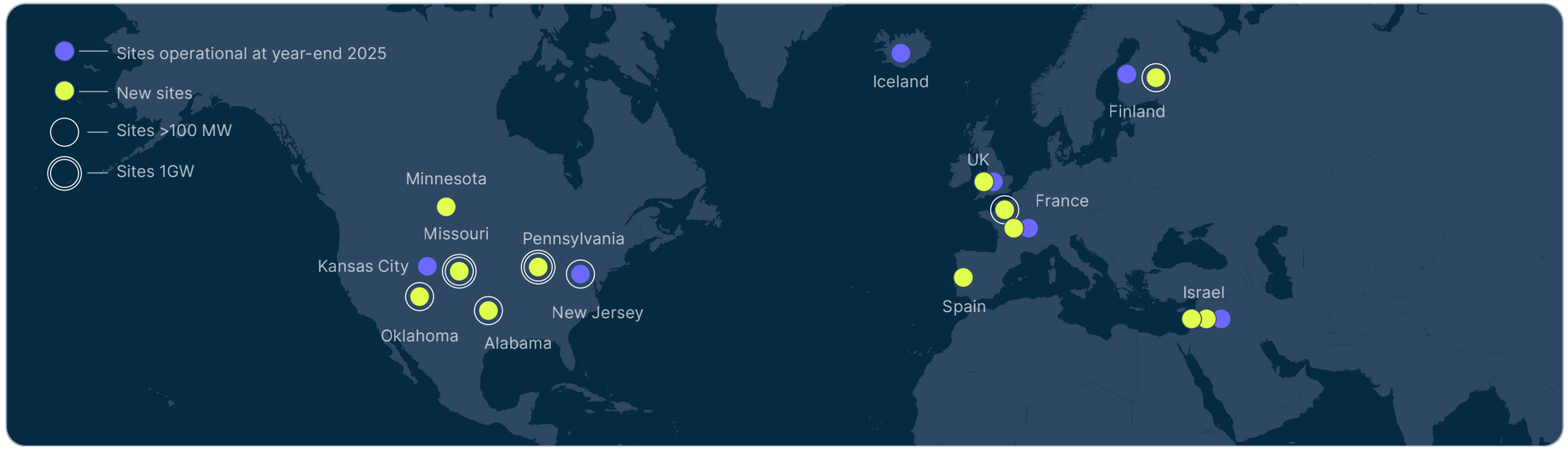
#10 most energy efficient among top-15 most powerful supercomputers globally, June 2026

(1) Represents the share of e-waste generated that was processed and recycled. In 2025, e-waste volumes were not reported across other sites, primarily because volumes were zero or near zero, as those sites were either in their first year of operation or still being commissioned, with no legacy equipment requiring retirement.

Data centers

Our growing data center footprint reflects our commitment to serve customers where they are, while taking a mindful approach to site selection. This includes energy availability, integration with local grids, and dialogue with communities where we operate.

In 2025, we owned and operated a data center in Finland-1, managed colocations across the US, the UK, the Middle East, France, and Iceland, and brought the first capacity online in a build-to-suit facility in New Jersey.



How we select our sites

When choosing sites on which to build our own data centers, we assess multiple factors including:

1 **Reliable power access:** we give priority to locations with dependable electricity and clear pathways for sourcing or transitioning to low-carbon energy.

2 **Industrial zoning:** we prioritize land already zoned for industrial use to reduce environmental disturbance and avoid impacts associated with rezoning.

3 **Regional opportunities and community engagement:** we prioritize locations where our presence can support local economic activity. We engage with municipalities at an early stage to align these opportunities, which include implementing community benefit plans. See Unlocking opportunity chapter for more information.

4 **Project economics:** we assess overall cost viability, including development, energy, and operating costs.

5 Our firmware and system designs are geography agnostic, allowing us to operate efficiently across both warmer and cooler environments. What's more, our systems don't rely much on water intake, improving resilience to drought conditions and avoiding strain on local resources.

We use a different yet equally robust approach when choosing colocation partners. We always apply rigorous selection criteria to ensure they meet our high standards for performance, scalability and reliability.

Power usage effectiveness

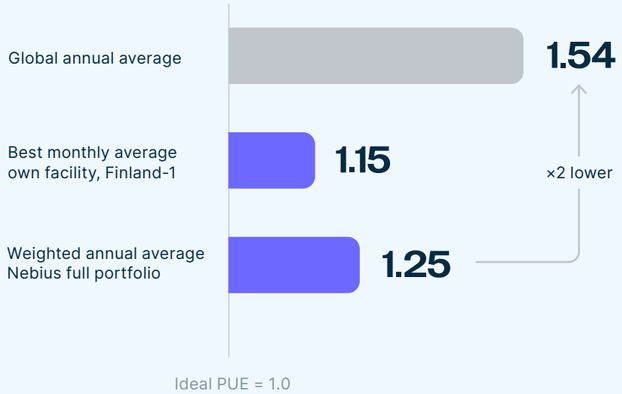
1.25
average PUE on 7 sites[1]. 2025

2× lower
energy overhead vs global average[2]. 2025

54 GWh
Total overhead savings across portfolio[3]. 2025

PUE performance

Full fleet and own facilities



Per-site PUE

Sites meeting full-year disclosure criteria[5]

Own sites Colocations

	Annual avg PUE	vs Global avg
Finland-1	1.20	-0.34
France-1	1.23	-0.31

Nebius's weighted average portfolio power usage effectiveness (PUE) of 1.25 is significantly better than the global average of 1.54, reflecting the energy efficiency embedded across our infrastructure — from in-house server and rack design to data center cooling systems in sites we design and control, to careful partner selection. Across the portfolio, this translates to over 54 GWh of energy consumption avoided in 2025 compared to operating at the global average PUE.

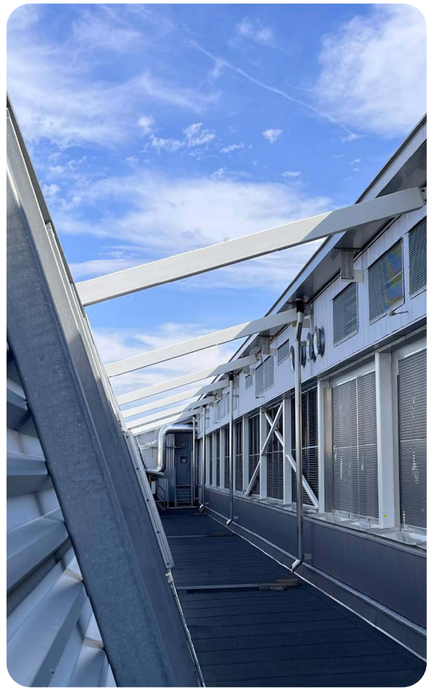
For own sites, Finland-1 is the only facility with a full calendar year of data in 2025. Its best monthly reported PUE of 1.15 is on par with the reported performance of major hyperscalers, which, according to publicly available disclosure as of Q1 2026, report global portfolio averages in the range of 1.09-1.16. Our colocation sites also outperformed the global average, demonstrating that efficiency is embedded into how we select and work with partners.

PUE at own sites in 2025 was influenced by two factors typical of a rapidly scaling infrastructure business: active commissioning of new buildings as part of Finland-1's ongoing expansion, and initial deployment phases at newly operational sites — both of which elevate the facility-to-IT power ratio while IT load is still ramping up. We control for these effects by excluding certain commissioning months from site averages. At stable, high IT loads, these sites are designed to operate at best-in-class PUE levels.

(1) **Weighted average methodology:** Portfolio PUE is calculated as the average of each site's annual PUE, weighted by IT load during the reporting period. Each site's annual PUE is calculated as facility energy consumption divided by IT equipment energy consumption for valid operational months, once IT load has stabilized. For colocation sites, allocated facility energy consumption is used rather than total facility consumption. Energy consumption and PUE from initial commissioning months are excluded from site-level annual PUE where load ramp-up makes facility-to-IT power ratios unrepresentative. Excluded months are documented per site in internal supporting records.
(2) **Global average benchmark:** Global average PUE of 1.54 sourced from the Uptime Institute Global Data Center Survey for the corresponding reporting period. Theoretical ideal PUE of 1.0 represents a facility where all energy consumed serves IT load directly, with zero overhead.
(3) **Overhead savings calculation:** Also referred to as "facility avoidance" in this report. Calculated as the difference between total energy that would have been consumed if each site operated at the global average PUE of 1.54, and actual energy consumed. Computed as: IT load (MWh) × (1.54 – actual PUE) for each site, summed across our portfolio for their valid operational months in 2025. We use the global average comparison because it is the most widely cited reference. A peer-specific benchmark based on AI cloud operators would be more precise comparison, but consistent and publicly available efficiency data for this segment does not yet exist of sufficient quality to support a comparison; we intend to update our methodology when data becomes available.
(4) **Best monthly PUE:** Reported for owned sites operational for a full calendar year. Reflects the lowest monthly PUE recorded during the reporting period.
(5) **Single-site disclosure criteria:** Single-site PUE is disclosed for our sites (owned and colocation) that had valid operational months for a full calendar year during the reporting period. Additional criteria apply to colocation sites, single-site PUE is reported when: (i) PUE is derived from independently received energy consumption data rather than contractual agreements; and (ii) sites represent more than 5% of Nebius's total connected power on 31 December of the reporting period. These criteria may be revised in future reporting periods as the fleet scales. Sites not meeting these criteria are included in the portfolio weighted average only. For 2025, those sites are: Iceland-1, UK-1, New Jersey-1, Missouri-1.
(6) **Site-reported PUE:** PUE figures for colocation sites are reported by the partner and have not been independently audited by Nebius. Site-reported PUE may reflect either metered facility performance or a contractual PUE parameter.

Efficient cooling systems

Cooling is designed around the hardware it serves. As GPU density increases, cooling becomes a key lever for reducing data center overheads. In our owned and designed facilities, we use two complementary architectures: air-based free cooling where thermal density allows, and closed-loop liquid cooling for high-density workloads.



Air-based cooling: free cooling without water or refrigerants

Our air-based cooling system was first deployed at our Finland-1 data center and continues to be used where equipment density allows. It relies on outside air rather than mechanical chillers, avoiding water and refrigerants in the cooling process. This is enabled by Nebius servers' ability to operate reliably at elevated temperatures, allowing facilities to use ambient air across a wide range of geographies.

Closed-loop liquid cooling: differentiated by no water intake

Next-generation high-density GPUs generate heat at densities that air alone cannot dissipate. To address this, we started using a closed-loop, direct-to-chip liquid cooling system in 2025, running alongside our air-cooled systems across the fleet. Coolant circulates through sealed loops: one carries heat from GPUs to a cooling distribution unit, while another transfers it to rooftop dry coolers, where ambient airflow rejects the heat.

Unlike evaporative systems, this architecture uses airflow alone for heat rejection. The glycol-based coolant recirculates for years, with no water intake, no discharge, and no evaporation losses. The system also avoids chillers, compressors, and refrigerants, reducing energy draw and maintenance complexity.

Each cooling module is designed to accommodate a heat exchange unit, which, once installed, can transfer heat to a municipal or industrial partner before the fluid reaches the dry coolers. See the Heat recovery section for more information.

Key features of the liquid cooling system designed in-house

No water intake
No evaporation losses, discharge or replenishment.

Lower pump energy
Higher server operating temperatures allow slower coolant flow.

Dry coolers instead of chillers
No compressors or refrigerants.

Heat exchange readiness
Recoverable heat can be transferred before reaching dry coolers.

WUE 0.018 L/kWh
At our data center in Finland-1, where our closed-loop liquid cooling system was first deployed, the choice of non-evaporative cooling by design keeps water consumption close to zero

~25x lower
than US data center average of 0.46–0.48 L/kWh, and 2-15x lower than WUE figures reported by leading hyperscalers[1]

~3.3 bln liters equivalent water intake eliminated
annually by design for each 1 GW of portfolio[2]

(1) Water usage effectiveness (WUE) measures liters of water consumed per kWh of IT power. US data center average WUE of 0.46 L/kWh is provided in Lawrence Berkeley National Laboratory's 2024 United States Data Center Energy Usage Report. Peer range is based on publicly reported global WUE figures from major hyperscalers (0.03 – 0.3 L/kWh)
(2) Design-based projection, not a reported operational saving. Calculated as the difference between water consumption at the US average WUE and Nebius's Finland-1 2025 WUE of 0.018 L/kWh, applied to a 1 GW portfolio operating at 85% utilization over a full year.

Heat recovery

Nebius treats server heat as a recoverable energy stream. By capturing and exporting surplus heat from our data centers, we can provide a reliable source of energy for local communities where demand exists.

Effective heat recovery depends less on high temperatures than on stable, continuous heat flows. This is where our sites are well suited: they operate 24/7, 365 days a year, providing a consistent source of recoverable heat. The model is geography-agnostic and does not depend on a cold climate.

Our data center designs are built with heat recovery in mind and can be integrated with local heat users where the right demand, technical conditions and partners are available.

Introducing liquid-to-liquid heat exchange in Finland-1

As high-density GPU workloads came online in 2025, the data center added a second heat recovery pathway: liquid-to-liquid heat exchange. The new system works alongside the site's long-running air-to-liquid recovery system, expanding how much server heat can be captured and how directly it can be transferred to local heat supply. Both systems connect to the same district heating water loop feeding the municipal heat pump plant located on site and operated by local utilities.

The new liquid-to-liquid system captures heat directly from high-density GPU workloads. Heated server coolant passes through a heat exchanger before reaching the dry coolers, transferring part of its heat to the district heating network. This reduces dry-cooler load and cooling electricity use, while turning higher-density AI infrastructure into an additional source of local heat.

In data halls that continue to use air-based free cooling, the existing air-to-liquid heat exchange remains in operation, capturing over 90% of recoverable heat and transferring it to the district heating network, while some airflow is recirculated internally to maintain optimal supply air temperatures.

>250 GWh

maximum annual heat export capacity of the existing heat recovery infrastructure, based on engineering design parameters[1]

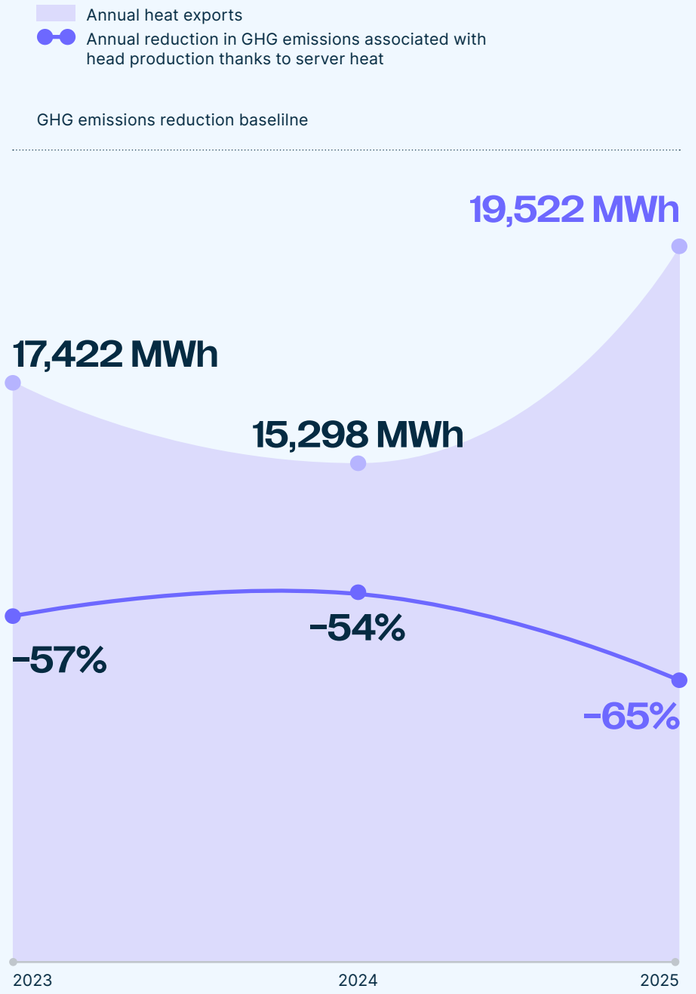
>90%

of recoverable heat captured in air-based systems

Up to 60% less

cooling energy per dry cooler in the liquid cooling system[2]

Annual heat export in Mäntsälä and decarbonization contribution



(1) Of which 175 GWh is the theoretical maximum capacity of the liquid-to-liquid heat exchange system. Represents the maximum that can be reached without structural redesign of data center systems. Actual heat exported depends on local demand and internal use of recovered heat.
(2) Based on the ratio of liquid-to-liquid heat exchanger capacity to dry cooler cooling capacity. When the heat exchanger operates at full capacity, it transfers heat to the district heating network before the fluid reaches the dry coolers, reducing the remaining cooling load by approximately 60%. Actual savings depend on heat exchanger utilization, which is driven by municipal demand.

Supporting the community in Mäntsälä

Our Finnish site is a leading example of data center heat recovery in action. In Finland-1, an integrated heat recovery system was set up between 2015-2016, making it one of the earliest large-scale implementations in Northern Europe.

For over a decade, the site worked with municipal authorities and local utilities to supply recovered heat to the district heating system, meeting up to 72% of the town's heating demand in recent years. From 2026, all recovered heat from the Finland-1 facility will be supplied to the municipality free of charge, increasing the local value of the system while supporting the decarbonization of heat production.

Turning server heat into local energy

In 2025, the Finland-1 data center exported 19.5 GWh of recovered server heat to the town's district heating network, 28% more than in 2024, driven by the scaling of data center operations. The facility's heat export capacity now exceeds current volumes, with actual delivery determined by local district heating demand.

~20%
of total electricity consumed recovered as heat and fed into district heating network in 2025

19.5 GWh
of heat exported (+27% YoY)

Creating direct value for residents

Heat recovery improves the economics of district heating in Finland-1 by lowering production costs per MWh and creating room to pass savings on to households.

10% reduction
in household heating costs in 2025

Lowering emissions from district heating

The collaboration strengthens local energy resilience by replacing conventional heat production. In 2025, recovered server heat helped avoid 4,000 tonnes CO_2e, equivalent to removing over 2,000 petrol-powered cars from the road for a year, and reduced the carbon footprint of Finland-1's district heating production by 65%.

4,000 tonnes CO_2e
avoided by the town of Mäntsälä

Energy sourcing

Power sourcing is managed site by site, with attention to grid impact, contract structure, local affordability and decarbonization potential. We secure power through contracts, PPAs, and targeted infrastructure investments, so the cost of serving our demand is covered directly rather than shifted to residents. Wherever possible, we aim to increase the share of low-carbon energy in our supply mix and work with energy providers on longer-term partnerships.

In 2025, 80% of Nebius's data center total energy consumption came from low-carbon sources and 62% from renewable sources resulting in a location-based carbon intensity of 0.12 tonnes CO_2e per MWh of energy consumed by data centers[1].

In locations where we build our own data centers, we often benefit from low-carbon grids. At the data center in Finland, the 2025 production electricity mix included a 95% from low-carbon sources, primarily nuclear, wind, and biofuels. To go further, in 2026 we completed a transition to long-term renewable electricity contracts designed to cover 100% of the site's electricity needs. The structure combines baseload PPAs with Finnish suppliers, primarily backed by hydropower, with Guarantees of Origin ensuring traceability of renewable attributes. It also provides multi-year price certainty, supporting predictable site operations and reducing exposure to electricity market volatility.

