

KI-Tagesbericht

Sonntag, 18.10.2026 · Trending-Modelle und die Analysen des Tages

Trending-Modelle Top 20 laut Hugging-Face-API · Stand 18.10.2026, 21:03 Uhr

#	MODELL	AUFGABE	LIKES	DOWNLOADS	TREND-SCORE	ERSTELLT
1	deepseek-ai/DeepSeek-V4.1-Flash-1017	image-text-to-text	1.282	10.233	1.192	2026-10-17
2	google/embeddinggemma-2	feature-extraction	2.649	164.339	812	2026-09-14
3	openbmb/MiniCPM-V-4.7-35B-A3B	image-text-to-text	986	57.935	804	2026-10-07
4	jialinyzz/humanizer	text-generation	1.933	103.709	727	2026-09-11
5	fishaudio/drama-3	text-to-speech	676	9.620	652	2026-10-01
6	black-forest-labs/flux-3-image-klein-comfy	image-to-image	778	338.036	614	2026-10-14
7	meta-models/Muse-Spark-Open-816B-A42B	image-text-to-text	536	1.809	497	2026-09-15
8	Lightricks/LTX-2.5	image-to-video	8.337	2.012.496	460	2026-07-23
9	LiquidAI/LFM2.5-24B-A2B	text-generation	461	13.007	446	2026-10-12
10	black-forest-labs/flux-3-image-klein	image-to-image	475	52.315	438	2026-10-13
11	abenzers/Qwen-Image-2.1-Uncensored-Quantized	text-to-image	4.544	2.661.687	399	2026-09-20
12	baidu/Unlimited-OCR	image-text-to-text	4.786	1.652.742	355	2026-06-19
13	Qwen/Qwen-Image-2.1-Turbo	text-to-image	980	23.131	324	2026-10-09
14	OpenMOSS-Team/MOSS-Image-SFT	text-to-image	324	2.087	314	2026-10-11
15	openbmb/MiniCPM-V-4.7-1B	image-text-to-text	384	22.689	290	2026-10-06
16	Qwen/Qwen3.8-27B	image-text-to-text	17.948	7.695.206	288	2026-08-05
17	OpenMOSS-Team/MOSS-Image-Base	text-to-image	280	1.791	269	2026-10-11
18	facebook/sam3	mask-generation	4.613	1.945.717	222	2025-11-07
19	nvidia/Qwen3.8-Flash-Next-FP4-Local	image-text-to-text	208	4.876	206	2026-10-08
20	LiquidAI/LFM2.5-24B-A2B-GGUF	text-generation	203	107.599	198	2026-10-09

So lesen: Reihenfolge nach dem Feld *trendingScore* der Hugging-Face-API, Stand 18.10.2026, 21:03 Uhr. Likes, Downloads und Trend-Score sind Zähler der API, keine Qualitätsbewertung und kein Benchmark-Ergebnis. „Aufgabe“ ist das Feld *pipeline_tag*. Quelle: [huggingface.co](#) (öffentliche API, ohne Schlüssel).

Die zwei Analysen des Tages

MODELLE · 18.10.2026 · 6 QUELLEN

Muse Spark Open: Was die offenen Gewichte technisch bedeuten

Metas 816-Milliarden-Modell im Check: Architektur, Speicherbedarf, Lizenzbedingungen und eine Benchmark-Bilanz, die eher gemischt ausfällt als überlegen.

- 816 Mrd. Parameter, 42 Mrd. aktiv, 1 Mio. Token Kontext, Eingaben aus Text, Bild, Audio und Video, Ausgabe nur Text.
- Speicherbedarf laut Modellkarte: mindestens 700 GB (NVFP4) bzw. 2 TB (BF16); unsere Überschlagsrechnung kommt auf etwa 674 GB bzw. 1,9 TB.
- Nach Metas eigener Tabelle liegt das Modell in 7 von 20 Fähigkeitstests vor allen Vergleichsmodellen, bei Terminal-Bench 2.1 und GPQA dahinter.

