NEBIUS

Nebius for life science

We provide life science teams with simple and elastic access to AI compute, empowering them to focus on resolving fundamental life secrets.

AICPA SOC — SOC 2 Type II
AICPA SOC — SOC 3
HIPAA
ISO — ISO/IEC 27001
ISO — ISO/IEC 27799
ISO — ISO 22301
ISO — ISO/IEC 27701
ISO — ISO/IEC 27018
ISO — ISO/IEC 27032
NIS 2
DORA

AI infrastructure for life science

AI has transformed bioinformatics and life science research, making powerful models accessible to diverse teams — from large pharmaceutical R&D departments to small startups and independent labs. AI-driven approaches in bioinformatics have shown promising results, while becoming increasingly accessible across the industry. These factors are intensifying competition and forcing organizations to deliver breakthrough results faster in order to maintain their competitive edge.

The challenges

Biomedical discoveries involve complex experimental workflows with highly unpredictable success rates, requiring researchers to run substantially more computational experiments to achieve life-changing breakthroughs. As a result, AI researchers must iterate across large datasets, complex models and extensive simulation pipelines. AI infrastructure requirements vary dramatically across teams and research phases — from single-node installations to large scale distributed cluster with thousand-GPUs — making it difficult to predict, provision and manage the AI infrastructure required to support their work.

Most research teams lack dedicated DevOps/MLOps expertise, creating significant frustration around infrastructure scaling, management and troubleshooting. This creates a critical need: life science researchers require straightforward access to scalable AI infrastructure and integrated AI tooling paired with expert technical support, allowing them to focus on scientific discovery rather than system administration.

Use cases

- Large-scale model pre-training and fine-tuning
- Foundation model training
- Genomics, multi-omics and genome editing design
- Next-Generation Sequencing (NGS) and variant calling
- Population-scale studies
- Protein engineering at scale
- Generating novel molecules
- Single-cell and spatial transcriptomics
- Biomarker identification
- Specialized natural language models
- Cryo-EM reconstruction
- Extreme-scale virtual screening
- Graph neural networks for molecular and cellular graphs
- Extreme-scale molecular dynamics simulations
- Quantum mechanics/molecular mechanics (QM/MM) simulations
- Free-energy perturbation (FEP) and binding-affinity prediction

Stanford

"Nebius enabled rapid model screening and fine-tuning, which was very important for the CRISPR-GPT project."

Professor Le Cong
Head of Cong Lab
Stanford University

How Nebius can help

Faster time to market. Start training your models faster and deliver AI results ahead of the competition. With Nebius, you can access large-scale, production-ready GPU clusters in days, not weeks.

Elasticity for unpredictable research pipelines. Launch and scale your experiments and batch inference workloads with pay-as-you-go pricing. Quickly scale up your cluster to handle burst workloads when needed.

Reproducible workflows, quick to deploy. Scale functional pipelines with fully supported industry-leading orchestration tools. Turn prototypes into production-ready pipelines in days, not months.

Fully managed cluster environment. Forget about Kubernetes and Slurm configuration. We provision AI clusters with fully managed orchestration and granular observability, available out-of-the-box — your researchers can schedule jobs immediately after provisioning.

Top-performing models, easy to find. Visualize all your experiments in one place: compare metrics, track experiment metadata and model versions, and select the best ones for production with managed MLflow.

Smooth transition, seamless integration. Rely on our solution architects and engineers to guide you through the migration process. We'll help you connect Nebius AI Cloud to your existing infrastructure, minimize disruption to current workflows.

Dedicated onboarding for every customer. From AI builders to enterprises, your team can count on our 24/7 MLOps support to move from proof-of-concept to production-grade deployments.

Co-engineered with NVIDIA. Run evolving life science AI workflows from a single platform. Accelerate drug discovery with NVIDIA BioNeMo frameworks in one click, and boost any life science workload with full support for custom model training and fine-tuning.

Growing ecosystem of tools and services

Bioinformatics frameworks

NVIDIA BioNeMo and NVIDIA Parabricks

Biocypher

AI orchestration platforms

seqera · SkyPilot · dstack · Outerbounds · Flyte

Experiment tracking and MLOps

mlflow · ClearML · Weights & Biases · DVC

Third-party storage options

WEKA · snowflake

Proven across diverse research workflows

Population-scale multiomics foundation training

Train modality-agnostic foundation models from large cross-cohort data-sets. Scale seamlessly from a few nodes to thousand-GPU clusters, and manage costs with autoscaling and preemptible capacity.

Extreme-scale molecular dynamics

Run microsecond-scale simulations for systems exceeding 100 million atoms with a single command, supported by resilient execution and checkpointing.

Ultra-high throughput docking

Screen the entire ZINC database of nearly 10 billion small molecules in just five days by using a few lines of code, featuring automated retry and error recovery for unin-terrupted runs.

Trusted by industry experts

BCR
BASECAMP RESEARCH

"Basecamp Research trained the world's largest genomic foundation model (by token count) on Nebius's GPUs — 100% of the data is fully proprietary, sourced from five continents. The model will be used to tackle unsolved frontier challenges in the Life Science industry to design programmable medicines."

Phil Lorenz
CTO
Basecamp Research

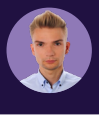
PRIMA MENTE

"Nebius' GPU Cloud and AI-centric infrastructure give Prima Mente the elastic, cost-efficient HPC we need to train and fine-tune our 90M to 7B-parameter epigenetic foundation models at scale. Their life-sciences-ready platform lets us spin up secure clusters in minutes, compressing model-development cycles and accelerating delivery of next-generation brain-health diagnostics to researchers and clinicians worldwide."

Ravi Solanki
CEO and Co-Founder
Prima Mente

Ardigen

"Nebius AI Cloud provided the scalable compute we needed to run high-resolution DNA sequence embedding and regulatory variant scoring at genome-wide scale. It allowed us to focus on extracting meaningful biological insights from non-coding regions without being constrained by compute resources or workflow limitations."



Krzysztof Kotlarz
Bioinformatician
Ardigen

compugen
FROM CODE TO CURE®

"Nebius gave us powerful GPUs, a smooth interface and competitive pricing to train large-scale models on sensitive biological data. We developed models that reveal spatial immune features in tumors — capabilities we haven't seen elsewhere. It's a great match for AI in biotech".

Dr. Roy Granit
Sr. Director & Head of Computational Discovery
Compugen

lynxanalytics

"Nebius gives us exactly what we need — powerful GPUs, fast provisioning and simple scaling without DevOps overhead. Paired with LynxKite: 2000MM's zero-code AI orchestration workflows, we help clients such as Hummingbird Bioscience go from one up to one thousand GPUs as needed, keeping utilization high and delivery fast."

Gyorgy Lajtai
CEO
Lynx Analytics

Nanyang Biologics

"With Nebius, we reduced our screening pipeline from days to hours. The scale and reliability of their infrastructure with Ray allowed us to push the limits of molecular docking and deliver faster results to our partners in pharma and consumer health."

Duy Trieu
CTO
Nanyang Biologics

Get started with a free proof of concept