Some of the other sites we operate, including in Iceland, France, and the UK, also have 100% of their electricity supply covered by renewable sources. This includes direct supply as well as a combination of PPAs and renewable energy certificates (EACs). The facility in Iceland, home to our ISEG2 supercomputer, is powered entirely by hydroelectric and geothermal energy.

80% low carbon

energy used across Nebius data center operations in 2025[1]

100% renewable

electricity to cover Nebius's Finnish data center's electricity needs from mid-2026 under a signed power purchase agreement (PPA)

Onsite power to reduce grid pressure

AI infrastructure is scaling faster than grid capacity can be built in some markets. In 2026, we signed a long-term agreement with Bloom Energy to deploy 328 MW of on-site fuel-cell capacity for its first deployment in the US, supplying behind-the-meter electricity directly to AI infrastructure.

Fuel cells generate electricity through an electrochemical process rather than combustion. This can reduce local air pollutants and water use compared with combustion-based onsite generation, all while providing stable power.

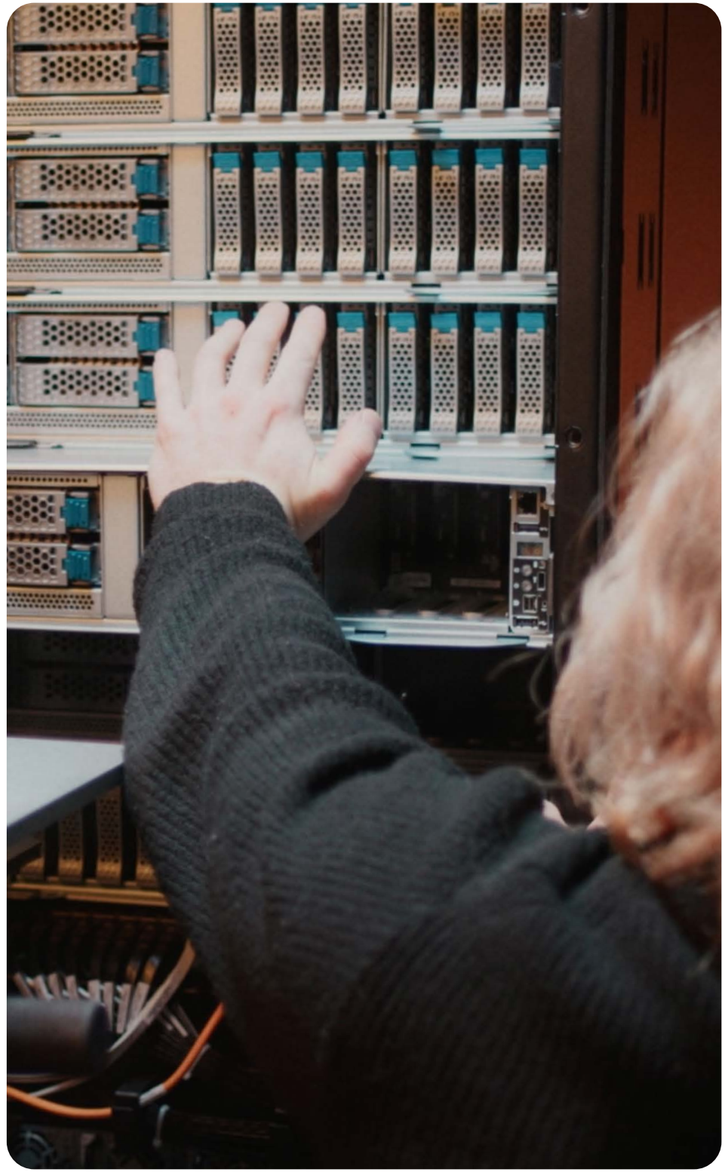
For Nebius, this keeps the low-carbon principle in view while reducing pressure on local grids. It also preserves fuel flexibility: solid oxide fuel cells can operate on natural gas today and are technically compatible with lower-carbon fuels such as biogas or hydrogen as supply becomes available.

Non-combustion generation

lower local pollutants, lower water use

Behind-the-meter power

stable electricity supplied directly to AI infrastructure



(1) For calculation details, refer to ESRS E1 disclosures.

Software

While hardware determines how much energy is consumed, software is the key to getting more compute out of every watt.

Any idle time can translate into substantial energy waste. Common causes include overprovisioning (allocating more compute, memory, storage, or network capacity than is physically needed at a given moment), leading to machines consuming electricity without doing useful work, as well as task bottlenecks, uneven loads or tenants simply finishing jobs at different times. In fact, we've observed idle servers or racks drawing 12–25% of their active power consumption, depending on the GPU model.

Energy-to-AI output efficiency lever

Our software works across model, cluster, and fabric layers to find efficient ways of transforming power into AI output throughout the stack. We have developed tooling to scale compute intelligently, speed up workload execution while optimizing resources needed for that, such as by reducing idling and overprovisioning.

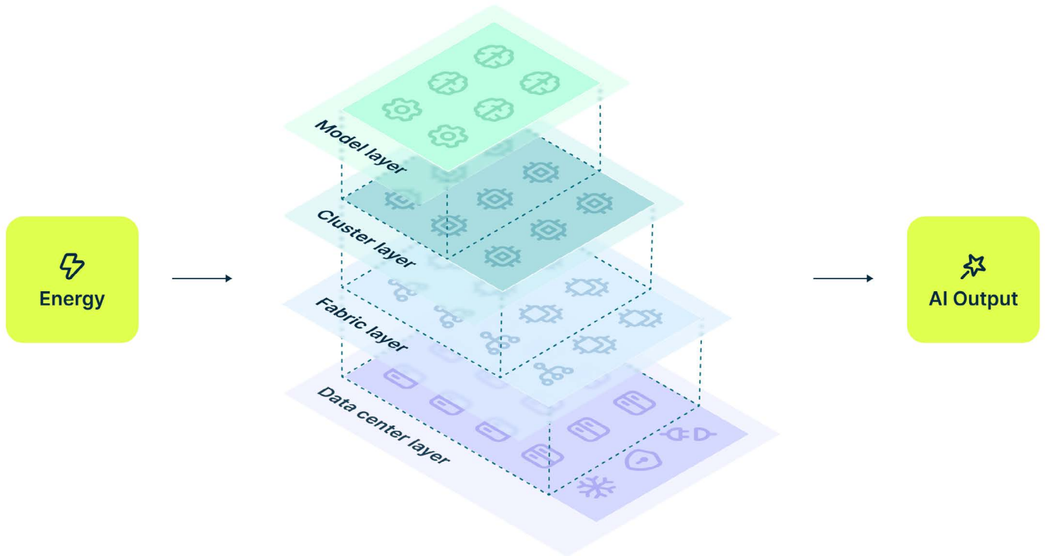
Explore the full framework in the white paper "The energy behind AI: how Nebius builds an energy-efficient cloud"

Efficiency layers

Model-layer efficiency: how well a model runs and interacts with the underlying AI hardware. Model parameters, hardware characteristics, and runtime configuration together determine how much energy is needed to generate each token, meaning the same hardware can yield very different output depending on how the model is deployed.

Cluster-layer efficiency: how well provisioned GPU clusters are put to use. A cluster that fails frequently or performs unpredictably wastes energy on retries, idle time and incomplete work. A stable, high-performing cluster completes the same job in less time, which means fewer GPU-hours and proportionally less energy needed to reach the same result.

Fabric-layer efficiency: how much AI work is generated per unit of available fabric capacity. This is shaped by tenants' workload patterns — jobs starting and stopping on different schedules; GPU requests varying in size and duration. Smart orchestration and workload allocation by the AI cloud provider determine how well these patterns are managed.



Model layer

In 2025, our focus was on bringing inference to production scale. We launched Nebius Token Factory in November — a managed inference platform and full model-lifecycle management, from fine-tuning to governed production deployment at scale.

Cluster layer

Every job interruption wastes the energy already consumed on incomplete work. Improving cluster reliability and performance therefore has tangible environmental benefits. In 2025, we reached significant reliability and performance improvement milestones, including an average MTBF of 168,900 GPU-hours on 3,000-GPU clusters[1], fast recovery times[2], topology-aware scheduling extended across all major job schedulers, and the launch of fully managed Slurm clusters.

Selected approaches and tools Efficiency contributions

Inference engine optimization

The platform optimizes interactions between models and hardware, without changing either, to maximize throughput per unit of compute.

→ Customers get more AI work done without additional energy input.

Reliability features

Combine preventive health monitoring in real time, automated node isolation and workload migration to healthy nodes. Known failure patterns trigger automated repair, while novel issues are escalated to engineers.

→ Reduces idle GPU time during failures, keeping compute converting energy into AI output rather than sitting unused.

Post-training optimization

Tooling to adapt open-source models to specific datasets and tasks, producing task-specific variants that are smaller and more accurate than general-purpose alternatives.

→ Customers achieve target output quality with a smaller base model, reducing compute and energy cost per query.

Performance enhancement

Topology-aware job scheduling assigns distributed training jobs to GPUs that are physically close to one another within the cluster, keeping inter-GPU communication on the shortest network path and reducing overhead. Available across Slurm, Kubernetes, Volcano, Kueue and other schedulers.

→ Benchmark tests run on Nebius infrastructure showed average communication performance improvements of 10–25% compared to jobs scheduled without topology awareness, depending on cluster size — shorter job runtime directly reduces GPU-hours and energy consumed per training run.

56.6 hours

a Nebius cluster runs before a failure occurs

3,000-GPU cluster in Q2 2025, compared with ~ 9.8 hours for an industry-typical cluster of similar size[1]

12 min

Average time from failure detection to recovery, enabled by end-to-end automation of the recovery process

Across most of Nebius installations in Q2 2025, compared to 1 hour assumed for most peers in industry research[2]

(1) Represents Mean Time Between Failures, or MTBF, shown in wall-clock hours. This measures how long the cluster runs continuously before an interruption, regardless of how many GPUs are in use. MTBF reflects the inherent stability of the infrastructure: a rising MTBF may indicate better component reliability, improved firmware or driver behavior, or more effective failure-prevention strategies. The underlying Nebius figure is 168,900 GPU-hours per week on a 3,000-GPU cluster, which converts to 56.6 wall-clock hours between failures at that cluster size. See Fault-tolerant training: How we build reliable clusters for distributed AI workloads for the details of the measurement and source for industry data.
(2) Represents Mean Time to Recovery, or MTTR, which measures the average time needed to detect, isolate, and resolve infrastructure failures, returning the affected node or cluster segment to a healthy, schedulable state. It is calculated as total resolution time divided by the number of failures. Resolution time includes all steps to replace a broken node and provision a ready-to-use healthy node: node isolation, spare node provisioning and state reattachment. The 1-hour peer benchmark is based on the Semianalysis's How Much Do GPU Clusters Really Cost? study

Fabric layer

Selected approaches and tools · Efficiency contributions

Reducing server idling is one of the most direct ways to reduce the energy footprint of AI output. This places the fabric layer — the network interconnecting servers shared across multiple tenants — at the center of efficiency work. In 2025, engineering efforts concentrated on virtualization, autoscaling and storage to achieve tighter workload allocation across the fabric.

Virtualization with on-demand and preemptible instances

Virtualized GPU allocations — including short-lived preemptible instances that can be reclaimed when capacity is needed — enable tighter packing of workloads than fixed bare-metal allocations allow.

→ Harvests small unused GPU × hour slices that would otherwise sit idle.

Shared capacity buffer

A single shared buffer pool across all tenants, allocated dynamically during demand bursts or when auto-healing requires node replacement.

→ Drops idle reserve from a typical 10–20% per tenant in bare-metal environments to around 4–5% across the shared pool, cutting the GPU capacity that draws power without producing output.

Dual-scheduler workload orchestration

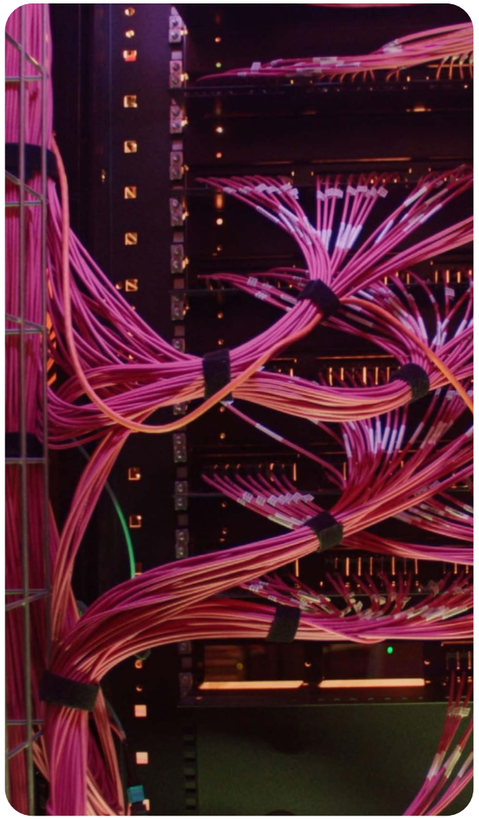
Training and inference have fundamentally different resource requirements — training needs sustained compute over hours or days, while inference demands rapid response to variable loads. Nebius runs two complementary systems: a managed Kubernetes solution for inference workloads, and Soperator — our open-source Slurm-based workload manager — for training jobs. Schedulers release GPUs when no longer needed, allowing other jobs to claim them immediately.

→ Reduces idle capacity that arises when a single scheduler must serve fundamentally different workload types simultaneously.

Software-defined storage

Block, file and object storage run on shared hardware, and the same SSD and HDD pools serve multiple protocols and workload types. SSDs serve active workloads (training datasets, inference serving and active checkpoints), while HDDs serve archival and cold data (completed artifacts, logs and infrequently accessed datasets). Customers manage data lifecycle through object storage lifecycle policies.

→ Shared hardware eliminates dedicated-but-idle storage nodes per workload type. Right-sizing storage to access patterns reduces energy per TB stored: HDDs consume minimal power when idle, keeping cold data off drives that would otherwise draw unnecessary energy.



Unlocking opportunity in the AI economy

Unlocking opportunity in the AI economy

Access to high-performance compute is fundamental to AI innovation, yet cost, availability and technical complexity can still create barriers for researchers, developers, startups and organizations. Alongside access to infrastructure, there is also a growing need to help people better understand, adopt and confidently use AI.

Nebius is working to both close the access and knowledge gap. For AI innovators, we provide the AI compute they need to build, test, and scale solutions. For people and organizations learning how to leverage AI, we support the tools, knowledge, and practical skills needed to apply it confidently and effectively.

This is where our role extends beyond infrastructure. By combining access to AI compute with practical enablement, we support a broader range of users in turning AI capabilities into real-world applications.

Nebius's initiatives for lowering barrier to adoption

Global infrastructure footprint

Support for startups

Critical sectors

Academic partnerships

Training and upskilling

Outcomes

- Equitable access
- Workforce readiness
- Reduced bias, more inclusion
- Innovation, breakthroughs
- Community benefits

Local community support

In 2025, we began formalizing our approach to community engagement as our AI infrastructure footprint expands.

Our approach to scaling starts with understanding how each site can contribute to local development. That means engaging early, before ground is broken, and staying present through construction and operations. It also means listening to local priorities, and assessing where our presence can create practical benefits. These may include workforce development, education partnerships, local hiring, energy infrastructure improvements, and support for stronger communities.

Community engagement is essential to how we operate: our business can only scale if communities consider us a good neighbour. As we continue to refine our approach, these efforts will be specific to each location, grounded in local needs, and connected to the way we design, power, build, and operate our data centers.

In practice, this means providing clear information about how facilities will operate, working with local leaders, educators, and residents, and identifying initiatives where Nebius resources can support long-term community value.

Responsible by Design

Early community engagement, transparent communication and long-term local value creation: this philosophy guides how we design, build and operate our facilities.

Priority areas under development

- Education and AI literacy
- Technical training
- Local workforce pathways
- Partnerships with schools, universities, and community organizations
- Contributions to energy systems

Strengthening local energy systems

Where local offtakers exist, data centers we design can be configured to export recovered server heat to local users. In Finland-1, this model is already in place, helping the municipality lower district heating emissions, reduce heating costs for residents, and improve energy resilience. See Heat recovery section for more information.

At sites announced in 2026, we are also investing in local power infrastructure. In Birmingham, Alabama, this includes funding a new switchyard required for our facility that is also expected to improve community power reliability without shifting the cost to ratepayers. In Independence, Missouri, our investment and power agreement supports the municipal utility in adding new generation capacity, helping strengthen the local electricity system at no cost to ratepayers.

Access for startups

We empower ambitious teams to turn their vision into real-world progress. Our flagship solution is Nebius for Startups, a go-to-market program for AI-native startups rolled out in 2025 to make Nebius the go-to infrastructure partner for emerging AI-native startups.

Nebius for Startups gives early-stage AI companies access to the compute capacity, technical support, and ecosystem connections they need to build and scale. The program provides AI cloud credits and discounts through a three-tier model, from experimentation to production, helping startups reduce infrastructure costs while improving workload efficiency.

Support goes beyond credits. Selected teams can work directly with Nebius AI experts, access our technical and enterprise networks, and use the program as a bridge to customers, investors, and partners. This is especially important for AI startups, where access to high-performance compute can determine how quickly a product can be tested, improved, and brought to market.

As demand for the program grows, selection is becoming more focused. We assess companies based on funding stage and investors, expected compute intensity, founder background, and the potential to build category-defining technology in areas such as robotics, agentic AI, and life sciences. In 2025, we received more than 1,000 applications and selected 440 startups.

The program is most established in the US, which accounted for around a quarter of supported startups in 2025. Around one fifth came from emerging markets. As Nebius expands, we are working to strengthen the program's presence in Europe, the Middle East, Asia-Pacific, and Japan, while exploring deeper engagement with startups in sectors such as financial services, retail, media, and entertainment

$2.1 million in AI cloud credits

allocated to startups in 2025

One fifth

of supported startups came from emerging markets

How we support founders

- Platform credits and collaborations with Nebius team
- Go-to-market support and fundraising introductions
- Events and venture capital partnerships
- Enterprise Readiness fast-track program for scale-ups graduating the core program (launched with NVIDIA, Insight Partners, Accel, and Fellows Fund in 2026)

Startup voices and founder lessons

Launched in 2025, the Nebius for Startups Podcast gives AI founders visibility and shares practical lessons from investors, builders, and Nebius experts on what it takes to scale.

Jua Featured story

The mission

To build a foundation model for Earth's physical systems, starting with weather forecasting for energy markets

Nebius's support

- Stable GPU clusters that outperformed three prior cloud providers on every benchmark.
- Enabled month-long training runs without mid-run failures.
- Freed engineers from watching infrastructure to focus on the model.

What this enabled

$1.5 million

estimated annual gain per megawatt from more accurate energy trading decisions

Hourly global updates

vs four times daily for every other system, including from the world's largest AI labs

Support for critical fields

Actively supporting In exploration

Healthcare and life sciences

Nebius works with healthcare, life sciences, and scientific AI teams across stages and sizes, providing GPU capacity, technical expertise, and sector-specific support for compute-intensive workloads. Our team draws on scientific backgrounds, technologies and partnerships to understand customer challenges and translate AI infrastructure into practical research and product outcomes.

