

OUR LEARNING EXCLUSIVE

- Custom exam prep software and materials
- Course-specific thinQtank® Learning publications to make the learning fun and engaging
- Hands-on labs fine-tuning, optimizing and deploying LLMs with NVIDIA tools
- **We do not teach the test.** We always deliver valuable hands-on experience
- All reading material and study guides sent to you when you register
- Every course taught by AI engineers experienced with NVIDIA platforms
- Six months mentoring access to our engineers after the course

TRAINING CAMP FORMAT

thinQtank® Learning delivers this course as a **four-day instructor-led training camp**, Monday to Thursday. About 50% lecture and 50% hands-on labs. Contact us for the schedule covering your dates.

PREREQUISITES

- No certification is required
- Recommended: Two to three years of practical AI/ML experience with large language models
- NCA-GENL or equivalent generative AI experience is a strong foundation

TARGET AUDIENCE

- Machine Learning Engineer
- Software Developer / Engineer
- Data Scientist
- Generative AI Specialist / Solutions Architect

CERTIFICATION

This course prepares you for the NVIDIA-Certified Professional: Generative AI LLMs (NCP-GENL) exam.

Exam	NCP-GENL, 60–70 questions in 120 minutes
Delivery	Online, remote proctored (Certiverse)
Exam fee	\$200, set by NVIDIA
Renewal	Valid two years; retake to renew

COURSE OVERVIEW

thinQtank® Learning is offering a unique four-day training camp comprised of four days of instructor-led learning that gives students the knowledge and skills needed to explain LLM architecture, engineer prompts, prepare data, fine-tune and optimize models, evaluate them, accelerate them on GPUs, deploy and monitor them in production and apply safety and compliance practices.

The camp follows the official NVIDIA NCP-GENL exam topics and study guide, with hands-on labs fine-tuning, optimizing and deploying LLMs with NVIDIA tools throughout.

This course prepares you for **the NVIDIA-Certified Professional: Generative AI LLMs (NCP-GENL) exam**.

WHAT YOU WILL BE ABLE TO DO

- Explain transformer-based LLM architectures and mechanisms
- Apply prompt engineering: zero/one/few-shot, chain-of-thought and output control
- Prepare, curate and tokenize data for pretraining, fine-tuning and inference
- Fine-tune and align LLMs with SFT, RLHF/DPO/GRPO and PEFT (LoRA, adapters, P-tuning)
- Scale training and inference across multiple GPUs with parallelism and profiling
- Optimize and deploy LLMs with containerized pipelines, Triton and Kubernetes
- Evaluate, monitor and maintain LLMs in production
- Apply guardrails, bias auditing and responsible-AI practices

EXCLUSIVE LEARNING PACKAGE

- NCP-GENL reading plan and review guide
- NCP-GENL practice exams
- Lab notebooks you keep after the course
- Course video recordings for review

DELIVERY METHODS

- Instructor-led training
- Immersive live-online training
- On-site and custom delivery

NVIDIA CERTIFICATION PATHS

AI Infrastructure

NCA-AIO · NCP-AII · NCP-AIO ·
NCP-AIN · NCP-ARIData center, operations and
networking

Data Science

NCA-ADS · NCP-ADS

GPU-accelerated analytics and ML

Generative AI

NCA-GENL · NCA-GENM · NCP-
GENL · NCP-AAI

LLMs, multimodal and agents

3D and Simulation

NCP-OUSD

OpenUSD development

COURSE MODULES

01 LLM Architecture and Transformers

03 Data Preparation and Tokenization

05 Parameter-Efficient Fine-Tuning (LoRA, Adapters, P-
Tuning)

07 GPU Acceleration and Distributed Parallelism

09 Model Optimization for Inference

11 Production Monitoring and Reliability

13 Exam Review

02 Prompt Engineering

04 Fine-Tuning and Alignment (SFT, RLHF, DPO)

06 Evaluation Metrics and Benchmarking

08 Performance Profiling and Memory Optimization

10 Model Deployment with Triton and Kubernetes

12 Safety, Ethics and Compliance

EXAM DOMAINS

LLM Architecture		6%
Prompt Engineering		13%
Data Preparation		9%
Model Optimization		17%
Fine-Tuning		13%
Evaluation		7%
GPU Acceleration and Optimization		14%
Model Deployment		9%
Production Monitoring and Reliability		7%
Safety, Ethics, and Compliance		5%

TERMS

© 2026 thinQtank® Global, Inc. All rights reserved. The product or information materials are protected by U.S. and intellectual property laws. thinQtank Global, thinQtank Learning and the thinQtank logo are registered trademarks of thinQtank Global, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies. NVIDIA, CUDA, RAPIDS, InfiniBand, Spectrum, NeMo, NIM and Omniverse are trademarks of NVIDIA Corporation. OpenUSD is maintained by the Alliance for OpenUSD. thinQtank Global, Inc. warrants that it will perform these training services in a reasonable manner using generally accepted industry standards and practices. THE EXPRESS WARRANTY SET FORTH IS IN LIEU OF ALL OTHER WARRANTIES, EXPRESS, IMPLIED, STATUTORY OR OTHERWISE INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A

PARTICULAR PURPOSE WITH RESPECT TO THE SERVICES AND DELIVERABLES PROVIDED BY THINQTANK GLOBAL, INC., OR AS TO THE RESULTS WHICH MAY BE OBTAINED THEREFROM. THINQTANK GLOBAL, INC. WILL NOT BE LIABLE FOR ANY THIRD-PARTY SERVICES OR PRODUCTS IDENTIFIED OR REFERRED TO CUSTOMER. All materials provided in this training are copyrighted by thinQtank Global, Inc. ("Learning Materials"). thinQtank Global, Inc. grants the customer of this learning a license to use Learning Materials strictly for the purpose of facilitating such company's internal understanding, utilization and operation of the technology covered herein. Except as set forth expressly in the sentence above, there is no transfer of any intellectual property rights or any other license granted under the terms of this training.