

LLMP® Local LLM Practitioner

What the exam asks, how it is scored, and what share of questions each domain carries.

At a glance

Questions	90
Time	120 minutes
Format	90 technical design questions, 120 minutes
Pass mark	70% scaled score
Retakes	1 retake included within 12 months
Refunds and rescheduling	Non-refundable; free reschedule up to 24h prior
Credential lasts	3 years. 45 TDUs every 3 years to renew.

Who it is for and who can sit it

Who it is for	Systems engineers, local AI engineers, security-focused developers (2–5 years)
With a degree	Experience deploying local inference engines (vLLM, llama.cpp, TensorRT-LLM) (2+ years)
Without a degree	36 months systems engineering, local AI hosting, and GPU/NPU acceleration

Share of questions by TPBOK domain

Domain	Share	What is covered
Architecture & Systems Design	30%  about 27 of 90 questions	Quantization formats, memory bandwidth calculations, KV-cache management, and latency
Engineering Governance & Agile Delivery	15%  about 14 of 90 questions	Air-gapped test validation, benchmark reproducibility, and model audit records

Domain	Share	What is covered
Applied AI & Emerging Tech Integration	35%  about 31 of 90 questions	Inference runtime optimization (vLLM, Ollama, llama.cpp), tool calling, and embeddings
Infrastructure & Security	16%  about 14 of 90 questions	Air-gapped server configurations, sovereign data compliance, and physical memory wiping
People Leadership in Engineering	4%  about 4 of 90 questions	Developer workstation enablement and internal privacy compliance education

Focus areas

- Quantization algorithms: GGUF, AWQ, GPTQ, EXL2, and precision degradation benchmarks
- Inference server architectures: continuous batching, PagedAttention, and KV-cache sizing
- Air-gapped deployment security, zero-outbound network enforcement, and sovereign privacy
- Local hardware topologies: Unified memory (Apple Silicon), multi-GPU rigs, and NPU offloading

Scoring. Tiered standard: 65% for foundational (ATP), 70% for CTP and specialty credentials, 75% for executive leadership (PgTP, PfTP). Evaluated via psychometrically calibrated scaled scoring with no negative marking.

Rescheduling. Reschedule-only. Examination voucher fees are non-refundable once registered. However, candidates can reschedule their exam sitting free of charge up to 24 hours prior to their scheduled exam window.

IICTP is a new independent credential and is not ISO/IEC 17024 accredited. Blueprint version 1.0, 21 October 2026. Domain shares are the share of questions on each exam form. See all credentials at <https://www.iictp.org/certifications/exam-facts>.