We see AI as a tool to accelerate scientific progress in health and medicine, by augmenting the work of researchers, clinicians, nurses, and care teams. This includes improving data processing, information flows, and productivity in systems facing capacity constraints. Academic partnerships include Stanford, Princeton and Google DeepMind, the London Institute for Mathematical Sciences, Harvard, and UCSF.

Building on two years of work in healthcare and life sciences, Nebius is expanding the vertical into Scientific AI and Healthcare, reflecting broader applications of AI across science.

Nebius AI cloud applications

Biopharma and drug discovery:
accelerating modeling and screening, cost-efficient discovery workflows

Medical devices:
improved surgical outcomes, tracking patient health

Genomics and disease:
faster genome analysis, rare disease detection, personalized treatment development

Medical imaging:
accelerated diagnostics, urgent case flagging

Digital health:
improved patient care, reduced clinical workloads

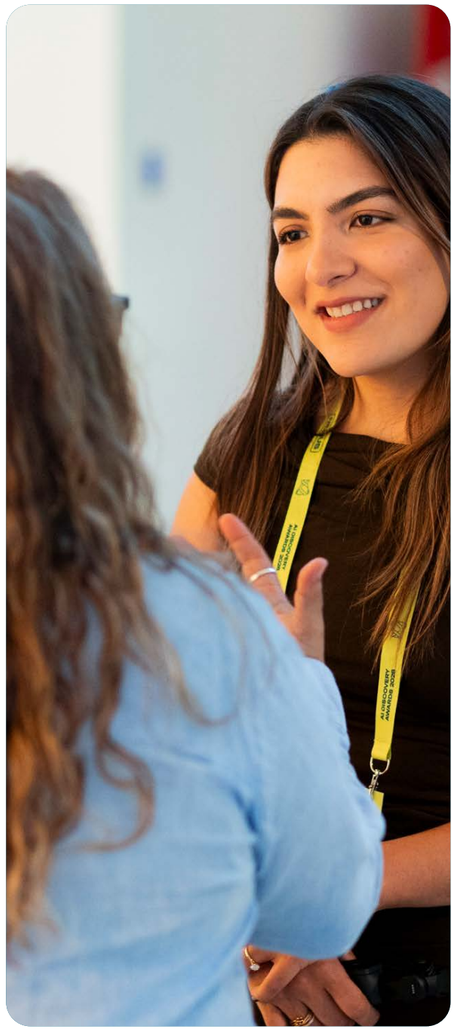
Scientific AI:
Broader applications across science, chemistry, physics, material sciences, animal and plant biology, agriculture, and cosmetics

Nebius AI Discovery Awards

In 2025, we held the inaugural Nebius AI Discovery Awards to spotlight emerging HCLS technologies, encourage collaboration between startups and potential investors, and support the next generation of health innovation. Selected from 257 applicants by independent expert judges, four winners working on cancer prediction, protein targeting, transcriptomic mapping, and precision diagnostics, along with eight runners-up, received GPU cloud and inference credits, mentorship, and investor networking opportunities.

The program's momentum continued into 2026, with the second edition attracting 647 applications from +70 countries, more than nearly 3x the inaugural year.

$850,000
in GPU cloud and inference credits awarded to AI Discovery winners in 2025



Featured story

sword

The mission

To make mental health support available anytime through an AI wellbeing specialist that continuously adapts its therapeutic approach.

Nebius's support

- Specialized inference infrastructure needed to deploy mental health solutions responsibly.
- Managed GPU clusters with automated training runs.

What this enabled

150,000
new enrollments just in Q1 2026 alone.

10+ million
AI guided care sessions delivered per year, bringing personalized mental health support closer to their goal of reaching 2 billion people worldwide.

Featured story

CRISPR-GPT

by Princeton, Stanford and DeepMind

The mission

To make the use of CRISPR gene editing more usable by any scientist without the need for weeks or months of intensive training.

Nebius's support

- Access to GPU clusters for fine tuning CRISPR-Llama3 (a model fine tuned on 11 years of expert gene data).
- The ability to provision compute immediately without the need for infrastructure setup overhead.

What this enabled

Successful experiments
run by researchers with no prior gene editing experience, achieving correct gene disablement for 80% of target cells on the first attempt.

Several months
of technical infrastructure development saved

Featured story

PRIMA MENTE

The mission

To detect Alzheimer's and Parkinson's disease from a routine blood test, by training AI on the epigenetic signals in cell-free DNA

Nebius's support

Multi-billion-parameter training on GPU clusters, at scale usually accessible to large tech companies

What this enabled

89% accuracy
detecting early Alzheimer's and Parkinson's from a blood sample alone, rising to 97% when combined with a standard protein test

20x more
disease signal extracted from a single blood draw

Days vs weeks
for same work

Robotics and physical AI

Whether they're training perception models, running digital twins, or deploying intelligent systems at scale, physical AI innovators need reliable, simple, and flexible cloud compute.

Our in-house, full-stack approach is designed for these needs, helping teams focus on product development rather than infrastructure complexity. Within this vertical, our key focus areas include foundation models and robot brains, vision AI and streaming video analytics, data engines and simulation, and industrial robotics.

Nebius Robotics & Physical AI Awards

In 2025, we launched the Robotics & Physical AI Awards to recognize early-stage teams building intelligent physical systems. 16 startups were selected from 254 applicants and awarded a combined $1.5 million in cloud compute and inference credits, supporting innovation across robotics, simulation, vision AI, and physical AI systems.

$1.5 million

in GPU cloud and inference credits awarded to robotics and physical AI startups in 2025

Research support and training programs

AI development depends on access to compute, but also on strong research institutions and people with the skills to use AI effectively. We support universities, researchers, and early-career talent through AI infrastructure access, research grants, technical learning, and career-oriented training.

NEBIUS ACADEMY

Nebius Academy

Launched in 2024, Nebius Academy is our AI education and research initiative, focused on providing practical training, supporting academic collaboration, and developing early-career talent for the AI industry.

The Academy focuses on applied learning. Its courses are designed around real workflows and are often delivered through a mix of self-paced content, live workshops, and hands-on assignments led by practicing AI engineers. This makes the learning more directly usable for technical teams, business users, and organizations adopting AI.

In 2025, Nebius Academy expanded its course offering and included free individual courses: Intro to ML from an LLM standpoint, AI-assisted programming, and AI in Mathematical discovery. It also launched open course materials on GitHub, making selected content available beyond formal cohorts.

Over 18,000

People benefited from Nebius Academy programs since its inception

12 cohorts & 9 courses

Were delivered

How Nebius Academy builds capability

Practical AI training

Nebius Academy provides applied courses for technical professionals, business users, companies, and communities. These programs help learners understand how AI systems work, how to use AI tools productively, and how to apply AI infrastructure in real workflows.

Academic partnerships and research collaboration

The Academy works with universities and research institutions on AI education, applied research, and research access to compute. In 2025, active and new partners included UCL, Stanford, Caltech, UCSF, USC, LIMS and others. In 2026, we also launched Nebius Science, an editorial platform featuring research across the Nebius ecosystem, from our own R&D teams to researchers building on our platform and wider scientific work.

Early talent and community learning

Nebius Academy is also building a talent pipeline for AI and infrastructure roles. In 2025, it launched the Early Talent Program for recent graduates, creating a structured pathway into applied infrastructure and software roles at Nebius. Going forward, the Academy will also focus on supporting AI literacy initiatives for communities, including local education programs in regions where Nebius operates or plans to operate.

Research grants

Access to GPU compute is one of the biggest constraints facing academic researchers. Many compete for the same infrastructure as commercial AI labs and well-funded startups, which can slow down scientific work or limit the scope of what researchers can test.

Nebius Academy helps remove this barrier through research grants. Successful applicants receive access to GPU resources for training workloads or AI model tokens for inference tasks, allowing research teams to move faster on compute-intensive projects.

By the end of 2025, Nebius had awarded credits to researchers and academics at 16 institutions across four continents. Supported fields included genomics, drug discovery, mathematics, climate science, and AI safety. The program is designed to support research with scientific novelty and public impact, regardless of whether it comes from a globally known institution or a smaller research group with a strong project.

Stanford Caltech

RÉPUBLIQUE FRANÇAISE TELECOM PARIS INSTITUT POLYTECHNIQUE DE PARIS

WEIZMANN INSTITUTE OF SCIENCE Utrecht University University of Sunderland

London Institute for Mathematical Sciences

25 YEARS CONSTRUCTOR UNIVERSITY

tripleten

TripleTen: reskilling for a changing tech job market

TripleTen, an edtech business within Nebius Group, is a career learning platform for people in a world being reshaped by AI. TripleTen helps people move into higher-paying tech and AI roles through online, part-time, project-based, intensive programs built for working professionals. Its mission is to give anyone the skills, experience, and confidence to build a lasting, AI-ready career.

As of May 2026, TripleTen's worldwide alumni network stands at over 7,500 graduates across the United States, Brazil, and Spanish-speaking Latin America. Students come to TripleTen from a wide range of ages and backgrounds — people in their late teens through their fifties, from a range of industries and levels of education.

In 2025, the tech job market underwent a sea change as AI reshaped what entry-level work looks like. AI skills became a necessity for workers hoping to advance in their careers.

TripleTen evolved quickly to meet this change. As of May 2026, every program has been updated with current AI tools and workflows, so students learn the skills employers currently need. Data Science was rebuilt to AI & Machine Learning, Software Engineering was relaunched as AI Software Engineering, with focus on working effectively with AI tools: prompting, reviewing, improving, and improving AI-assisted output. New programs were added: AI Automation, Cybersecurity, and UX/UI Design, bringing its offer to seven programs. In 2026, TripleTen is launching its eighth program, AI Systems Engineering, for people already working in tech who want to step up into senior AI roles.

The learning model emphasizes hands-on work. Students apply their skills in projects at the end of each learning sprint. They also have the opportunity to participate in externships (shipping projects for partner companies) and hackathons. Each project gives students portfolio-ready experience, and all students receive career coaching and job search support. Nearly two-thirds of TripleTen graduates land their new roles within 10 months after graduation.

In 2025, Nebius Academy expanded its course offering and included free individual courses: Intro to ML from an LLM standpoint, AI-assisted programming, and AI in Mathematical discovery. It also launched open course materials on GitHub, making selected content available beyond formal cohorts.

88% of graduates
still thriving in their new roles 12+ months later

80% of students
start with no prior tech background

250+ companies
in TripleTen's employer network

60+ projects
delivered to partners and 56 projects developed during TripleTen hackathons in 2025

49% of employed graduates
credit their success to their career coach, who helped them change their mindset

Dot: AI that builds confidence and independence

One challenge in online learning is that students often get stuck and hesitate to ask for help. Some feel insecure, others are unsure how to explain the problem. TripleTen developed Dot, an AI learning assistant built into the platform, to intervene at that moment.

Dot proactively detects when someone is struggling and steps in. It is deliberately designed to give guidance and not answers. It offers hints and asks questions so students reason through the problem themselves. This approach helps build confidence and independence, while reducing the risk that a temporary blocker turns into disengagement.

Early results show that students use Dot roughly 15 times more often than traditional tutor channels, report 90% satisfaction with its help, and were more likely to complete and submit their first project when instant AI support was available.

The same AI layer also reduces repetitive work for human reviewers. An AI review step now runs before a human review, cutting time-to-first-feedback from about a day to minutes and freeing reviewers to focus on higher-value guidance. Savings from these efficiencies are being reinvested into the product, including more advanced capabilities such as a voice AI tutor.

~33% more likely
for a student to complete and submit their first project, feeding through to a clear lift in overall course completion[1]

~2× engagement
engagement nearly doubled once students were able to access instant help within the learning platform

(1) Results for a controlled experiment

Reliability, resilience, and trust

Reliability, resilience, and trust

For AI teams, reliable infrastructure means workloads can run securely, continuously and predictably. This is especially important for large-scale training and inference, where a single hardware, node or network failure can interrupt jobs.

Nebius approaches reliability across resilient physical infrastructure, secure platform controls and workload-level continuity mechanisms.

2025 milestones

Formalized security framework

With 50+ robust policies introduced

56.6 hours

a Nebius cluster runs before a failure occurs[1]

Enterprise-grade certifications

including 6 ISO cybersecurity certifications, SOC2 incl. HIPAA and more

Certifications[2]

As Nebius adds more enterprise and healthcare AI workloads to its platform, certifications have become part of how we get ready for customers. They help clients move faster through procurement, security review and onboarding by showing which controls have been independently assessed.

In 2025, Nebius obtained certifications covering information security management, privacy, cloud data protection, healthcare data handling, and business continuity. In the United States, HIPAA-aligned controls help Nebius support healthcare AI applications involving sensitive data, enabling us to work with partners such as Hippocratic AI, Sword Health, and Slingshot AI.

In June 2026, aligned controls with NIS2 and DORA regulations, supporting European customers use our platform with greater confidence.

ISO 27001
Information security management.

ISO 27701
Privacy information management.

ISO 27018
Protection of personal data in cloud environments.

ISO 27032
Cybersecurity guidelines for internet security.

ISO 27799
Management of healthcare data and health informatics.

ISO 22301
Business continuity management.

SOC 2 Type II + HIPAA
Audited security controls, including a public summary of control environment.

NIS2
EU cybersecurity rules for critical sectors.

DORA
EU operational resilience rules for financial-sector ICT risks.

CSA STAR level 1
Cloud provider's security practices and controls.

ISO, SOC, and HIPAA certificates were obtained in 2025
NIS2, DORA, CSA were obtained in 2026

(1) 3,000-GPU cluster in Q2 2025, compared with ~ 9.8 hours for an industry-typical cluster of similar size. Represents Mean Time Between Failures, or MTBF, shown in wall-clock hours. This measures how long the cluster runs continuously before an interruption, regardless of how many GPUs are in use. MTBF reflects the inherent stability of the infrastructure: a rising MTBF may indicate better component reliability, improved firmware or driver behavior, or more effective failure-prevention strategies. The underlying Nebius figure is 168,900 GPU-hours per week on a 3,000-GPU cluster, which converts to 56.6 wall-clock hours between failures at that cluster size. See "Fault-tolerant training: How we build reliable clusters for distributed AI workloads" for the details of the measurement and source for industry data.
(2) Reports and certificates are available at our Trust Center page.

Operational and information security

We operate under an end-to-end security framework to match the scale and complexity of modern AI cloud environments and regulatory requirements, protect personal data, and support compliance with applicable regulatory requirements, including the General Data Protection Regulation (GDPR).

In 2025, we completed a comprehensive review and formalization of our security frameworks, designed to support the confidentiality, integrity, and availability of information assets while ensuring the responsible processing and protection of personal data and now includes 50+ policies across information security, privacy, physical security and business continuity.

Physical resilience

Across our owned and principal colocation sites, facility designs are generally built around resilient infrastructure features with N+1 or 2N redundancy[1], including dual utility power feeds[2], uninterruptible power supply (UPS)[3] and generators, and cooling systems sized for continuous high-density AI workloads. Site-specific configurations vary based on facility architecture.

Physical access is controlled through layered security measures, which may include mantrap entry, card and PIN or biometric readers, and on-site security personnel. Continuous CCTV monitoring and environmental controls, including fire suppression and leakage detection systems, are maintained across our data center portfolio.

Security framework

Oversight: designated executive-level expert reporting to the CEO

Coordination: Information Security Office, a cross-functional committee composed of senior security leadership

Core governance

Security roles, oversight, risk management, control testing, certifications, audit readiness, privacy governance and regulatory compliance

- Information Security and Privacy Policy
- Security Organization Policy
- Information Security Compliance Policy
- Data Protection and Privacy Policy

and more

Data protection and access control

Data classification, privacy controls, lawful processing of personal data, encryption, secure transfer, user access, least-privilege principles, retention, and protection of data subject rights

- Authorization and Access Control Policy
- Encryption and Cryptographic Controls Policy
- Data Classification and Handling Policy
- Transfer of Information Policy

and more

Product and AI cloud

Secure development, application security, cloud infrastructure controls, change management, patching, AI tool use

- Secure Software Development Lifecycle Policy
- Application Security Policy
- Cloud Security Policy

and more

Monitoring and incident response

Logging, monitoring, threat intelligence, vulnerability management, incident response, backup, business continuity

- Audit Logging and Monitoring Policy
- Incident Response Plan
- Business Continuity Policy

and more

Third-party, employee and physical security

Vendor security, employee lifecycle controls, awareness training, endpoint and mobile device security, physical access

- Third-Party Risk Management Policy
- Employee Security Lifecycle Policy
- Training and Awareness Program

and more

GDPR compliance and privacy

Transparency, accountability, the protection of individual privacy rights in line with GDPR

- Procedures for responding to requests
- Privacy by Design Implementation Process
- Privacy Risk Assessment Procedure

and more

(1) N+1 means one backup unit is available beyond the minimum required. 2N means a fully independent parallel system is in place.
(2) Two independent power feeds from the utility grid, so that a failure on one line does not interrupt supply.
(3) UPS systems provide battery-backed power to bridge any gap between a utility outage and generator startup.

Risk and incident management

We monitor systems, networks and applications for suspicious activity and use structured incident response to contain, investigate and remediate issues. In 2025, Nebius expanded risk and incident management model to cover AI-specific risks and third-party dependencies.

1. Prevent & detect

Development: Secure Software Development Lifecycle, architecture security reviews, SAST/SCA and external penetration testing

Infrastructure: Architecture reviews, physical security assessments and external testing where relevant

Production: Real-time monitoring through SIEM, on-call first response and technical escalation

Customer channels: Customers can report security incidents through established reporting routes

2. Respond

Dedicated on-call first responders, multidisciplinary technical escalation and rapid containment. Risks tracked in the GRC system, with remediation plans, owners and resolution timelines

3. Learn

C-level reporting, root-cause analysis and structured post-incident reviews

Security services

Security in cloud environments is a shared responsibility. Nebius secures the underlying platform, while customers control how their workloads, data, identities, permissions, and credentials are configured and used. To support this, we provide built-in services that help customers manage access, trace activity, and protect sensitive information.

Identity and access management. Customers can manage users, service accounts, roles, and permissions, applying least-privilege access to cloud resources.

Audit logs. Record key activity in the customer environment, including authentication, configuration changes, and resource access.

MysteryBox. Stores sensitive credentials, such as API keys, passwords, tokens, and certificates, in encrypted form, reducing the risk of exposure in code, scripts, or configuration files.

Customer data

Customers need clarity on where data resides, how access is managed, and how requirements are handled across jurisdictions. We treat customer data, including personal and non-personal datasets used for AI training and inference, with the same care as we would our own information.

Encryption by default. Customer data is encrypted at rest and in transit. For sensitive workloads, we also support confidential computing and in-use encryption to help protect data while it is being processed.

Customer-held controls. Customers choose the cluster location where their data is stored and processed. Through Nebius Key Management Service, they can manage encryption keys directly, including using their own keys through BYOK[1].

Limited access. Access to customer data is governed by role-based permissions, multi-factor authentication and least-privilege principles. Key operations and relevant activity are logged to support auditability and accountability.

