

NEBIUS

Nebius for healthcare

We provide purpose-built AI infrastructure for healthcare innovators to build, validate and deploy clinical AI with full data sovereignty and built-in compliance.

AICPA
SOC

SOC 2
Type II

AICPA
SOC

SOC 3



HIPAA



ISO/IEC
27001



ISO/IEC
27799



ISO
22301



ISO/IEC
27701



ISO/IEC
27018



ISO/IEC
27032



NIS 2



DORA

AI-enabled healthcare

Boosted by the rapid growth of medical data volume and diversity, AI is amplifying the scale and reach of care delivery.

From medical imaging models to digital health solutions and smart examination devices, AI expands early diagnostics, enables personalized therapies and broadens access to cutting-edge treatment. Ambient agents are saving hours in administrative work so clinicians can focus more on caregiving, while advancements in robotics are enhancing surgical precision to improve patient outcomes.

Scaling AI care innovation

While emerging solutions grow more capable and widely deployed, they also become more challenging to develop and scale. From building foundation models and agentic workflows to delivering responsive inference at scale, AI innovators need to experiment with evolving models, privacy-sensitive datasets and reproducible pipelines while ensuring patient safety.

As AI infrastructure needs range from single-node experimentation setups to thousand-GPU production clusters, healthcare AI's demanding workflows require high-performing infrastructure purpose-built as secure, compliant environments ready to match the speed of iteration. Direct access to scalable AI clusters, expert MLOps support and integrated AI tooling have become essential for innovators to focus on patient outcomes instead of managing systems.

Use cases

Large-scale, multimodal model training

Real-time clinical AI inference

Clinical AI safety evaluation at scale

Large-scale inference for DICOM data

Medical image reconstruction

Ambient patient and provider agents

Physical AI and robotics

Inside the AI transformation in healthcare

As AI reshapes clinical realities, solution providers face not only technical hurdles but also implementation challenges around fragmented data, integration complexities and regulatory uncertainty.

To help address systemic barriers, experts from NVIDIA, AstraZeneca, Nebius and healthcare AI builders discussed adoption success factors and ecosystem collaboration strategies that center care teams' most pressing challenges.

Explore their views. Click to get the [whitepaper](#):

AI's potential to reduce Potentially Avoidable Hospitalizations (PAH): Opportunities for solutions to lower adoption thresholds and amplify evidence-based care



How Nebius can help

Faster time to compliance-ready workloads. Skip months of configuration hurdles and launch production-grade GPU clusters in days with a Business Associate Agreement (BAA) in place and HIPAA-compliant object storage.

Elastic capacity to match development cycles. Scale up your clusters for large training jobs and bursty evaluation runs with pay-as-you-go pricing and save with committed usage discounts for predictable demands.

Data never leaves your region. Maintain full control over your workloads. With data centers and hardware in the EU, UK, US and Israel, Nebius offers region-specific clusters and storage.

High-performing inference at scale. Reliable, low-latency inference in secure environments. Serve your models privately on Nebius infrastructure with single-tenant endpoints, centralized access controls, EU- or HIPAA-compliant data residency and no data leaving your region. Purpose-built for healthcare AI teams that cannot use shared public inference endpoints.

Predictable, reproducible ML pipelines. Stay on top of model versions, metrics and metadata throughout the ML lifecycle with fully managed industry-leading experiment tracking tools.

Frictionless cluster setup. Reduce Kubernetes and Slurm overhead. Your engineers can launch jobs instantly after clusters are provisioned with managed orchestration and built-in observability.

On-demand AI expertise. Count on our solution architects for hands-on support 24/7, from integrating Nebius with your existing infrastructure to helping your team scale from proof of concept to production-grade systems.

Accelerated by NVIDIA. Build and optimize agentic systems with NVIDIA NeMo and future-proof your AI stack with expanding NVIDIA integrations in Nebius.

Proven across diverse research workflows



Model training for medical imaging

Train foundation models on noisy, high resolution 3D datasets on self-healing infrastructure purpose-built to sustain long, uninterrupted runs.



Guardrail-protected clinical inference

Deploy custom fine-tuning strategies to scale models up to hundreds of billions of parameters for multi-step, validated reasoning within patient-facing response times.



End-to-end AI journeys

Optimize every step of the model development lifecycle with AI-native infrastructure. Keep costs under control and simplify workflows with managed SkyPilot.

Growing ecosystem of tools and services

Model frameworks



NVIDIA BioNeMo



Biocypher

AI orchestration platforms

Outerbounds  dstack  SkyPilot

Experiment tracking and MLOps



 Weights & Biases  DVC

Data and storage



Trusted by healthcare innovators

Digital health



"The hardest part of healthcare AI is building a complete system that can support real interactions every day, from everywhere. Designing Dawn requires access to specialized infrastructure, deep health-specific governance and a lot of specialized data, but also system stability, reliability and consistency.

What I love most about working with Nebius is that it's been super stable. The support teams are quite good, so usually it's just a couple of hours for things to be solved."



Ricardo Rei
Head of AI Research
Sword Health

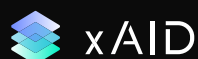


"With Nebius, we've scaled large-model training in mental health — using cutting-edge infrastructure to build something entirely new: AI that guides your therapeutic journey. Their performance, flexibility and reliability have been key to accelerating our progress."



Alexey Bukhtiyarov
ML Infrastructure Lead
Slingshot AI

Medical imaging



"As the founder of xAID, I can confidently say that partnering with Nebius has been transformative. Their intuitive cloud platform took care of the infrastructure, allowing us to stay laser-focused on building our AI foundation model for CT chest and abdomen.

With DevOps off our plate, we accelerated development, shipped faster and stayed aligned with our mission to improve patient outcomes. Their team has been consistently fast, transparent and a true partner in our journey."



Kirill Lopatin
CTO and Co-Founder
xAID

Get started with a free
proof of concept

Contact us