

The AG Gazette

Models, machines, and matters of consequence

PARIS, FRIDAY 16 OCTOBER 2026

AFTERNOON EDITION · No. 2026-10-16-AFTERNOON

ISSUED 16:09 CEST

PRICE: COMPLIMENTARY

AFTERNOON EDITION — FRIDAY 16 OCTOBER 2026

Reporting window: 16 Oct 2026, 06:12 CEST to 16 Oct 2026, 16:09 CEST (Paris time). Issued 16 Oct 2026, 16:09 CEST.

One material publication milestone cleared the verification gate since the actual morning edition at 06:12 CEST: Beam's public weights availability. Its October 5 preview is explicitly distinguished from today's repository publication milestone. We do not repeat this morning's six stories or relabel older model announcements, research papers and security reports as new because they resurfaced in discussion.

1 NEW OR UPDATED DISPATCH · 0 BACKGROUND ITEMS · PREVIOUS: [2026-10-16-MORNING](#)

THE NEWS

REFLECTION MAKES BEAM WEIGHTS AVAILABLE UNDER APACHE 2.0, WITH UNEVEN AGENT BENCHMARK RESULTS

PUBLISHED: 16 Oct 2026, 12:34 CEST

Reflection has made its Beam-1.0-501B-A23B weights publicly available on Hugging Face. The repository commit labeled "Beam is available" is timestamped October 16 at 12:34 CEST, after this morning's edition. The repository is public and ungated at verification. That commit is the timing anchor for this dispatch, not the model repository's earlier creation date.

The model card describes a text-only sparse mixture-of-experts model with 501 billion total parameters, 23 billion active parameters and a one-million-token context window. Instruction-tuned BF16, FP8 and NVFP4 releases are listed under Apache 2.0. This is a weights-availability development, not the first announcement of Beam: Reflection previewed it on October 5 and said the weights would follow later in the month.

In its October 16 evaluation, Artificial Analysis gives Beam an Intelligence Index score of 30, behind Meta's Muse Spark Open at 39. Its results are mixed: 48.5% on AutomationBench-AA but 3.0% on Terminal-Bench 4.0. The same evaluator reports 33% accuracy and a 47% hallucination rate on AA-Omniscience. These are results under specific evaluation setups, not a verdict that Beam is best for every agent workload.

Reflection's own model card reports different benchmarks, including Terminal Bench v2.1; those numbers are not directly comparable with Artificial Analysis's Terminal-Bench 4.0 result. The company's model-development technical report is still marked as forthcoming in the card, while a separate safety technical report is already accessible.

ANALYSIS — WHY IT MATTERS. *Downloadable, permissively licensed weights give researchers and operators a new option they can inspect and adapt. The independent evaluation argues for testing the intended workflow rather than equating model size, one score or token efficiency with reliable coding performance.*

CAVEAT. We verified public artifacts, their provenance and reported benchmark results; we did not download the large weights or run the model. A repository commit is a directly observed publication milestone, not proof of the exact first instant anyone could obtain access. The October 5 preview is dated background within this story, not fresh news. The evaluated terminal benchmarks and harnesses differ, so their percentages must not be compared as like-for-like.

SOURCES

- Beam model card** <<https://huggingface.co/reflection/Beam-1.0-501B-A23B/raw/main/README.md>> (primary source) **CHECKED: 16 Oct 2026, 16:08 CEST**
- Beam model commits** <<https://huggingface.co/api/models/reflection/Beam-1.0-501B-A23B/commits/main>> (primary source) **CHECKED: 16 Oct 2026, 16:08 CEST**
- Beam model metadata** <<https://huggingface.co/api/models/reflection/Beam-1.0-501B-A23B>> (primary source) **CHECKED: 16 Oct 2026, 16:08 CEST**
- Reflection Beam analysis** <<https://artificialanalysis.ai/articles/reflection-beam>> (primary source) **CHECKED: 16 Oct 2026, 16:07 CEST**
- Reflection launch** <<https://reflection.ai/blog/introducing-beam>> (primary source) **CHECKED: 16 Oct 2026, 16:08 CEST**
- Reflection safety** <<https://reflection.ai/beam-safety-report>> (primary source) **CHECKED: 16 Oct 2026, 16:08 CEST**

KEY PATTERNS — EDITORIAL ANALYSIS	
— Model announcement and public weights availability are separate events: the October 5 preview is not presented as today's launch. — Benchmark versions and harnesses matter: strong SaaS-workflow performance does not establish reliable terminal coding.	
SOURCE COVERAGE & COLOPHON	
COLLECTION ENGINE VERSION: 3.28.0 RAN THIS EDITION: yes ENGINE LOOK-BACK: 2 days	SOURCE STATUS REDDIT: ok YOUTUBE: ok HACKERNEWS: ok POLYMARKET: unreachable GITHUB: ok DEVTO: ok X: skipped-unconfigured
COVERAGE NOTES — The edition interval begins with the actual October 16 morning issue at 06:12 CEST. One distinct new story is included; no background dispatches were added to fill a quota. — The installed engine completed successfully with 71 evidence items across five returning sources. These are discovery leads, not 71 verified news developments. Models, coding and agents, research, infrastructure and policy were searched. — Polymarket was unreachable because TLS hostname verification failed; TLS verification was not bypassed. X was unconfigured and omitted. YouTube supplied two transcripts for three videos; the remaining subtitle request was rate-limited. — Three post-engine native web supplements were completed. The configured search backend was unavailable and its automatic keyless fallback supplied results. The text extractor was unavailable, so official pages, APIs and the research paper were retrieved directly with normal TLS. — New discussion does not prove a new publication. Hybrid Reasoning Agents was submitted October 13; the Swarmchasers report is dated October 14; NVIDIA Nemotron announcements and the Brin and Apple/Meta coverage predate this morning. Timestral metadata last changed October 15. No separate coding, research or policy story cleared the post-morning timing gate. — The engine's available community comments concern other or older stories, not Beam. They are not attached to this release as if they were reactions to it. No representative sentiment or community consensus is claimed. — No browser cookies, optional credentials or services, paid enrichment, TLS bypass, job-alert edits or schedule changes were used. Raw evidence, source snapshots, query plan and provenance remain private.	ENGINE FOOTER --- All agents reported back! └─ Reddit: 14 threads 892 upvotes 407 comments └─ YouTube: 3 videos 900,606 views 2/3 with transcripts └─ HN: 21 stories 191 points 81 comments └─ dev.to: 12 articles 62 reactions 23 comments └─ GitHub: 21 items 58 reactions 224 comments └─ Top voices: r/OpenAI, r/LocalLLaMA, r/ClaudeAI ---
THIS EDITION — <u>Exact input JSON</u> — <u>Markdown text</u> — <u>Permanent page</u>	

The AI Gazette is a locally rendered digest. Each dispatch separates reported facts, labeled analysis and caveats, and cites the sources checked at the times shown. Background items are older or undated context, not new reporting.

Rendered as static HTML with inline styles and no scripts, cookies or analytics. Typefaces are requested from Google Fonts (fonts.googleapis.com, fonts.gstatic.com); if that request is blocked, Georgia, Times or another installed serif is used.