Employee training. All employees complete annual security awareness training covering phishing, access management, and data handling.

Building trust through AI sovereignty

For customers deploying AI, **sovereignty means control:** where data is stored, where it is processed, who can access it, and which legal and regulatory requirements apply.

Nebius supports this by giving customers choice over data location and processing environments, helping them align AI deployments with internal policies, data residency requirements, and governance expectations. As a European company headquartered in the Netherlands, we see this control as part of digital trust: customers should be able to use advanced AI infrastructure with confidence over data, compliance, and continuity of access.

Beyond data governance, sovereignty now extends compute: whether organizations can continuously access the infrastructure and advanced models they need, in a jurisdiction of their choice. Nebius supports this through owned GPU clusters, multi-region deployment, including infrastructure operated under European legal frameworks.

(1) BYOK (Bring Your Own Key): customers can use their own encryption keys to maintain direct control over key ownership, rotation and access

Workload continuity

For AI customers, workload continuity means more than keeping infrastructure online. A failed or interrupted training job can waste hours or days of compute, delay product development, and require teams to restart or manually recover work.

We address this through a combination of fault-tolerant infrastructure, checkpoint-ready storage, and managed orchestration. The same capabilities also contribute to reduce idle GPU time, with their energy-efficiency impact described in the Efficiency by design chapter.

Fault-tolerant training infrastructure

Nebius infrastructure is designed to automatically detect failures, isolate affected nodes, provision replacement capacity, and recover quickly when interruptions occur. In Q2 2025, Nebius recorded an average MTBF of 168,900 GPU-hours on 3,000-GPU clusters, equivalent to 56.6 wall-clock hours between failures, compared with approximately 9.8 hours for an industry-typical cluster of similar size. For detected failures, average MTTR was 12 minutes across most Nebius installations[1].

Checkpoint-aware storage and data persistence

For training continuity, recovery only matters if the job can resume from a saved state. Training data, model checkpoints, logs, and artifacts need to remain accessible when compute nodes are replaced or jobs restart. Nebius supports this through persistent storage options, including shared file systems and Object Storage.

Managed orchestration

Recovery also depends on how workloads are scheduled back onto available infrastructure. Nebius supports Slurm-based and Kubernetes-based orchestration, helping training and application workloads continue or resume once healthy capacity is available.

(1) These metrics are discussed in detail in he Efficiency by design chapter.

AVRIDE

Avride: safety discipline in autonomous systems

Avride, part of Nebius Group, develops autonomous passenger vehicles and delivery robots for controlled, real-world environments. As of the date of this report, Avride's autonomous vehicle fleet has driven more than 1.3 million autonomous miles overall, completed over 60,000 autonomous rides on the Uber platform, and its delivery robots have completed more than 600,000 deliveries.

Avride's approach to safety is purpose-built: vehicles and robots operate within defined use cases and service areas, supported by remote monitoring and human-in-the-loop oversight so human judgment remains part of the safety net. In 2025, Avride continued moving from research-led autonomy toward production-grade safety engineering.

For autonomous vehicles, 2025 improvements built on existing safety systems and strengthened four areas.

Hardware

Added stronger redundancy across power, network, connectivity, chassis control, and compute, including independent compute modules designed to preserve 360° coverage if a single component fails.

Why it matters: Reduces single points of failure and helps the vehicle maintain safe operation when hardware issues occur.

Software & failsafes

Formalized a tiered fallback hierarchy, strengthened driver monitoring, and introduced a Minimal Risk Maneuver designed to reach a safer stop during faults.

Why it matters: Gives the system clearer, pre-defined responses when conditions move outside normal operation.

Perception, prediction & planner

Improved long-range detection, adverse-weather performance, recognition of vulnerable road users, school buses, emergency vehicles, and planning under occluded visibility.

Why it matters: Helps the vehicle understand and respond to more complex real-world road conditions.

Governance

Established a formal safety framework, strengthened incident review, tested each production release against 9,400+ simulation scenes before rollout, and introduced recurring driver recertification.

Why it matters: Reduces release risk and turns field experience into systematic safety improvements.

For delivery robots, Avride introduced the A1.5 model, adding improved suspension, a taller LED flag for visibility, and two-way communication. Field events are reviewed, root-caused, and fed back into model validation and product roadmaps.

Disclosures referencing ESRS

General disclosures

ESRS 2. General disclosures

Basis for preparation

ESRS 2 BP-1

General basis for preparation

Frameworks

Sustainability is an integral part of how Nebius operates. Nebius is not yet required to report under the European Sustainability Reporting Standards (ESRS) and does not claim formal compliance with ESRS or CSRD. We have chosen to use the ESRS framework as a voluntary reference, ahead of mandatory requirements expected to apply for 2027 disclosures.

In 2025, we conducted a formal double materiality assessment (DMA) in line with ESRS methodology, identifying material topics and associated impacts, risks and opportunities. This marks a significant step forward in the maturity of our sustainability reporting, as we build the processes, data infrastructure, and governance needed to support future mandatory disclosures.

This report is prepared on a voluntary basis and does not constitute a sustainability statement under Article 19a or 29a of the CSRD. Nebius does not claim full compliance with ESRS. Our disclosures are structured with reference to the November 2025 draft amended ESRS 2 framework, as part of our preparation for mandatory reporting.

Scope and consolidation

We report on a consolidated basis, using the same scope of consolidation as our financial statements, as detailed in our annual 20-F report.

This report does not yet include quantitative value chain metrics but provides an overview of how our value chain operates and the measures in place to manage impacts across upstream and downstream activities.

We have excluded sensitive information related to intellectual property, know-how and client relationships while ensuring our disclosures remain transparent and sufficiently detailed to provide a fair representation of our sustainability performance.

ESRS 2 BP-2

Disclosures in relation to specific circumstances

Time horizon and base year

This report follows a structured approach to time frames, defining short-term as one year, medium-term as two to four years, and long-term as more than five years.

This is our second sustainability report and the first one for the full year operating as Nebius. Where comparable 2024 data is available, we provide year-on-year context. Where data is not yet comparable due to structural changes or evolving measurement processes, 2024 figures are not provided side by side. In the next reporting period, disclosures will track progress against 2025 figures, while 2024 data will be mostly excluded from historical comparisons.

We continue to monitor the relevance of the EU Taxonomy Regulation to our business. At this stage, we have opted not to include Taxonomy disclosures, as the classification of our activities requires further analysis to ensure any disclosures would be reliable and meaningful.

Incorporation by reference

For some disclosures, we reference other documents where relevant.

Sustainability governance

ESRS 2 GOV-1

The role of the administrative, management and supervisory bodies

Sustainability is embedded in how we make decisions at every level, and our governance structure ensures that it is aligned with long-term value creation.

Board of Directors oversight

As of December 31, 2025, the Nebius Board of Directors consists of 8 members, of whom 5 (63%) are independent in line with Nasdaq independence criteria. The Board comprises 2 women and 6 men (25% female). Detailed information on the Board's composition, expertise, gender diversity and independence is available in our 20-F annual report, Item 6.

The Board brings deep industry expertise, with members who have experience in fast-growing sectors and a strong understanding of the risks and opportunities in our field. One Board member also has a background in IT education and tech-driven skill development, reinforcing our focus on leveraging the technological stack and expertise to create social impact.

The Board contributes to overseeing our sustainability performance, shaping our sustainability strategy, and ensuring that our sustainability commitments translate into real progress. Regular reviews are conducted to assess the impact of ongoing initiatives. The final draft of this sustainability report is reviewed by the Board, ensuring a critical evaluation before publication.

Executive decision-making

The executive leadership team is responsible for shaping and communicating our business vision and priorities across the organization, enabling leaders at all levels to align their efforts accordingly. Decisions are based on various inputs and key business results, which already incorporate sustainability gains. One of the key goals of this process is to ensure that major impacts, for example from energy-efficient innovations

or cloud optimization tools, are amplified and translated into competitive advantages.

Senior management & its role in sustainability

As of December 31, 2025, our executive management team — comprising the CEO and C-level executives — consists of 9 members, including 2 women (22%). Our broader senior management team, which also includes heads of department, comprises 114 members, including 26 women (23%).

Members of the senior management team are responsible for managing sustainability risks and opportunities within their domain. Since its inception, Nebius has appointed a Head of Sustainability who is responsible for driving alignment across teams, identifying new opportunities, embedding best practices to stay ahead of relevant legal and industry developments, and leading non-financial reporting.

Information flow to governance bodies on sustainability matters

Sustainability is part of decision-making, from day-to-day teamwork to Board-level oversight.

For example, matters related to energy use in our data center expansion are first identified at an operational level, by the infrastructure team. From there, they are evaluated in collaboration with subject-matter professionals from functions including finance, procurement and sustainability.

Once assessed, they are presented to the executive team for input and decision-making. Depending on how urgent or material a topic is, these discussions can happen more frequently than weekly.

In 2025, we continued to mature our sustainability governance. This included ongoing work to formalize ESG data monitoring processes, completing both the double materiality assessment and the climate risk assessment, and strengthening the data collection and validation procedures that underpin this report. These systems are still evolving, and we are committed to improving them as we grow.

The Board of Directors receives regular sustainability updates, typically on a quarterly basis — either as part of a general business update or as a dedicated session — with the annual sustainability report serving as the primary structured report for the Board's review and discussion.

ESRS 2 GOV-2

Integration of sustainability-related performance in incentive schemes

As of December 31, 2025, Nebius has not defined formal sustainability-specific KPIs within its incentive schemes for members of the management and supervisory bodies. The way we think about management incentives, however, goes beyond prescribed targets.

We treat building sustainably and building successfully are the same objective. Designing energy-efficient infrastructure, maintaining operational resilience, and governing the business responsibly are fundamental to what makes our business competitive, and this report illustrates that connection throughout.

Nebius relaunched its equity incentive plan in October 2024, which has been active throughout the 2025 reporting year. Through this plan, total compensation for management is directly linked to the long-term performance of the company. As a publicly listed company, Nebius operates in a market environment where capital markets increasingly factor sustainability-related issues into their evaluation of technology businesses. Building a more efficient, resilient and responsibly governed business therefore has a direct bearing on the outcomes that matter to management as shareholders.

We will continue to assess how sustainability considerations can be more explicitly integrated into our incentive framework, where relevant.

ESRS 2 GOV-3

Statement on due diligence

Nebius embeds due diligence for impacts on people and the environment through a structured, risk-based compliance framework covering both its own operations and third-party relationships. The framework is built on formal policies, monitoring mechanisms, and training programs designed to ensure that due diligence is comprehensive, proportionate, and aligned with regulatory expectations.

Our due diligence system covers both human rights and environmental aspects and is embedded through a number of key policies and procedures, including:

- **Code of Conduct** — sets ethical standards applicable to all employees and stakeholders, including commitments to human rights, non-discrimination and responsible business conduct.
- **Supplier Code of Conduct** — establishes expectations for suppliers regarding labor standards, human rights, environmental responsibility and legal compliance.
- **Speak-up Policy** (updated July 2025) — provides confidential reporting channels for concerns relating to unethical conduct, human rights violations or environmental risks.
- **Anti-Bribery and Anti-Corruption Policy** (updated July 2025) — promotes integrity and supports the prevention of misconduct that could lead to broader harm.
- **Trade Controls Policy** (introduced November 2025) — sets out the approach to managing sanctions and export control risks across jurisdictions, replacing the previous Sanctions Compliance Policy.
- **Third-Party Due Diligence Guidance** (internal document, introduced November 2025) — formalizes risk-based due diligence methodology for all third parties, covering trade controls, anti-corruption, human rights, environmental and geopolitical risks.
- **Conflict of Interest Policy** (internal document, introduced May 2025) — helps identify and manage situations where personal interests could interfere with business decisions.
- **Charitable Donations Policy** (internal document) — applies due diligence to donations, ensuring responsible giving.

How we identify adverse impacts

We have built several mechanisms into our day-to-day operations to help us catch potential negative impacts before they escalate. We stay alert in a number of ways:

- **Knowing who we work with:** Before we partner with anyone — a supplier, customer, or service provider — we run thorough due diligence checks. We look at their track records, policies and practices and overall business ethics to flag potential risks, including those related to sustainability performance. In 2025, we transitioned to a structured Third-Party Risk Management approach, with all third parties now assessed across sanctions, anti-corruption, human rights, environmental and geopolitical

risk dimensions and classified as high, medium or low risk accordingly. As of December 31, 2025, a total of 4,000+ third parties had been assessed.

- **Auditing our own house:** Our internal audit function helps us step back and evaluate how well our safeguards are working. These regular reviews allow us to identify gaps, take corrective action and continually improve.
- **Protecting data and privacy:** Our security and privacy teams monitor for risks such as data breaches or privacy concerns. They follow clear procedures to act on potential vulnerabilities — from logging an issue, identifying who needs to act to reporting and resolving it.
- **Through our people:** Our HR team plays a frontline role in identifying issues that could affect employee well-being and workplace practices. From regular check-ins to performance reviews, these touchpoints help us stay closely connected to employee experiences, including satisfaction levels, team dynamics and emerging concerns.
- **Listening when something is not right:** Our confidential Speak Up channel is open to employees and external stakeholders to report concerns related to unethical behavior, environmental or human rights. Every report is tracked through a secure case management system, helping us ensure proper oversight and handling.

How we address adverse impacts

We take a structured, practical approach to addressing adverse impacts.
The goal is to ensure that concerns are taken seriously, handled transparently and lead to improvements across the business.

- **Dialogue with stakeholders:** Whether it's customers, employees or partners, we make it a point to stay engaged. Their input helps us understand where potential issues may arise, and that actions we take are effective and reflect real-world concerns rather than internal assumptions. For employees, this includes engagement via the Works Council where employee representatives can raise workplace concerns directly and take an active role in ensuring they are acted upon.
- **Speak Up process:** For concerns raised through our Speak Up channels, we have developed a clear end-to-end process: Concerns are assessed, triaged, investigated, and remediated. As appropriate investigations are conducted independently, by qualified and skilled (external) investigators — bound by confidentiality requirements.
- **Team-level safeguards:** Each team has its own set of internal checks and escalation paths. For example, the HR team has a process to address employees' disciplinary issues that are built to be fair, consistent, and transparent. All the concerns are investigated first, and employees are invited to share their perspective. To protect privacy, the whole process stays confidential, and outcomes are shared only with those directly involved on a need-to-know basis.
- **Awareness building:** Raising awareness around ethical and accountable conduct is a core part of how we equip our teams to make informed decisions every day. Our program is structured around an annual Ethics and Compliance Training and Awareness Plan, combining mandatory company-wide training with targeted, role-based sessions for higher-risk functions. In 2025, a mandatory Code of Conduct and Speak Up training was deployed to all employees. In addition, role-based anti-bribery and anti-corruption training was delivered to Sales, trade controls training to Legal, Infrastructure and Security, and third-party risk management training to Procurement and Sales.
- **Leadership at the table:** Regular management meetings provide a forum to surface and resolve concerns, track progress on our initiatives, and make sure that addressing impacts is a shared responsibility across the business

What we did in 2025

In 2025, our Ethics and Compliance work moved from building foundations to putting them into practice. A focus was strengthening our policy framework, with the Board approving updated and newly introduced policies covering anti-bribery and anti-corruption, trade controls, conflict of interest, speak up and third-party risk management. On the third-party side, we procured and implemented a dedicated risk management tool and expanded our due diligence scope to include environmental and social risk factors and CSDDD considerations. We also introduced a Gifts and Entertainment register, launched an annual Conflict of Interest questionnaire, and integrated conflict of interest screening into the hiring process. Looking ahead, work is underway on a structured Ethics and Compliance KPI dashboard, planned for implementation in 2026.

Core elements of due diligence	Described in
Embedding due diligence in governance, strategy and business model	ESRS 2: GOV-1, GOV-2, SMB-3
Engaging with affected stakeholders in all key steps of the due diligence	ESRS 2: GOV-1 SMB-2, IRO-1; E1; E5; S1; G1
Identifying and assessing adverse impacts	ESRS 2: GOV-3, IRO-1, SMB-3
Taking actions to address those adverse impacts	ESRS 2, GOV-3, E1; E5; S1; G1
Tracking the effectiveness of these efforts and communicating	E1; E5; S1; G1

ESRS 2 GOV-4

Risk management and internal controls over sustainability reporting

In the 2025 reporting cycle, the quality of sustainability data is upheld through multiple layers of control. At the operational level, data owners are responsible for sourcing relevant data from internal systems and, where applicable, external sources, and perform accuracy checks to ensure reliability. As the scope of reporting expanded in 2025 to include additional operational sites, including newly active data centers, data collection guidelines were refreshed to ensure consistency in how data is captured across all locations. Data collection forms were updated to reflect both the amended ESRS requirements and a broader set of metrics identified as operationally relevant. Where data is spread across geographies, consolidation points were established to support a more structured flow of information into the reporting process.

The Head of Sustainability oversees the data collection, consolidation and reporting process, reviewing submissions for completeness and consistency, enabling timely corrections and process improvements. The Head of Sustainability also consults with Internal Audit to validate key material data points, and engages Legal and Ethics and Compliance to ensure that sustainability claims are substantiated, supporting future audit readiness.

The final draft of the sustainability report is presented to the board of directors for critical review and feedback, and is signed off by the Head of Sustainability and the executive team responsible for reporting oversight.

The 2025 reporting process remains largely manual in terms of data consolidation and verification. Going forward, the focus is on building a data foundation that reduces reporting effort and makes sustainability data available in support of business decision-making on an ongoing basis.

Strategy

ESRS 2 SBM-1

Strategy, business model and value chain

Detailed information on our strategy, market position, products and services, and financial performance can be found in our annual 20-F report, Items 4 and 5. For additional detail on sustainability agenda, see the respective chapter of this sustainability report. For details on employee composition, see section S1-5 Characteristics of the undertaking's employees.

Nebius operates in the technology sector, providing AI infrastructure and cloud computing services. Nebius is not active in the fossil fuel sector, chemicals production, the controversial weapons industry, or the cultivation and production of tobacco.

ESRS 2 SBM-2

Interests and views of stakeholders

Understanding the perspectives of our stakeholders helps us shape a business that's competitive as well as grounded in real-world needs and expectations. As a provider of AI infrastructure, our stakeholder landscape is diverse — from developers training next-gen models to the communities where our data centers operate.

Who we listen to:

Our key stakeholders include existing and potential customers, employees, partners and suppliers, investors, regulators, local communities, and members of the tech and research communities.

How we engage:

We rely on a mix of direct dialogue and indirect insights. Customer feedback loops, employee meetings, investor updates, supplier reviews, audits, continuous dialogue with regulators, media monitoring and research collaborations all help us stay informed.

How this shapes what we do:

This input influences how we evolve our services, where we invest, and how we act. Customer feedback shapes how we develop our platform, from adding tools and managed services to pursuing certifications that let us serve customers in regulated sectors. Staying connected to the research community ensures our platform evolves in step with where AI is heading. Ongoing dialogue with chipmakers and suppliers gives us early visibility into next-generation architectures, so we can plan ahead on infrastructure and performance. Conversations with local municipalities inform where we expand and how we operate responsibly in new regions. Investor and analyst engagement keeps our capital allocation grounded in market reality. By staying aware of broader industry concerns, from power availability and evolving AI regulation to the competition for specialized talent, we are able to act early and build resilience into how we operate.

