



Mistral Global Hackathon – NVIDIA briefing

Ziv Ilan, Rachel Oberman

NVIDIA in the hackathon

- Best team to develop locally with under \$100 compute – **London, Paris, NY**
 - Powered by NVIDIA Brev platform
 - Best team will win an RTX 5090
- NVIDIA build.nvidia.com – **All Locations**
 - End-points of Mistral Large 3 Instruct, Ministral 14B Instruct, and Devstral-2-123B
 - Blueprints - Inspiration for different use cases
 - Documentation and additional resources
- Mentors on-site – **London, Paris, NY, SF**
 - We will have NVIDIA Solution Architects available throughout the day-time ☺
 - Available for questions and brainstorming covering accelerated computing, agentic AI, and a variety of technical topics (**not only HW -> full stack**)

City	PIC
London	Zenodia Charpy
Paris	Dora Csillag
NY	Rachel Oberman
SF	Aastha Jhunjunwala

NVIDIA Nemotron Models Available Today

Reasoning



Nemotron 3 Nano

Highest efficiency SLM

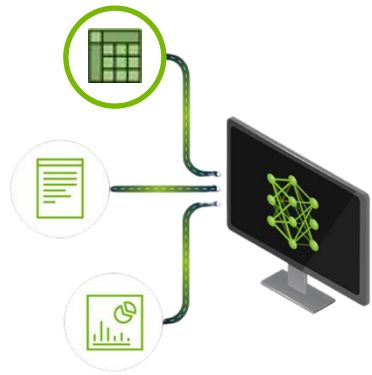
Llama Nemotron Super

Optimal accuracy-efficiency LLM

Llama Nemotron Ultra

Highest accuracy powered by world knowledge

Vision



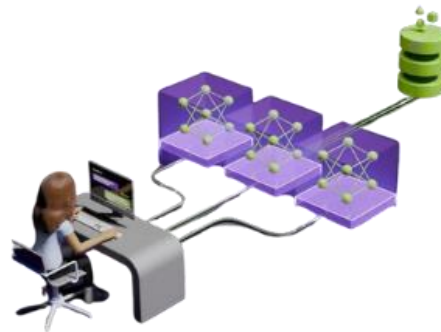
Nemotron Nano 2 VL

Video understanding and Image reasoning

Nemotron Parse 1.1

Document extraction

Information Retrieval



Nemotron Page Elements

Structured images detection & classification

Nemotron Table Structure

Rows, columns detection & preservation in markdown

Nemotron Graphic Elements

Chart component detection & classification

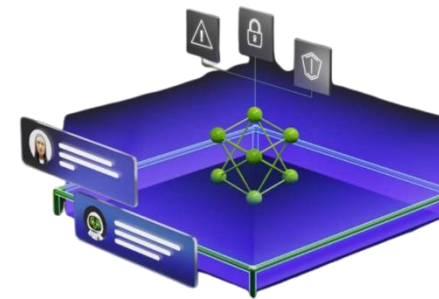
Llama Nemotron Embed

Multilingual text question-answering retrieval

Llama Nemotron Rerank

Multilingual fine-tuned reranking model

Content Safety



Llama Nemotron Safety Guard

Multilingual content safety

Llama Nemotron Content Safety

Multilingual content safety

Nemotron Content Safety Reasoning

Domain-specific safety policies

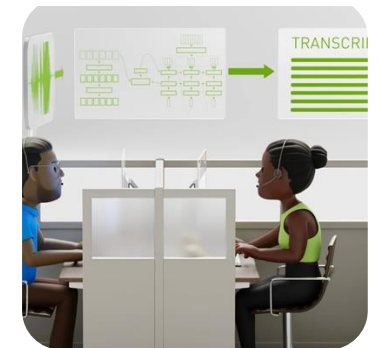
Llama Nemotron Topic Control

Prevent off-topic / unsafe deviations

Nemotron Jailbreak Detection

Detect adversarial or jailbreak prompts to block unsafe behavior

Speech



Nemotron ASR

Real-time English speech recognition

Magpie TTS Multilingual

Research-ready, text-to-speech for natural, fine-tunable voices across 7 languages

