

7 May 2024

WHAT'S TRENDING NOW

AI SPICE

just for you @dhanushreddy291

CLOUD GPU BOOM

Alternative clouds are surging in popularity due to the rapidly growing demand for affordable GPU access to train and run generative AI models. CoreWeave, Lambda Labs, Voltage Park, and Together AI are among the well-funded startups offering alternative cloud GPU services. These providers typically offer lower costs for GPU instances compared to cloud giants AWS, Google Cloud, and Microsoft Azure. While alternative clouds are expected to face competition from incumbents investing in custom...



MICROSOFT LAUNCHES PHI-3

Phi-3-mini is a highly capable 3.8 billion parameter language model that can be deployed on a phone. It is trained on a scaled-up dataset of heavily filtered web data and synthetic data, and aligned...

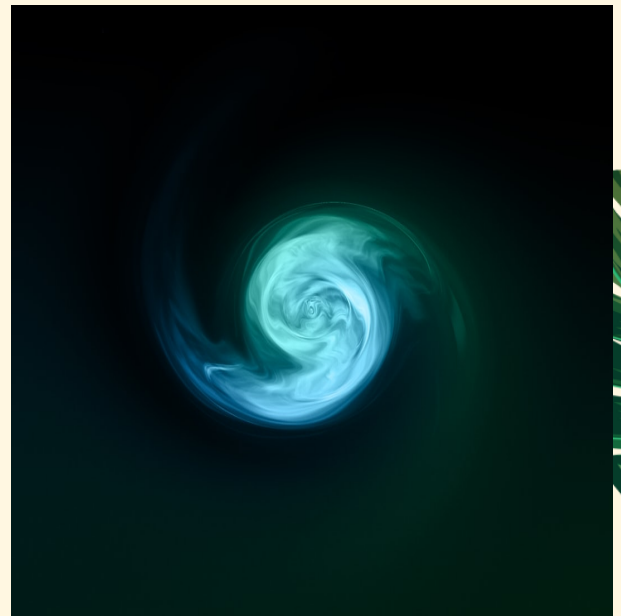


MICROSOFT'S AI MODEL

Microsoft is reportedly developing MAI-1, an in-house AI model with 500 billion parameters to compete with models from Google, OpenAI, and others. The model is led by Mustafa Suleyman, former CEO of DeepMind, and may be unveiled at Microsoft Build on May 21-23. Microsoft also emphasizes the use of AI models across its products and services...

SHOW HN: KUBERNATES IN NODE.JS

nodejs-k8s recreates core Kubernetes functionality in NodeJS, compatible with the kubectl CLI. It partially implements resources like Pods, Services, and Ingress, and fully implements others like ConfigMap, Deployment, and Namespace. To use it, install NodeJS, Docker, and MongoDB...



MEMORY-FIRST CLOUD DATABASE

Amazon MemoryDB is a cloud database built on open-source Redis, offering 11 9s of durability and in-memory performance. It uses a separate transaction log service to handle durability, enabling independent scaling of performance, availability, and durability. This architecture maintains Redis compatibility...