Stakeholder group	How we engage	Key interests	
Customers Current and potential. AI developers, startups, enterprises	Account check-ins, customer support channels, industry events, market research	Flexibility & cost-efficiency of compute	Responsive customer support
		Reliability & security	Data center sustainability
		Workload optimization	
Employees	Team all-hands, performance reviews, open-door policy, Works Council representation	Opportunities for career growth	Transparent communication
		Work-life balance	Inclusive workplace culture
Supplier and partners Hardware vendors, energy providers, and other	Contracts with review points, planning sessions, regular ops syncs	Early visibility into demand forecasts	Stable financial performance
		Responsible business conduct	
Local communities	Local partnerships, community open houses, local government dialogue, dedicated project websites	Transparency on facility operations	Local development and benefit programs
		Responsible resource use	Job creation
Investors and shareholders	Quarterly updates, investor Q&A sessions, industry events	Growth and profitability	Competitive differentiation
		Efficient capital utilization	Business model resilience
Regulators and policymakers	Reporting, consultation, roundtables	Facilitation of AI innovation	Efficient energy usage
		Compliance with applicable regulations Incl. data security, privacy, labor, climate change mitigation	Job creation
		Transparency	
Tech and research communities	Research partnerships, industry events	Access to top-tier GPUs for research and innovation	Responsible AI
		Open-source contributions	

Approach to materiality assessment

ESRS 2 IRO-1

Description of the processes to identify and assess material impacts, risks and opportunities

In 2025, Nebius updated its double materiality assessment approach in line with ESRS methodology, applying the most recent available standards at each stage of the process and monitoring for updates to adjust the approach accordingly. In parallel, we conducted a climate risk assessment, with further details reported under E1-2, Identification of climate-related risks and scenario analysis. The assessment focused primarily on the Group's core business, Nebius, though a broader set of impacts was considered across the Group during the context-setting phase.

Full DMA updates are intended on a two-year cycle, with the next full refresh planned for 2027. Trigger-based reviews will be conducted in the event of significant changes in business priorities, the external environment, or reporting standards. Materiality scoring relied on qualitative judgment rather than quantitative measurement. We continue to monitor EFRAG's guidance to ensure our approach stays aligned as we progress toward mandatory reporting.

The 2025 approach to materiality assessment

Steps 1-2. Defining topics. Structured interviews with key internal stakeholders were followed by a review of Nebius's public disclosures, peer analysis spanning AI cloud providers and hyperscalers, industry trends, sustainability frameworks and a media scan.

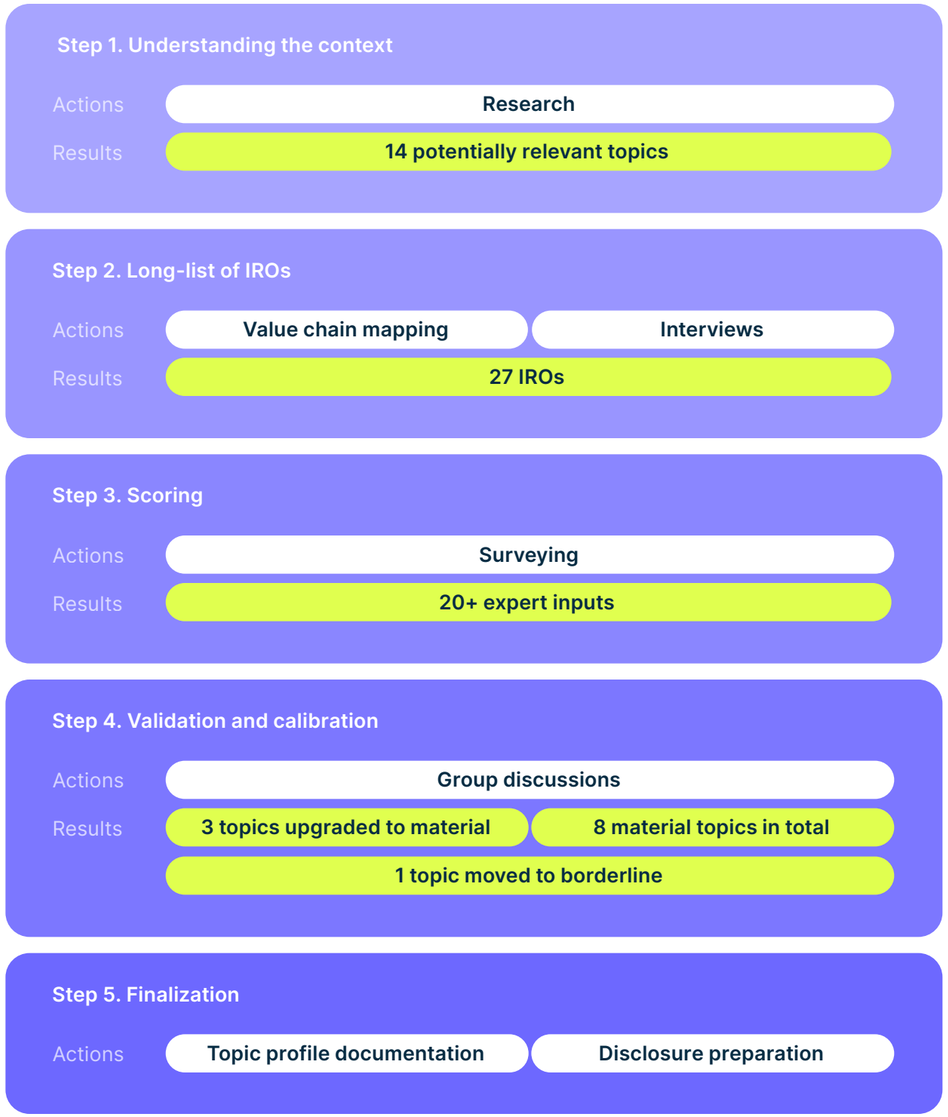
Step 3. Stakeholder engagement and scoring. Materiality was assessed by a representative group of C-level officers and heads of department with decision-making authority across the business, spanning functions including Infrastructure, Product, Finance, Legal, HR, Sales, Security and others. At least two representatives were invited per topic area to allow views to form and be tested independently.

The process was structured to enable independent judgment while allowing for clarification, to ensure responses reflected genuine assessment. Stakeholders assessed the severity and likelihood of impacts, and the financial magnitude and likelihood of risks and opportunities, over short-, medium- and long-term time horizons.

Step 4. Validation and calibration. Validation sessions were held with stakeholder groups by topic area to sense-check scores, surface disagreements, align on interpretation and ensure conclusions reflected a considered collective view rather than an aggregation of individual responses.

Step 5. Finalization and borderline radar. Material topics were confirmed alongside a full picture of borderline topics that will not be reported on in detail this cycle.

Some of these are widely considered material across the broader industry. Water stewardship is one such topic: for most data center operators, water use for cooling is a pressing concern. Nebius's closed-loop liquid cooling design means we do not rely on water intake for cooling, which is why this topic does not register as a material impact for us. Given its relevance to the industry, we have included an explanation of our approach in this report.



ESRS 2 SBM-3

Interaction of material impacts, risks and opportunities with strategy and business model

● Positive impact ● Negative impact ● Opportunity ● Risk ☑ Material ☐ Borderline ☒ Not material

ESRS	Topic	Impact materiality	Financial materiality	IROs	How it connects to our business model and strategy	Financial effect (current / anticipated)
Material						
E1	**Climate change and energy efficiency** Energy · Climate change mitigation	☑	☑	● ● ● ● ● ● ● ● ●	Energy is our largest operational input. How efficiently we use it shapes our cost per GPU, our carbon footprint, and our ability to scale. Innovations in server design, cooling and workload orchestration are central to how we compete and to how we help customers reduce their own compute emissions.	Energy costs represent a material share of operating expenses, with direct exposure to electricity price volatility. Continued efficiency improvements are expected to support cost competitiveness. Evolving carbon pricing requirements may introduce additional cost exposure, monitored as part of our climate risk assessment.
E5	**Resource use and circular economy** Resource inflows · Resource outflows (waste)	☑	☑	● ●	We design our servers and racks for longevity, repairability and component reuse, reducing replacement frequency, demand for virgin materials and e-waste, while keeping total cost of ownership competitive.	Circular design contributes to lower capital expenditure over equipment lifecycles. Extended useful life of GPU generations has supported revenue beyond typical refresh cycles, with further procurement cost benefits expected over the medium term.
S1	**Attraction, retention and employee well-being** Working conditions · Social dialogue and freedom of association · Training and skills development	☑	☑	● ● ● ●	Our ability to build and operate AI infrastructure at scale depends on specialized technical talent. We focus on culture, working conditions, and meaningful work as the primary drivers of retention.	Talent availability affects our capacity to execute on infrastructure expansion. High turnover in specialized roles carries measurable recruitment and onboarding costs. A stable workforce reduces execution risk as we expand into new markets and market segments.
S1	**Human capital development** Training and skills development	☑	☑	● ● ●	Skilled, technically current teams enable Nebius to move fast, build reliably, and stay ahead. Investing in development keeps our people at the frontier of AI infrastructure, translating into better products, faster execution and the ability to capture emerging opportunities.	Technically strong teams are a direct driver of competitive position and growth. Investing in development also supports retention, reducing dependence on external hiring and the costs that come with it.

ESRS 2 SBM-3

Interaction of material impacts, risks and opportunities with strategy and business model

Positive impact · Negative impact · Opportunity · Risk · Material · Borderline · Not material

ESRS	Topic	Impact materiality	Financial materiality	IROs	How it connects to our business model and strategy	Financial effect (current / anticipated)
S1	**Diversity, equality and inclusion** Gender equality and equal pay	☑	~	● ●	A workforce reflecting a wide range of perspectives is better equipped to solve complex problems and innovate. Inclusion drives engagement and retention in a talent market where the best people have many options.	Retention across all roles has a direct effect on recruitment and onboarding costs. A reputation as an inclusive employer strengthens our ability to attract talent over the long term.
G1	**Business ethics** Corporate culture	☑	☑	● ● ● ● ● ●	A strong ethics and compliance framework protect our reputation, keeps us in good standing with regulators, and supports the trust of the customers and partners we depend on.	Partners and customers who trust us tend to deepen their relationships with Nebius, supporting revenue growth. With suppliers, ethical conduct supports better commercial terms and early access to high-demand resources. Non-compliance could result in financial penalties or reputational damage.
Company-specific	Data protection, privacy and cybersecurity	☑	☑	● ● ● ● ● ●	Our customers trust us with their workloads and sensitive research. Robust data protection and cybersecurity are core to what makes Nebius a reliable infrastructure partner, and security certifications open access to enterprise and regulated-sector customers.	A material cybersecurity incident could have significant financial and reputational consequences. Strong security practices and certifications support access to regulated markets, with a positive expected effect on revenue over the medium term.
Company-specific	Products and services with social benefits	☑	☒	●	Our platform makes high-performance compute accessible to startups, researchers and enterprises that would otherwise be priced out. In the communities where we operate, we create local value through job creation, skills development and partnerships.	Certifications for regulated sectors may open addressable markets currently inaccessible. Community partnerships and local hiring support site approvals and permitting, with practical value for expansion.

ESRS 2 SBM-3

Interaction of material impacts, risks and opportunities with strategy and business model

Positive impact · Negative impact · Opportunity · Risk · Material · Borderline · Not material

ESRS	Topic	Impact materiality	Financial materiality	IROs	How it connects to our business model and strategy	Financial effect (current / anticipated)
Borderline						
S3	Community impact	Borderline	Borderline		Our data center expansion creates opportunities for local economies through job creation, skills development and local procurement. Our buildout also interacts with local energy systems, land use and infrastructure, which we actively seek to manage responsibly. Assessed as borderline for 2025 given the operational footprint for that year. As we grow, the materiality of this topic is expected to grow with it, and we are building the relevant management practices ahead of that.	Local support contributes to predictable permitting and site development timelines, supporting on-schedule delivery of new capacity. As Nebius scales into new regions, proactive community engagement is expected to reduce execution risk.
E3	Water stewardship	Not material	Borderline		This topic is not material from an impact perspective: Nebius does not rely on water intake for cooling. Our closed-loop systems operate independently of local water sources. This independence is a financial advantage as water availability becomes a regulatory and operational constraint across the data center industry.	Our cooling approach supports site approvals, reduces regulatory risk, and differentiates Nebius with customers who evaluate the environmental footprint of their infrastructure providers.

The DMA was conducted at sub-topic level. Material topics in this table reflect the aggregate result across sub-topics. Reporting in this sustainability statement covers only the sub-topics where material impacts, risks or opportunities were identified, consistent with ESRS 1, paragraph 30. The following topics were assessed and found not material on both dimensions: Pollution (E2), Biodiversity (E4), Occupational health and safety (S1), Workers in the value chain (S2). Financial effects have not been quantified. The descriptions reflect a high-level qualitative assessment conducted by internal teams as part of the 2025 double materiality assessment process.

ESRS 2 IRO-2

Disclosure Requirements in ESRS covered by the undertaking's sustainability statements

Please refer the ESRS Index in Annexes. The index has been updated to reflect the material topics identified in the 2025 cycle and to align with the November 2025 draft amended ESRS 2 structure.

Environmental disclosures

ESRS E1. Climate change

Material impacts, risks and opportunities (IROs)

Climate change and energy efficiency

Electricity consumption across our data center operations is the primary driver of our GHG footprint, making energy efficiency both an environmental and business imperative. As demand for AI compute increases, efficiency determines how much useful compute can be delivered within available power, space, cooling, and hardware capacity. It also shapes customer value by lowering energy intensity per workload and supporting customers' own climate commitments. We recognize that our broader footprint also includes upstream emissions from hardware manufacturing, data center construction, cooling fluid procurement, and logistics. Design choices that extend hardware lifecycles, reduce material intensity, and improve supplier-side efficiency help limit these impacts as we scale.

How we manage the IROs

Through energy-efficient infrastructure design, smart workload orchestration, and low-carbon energy sourcing. Read more in the Efficiency by Design chapter of this report.

Why it matters

Energy use is the main driver of data center emissions and one of the core physical constraints on AI infrastructure growth. As compute demand expands, providers need to secure power in increasingly constrained markets while managing the carbon intensity, cost, and regulatory exposure of that consumption.

For Nebius, energy efficiency has a direct business impact. It affects cost of revenue, margin resilience, speed of deployment, access to power in constrained markets, and the carbon profile of customer workloads. Our approach is to optimize infrastructure before demand arrives, so we can scale capacity faster and more efficiently as AI compute becomes more widely used.

Efficient infrastructure allows more compute to be delivered within available power, space, cooling, and hardware capacity. It also helps customers reduce the energy intensity of their AI workloads while supporting Nebius's ability to grow with stronger margins and lower exposure to energy price volatility and carbon costs.

Our approach, policies and targets

E1-1

Transition plan for climate change mitigation

Nebius does not yet have a formalized climate transition plan. Following the formation of Nebius Group in mid-2024, our current focus is on establishing the foundations for target-setting, including methodology and governance structures.

Our primary climate mitigation approach at this stage is infrastructure design. The aim is to build infrastructure with low emissions per unit of compute from commissioning, avoiding high-impact assets that would require subsequent decarbonization. Current decarbonization levers include renewable energy procurement, liquid cooling deployment, hardware efficiency optimization, workload orchestration, and heat recovery where local demand exists. This approach reflects both our sustainability objectives and our business model, where resource consumption is a direct operational cost driver.

Further decarbonization pathways will need to be assessed as part of transition plan development. The nature and magnitude of locked-in emissions will also need to be assessed to ensure they are adequately reflected in the plan.

E1-2

Identification of climate-related risks and scenario analysis

In 2025, we completed our first climate risk and opportunity assessment. The assessment identified five material and borderline risks and opportunities across transition and physical risk categories. All borderline risks and opportunities are expected to require closer monitoring and potential reclassification in the next assessment cycle.

Transition · Physical · Opportunity · Risk · Material · Borderline

2025 climate risk assessment outcomes

Name	Transition / Physical	Risk / Opportunity	Materiality	Description
Internal technological innovations	Transition	Opportunity	Material	Full-stack control allows Nebius to reduce energy use per unit of compute across hardware, cooling, and orchestration. This lowers emissions intensity and cost of revenue, making efficiency a core advantage as capacity scales.
Heatwave	Physical	Risk	Borderline	Increasing frequency and intensity of extreme heat events poses operational and financial risks to data centers by raising cooling demand and stressing local electricity grids. Prolonged high temperatures can accelerate equipment degradation, increase energy costs, and risk service interruptions. This risk is classified as borderline material, reflecting current design resilience.
ESG-aligned market demand	Transition	Opportunity	Borderline	As scrutiny of AI workload emissions increases, customers are increasingly selecting providers who can demonstrate low-carbon operations, transparent reporting, and credible climate commitments.
Regulatory and carbon cost exposure	Transition	Risk	Borderline	Tightening climate regulation and rising carbon costs represent a financial and compliance risk as an energy-intensive technology company.
Renewable energy sourcing	Transition	Risk	Borderline	Dependence on fossil-based electricity at sites where sufficient renewable capacity is unavailable represents a cost and emissions risk. Exposure varies across our operating regions, with certain markets presenting challenges in securing renewable supply.

Risks and opportunities assessment coverage

The assessment covered data center sites operational as of December 31, 2025. It assessed transition and physical climate risks across short-, medium-, and long-term horizons using stakeholder input, industry benchmarking, climate data, and scenario analysis.

Risks were evaluated based on likelihood and financial magnitude, using stakeholder input, industry benchmarking, climate data, and site-level physical risk modeling. As part of the physical risk review, 26 EU Taxonomy climate hazards were screened, with nine carried forward for site-level analysis. The results will inform future planning and action development.

Climate scenarios

Scenario	Low emission	High emission
Description	Assumes aggressive global decarbonization with strict climate policy, rapid renewable deployment, and rising carbon prices. Under this scenario, Nebius faces increased compliance costs and transition pressures.	Assumes limited policy intervention and continued fossil fuel dependence, leading to global temperatures exceeding 2.5–3°C by 2100. This scenario increases exposure to physical climate hazards such as extreme heat and weather disruptions, which potentially raise long-term operational costs for data center resilience and limits access to ESG-aligned capital.
Reference	NGFS Net Zero 2050 / IEA NZE2050 / IPCC SSP1-1.9	NGFS Current Policies / IEA STEPS / IPCC SSP5-8.5
Warming pathway	1.4–1.8°C by 2100	>3°C by 2100
Risk Category Assessed	✓ Physical risks ✓ Transition risks	✓ Physical risks ✗ Transition risks[1]

(1) Do not materialize as this scenario does not assume policy, market, and technological shifts.

E1-3

Resilience in relation to climate change

Our resilience approach is built into how we select, design, power, and operate our infrastructure.

Site selection and building design

New data center sites are screened for physical climate hazards, including flooding, wildfire, subsidence, and extreme heat, alongside energy access and infrastructure stability. Geographic diversification also reduces concentration risk from localized weather events and grid disruptions. Read more about site selection and building design in Data Centers section of this report.

