LLMScan.online
AI Model Security Scan Report

A
Score 94/100

Report summary

GGUF v3 model | architecture: gpt2 | size: 124M | quantization: F16 | 148 tensors | 2 warnings / 1 informational

Model

Filename	`john-before.gguf`
Format	`gguf`
Architecture	`gpt2`
Size	`240.8 MB`
SHA-256	`1809955eb25d7667f2c2cd8adc269677866ff244467cd4c83da6d04074e67ec1`
Scan mode	`static-deep`
Scanned at	`2026-08-23T12:12:55+00:00Z`
Report ID	`6beedd7f0eb6479f897bbbfe297fabdb`

Findings (3)

SEVERITY	CATEGORY	DETAIL
LOW	Supply Chain	**No license field in GGUF metadata (general.license)** Absence of license information makes it impossible to determine usage rights **Recommendation:** Add general.license to model metadata before distributing; use an SPDX identifier
LOW	Supply Chain	**No embedded provenance metadata** None of general.author, general.source.url, general.base_model.0, or general.source.huggingface.repository are embedded in the GGUF file. Without any origin signal the model cannot be traced to a trusted source or compared against a reference checkpoint for tampering detection. For files uploaded directly, provenance is especially critical because the scanner cannot infer repository origin from the file alone. **Recommendation:** Add at least one of: general.author, general.source.huggingface.repository, or general.source.url
INFO	Embedding Size	**Large embedding tensor 'token_embd.weight': 77,194,752 bytes** vocab=50,257 x emb_dim=768 x 2.0 B/elem -> expected 77,194,752 B — size is consistent with architecture **Recommendation:** No action needed; large embeddings are expected for large-vocabulary models

Runtime compatibility

RUNTIME	SUPPORTED	NOTES
llama.cpp	Yes	GGUF v3 container: supported | quant F16: supported
Ollama	Yes	GGUF via llama.cpp backend: container and quantization supported
LM Studio	Yes	GGUF container and quantization supported; GPU/CPU inference
vLLM	No	GGUF not supported; requires safetensors format

HuggingFace Transformers	No	GGUF not supported; requires safetensors or PyTorch checkpoint

Generated 2026-09-21 06:03 UTC by LLMScan.online (engine v2.0.0, report schema v1.0). Full interactive report: https://llmscan.online/report/6beedd7f0eb6479f897bbbfe297fabdb — Static analysis only; not a guarantee of safety. See llmscan.online/docs/how-it-works for methodology.