<https://kifachblatt.de/muse-spark-open-offene-gewichte-analyse>

Quellen

1. [Introducing Muse Spark Open, The Most Capable US Open-Weight Model to Date](#) (research.meta.ai)
2. [Welcome Muse Spark Open: open-weight frontier multimodal mixture-of-experts model from Meta](#) (huggingface.co)
3. [meta-models/Muse-Spark-Open-816B-A42B: Muse Model License und Model Card](#) (huggingface.co)
4. [Muse Spark Open \(Collection\)](#) (huggingface.co)
5. [Muse Spark Is Now an Open-Weights Model](#) (pytorch.org)
6. [Muse Spark Open, an open-weight model for coding agents](#) (baseten.co)

EVALS · 18.10.2026 · 6 QUELLEN

Long-Horizon-Terminal-Bench 2.0: Was drei Stunden Agentenarbeit zeigen

Ein Preprint testet 15 KI-Modelle drei Stunden lang auf 77 Terminalaufgaben. Was Methode, Zahlen und Kosten belegen und wo die Grenzen der Studie liegen.

- Umfang: 77 Aufgaben, 13 Domänen, acht Aufgabentypen, 1.145 gewertete Läufe von 15 Modellen.
- Spitze: GPT-6-astra 37 von 77 Aufgaben (48,1 %), GPT-5.6-sol 32 (41,6 %). Eigene Rechnung: Die 95%-Intervalle beider Modelle überlappen weit, der Abstand von fünf Aufgaben ist statistisch nicht abgesichert.
- Binär gegen dicht: Auf den 30 Rekonstruktionsaufgaben liegt der binäre Durchschnitt bei 0,221, der Durchschnitt mit Fortschrittsleiter bei 0,653.

<https://kifachblatt.de/lhtb-2-langzeit-agenten-benchmark>

Quellen

1. [Long-Horizon-Terminal-Bench 2.0: Automating Diverse Task Construction and Measuring Sustained Agent Progress](#) (arxiv.org)
2. [Long-Horizon-Terminal-Bench: Testing the Limits of Agents on Long-Horizon Terminal Tasks with Dense Reward-Based Grading](#) (arxiv.org)
3. [Long-Horizon Terminal-Bench: Where Agents Run Out of Steam \(Projektseite\)](#) (zli12321.github.io)
4. [Terminal-Bench: Benchmarking Agents on Hard, Realistic Tasks in Command Line Interfaces](#) (arxiv.org)
5. [Terminal-Bench 2.0 and Harbor](#) (tbench.ai)
6. [Measuring AI Ability to Complete Long Software Tasks](#) (arxiv.org)

Vollständige Texte mit Quellen, Methode und Grenzen auf kifachblatt.de. „Eigene Rechnung“ kennzeichnet Berechnungen der Redaktion auf Basis der genannten Quellen.

Impressum & Redaktionsschluss

Anbieter nach § 5 DDG

Philipp Paulik
Mozartstraße 1
53819 Neunkirchen-Seelscheid
Nordrhein-Westfalen

Verantwortlich nach § 18 Abs. 2 MStV

Philipp Paulik, Anschrift wie links

Redaktion

KI Fachblatt, Redaktion Philipp Paulik

Kontakt

E-Mail: philipp.l.paulik@gmail.com
Web: kifachblatt.de

Redaktionsschluss

Diese Ausgabe wurde am 18.10.2026, 21:03 Uhr automatisch erstellt. Später veröffentlichte Beiträge und Zählerstände sind nicht berücksichtigt.

Hinweis zu KI

Die Inhalte des KI Fachblatts werden mit Unterstützung von KI erstellt und redaktionell geprüft (siehe kifachblatt.de/transparenz). Dieser Bericht wird automatisch aus öffentlichen Daten und den veröffentlichten Beiträgen zusammengestellt. Trotz Prüfung können Fehler nicht ausgeschlossen werden.

Datenquellen

Modellliste: Hugging Face, öffentliche API (huggingface.co/api/models, Sortierung trendingScore), Stand 18.10.2026, 21:03 Uhr. Analysen: eigene Beiträge des KI Fachblatts.

© 2026 KI Fachblatt · Philipp Paulik · Stand 18.10.2026, 21:03 Uhr