IT design

Nebius hardware is designed to operate at elevated temperatures: up to 40°C (104°F) cooling air temperature and 45°C (113°F) liquid temperature for liquid-cooled loops. This provides a thermal buffer during heat events, reduces cooling sensitivity, and expands the range of viable site locations. Read more about how we design our hardware in Hardware section of this report.

Cooling systems

Our cooling architecture matched to workload density and local conditions. At sites we own or design, we use air-based free cooling where thermal density allows, closed-loop liquid cooling for high-density workloads, or a hybrid of the two. These designs reduce reliance on chillers, refrigerants, and external water loops while helping maintain performance under changing climate conditions. Read more about key features of our cooling systems in Efficient cooling systems section of this report.

Energy supply

Energy resilience is supported through backup power, long-term energy contracts, renewable PPAs, and geographic diversification. We are also exploring advanced energy systems as a lower-emission alternative to conventional grid and combustion-based generation, including the 2026 Bloom Energy partnership for fuel cell technology at one US site. Read more about our approach to energy sourcing in Energy sourcing section of this report.

E1-4

Policies related to climate change mitigation and adaptation

To manage our impacts, risks and opportunities related to climate change, we rely on a set of internal policies and operational guidelines that shape how we design infrastructure, expand its geographical footprint, and engage with third parties. As we introduce more structured processes for identifying and assessing climate risks, we expect to revisit and expand the current policy scope to ensure we have the right mechanisms and guidance in place to address the most material climate impacts.

Code of Conduct

Our Code of Conduct affirms our commitment to minimizing environmental impact. It includes expectations around compliance with environmental laws, responsible resource use, and emissions reduction.

Infrastructure design & energy management guidelines

We maintain best practices for designing our own hardware and data center systems to improve energy efficiency across the stack. A dedicated R&D team works on optimizing energy loads, testing hardware performance, and developing smarter orchestration. Our engineers also design and implement custom cooling and heat recovery systems, which reduce operational emissions and improve energy reuse in nearby communities.

Siting and partner selection criteria

Our long-term focus is on designing our own data centers, where we can maximize efficiency gains and drive stronger sustainability outcomes. For these projects, we apply additional screening criteria related to energy infrastructure and climate change mitigation — including cooling potential (e.g. ambient temperature), opportunities for heat reuse, and access to low-carbon energy.

To support rapid growth and ensure capacity is available where and when it's needed, we work with a number of colocation sites across regions. When selecting new colocation sites, we apply a clear set of criteria to reduce exposure to both operational and environmental risks. This includes avoiding areas with high risk of natural disasters (such as floods, wildfires, earthquakes) and considering broader stability factors such as energy security and physical safety.

Supplier Code of Conduct

Suppliers are required to comply with relevant environmental laws and encouraged to actively reduce their emissions and material footprint. The Code promotes transparency and ongoing collaboration to improve sustainability performance across the value chain.

E1-5

Actions and resources in relation to climate change mitigation and adaptation

Climate change mitigation

Our mitigation actions focus on reducing emissions intensity per unit of compute. They are built into how we design and operate the full technology stack, so energy efficiency is applied from commissioning rather than added later through retrofits. Our mitigation levers are:

Designed energy efficiency. Servers, racks, cooling systems, and data center infrastructure are designed to reduce energy demand from the outset.

Operational optimization. We use software features to reduce avoidable energy use during operation. Workload orchestration reduces idle GPU time and overprovisioning, while thermal-aware controls align cooling with actual load as demand patterns, GPU generations, and fleet capacity change.

Renewable and low-carbon energy sourcing. We work to increase the share of low-carbon and contractually secured renewable energy, including through long-term PPAs where feasible.

Quantification of GHG reduction

As Nebius is scaling rapidly, absolute emissions and energy use are still primarily driven by growth in deployed capacity. We are not yet quantifying achieved or expected GHG reductions by individual mitigation lever, as our operating baseline is still stabilizing.

Climate change adaptation

Our adaptation actions focus on resilience to physical climate risks and energy system constraints. We elaborate on them under E1-3 Resilience in relation to climate change.

Resources

Climate-related actions are funded through core capital and operational decisions rather than a standalone budget. These include hardware R&D, data center development, energy procurement, and infrastructure engineering. The sustainability team supports implementation with engineering, infrastructure, finance, and legal teams.

E1-6

Targets related to climate change

Our key target areas span the full technology stack and reflect how we measure the impact of our decarbonization actions. These include infrastructure efficiency, including PUE and energy-to-AI output efficiency, and energy sourcing, including renewable and low-carbon energy coverage. We are also tracking GHG emissions intensity, recognizing that absolute energy consumption will grow as we scale and that carbon efficiency per unit of output is a more meaningful indicator of progress.

We are not disclosing specific targets at this time, as they are currently undergoing calibration in the context of continued business scaling and recent organizational changes. We expect to be able to provide more details on targets in future reporting periods.

Insights

In 2025, energy consumption was almost entirely driven by data center operations, which accounted for 99% of total Group energy use. Consumption grew as the infrastructure footprint expanded. Finland-1 remained the largest single site, accounting for the biggest individual share of total data center energy consumption. Both Finland-1 and sites outside Finland-1 were adding capacity during the year.

Energy data covers seven data center sites operational during 2025. Finland-1, France-1, Iceland-1, and Missouri-1 operated for the full year, while Israel-1, UK-1, and New Jersey-1 came online during the year and contribute partial-year figures. Iceland ramped to full operational load by mid-year. As sites operate for the full year and new capacity comes online, we expect energy use to increase in 2026.

80% of Group energy consumption came from low-carbon sources and 62% from renewable sources, based on documented sourcing instruments and local grid production mixes. This includes PPAs, renewable energy attribute certificates, on-site solar generation, and the low-carbon share of electricity grids where contractual instruments were not in place.

In Finland-1, on-site solar generated approximately 100 MWh in 2025, supplying the associated office load. Across sites where Nebius sourced electricity directly from the grid, the weighted average low-carbon share was 85%, reflecting the different generation mixes of the regions where we operate. Operating across multiple geographies helps balance our energy sourcing profile and allows us to benefit from cleaner grids where they are available.

E1-7

Energy consumption and mix, 2025

Energy consumption related to own operations, MWh	
Total energy consumption	249,842
Total energy consumption from **non-renewable** sources	95,273
Fossil fuels sources	50,412
Direct combustion	18,789
Grid (production mix)	31,623
Nuclear sources	44,861
Grid (production mix)	44,861
Total energy consumption from **renewable** sources	154,569
Direct combustion	31
Onsite generation	100
Contractual instruments	80,771
Grid (production mix)	73,666

Energy consumption share, %	
Share of renewable energy consumption	62%
Share of low-carbon energy consumption	80%
Share of non-renewable energy consumption	38%

How data is compiled

Energy consumption related to own operations

Includes energy consumed by assets within the reporting boundary: electricity and fuel used by data centers operated as of the end of the year; electricity, heat and cooling (where applicable) used by offices covering the majority of total office floor area; and electricity and fuel used by Avride-operated vehicles, including autonomous and other company vehicles.

Energy attribute accounting

Energy attributes are assigned using a hierarchy based on the most direct available link to the energy source. Contractual instruments are applied first, followed by on-site generation, supplier-specific factors where available, and then the relevant country or regional production mix. Production mix data is sourced from AIB, RTE for France, DUKES for the UK, and EPA eGRID for the US, with preference given to the most granular geographic data available.

Share of renewable energy consumption

Renewable energy consumption is calculated by classifying energy in the following order: contractual instruments that specify renewable supply, on-site renewable generation, supplier-specific renewable factors where available, and the renewable share of the relevant local grid mix. Renewable sources include wind, solar, hydro, geothermal, and biofuels. Nuclear and fossil sources are excluded.

This differs from the default ESRS method, which counts only contractually secured renewable energy. We use this approach because it reflects both documented renewable sourcing and the renewable content of local grids, including the impact of locating infrastructure in regions with cleaner electricity systems.

(continues on the next page)

(continued from the previous page)

Share of low-carbon energy consumption

Low-carbon energy consumption represents the combined share of renewable and nuclear energy sources. Renewable energy is calculated as described above. Nuclear energy is sourced from the relevant country or regional production mix where applicable. Nuclear is considered to have a zero greenhouse gas emission factor, reflecting its negligible direct greenhouse gas emissions relative to fossil fuel-based energy generation.

Share of non-renewable energy consumption

The share of non-renewable energy consumption is calculated by applying the same attribute accounting hierarchy as for renewable energy, counting energy sourced from non-renewable origins at each level.

This includes direct fuel combustion from stationary sources such as backup generators and onsite power generation, as well as grid-sourced electricity attributed to non-renewable origins based on contractual instruments, supplier-specific factors, or country and regional production mix data where no more specific source is available. Nuclear energy is classified as non-renewable under this methodology, and included in the total non-renewable as well as broken out as a sub-category separately.

Energy sourced from renewable origins at any level of the hierarchy is excluded from the non-renewable share calculation.

E1-8

Gross Scopes 1, 2, 3 and Total GHG emissions, 2025

Greenhouse gas (GHG) emissions, tonnes CO2-eq	
Direct GHG emissions (Scope 1)	3,840
Indirect GHG emissions (Scope 2), market-based	61,161
Indirect GHG emissions (Scope 2), location-based	27,074
Indirect GHG emissions (Scope 3)	N/A
Direct biogenic GHG emissions (Scope 1)	8
Total GHG emissions (Scope 1 and Scope 2), market-based	65,001
Data centers	64,095
Other assets	906
Total GHG emissions (Scope 1 and Scope 2), location-based	30,914
Data centers	30,028
Other assets	886

Insights

In 2025, the majority of the Group's carbon footprint (99% of market-based and 97% of location-based Scope 1 and 2 emissions) was attributable to data center operations. The remainder was associated with office operations and vehicle fuel and electricity use by Avride-operated vehicles.

Scope 1

In 2025, Scope 1 emissions came from backup diesel generator testing, limited natural gas use for electricity generation toward year-end, and vehicle operations. Fugitive emissions from refrigerants are not included due to incomplete data coverage across the portfolio. Direct fuel emissions remained low relative to total energy consumption, supported by limited generator use and Avride's electric autonomous fleet.

Scope 2

Scope 2 emissions are reported under both location-based and market-based methods, which produced materially different results in 2025. The location-based method reflects the average carbon intensity of electricity generated in the grids where we operate. The market-based method reflects the electricity attributes we can contractually claim. Where renewable energy attestations are in place, emissions are treated as zero; where they are not, residual mix factors are applied, representing the remaining grid mix after renewable attributes have been claimed by other buyers.

The difference was mainly driven by Finland-1, our largest facility by energy consumption as of 2025. Finland-1's production mix factor is low at 33 gCO_2e/kWh, reflecting a grid that is more than 95% low-carbon, while its residual mix factor is much higher at 406 gCO_2e/kWh. In 2025, Finland-1 purchased electricity from the grid. Although the grid is very low-carbon, the market-based methodology requires the residual mix factor to be applied where renewable attributes are not held. This effect is particularly pronounced in the Nordic region, where substantial low-carbon electricity is generated but renewable attributes are also widely exported through the European Guarantees of Origin system.

Looking ahead, the PPA for Finland-1 we signed in 1H 2026 is expected to cover 100% renewable electricity for the site and reduce the gap between location-based and market-based emissions in future periods. Other sites sourcing renewable electricity include France, Iceland, and the UK, where colocation partners cover Nebius's electricity use with renewable electricity sources.

Scope 3

In 2025, Nebius has preliminarily identified the Scope 3 categories expected to be most material to the Group. However, Scope 3 reporting boundaries, data systems, and processes for collecting sufficiently reliable supplier and activity data are still being developed. Therefore, we did not calculate Scope 3 emissions for 2025, as the available data would not support a representative inventory. We are strengthening our data coverage, calculation methodologies, and controls to report material Scope 3 emissions in the coming reporting periods.

Greenhouse gas (GHG) intensity	
Scope 1 + 2 over revenue, tonnes CO_2e/ mln USD	
market-based	123
location-based	58
Scope 1 + 2 over total power consumed, tonnes CO_2e/ MWh	
market-based	0.26
location-based	0.124
Scope 2 over total electricity consumed, tonnes CO_2e/ MWh	
market-based	0.265
location-based	0.118

How data is compiled

Greenhouse gas (GHG) emissions

Scope 1 and Scope 2 emissions are reported for the consolidated accounting group aligned with Nebius annual financial report. The scope of assets used for emissions calculations corresponds to those used in energy consumption reporting.

The calculation of GHG emissions is based on conversion factors provided by internationally recognized authorities, including the US EPA, the UK DEFRA, the IPCC Guidelines for National Greenhouse Gas Inventories, and EEA. We report our GHG emissions in CO_2 equivalents, encompassing the impacts of methane (CH_4) and nitrous oxide (N_2O) within our Scope 1 and Scope 2 emissions. The global warming potential factors are derived from the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Emissions factors are reviewed annually, and updates are made when necessary.

Direct GHG emissions (Scope 1)

Scope 1 emissions are reported in accordance with the GHG Protocol and cover direct emissions from sources within the consolidated reporting boundary. Stationary sources include on-site backup generators at data centers and limited natural gas generation that came online toward the end of the reporting period. Mobile sources include fuel use from Nebius-operated vehicles, primarily associated with Avride operations. Fugitive emissions are not included in the 2025 Scope 1 inventory because reliable data on potential sources, including refrigerants, was not available across all operated facilities during the reporting period.

Fuel consumption data is collected from site-level energy management systems, facility operators, and fleet management teams. Data is reviewed for completeness and consistency before emissions are calculated, with estimates used where direct metering data is unavailable. We are working to improve data coverage in future reporting periods.

Indirect GHG emissions (Scope 2)

Scope 2 emissions are reported in accordance with the GHG Protocol Scope 2 Guidance and cover indirect emissions from purchased electricity, heat, and cooling consumed across operated facilities. Steam consumption was not applicable in 2025.

Electricity data is collected from utility invoices, site-level metering, and facility operator reporting. For colocation sites, consumption is based on Nebius's allocated share of facility electricity use. Data is reviewed for completeness prior to emissions calculation. Purchased heat and cooling represented a minor share of total energy consumption. Location-based emissions are calculated using average production mix factors for the relevant country or region. Market-based emissions use supplier-specific factors where available and residual mix factors otherwise, reflecting renewable energy attestations where held. The methodology for assigning energy attributes and calculating renewable / non-renewable shares is described under E1-7 Energy consumption and mix.

GHG emissions intensity (Scope 1+ 2 over revenue)

Total emissions divided by annual group revenue in 2025, as reported in the 20-F consolidated statements of operations.

GHG emissions intensity (Scope 1 + 2 over total power consumed)

Total emissions divided by total energy consumption in 2025 (electricity, fuel, and heat / cooling) in MWh.

GHG emissions intensity (Scope 2 over total electricity consumed)

Total Scope 2 emissions divided by total electricity consumption in 2025 (purchased, generated onsite) in MWh.

ESRS E3. Water

Borderline materiality in 2025

Water is a closely watched issue for data centers, especially where evaporative cooling increases freshwater use and exposure to local water stress.

For Nebius, the topic is assessed as borderline material because our infrastructure is designed to reduce or eliminate water dependency. At sites we own or design, we use air-based free cooling where thermal density allows, closed-loop liquid cooling for high-density workloads, or a hybrid of the two. This reduces the need for water intake and discharge, helps avoid competition with local communities for water access, and lowers exposure to drought and water-stress risks, consistent with our climate resilience approach described under E1-3 Resilience in relation to climate change.

We continue to report water metrics transparently even when materiality is low, and monitor the topic as our footprint expands. More detail on our cooling architecture is provided in the Efficiency by design chapter.

E3-4

Water metrics

Water withdrawal in data centers, 2025	m³
Total data center water withdrawal	23,909
Finland-1	1,459
Other DCs (incl. colocations)	22,450

How data is compiled

Water withdrawal

Figures represent known water withdrawal reported across facilities. Finland-1 reflects our owned site operations for the full 12 months. Other data covers four of the six remaining sites operational as of December 31, 2025 based on metered readings, utility invoices, and provisional estimates where metered data was not yet available. Water withdrawal covers water sourced from municipal utility connections as a primary source with some rainwater input as a supplementary source. No operated site draws directly from surface water or groundwater as a primary source.

Other categories

The following ESRS water categories are not reported for 2025 due to data limitations: water consumption, water discharge, water recycled and reused, water stored, water withdrawal in water-stressed areas, and water consumption in water-stressed areas. Consumption and discharge data had limited site coverage, while recycled, reused, and stored water data was not collected. Water-stressed area classification is under review. These metrics will be expanded as data collection matures across the asset portfolio.

Insights

Water use across Nebius operations is primarily associated with sanitary and domestic use at office and data center facilities, and with cooling water intake at colocation sites where conventional cooling systems are operated by facility partners outside Nebius's direct operational control. Water withdrawal associated with office locations is excluded from this disclosure on the basis of non-materiality.

At Finland-1, which is the only owned site operational for the full 12 months in 2025, closed-loop liquid cooling minimizes water intake by design. Finland-1's WUE of 0.018 L/kWh (or 0.018 m³/MWh) is approximately 25 times lower than the US data center average of 0.45–0.48 L/kWh projected for 2025 by Lawrence Berkeley National Laboratory, and 2 to 15 times lower than WUE figures publicly reported by leading hyperscalers.

As our operational footprint expands, we are working to improve the granularity and consistency of water-related data collection across our operations.

ESRS E5. Resource use and circular economy

Material impacts, risks and opportunities (IROs)

Resource use and circular economy

Designing for resource efficiency across the hardware lifecycle directly reduces the carbon footprint associated with our infrastructure. By extending component lifespans through modular design, thermal-aware controls, and workload shifting, we defer the embodied emissions cost of hardware replacement. Circular economy principles embedded in rack and server design reduce demand for virgin materials and lower e-waste generation, easing pressure on carbon-intensive manufacturing supply chains. Looking ahead, next-generation high-performance GPU deployments may require more resource-intensive cooling solutions, potentially increasing material footprint, though our hybrid cooling development program is designed to address this risk.

How we manage the IROs

We focus on smart design choices and continuous optimization to make the most of every component and system we use, from inception through end of life. Read more in the Efficiency by Design chapter of this report.

Why it matters

AI infrastructure depends on hardware that is material- and carbon-intensive to produce. As we scale, extending the useful life of hardware components helps reduce replacement needs, embodied emissions, e-waste, and exposure to supply chain constraints.

Resource-efficient design also supports business performance and growth. Durable, modular, and serviceable hardware can lower maintenance complexity, reduce downtime, and preserve value at end of life through material recovery. Smart resource management reduces pressure on our own operations as well as helps ease strain across the broader value chain.

Our approach, policies and targets

E5-1

Policies related to resource use and circular economy

While we have not yet adopted a standalone policy on resource use, our commitment to resource efficiency is reflected in our Code of Conduct and reinforced through technical guidelines that shape our infrastructure design and operational practices.

E5-2

Actions and resources related to resource use and circular economy

Circularity principles are embedded across three phases of our hardware lifecycle: design, use and end-of-life. See the Efficiency by Design chapter for additional detail.