Nemotron Data

Open, transparent training data for every stage of AI development

Foundations



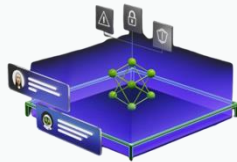
Nemotron-Pretraining

Code, STEM, Safety, Chat,
Multilingual

Nemotron-ClimbLab

Optimized pre-training mix

Alignment



Nemotron-Post-Training

Code, STEM, chat, reasoning

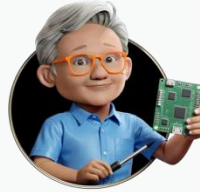
HelpSteer

Alignment: helpful, factual,
coherent

Nemotron-VLM2

OCR training

Agentic AI



Nemotron-MIND

Structured mathematical
dialogues

Safety



Nemotron-Content-Safety 2

Human-LLM interactions

Nemotron-AIQ-Agentic-Safety

Contextual safety risks

Nemotron-Content-Safety 4

Multilingual safety training

Nemotron-Content-Safety-Audio

Multimodal content safety

Sovereign AI



Nemotron-Personas

Grounded national demographics

Nemotron-PII

Sensitive data mitigation

Leverage Open Techniques to Build Custom, Efficient Models

Nemotron Advanced Using NVIDIA Techniques

Architecture



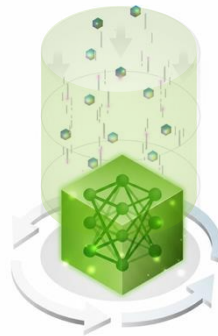
Hybrid Mamba-Transformer

Efficient arch for best accuracy, throughput

Mixture of Experts

Improve compute and latency

Accuracy



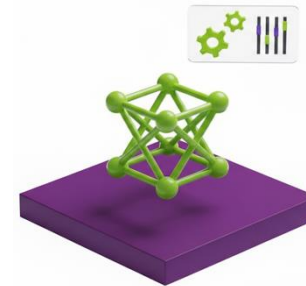
NeMo RL

Align models with human preferences

NeMo

Data curation accelerates training and improves accuracy

Efficiency



Minitron, Neural Arch Search
Right-size models for higher efficiency

Thinking Budget

Minimize model overthinking

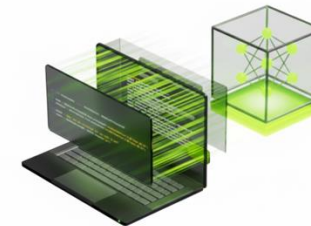
Efficient Video Sampling

Process longer videos

Quantization

Lowers inference cost while maintaining accuracy

Recipes



Technical Reports

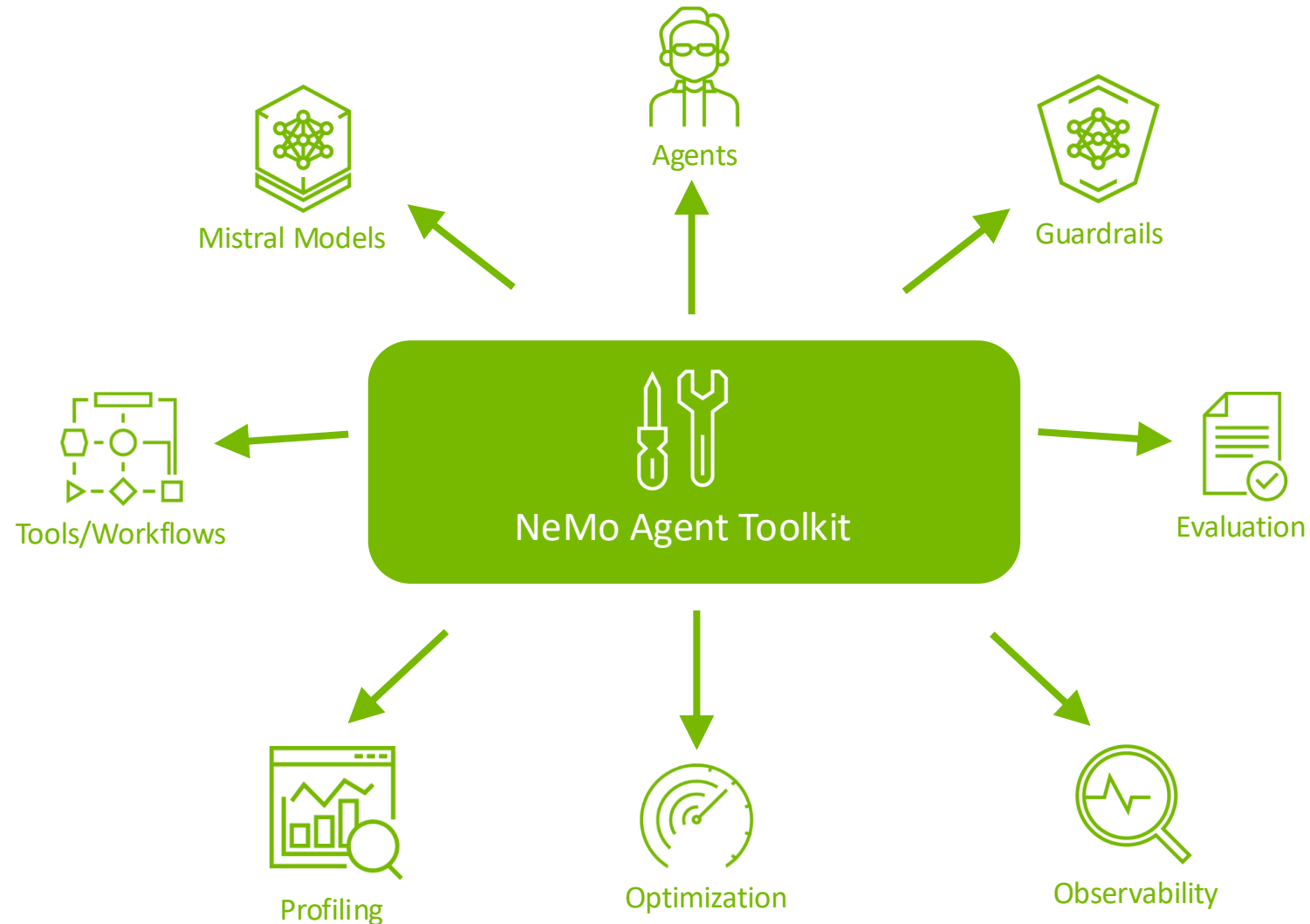
Detailed model development recipes

Starter Kits

Instructions to build custom agents with sample Jupyter Notebooks

NeMo Agent Toolkit Unifies Agent Development & Optimization

Framework-Agnostic Platform for Enterprise AI Agents



How to Get Started

Connect, Observe, Optimize AI Agents

Try

Use NeMo Agent Toolkit in [GitHub](#), view the roadmap and contribute to the community



Learn

Watch Tutorial Videos and access resources and [documentation](#) from our Learning library on the [Developer](#) page



Join

The community by participating in the [Forum](#) or [Discord](#)

