Ubuntu 26.04 Is Built for Local AI — What Actually Changes

February 26, 2026 · by Mark Bartlett

Download this guide as PDF

> **Quick Answer:** Ubuntu 26.04 LTS (April 23, 2026) packages both NVIDIA CUDA and AMD ROCm directly in Ubuntu's official repositories. No more downloading .deb files from NVIDIA's website. No more fighting AMD's external ROCm repos. `sudo apt install rocm` or a single apt command for CUDA. The ubuntu-drivers tool already auto-detects your GPU and installs the right NVIDIA driver — now the compute stack behind it is just as easy. If you're building a new local AI machine, wait for 26.04. If 24.04 works for you today, there's no rush.

📚 **Related:** ROCm Not Detecting GPU: AMD Fix Guide · CUDA Out of Memory Fix · Local AI Troubleshooting · Budget AI PC Under $500

The number one thing that stops people from running AI locally on Linux isn't the models, the VRAM, or the software. It's the GPU driver.

You install Ubuntu. You install Ollama. You type `ollama run llama3.3:8b`. And then you get a wall of errors because CUDA isn't installed, or ROCm can't find your AMD card, or the kernel module didn't build because Secure Boot blocked it. You spend the next two hours on Stack Overflow instead of running models.

Ubuntu 26.04 LTS, due April 23, 2026, is Canonical's attempt to fix this. Both NVIDIA CUDA and AMD ROCm will ship in Ubuntu's official package repositories. The goal: get from fresh install to running models without touching a browser.

What's actually confirmed

Two separate announcements, both from Canonical:

NVIDIA CUDA in Ubuntu repos

Announced September 2025. Canonical is packaging and distributing the CUDA toolkit and runtime directly in Ubuntu's repositories. Their statement: "Once CUDA redistribution is fully integrated into Ubuntu, the current multi-step installation process becomes a single command."

No specific Ubuntu version was named in the announcement, but 26.04 LTS is the logical target.

AMD ROCm in Ubuntu repos

Announced December 2025. Canonical confirmed that ROCm will be available in Ubuntu 26.04 LTS repositories. Installation becomes `sudo apt install rocm`. AMD's Senior VP Andrej Zdravkovic: "Working with Canonical to package AMD ROCm for Ubuntu makes it easier for developers and enterprises to deploy AMD solutions on supported systems."

ROCm packages will also ship in every Ubuntu release after 26.04 (26.10, 27.04, etc.), not just LTS versions.

What's NOT included

To be clear about what these announcements do and don't say:

- ROCm and CUDA are **not installed by default**. They're in the repos, but you still run `apt install`. This is optional software, not pre-loaded.
- No specific CUDA or ROCm **version numbers** have been announced for 26.04.
- `ubuntu-drivers autoinstall` already handles NVIDIA driver detection on current Ubuntu. That's not new — what's new is the compute stack (CUDA/ROCm) being available the same way.

Why this matters: the before and after

Here's what GPU setup looks like today on Ubuntu 24.04 vs what it will look like on 26.04.

NVIDIA (current: 7+ steps)

```
# Step 1: Install the NVIDIA driver
sudo ubuntu-drivers autoinstall
sudo reboot

# Step 2: Go to developer.nvidia.com, find your Ubuntu version
# Step 3: Download the .deb repo file
wget https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2404/x86_64/cuda-keyring_1.1

# Step 4: Install repo package and GPG key
sudo dpkg -i cuda-keyring_1.1-1_all.deb
```

```
# Step 5: Update and install
sudo apt update
sudo apt install cuda-toolkit-12-6

# Step 6: Set environment variables in ~/.bashrc
export PATH=/usr/local/cuda/bin:$PATH
export LD_LIBRARY_PATH=/usr/local/cuda/lib64:$LD_LIBRARY_PATH

# Step 7: Verify
nvcc --version
```

NVIDIA (26.04: 2 steps)

```
sudo ubuntu-drivers autoinstall && sudo reboot
sudo apt install cuda
```