Design phase. We design in-house hardware for durability, modularity, and serviceability. Key components can be replaced without specialized tools, while board-level cable-free connections reduce selected material use, failure points, and signal loss. Where relevant, we engage with chip manufacturers on reference architectures before finalization.

Use phase. Built-in hardware monitoring adjusts thermal conditions in real time, reducing thermal stress and extending component lifespan. Modular design enables individual components to be upgraded or replaced without retiring the full system. Where workload profiles allow, hardware is shifted from training to inference workloads to extract further useful life before decommissioning.

End-of-life phase. Retired components are processed by certified recycling vendors, with recoverable materials extracted and returned to productive use. We are developing a more structured approach to e-waste management as our operational footprint expands.

Resources for actions. As with our climate-related actions, resource efficiency and circularity are funded through core capital and operational decisions rather than a standalone budget. Hardware R&D investment, infrastructure engineering, and procurement criteria are the primary resource allocation mechanisms. Our sustainability team supports implementation across relevant functions, including engineering and infrastructure.

E5-3

Targets related to resource use and circular economy

Our key target areas include hardware lifetime extension and material recovery through responsible recycling. As 2025 is our first full-year data cycle, specific targets are currently under development in the context of continued business scaling. We expect to provide more detail on targets in future reporting periods.

E5-5

Waste, 2025

Total data center waste	184 tonnes
Hazardous waste	2.3%
Diverted from disposal	2.0%
Reuse	1.2%
Recycling	0.6%
Other recovery operations	0.3%
Directed to disposal	0.3%
Incineration	0.0%
Landfill	0.0%
Other disposal operations	0.3%
Non-hazardous waste	97.7%
Diverted from disposal	90.7%
Reuse	0.2%
Recycling	10.7%
Other recovery operations	79.8%
Directed to disposal	7.0%
Incineration	0.0%
Landfill	7.0%
Other disposal operations	0.0%

How data is compiled

Total waste

Reflects total *recorded* waste volumes from data center operations that provided waste generation data for the reporting period. Office and garage sites are excluded due to data non-materiality or lack of access to third-party data (e.g. from office landlords). We are working to improve waste data coverage across operated assets going forward.

Hazardous waste

Accounts for waste defined as hazardous under local regulations, such as waste oil (solid, sludge and liquid) and brake and coolant waste (glycol). Depending on jurisdiction, some categories of e-waste can also be classified as hazardous waste.

Non-hazardous waste

Accounts for waste defined as non-hazardous under local regulations. Includes general and packaging waste (plastics, cardboard), mixed wood from pallets and shipping crates, and various e-waste types including retired electronics, parts and cables. Retired electronics, electronic parts and cables are shredded into small fragments. Recoverable materials are extracted for recycling, while non-recyclable residues are processed in waste-to-energy facilities. Based on partner reporting, the full volume is included under “Recycling.”

Other recovery operations

Under this category, we report waste incineration with energy recovery. Waste incineration without energy recovery, if applicable, is reported under “Incineration”.

Insights

In 2025, recorded waste from data center operations totaled 184 tonnes. Office waste is excluded due to non-materiality. Waste data was not available for all sites: France-1 and Missouri-1 did not report waste in 2025, reflecting data availability constraints at colocation facilities. Reporting coverage is expected to improve as data collection matures.

Non-hazardous waste accounted for 97.7% of total recorded waste and mainly comprised *e-waste* and *general operational waste*, including packaging materials, primarily wood, cardboard, and plastics, from hardware delivery and site commissioning. Where disposal routes were not specified, waste was conservatively classified as landfill, although actual treatment may differ. Overall, 92.7% of recorded waste was diverted from disposal, primarily through energy recovery, including waste-to-energy, anaerobic digestion, and refuse-derived fuel processing.

E-waste from retired servers, racks, and electronic components is the most significant stream by material value. In 2025, recorded e-waste was concentrated at Finland-1, which operated for the full year and had equipment reaching scheduled replacement or maintenance cycles. Most other sites were newly operational or still ramping up, with limited or no hardware replacement activity. 100% of recorded e-waste at Finland-1 was diverted from disposal.

Hazardous waste represented 1.6% of total recorded waste, primarily from vehicle operations. It was managed in accordance with local regulations, with 0% sent to landfill.

Social disclosures

ESRS S1. Own workforce

Material impacts, risks and opportunities (IROs)

Attraction, retention and employee well-being

We create high-skilled jobs in AI infrastructure — a field that is actively shaping how the next generation of computing gets built. The opportunity is to become one of the most attractive employers in the AI space, drawing in strong talent through interesting work, a collaborative culture, and support that keeps people engaged long-term. The risk is that competition for specialized talent is intense, and that fast growth creates pressure on people if support systems do not keep pace.

How we manage the IROs

Through researching and adopting best HR practices, listening to employee feedback and offering flexibility where we can, and investing in well-being support. We track retention carefully and use it to guide hiring and team-building decisions.

Human capital development

We build and deepen expertise in AI infrastructure engineering, inference optimization, agentic AI, and ML research — fields where deep capability is scarce and competitively valuable. The opportunity is to stay at the cutting edge through both organic development and strategic talent additions. The risk is that skills do not keep pace with technology if learning is not actively managed.

How we manage the IROs

Primarily through the work itself: R&D, live infrastructure, and hands-on innovation. Supported by a formal Learning and Development Policy and a dedicated learning budget for every employee.

Diversity, equality and inclusion

Building a workplace where people from different backgrounds, nationalities, and disciplines work well together is both a positive impact and a competitive advantage — diverse teams make better decisions and reach a wider talent pool. The risk is underrepresentation in technical and senior roles, and pay inconsistencies across groups, if not tracked and corrected over time.

How we manage the IROs

Fairness and inclusion principles embedded in hiring, evaluation, and development. Gender representation and pay consistency monitored. Code of Conduct sets the baseline.

Why it matters

Talent is the backbone of our business. Nebius has top AI engineers and deep proprietary expertise across every layer of the stack, and we are quickly becoming a magnet for top talent. Teams with shared context and established trust are hard to build and hard to replace.

As we grow organically and through acquisitions, the opportunity is to keep expanding what we can build while maintaining the culture that defines how we work. That requires deliberate HR work: investing in development, keeping processes fair and transparent, and making sure every employee has what they need to thrive.

Our approach and policies

S1-1

Policies related to own workforce

The policies below cover our principal workforce frameworks; the list is not exhaustive. All policies apply equally to all employees and are reviewed at least annually. Key policies require approval by the Works Council (where applicable), C-level leadership, and/or the Board. Implementation is owned by HR or Legal, depending on the policy, with oversight to ensure consistent application across teams and geographies. Employees can raise concerns or propose changes through their HR Business Partner (HRBP) or, in the Netherlands, through the Works Council.

Code of Conduct

Sets standards for how we work together, grounded in the UN Guiding Principles on Business and Human Rights and the ILO Declaration. Covers human rights (including prohibition of child labor, forced labor,

and trafficking), safe working environments, equal treatment, and a clear stance against discrimination and harassment.

Speak Up Policy

This policy complements our Code of Conduct by providing a safe and confidential channel to report concerns or violations, with protection from retaliation. It is open to employees, contractors, job applicants, former employees and other stakeholders.

Learning and Development Policy

Introduced in 2025, this policy formalizes our approach to employee skills development. It covers technical, professional, soft skills, and managerial programs, and defines how learning needs are identified and supported across the business.

People Managers' Handbook

A practical, living guide for anyone managing people at Nebius. It covers the full people management lifecycle — from setting expectations and giving feedback to navigating team dynamics and development conversations — and gives managers at every level a shared framework for leading fairly and consistently.

Compensation Principles

Our compensation principles set out how we approach pay decisions at Nebius, grounding them in market benchmarking, internal equity, role scope, and individual performance. They are designed to be transparent and consistently applied across the organization, while allowing for role-specific design where the nature of the work requires it.

Equity Incentive Plan

Our Equity Incentive Plan provides eligible employees with RSU grants designed to strengthen alignment between employees and the long-term success of the company. Employees can access clear information on eligibility, vesting conditions, and how grants are affected by changes such as promotions, contract status, or terminations.

S1-2

Engagement with own workforce and workers' representatives

HRBPs embedded across functions serve as the primary point of contact for employees on matters from workload and development to compensation and interpersonal concerns. Regular updates and community discussions are shared via our internal communication platform. Every two weeks, company-wide All-Hands meetings give teams a forum to share progress, with dedicated time reserved for employees to raise questions directly. Onboarding sessions help new employees connect with our culture and understand their rights, responsibilities, and available support.

In 2025, we launched our first company-wide engagement survey, giving employees a structured channel to share feedback on their experience and the direction of the business. Employees can also raise concerns at any time through our Speak Up line, an independent and confidential reporting channel.

In the Netherlands where Nebius Group is headquartered and had the largest employee base by country in 2025, employees are represented through a Works Council in accordance with Dutch law. The Works Council is consulted on, or gives its consent to, key employment-related decisions, including those affecting working conditions, restructuring, or changes to employment terms, and serves as a formal channel for the workforce to bring forward concerns and ideas. Employees who prefer not to raise concerns through HR can also engage an independent Trusted Advisor, who operates under strict rules of confidentiality. Similar employee representation bodies exist in other countries where Nebius operates, in line with local legal requirements and practice.

Topics most commonly engaged include organizational changes, work-life balance, compensation, career development, and upskilling.

S1-3

Actions and resources related to own workforce

Bringing in talent and building teams

546 new employees joined in 2025 — a 12% net increase — with expansion across Europe and the United States. To support high-quality onboarding at this pace, we invested in structured programs and HRBP capacity.

In early 2026 we also launched the Early Talent Program, a structured entry point for recent graduates to contribute to live products and grow alongside experienced teams. As part of efforts to support women navigating careers in AI, we also began building an internal women in tech community and hosted our first external meetup at our Amsterdam office, connecting female employees and the wider community of women in tech to shared perspectives and experiences.

Developing skills and expertise

Skills develop primarily through the work itself: R&D, live infrastructure, and hands-on opportunities to drive innovation. This is supported by the Learning and Development Policy introduced in 2025, which gives every employee access to a dedicated learning budget.

Skill needs are identified through a combination of inputs: people managers monitor their teams' development areas and discuss growth paths with employees and HRBPs, and senior management determines business direction and what capabilities are needed to execute on it. Where needed, bringing in specialist teams through acquisition is an effective route to adding depth, which is the ambition we have pursued into 2026.

Maintaining fair and transparent compensation

Role evaluation and compensation-setting follow structured frameworks, with distinct approaches for corporate, technical, and sales functions, each reflecting the nature and incentive structure of the work. Eligible employees receive performance-based bonuses, and the bonus matrix is transparent and accessible. Salaries are benchmarked per country and role against market rates and peer companies in the AI infrastructure and cloud sector.

Supporting well-being in fast-paced teams

Working at the pace Nebius moves creates both real energy and real demands. All employees globally have access to flexible working hours and arrangements, professional mental health support including burnout prevention counseling, and — introduced in 2026 — birthday leave, a day off within a seven-day window around each employee's birthday.

Beyond these, a number of benefits apply based on working arrangements or location. Remote employees receive a home office allowance, which can

be used flexibly toward coworking spaces or home office needs. Office and hybrid employees in the Netherlands and a number of other locations receive a meal allowance; employees working from our Amsterdam headquarters also have access to free catering in the office. Employees are encouraged to raise workload or stress concerns with their manager or HRBP, who are supported through the People Managers' Handbook on workload planning and team dynamics.

Scaling productivity

We equip employees with tools that free up time for the most meaningful parts of their work. A range of enterprise-grade AI assistants, AI coding tools, meeting transcription tools, and AI-powered knowledge tools are also available across functions. Access is governed through a formal AI Tools Inventory maintained by Legal and Privacy. Employees also have access to Nebius Token Factory internally for tasks including document analysis, code assistance, and knowledge retrieval.

Scaling the HR function

Nebius's growth pace placed significant demands on HR operations in 2025. Revising processes, deploying teams across geographies, and introducing platforms to operate HR at scale was a substantial undertaking that continues into 2026. This includes optimizing processes through simplification and tooling and deploying AI-enabled tools that reduce administrative burden.

S1-4

Targets related to own workforce

Key focus areas include attracting and retaining top talent, building an engaging and inclusive workplace, and strengthening work-life balance support as we operate in a hyper-growth phase. We are not disclosing specific targets at this time. As we build our people management practices to match the scale and ambition of the business, we are developing the frameworks and measurement approaches that will underpin meaningful target-setting. We expect to provide more details in future reporting periods.

S1-7

Collective bargaining coverage and social dialogue

As of of December 31, 2025, no employees were covered by collective bargaining agreements, as no such agreements were requested or required under applicable law. Employees are free to enter such agreements if they choose — our Code of Conduct affirms the right to freedom of association. In the Netherlands, employee interests are represented through a Works Council, in line with local law. The Works Council is consulted on or gives its consent to key employment-related decisions and serves as a formal channel for workforce input.

S1-5

Characteristics of the undertaking's employees

Number of employees		Unit	2024	2025
Total headcount		#	1,379	1,542
By country	The Netherlands	#	501	590
	USA	#	190	353
	Other (incl. remote workers)	#	688	599
By business unit	Nebius	#	741	1,064
	Other businesses	#	638	478
By role type	STEM roles	%	64%	63%
By employment type	Full-time	#	1,371	1,535
	Part-time	#	8	7
By contract type	Permanent	#	1,264	1,297
	Temporary	#	115	245
	Non-guaranteed hours	#	0	0

How data is compiled

Total headcount

Represents employees on contract as of December 31 of the reporting period, including those under notice of redundancy regardless of whether they have been released from duties during their notice period.

Employees by country

We report headcount for the two countries with the highest employee concentration. Other countries with more than 50 employees in 2025, including Germany and Israel, are not shown separately for simplification. We track this data internally and plan to report by individual country in future ESRS-aligned reporting.

STEM roles

Covers employees in technical and tech-adjacent roles including software engineering, ML/AI, cloud infrastructure, R&D, data analysis, and product management, among others.

Insights

Total headcount increased by 12% in 2025. The US team grew by 86%, reflecting growth of the US market, scaling of go-to-market teams to serve US-based clients, and the buildout of new facilities; the US now represents 23% of total headcount, up from 14% in 2024. The Netherlands grew by 18%. The workforce remained predominantly full-time (over 99%) and permanent (84%). STEM roles accounted for 63% of total headcount, broadly consistent with the prior year. 546 employees joined Nebius in 2025, nearly twice the number who joined in 2024.

Turnover fell from 16.2% to 11.7%, reflecting stronger retention across the business. Undesirable turnover was 1.3%, remaining low despite growing competition for top talent.

Equity Incentive Plan coverage reached 65% of employees completing a full year, up from 46% in 2024, when many employees had recently joined and not yet met the full-year threshold. Across all employees, almost half held RSU grants as of end of 2025.

S1-5

Characteristics of the undertaking's employees

Number of employees		Unit	2024	2025
By gender	Male	#	958	1,084
	Female	#	377	395
	Genderqueer / non-binary	#	7	6
	Not reported	#	37	57
Turnover				
	Total turnover	%	16%	12%
	Total undesirable turnover	%	0.8%	1.3%
	Number of employees who left the company	#	210	175
Employees covered by Equity Incentive Plan				
	Employees receiving RSUs	%	46%	65%

How data is compiled

Employees by gender

Based on self-reported data. In addition to male and female, some employees identify as genderqueer / non-binary or choose not to disclose their gender. All identities are reflected in workforce data.

Turnover

Calculated as the number of employees who left during the year divided by average headcount for the period. Rotations between business units within Group is excluded from terminations. Undesirable turnover represents employees who left during the reporting year and whose departure was identified as regrettable based on internal criteria, including performance and business-criticality. Criteria were updated in October 2025 and applied from that point onward, so the calculation reflects the criteria in place at the time of each termination.

Employees receiving RSUs

The percentage of employees who were employed for the full year and, as of December 31, received equity-based incentives (RSUs) under the company's Equity Incentive Plan.

S1-9

Adequate wages

As of December 31, 2025, all employees were remunerated above the applicable statutory minimum wage in each country of operation. Nebius does not employ any workers at or below the minimum wage threshold in any jurisdiction where it operates.

S1-12

Training and skills development metrics

All employees at Nebius participate in regular performance and career development reviews, with 100% coverage in 2025.

Learning and development at Nebius is structured primarily around on-the-job experience. This includes R&D projects, live infrastructure work, and direct collaboration with specialists in the field. Formal trackable training programs complement this and are facilitated through the Learning and Development Policy. Tracking training hours across the full spectrum of learning modalities, including self-directed study, mentoring, and project-based skill acquisition, is not in place in the reporting period. Formal training hours are therefore not disclosed as a per-employee metric at this time. We expect to develop a more comprehensive learning measurement approach in future reporting cycles.

Insights

The majority of employees at Nebius fall within the 30–50 age group, which is the most represented across both specialists and management roles.

Women represent 26% of total headcount, a slight decrease from 27% in 2024, reflecting faster headcount growth among male employees driven by hiring across technical roles. Employees identifying as non-binary or preferring not to disclose their gender represented 4% of total headcount. Women represented 11% of STEM employees at end 2025, down from 2024 due to the deconsolidation of Toloka. Within Nebius's core business the share grew by two percentage points. Within Nebius's core business, the share of women in STEM roles grew by two percentage points in 2025. Representation of women in senior management held steady at 23%. These figures are consistent with available peer reports[1] on the share of women in the total workforce and with industry data on the share of women in tech roles. According to a McKinsey & Company study focused on Europe, women account for an average of 13–17% of roles in software engineering, architecture and infrastructure.

S1-8

Diversity metrics

Employees by seniority level and age group	Unit	2024	2025
Directors and managers	#	377	433
<30	#	44	47
30-50	#	319	356
>50	#	14	29
Specialists	#	1,002	1,109
<30	#	283	262
30-50	#	702	817
>50	#	17	25
Unique nationalities among employees			
Unique nationalities	#	52	72
Women in the governance bodies and workforce			
Share of women on the board of directors	%	38%	25%
Share of women in total headcount	%	27%	26%
Share of women in senior management	%	23%	23%
Share of women in STEM roles	%	14%	11%

How data is compiled

Employees by seniority level and age group

Employee age is determined as of December 31 of the reporting period. Totals include six employees whose date of birth was not recorded in internal systems and who are therefore excluded from the age breakdowns: one director / manager and five specialists.

Unique nationalities

Represents the total number of nationalities reflected in the workforce. Employees with dual or multiple citizenships are counted once under each nationality, meaning a single individual may be represented across multiple national categories

Senior management

Includes Chief Officers and Heads of Department. The share of women among top managers, which excludes Heads of Department and includes executive management roles only, was 22% in 2025.

S1-14

Work-life balance metrics

In 2025, family-related leave was tracked across multiple HR systems. Aligning leave categories used in ESRS with local market classifications — in particular in the United States, where statutory leave frameworks differ significantly from ESRS definitions — requires additional reconciliation work. This disclosure is therefore omitted in this reporting cycle on the basis of undue cost or effort, in accordance with ESRS 1.

S1-16

Incidents, complaints and severe human rights impacts

Nebius operates an independent Speak Up channel through a confidential platform. Cases, investigation results, and remedial actions are recorded and tracked within the platform. In 2025, 13 cases were created through the channel. Of these, 6 were substantiated. All cases were reviewed and addressed appropriately: seven were resolved with no further action required, and remediation measures are being implemented for six cases.

No severe human rights incidents were identified in our workforce during 2025. No fines, penalties, or compensation were paid in relation to any of the reports.

(1) For example, NVIDIA reported that women represented 27% of its workforce. Companies in our peer group had either not updated or did not disclose diversity data as of the publication date of this report.

ESRS S3. Affected communities

Borderline materiality in 2025

Our data center expansion creates opportunities for local economies through job creation, skills development, and local procurement. Our buildout also interacts with local energy systems, land use, and infrastructure, which we actively seek to manage responsibly.

Community impact was assessed as borderline material for 2025, reflecting the operational footprint for that year. As Nebius scales into new regions and data center capacity grows, the materiality of this topic is expected to increase, and we are building the relevant management practices ahead of that. We talk more about that in the Unlocking opportunities chapter of this report.

From a financial perspective, constructive community engagement supports predictable permitting and site development timelines, contributing to on-schedule delivery of new capacity. As Nebius expands into new regions, proactive engagement is expected to reduce permitting risk and support continuity of the buildout program.

CSD. Products and services with social benefits

Material impacts, risks and opportunities (IROs)

Products and services with social benefits

Nebius's AI infrastructure enables developers, startups, enterprises, and research institutions to build and deploy AI applications at scale. The positive impact extends beyond our own operations: by making frontier compute and inference capabilities broadly accessible, we help more teams bring AI-driven solutions to sectors where they create real value, including healthcare, scientific research, and education. Our embedded workload optimization tools also help customers reduce their own energy use.

How we manage the IROs

Developing AI infrastructure verticals tailored to specific industries, partnering with research institutions, running grant programs for innovators, and tracking outcomes.

Why it matters

Our infrastructure exists to power ideas, from training advanced AI models to running production workloads at scale. By providing infrastructure that serious technical teams outside big tech can actually use and afford, Nebius expands who can build AI applications and where those applications are deployed. That has a direct bearing on the range and diversity of AI-driven solutions that reach the world.

Our approach and policies

As of the date of this report, Nebius does not have formal policies specific to this topic. Work in this area is governed by principles followed by the relevant teams. Grant-giving activities are subject to internal procedures. Actions and target areas related to products and services with social benefits are discussed in the Unlocking opportunities chapter of this report.

Governance disclosures

ESRS G1. Business conduct

Material impacts, risks and opportunities (IROs)

Business ethics

Positive impacts relate to our opportunity to build leading ethical practices across our AI cloud operations, strengthening client trust and our long-term reputation. At the same time, we recognize the risk that insufficient oversight of business conduct, human rights, or supply chain risks could lead to reputational or legal exposure.

How we manage the IROs

Through our ethics and compliance framework: policies, training, third-party due diligence and a Speak Up channel. This section covers all of it.

Why it matters

Nebius is scaling fast, entering new markets, and serving an increasingly complex customer base. Choosing an AI infrastructure provider is a long-term decision: customers need to trust that sensitive workloads, data, and critical systems are handled with discipline. That trust depends on consistent accountability and real transparency. At the same time, AI development is facing growing public and regulatory scrutiny. Companies that set high standards now will be better positioned to operate, compete, and earn trust as expectations continue to rise.

Our approach and policies

G1-1

Policies related to business conduct

Our Ethics and Compliance (E&C) function sets the framework for ethical business conduct and ensures its consistent application across all Nebius businesses. Led by the Group Head of Ethics and Compliance, the function reports into the Group General Counsel and the Board. It is responsible for designing, deploying and improving the Group E&C framework, overseeing implementation across business units, identifying emerging risks, and managing the end-to-end Speak Up process.

Our business conduct policies are listed under ESRS 2 GOV-3 Statement on due diligence. Additional internal procedures and guidelines supporting the implementation of those policies are also in place. Our Anti-Bribery and Anti-Corruption Policy is consistent with the United Nations Convention against Corruption. Our Speak Up Policy meets the requirements of the EU Whistleblower Directive (2019/1937/EU). Functions considered most at risk of corruption or bribery, including Sales, Procurement and roles with exposure to public authorities or high-risk geographies, are subject to role-based training described under G1-2.

G1-2

Actions related to business conduct

2025 priorities and progress.

In 2025, E&C shifted from framework launch to structured execution.

- Governance and policy suite expanded. The Board approved an updated Anti-Bribery and Anti-Corruption Policy and Speak Up Policy, a new Conflict of Interest Policy, and a new Trade Controls Policy replacing the previous Sanctions Compliance Policy. A Gifts and Entertainment Guideline was also introduced, establishing monetary thresholds, approval requirements and monitoring mechanisms.
- Third-party risk management formalized. A dedicated TPRM Policy and Procedure, adopted in November 2025, established a risk-based methodology covering sanctions and trade control exposure, anti-bribery and corruption risk, human rights and labor-related risks, sustainability-related considerations, and geopolitical risk. All third parties are now classified by risk level, with enhanced due diligence applied where higher risk is identified.
- Training moved to an annual structured plan. Mandatory Code of Conduct and Speak Up training was rolled out to cover all employees, with role-based sessions on anti-bribery and corruption, trade controls and third-party risk management delivered to higher-risk functions including Sales, Legal, Procurement and Internal Audit.
- Compliance control framework in progress. In 2025, key compliance risks were mapped and foundational controls introduced across anti-bribery and corruption, trade controls, third-party risk management, conflicts of interest and Speak Up. Building out the full monitoring capability is ongoing.

Management of supplier relationships

Nebius works with a broad supplier base, from global hardware manufacturers to local energy providers and service partners. As we scale, we continue to value long-term relationships built on transparency, mutual respect, and clear standards.

Our procurement approach is designed to manage risk, support operational speed, and avoid unnecessary complexity. We aim to keep processes fair, consistent, and practical, giving qualified partners a clear path to work with us regardless of size.

Sustainability-related factors are integrated into supplier selection. Under the TPRM framework introduced in 2025, all prospective suppliers are assessed against criteria that include human rights and labor-related risks and environmental and sustainability considerations, alongside anti-corruption, sanctions and geopolitical risk. Procurement teams receive dedicated training on these assessment criteria and procedures. For suppliers rated medium or high risk, E&C recommends mitigation measures that must be implemented before the engagement proceeds and may, in some cases, block it altogether.

Our Supplier Code of Conduct sets clear expectations on labor standards, human rights, environmental responsibility and legal compliance. Due diligence questionnaires are completed for all new engagements.

Prevention and detection of corruption and bribery

We maintain a clear, zero-tolerance position on corruption and bribery, backed by policies, controls and a culture where concerns can be raised safely and acted on.

Setting clear expectations

The Code of Conduct and Anti-Bribery and Anti-Corruption Policy define what is and is not acceptable, for employees and external partners alike. The Supplier Code of Conduct carries the same expectations into our supply chain. These documents are accessible on the intranet, covered in onboarding, and reinforced through training.

Training and awareness

In 2025, 75% of the total workforce completed mandatory Code of Conduct and Speak Up training. The completion rate reflects two factors: some employees due to refresh their annual training had not yet completed it by year-end, and some employees who joined in the second half of the year had not completed training before the reporting period closed. Training extends to management and covers anti-bribery and corruption. Higher-risk functions, including Sales, Legal, Procurement, and Internal Audit, received additional role-based sessions on anti-bribery and corruption, trade controls, and third-party risk management.

Proactive detection measures

Our Speak Up channel is open to employees and external stakeholders, independently managed and compliant with the EU Whistleblower Directive.

Third-party due diligence screens for bribery and corruption risk before any supplier, contractor or agent relationship is established. Gifts and Entertainment declarations are mandatory across the Group, with thresholds and a monitoring register introduced in 2025.

Conflict of Interest declarations are required annually from all employees, with the process now integrated into hiring. Internal Audit reviews controls independently, identifying gaps before they become issues.

Handling reports and acting on concerns

Reports submitted through the Speak Up channel are reviewed by E&C. Where an investigation is warranted, the Ethics Committee leads the process, drawing on Legal, HR, and Internal Audit as needed and working under strict confidentiality. External specialists may be engaged where appropriate. All cases are logged in a secure system, and the Board and senior leadership receive regular E&C updates.

In 2025, one confirmed incident of a violation of the corporate anti-corruption policy was recorded in the context of vendor management. The employee involved was terminated and the relevant vendor relationships were ended.

G1-3

Targets related to business conduct

Our E&C program is focused on building a compliance culture that can scale with the business. This includes effective controls and training, the ability to detect and respond to fast-evolving risks, readiness for upcoming regulation, and practical application of legal requirements to strengthen governance and streamline processes.

We are not disclosing specific targets at this time. As the compliance function matures, we are developing the KPI framework and monitoring mechanisms needed to support meaningful target-setting and expect to provide more detail in future reporting periods.

G1-4

Metrics related to corruption or bribery

	2024	2025
Total number of convictions and sanctions for violation of anti-corruption and anti-bribery laws	0	0
Amount of fines for violation of anti-corruption and anti-bribery laws	0	0

How data is compiled

Definitions for convictions, violation and fines are in line with ESRS. Data is sourced from the Legal function.

CSD. Data protection, privacy and cybersecurity

Material impacts, risks and opportunities (IROs)

Data protection, privacy and cybersecurity

Customers trust Nebius with sensitive workloads, data, and critical systems. A serious security incident or data breach could cause direct harm, disrupt customer operations, and undermine confidence in our services. Strong security is also a competitive differentiator: customers choosing an AI infrastructure provider treat security, privacy, and resilience as core selection criteria.

How we manage the IROs

Through ongoing investment in security governance, employee training and incident response. Read more in the Reliability and security chapter of this report.

Why it matters

Customers use our platform to process sensitive data and run business-critical workloads, including applications where disruption can have real consequences. Security, privacy, and resilience are therefore core requirements for operating in this market.

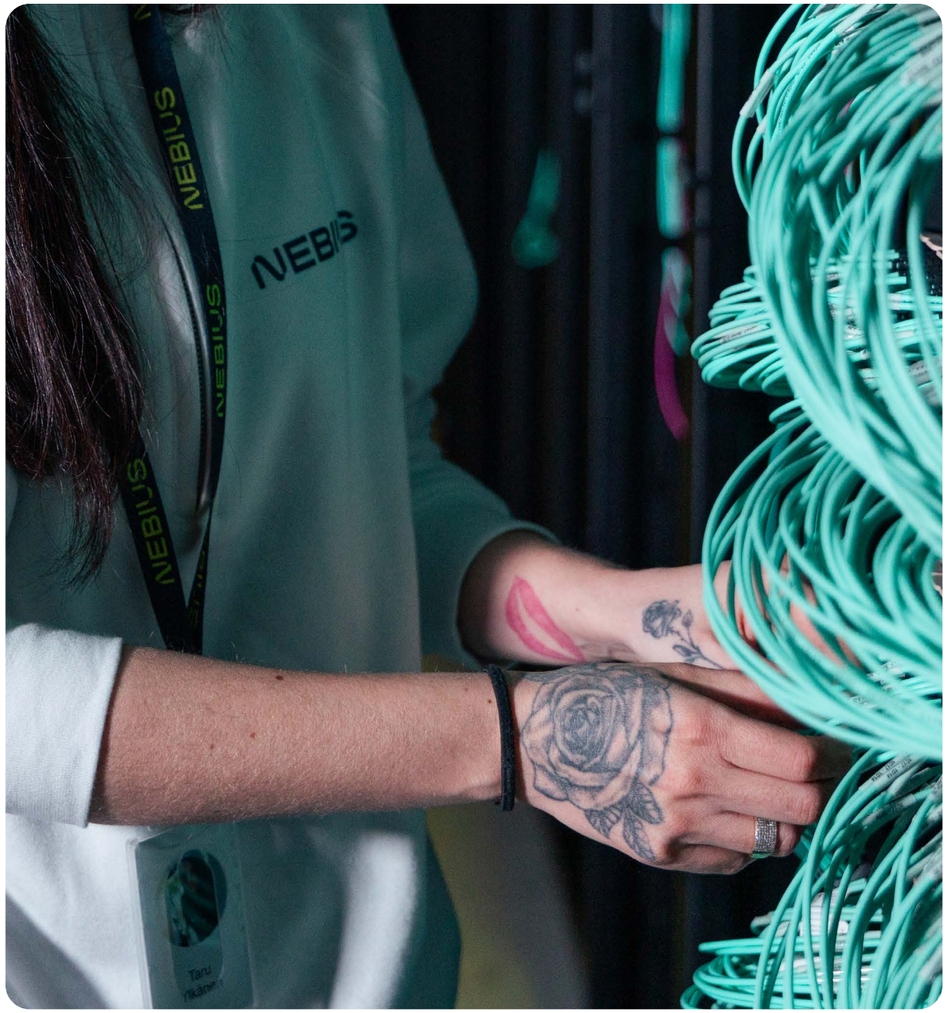
As we expand into new geographies and serves a more complex customer base, our obligations also grow. We operate across jurisdictions with evolving requirements on data privacy, data protection and information security for AI workloads. Customers are also assessing AI infrastructure providers on their ability to protect data, manage incidents, and maintain reliable operations.

Our approach and policies

Cybersecurity is integrated into our overall risk management processes. We use a combination of internal security monitoring, employee training, technical controls, incident response processes, and external assessments to identify, assess, and manage cybersecurity risks.

Security governance covers physical, digital, and cyber security, as well as compliance and risk management. Incident response includes detection and analysis, containment, eradication and recovery, post-incident review, and assessment of operational, business, and privacy impacts. Cybersecurity matters are reviewed by senior management, with Board-level oversight through the Audit Committee.

Details on security governance, certifications, incident management, and key 2025 developments are provided in the Reliability and security chapter of this report.



Disclosures referencing ESRS. ESRS Index

ESRS Index

Disclosure requirement	Description	Status	Reference
ESRS 2 General disclosures	BP-1 – Basis for preparation of the sustainability statement	D	p. 38
	BP-2 – Specific information if the undertaking uses phasing-in options	D	p. 38
	GOV-1 – The role of the administrative, management and supervisory bodies in relation to sustainability	D	p. 38–39
	GOV-2 – Integration of sustainability-related performance in incentive schemes	D	p. 39
	GOV-3 – Statement on due diligence	D	p. 39–40
	GOV-4 – Risk management and internal controls over sustainability reporting	D	p. 41
	SBM-1 – Strategy, business model and value chain	D	p. 41
	SBM-2 – Interests and views of stakeholders	D	p. 42
	SBM-3 – Interaction of material impacts risks and opportunities with strategy and business model, and financial effects	D	p. 44–46
	IRO-1 – Description of the process to identify and assess material impacts, risks and opportunities and material information to be reported	D	p. 43
	IRO-2 – Material impacts, risks and opportunities and disclosure requirements included in the sustainability statement	D	p. 46

Status

D — Disclosed
PD — Partially disclosed
NR — Not relevant
NS — Not stated
NM — Not material

Disclosure requirement	Description	Status	Reference
ESRS E1 Climate change	E1-1 – Transition plan for climate change mitigation	D	p. 47
	E1-2 – Identification of climate-related risks and scenario analysis	D	p. 48–49
	E1-3 – Resilience in relation to climate change	D	p. 49
	E1-4 – Policies related to climate change mitigation and adaptation	D	p. 50
	E1-5 – Actions and resources in relation to climate change mitigation and adaptation	D	p. 50
	E1-6 – Targets related to climate change	D	p. 50
	E1-7 – Energy consumption and mix	D	p. 51–52
	E1-8 – Gross scope 1, 2, 3 GHG emissions	D	p. 53–54
	E1-9 – GHG removals and GHG mitigation projects financed through carbon credits	NR	—
	E1-10 – Internal carbon pricing	NR	—
	E1-11 – Anticipated financial effects from material physical and transition risks and potential climate-related opportunities	NS	—
ESRS E3 Water (borderline material topic in 2025)	E3-1 – Policies related to water	NS	—
	E3-2 – Actions and resources related to water	NS	—
	E3-3 – Targets related to water	NS	—
	E3-4 – Water metrics	PD	p. 55

Status

D — Disclosed
PD — Partially disclosed
NR — Not relevant
NS — Not stated
NM — Not material

Disclosure requirement	Description	Status	Reference
ESRS E5 Resource use and circular economy	E5-1 – Policies related to resource use and circular economy	D	p. 56
	E5-2 – Actions and resources related to resource use and circular economy	D	p. 56
	E5-3 – Targets related to resource use and circular economy	D	p. 56
	E5-4 – Resource inflows	NM	—
	E5-5 – Resource outflows	D	p. 57
ESRS S1 Own workforce	S1-1 – Policies related to own workforce	D	p. 58–59
	S1-2 – Engagement with own workforce and workers' representatives, existence of channels for own workforce to raise concerns or needs and approaches to remedy	D	p. 59
	S1-3 – Actions and resources related to own workforce	D	p. 59–60
	S1-4 – Targets related to own workforce	D	p. 60
	S1-5 – Characteristics of the undertaking's employees	D	p. 60–61
	S1-6 – Characteristics of non-employees in the undertaking's own workforce	NM	—
	S1-7 – Collective bargaining coverage and social dialogue	D	p. 60
	S1-8 – Diversity metrics	D	p. 62
	S1-9 – Adequate wages	D	p. 61
	S1-10 – Social protection	NM	—
	S1-11 – Persons with disabilities	NM	—
	S1-12 – Training and skills development metrics	D	p. 61
	S1-13 – Health and safety metrics	NM	—
	S1-14 – Work-life balance metrics	D	p. 62
	S1-15 – Remuneration metrics	NS	—
	S1-16 – Incidents of discrimination and other human rights incidents	D	p. 62

Status

D — Disclosed
PD — Partially disclosed
NR — Not relevant
NS — Not stated
NM — Not material

Disclosure requirement	Description	Status	Reference
ESRS S3 Affected communities (borderline material topic in 2025)	S3-1 – Policies related to affected communities	NS	—
	S3-2 – Engagement with affected communities, existence of channels for affected communities to raise concerns or needs and approaches to remedy	PD	p. 24
	S3-3 – Actions and resources related to affected communities	PD	p. 24
	S3-4 – Targets related to affected communities	NS	—
ESRS S1 Business conduct	G1-1 – Policies related to business conduct	D	p. 64
	G1-2 – Actions related to business conduct	D	p. 64–65
	G1-3 – Targets related to business conduct	D	p. 65
	G1-4 – Metrics related to corruption or bribery	D	p. 65
	G1-5 – Metrics related to political influence, including lobbying activities	NS	—
	G1-6 – Metrics related to payment practices	NM	—

Status

D — Disclosed
PD — Partially disclosed
NR — Not relevant
NS — Not stated
NM — Not material

nebius.com/sustainability

Forward-looking statements. This report contains forward-looking statements within the meaning of the Private Securities Litigation Reform Act of 1995, which involve risks and uncertainties. All statements other than those of historical fact — including those about future performance, business strategy, and expected growth — are forward-looking in nature. These statements are based on information available as of the date of this report and involve assumptions that may change or prove inaccurate. Actual results may differ materially from those expressed or implied in forward-looking statements.