AMD ROCm (current: good luck)

```
# Step 1: Check kernel version matches AMD's supported list
uname -r  # pray it's on the list

# Step 2: Add AMD's external repository
wget https://repo.radeon.com/amdgpu-install/...
sudo apt install ./amdgpu-install_x.x.x.deb

# Step 3: Install ROCm (and debug dependency conflicts)
sudo apt install rocm

# Step 4: Fix DKMS build failures when it doesn't compile
# Step 5: Add yourself to render and video groups
sudo usermod -aG render,video $USER

# Step 6: If your GPU isn't officially supported (most consumer cards):
export HSA_OVERRIDE_GFX_VERSION=11.0.0  # RX 7900 spoof

# Step 7: Reboot
# Step 8: Verify (and troubleshoot when it doesn't work)
rocminfo
```

The ROCm install path on Ubuntu 24.04 is genuinely bad. Dependency conflicts between `amdgpu-dkms`, ROCm runtime versions, and kernel headers are common enough that we wrote an entire troubleshooting article about it.

AMD ROCm (26.04: 1 step)

```
sudo apt install rocm
```

Dependencies managed by Ubuntu. Security updates through normal apt channels. No external repos. No GPG key imports. No DKMS version mismatches.

What this fixes for local AI builders

The three-step dream

The pitch from Canonical is that this workflow becomes possible:

1. Install Ubuntu 26.04
2. `sudo apt install ollama-amd` or `sudo apt install cuda && curl -fsSL https://ollama.com/install.sh | sh`
3. `ollama run llama3.3:8b`

No driver hunting. No Stack Overflow detours. GPU-accelerated inference from a fresh install in under 20 minutes.

This is closer to the macOS experience where you install Ollama and it just works, because macOS handles Metal GPU access automatically. Linux has never had that smoothness for discrete GPU computing. 26.04 gets closer.

ROCm detection is the big win

NVIDIA driver detection on Ubuntu already works reasonably well. `ubuntu-drivers autoinstall` finds your card, picks the right driver, and handles Secure Boot signing. The pain point was CUDA toolkit installation — messy but solvable.

ROCm is a different story. The install process on AMD has been genuinely broken for many users. DKMS build failures on recent kernels, dependency conflicts between ROCm versions and

kernel headers, and a supported GPU list that officially only covers Radeon RX 7900 GRE and above (everything else requires `HSA_OVERRIDE_GFX_VERSION` hacks).

Canonical maintaining ROCm packages means:

- Dependencies stay compatible with the Ubuntu kernel version
- Security patches arrive through normal `apt upgrade`
- ROCm becomes a first-class dependency that other packages can declare (so `apt install ollama-amd` can pull in ROCm automatically)

For people who chose AMD GPUs for local AI and have been fighting ROCm, that fight is almost over.

Secure Boot stops being a landmine

Kernel module signing is one of the more obscure headaches. You install NVIDIA or AMD drivers, reboot, and the GPU isn't detected because Secure Boot won't load unsigned kernel modules. On current Ubuntu, `ubuntu-drivers` handles this for NVIDIA by installing signed drivers. But third-party DKMS modules (like AMD's `amdgpu-dkms` from their external repo) don't always get signed correctly.

With ROCm packaged by Canonical, the kernel modules are built against the Ubuntu kernel and signed by Canonical's key. Secure Boot works without manual key enrollment.

Should you upgrade?

If you're on 24.04 LTS and everything works

No rush. Seriously. If your NVIDIA or AMD GPU runs Ollama, llama.cpp, or whatever else you use, don't break a working setup for packaging improvements. 24.04 has support until 2029. Upgrade on your schedule, not Canonical's.

If you're fighting ROCm on AMD

Wait for 26.04 and do a fresh install. It will almost certainly be less painful than debugging your current ROCm setup. The release date is April 23, 2026 — mark it.

Or if you can't wait: Docker with ROCm support works today and sidesteps most install issues. Map `/dev/kfd` and `/dev/dri` into the container and let the container handle the ROCm runtime.

If you're building a new machine

Wait for 26.04 if you can. Especially if you're going AMD GPU. The difference between "fight ROCm for 3 hours" and "apt install rocm" is the difference between a new user sticking with local AI or giving up.

If you're buying hardware now and can't wait, NVIDIA is still the easier path on Linux. The existing `ubuntu-drivers autoinstall` plus NVIDIA's CUDA repo works — it's just more steps than it should be.

The kernel and driver version caveat

LTS doesn't mean bleeding edge. Ubuntu 26.04 will ship with **Linux kernel 6.20** (possibly renamed to 7.0) and whatever CUDA and ROCm versions are current at the time of release. If you need the absolute latest driver for a just-released GPU, you may still need PPAs or manual installs.

However, LTS point releases (26.04.1, 26.04.2, etc.) typically backport newer hardware enablement (HWE) kernels. By 26.04.2 or 26.04.3, support for newer GPUs will catch up.

What this doesn't solve

Ubuntu 26.04 makes the install easier. It doesn't change the fundamentals of running AI locally.

VRAM is still VRAM

Getting ROCm installed in one command doesn't give your RX 7900 XTX more than 24GB. A 70B model still needs 40GB+ at Q4. See our VRAM requirements guide for what fits where.

Unsupported AMD GPUs are still unsupported

ROCm in Ubuntu repos doesn't change which GPUs AMD officially supports. If you have an RX 6600, 6700 XT, or older card, you'll likely still need the `HSA_OVERRIDE_GFX_VERSION` workaround. The packaging is easier, but the support matrix is AMD's decision, not Canonical's.

Multi-GPU is still manual

If you're running dual RTX 3090s or a mixed GPU setup, you're still configuring tensor parallelism, NCCL, and device placement yourself. Ubuntu doesn't automate multi-GPU topology.

Model selection is still on you

The OS can get your GPU ready. Picking the right model for your hardware and use case is a separate problem. Start with our troubleshooting guide if things aren't working after drivers are installed.

Apple Silicon doesn't apply

Ubuntu on Apple Silicon exists but is not a serious path for local AI. If you have a Mac, use macOS with MLX or Ollama directly.

Ubuntu 26.04 specs at a glance

	Ubuntu 24.04 LTS	Ubuntu 26.04 LTS
Release date	April 2024	April 23, 2026
Codename	Noble Numbat	Resolute Raccoon
Kernel	6.8	6.20 (likely 7.0)
NVIDIA CUDA	External NVIDIA repo required	In Ubuntu repos
AMD ROCm	External AMD repo required	In Ubuntu repos (`apt install rocm`)
GPU driver detection	`ubuntu-drivers autoinstall`	Same (already works)
Desktop	GNOME 46	GNOME 50
Python	3.12	3.14
Standard support	5 years (to 2029)	5 years (to 2031)
Extended (Ubuntu Pro)	12 years	Up to 15 years

The bigger picture

Other distros are heading the same direction. Fedora has been improving its NVIDIA driver story through RPM Fusion. NixOS has declarative CUDA configs. Arch has streamlined NVIDIA packages. Ubuntu packaging both CUDA and ROCm in official repos matters more because Ubuntu is what most people install when they Google "Linux for AI."

Docker remains the most reproducible way to run GPU workloads, and it will stay that way for production deployments. But native GPU support through apt means Docker isn't required anymore. For someone building a budget AI PC and installing Linux for the first time, the path from "I have a computer" to "I'm running a local LLM" gets shorter.

People who would never have compiled a kernel module are starting to run local AI. Ubuntu 26.04 makes that easier.

Bottom line

Ubuntu 26.04 LTS doesn't do anything magical. What it does is remove the single biggest friction point for new local AI users on Linux: GPU compute stack installation.

CUDA in the repos means no more downloading .deb files from NVIDIA's developer site, importing GPG keys, and pinning repos. On the AMD side, ROCm dependency nightmares and DKMS build failures go away.

The practical impact: fresh Ubuntu install, `apt install rocm` or `apt install cuda`, install Ollama, run models. Three commands between a new machine and local inference.

If you're setting up a new system for local AI, April 23 is worth waiting for. If your current setup works, keep using it — this is a quality-of-life upgrade, not a performance upgrade. And if you're on AMD and have been fighting ROCm, you finally have a date when it gets better.
