

Emerging

Smaller players and specific developments to spot early · week 43, as of 24 October 2026

What is emerging outside the four large providers: smaller players, open models, protocols, tools and rules that may change how work is done. This report signals early and recommends nothing; most of it can only be used by developers and early adopters today. What already works for everyone is in the reports *The landscape* and *This month*.

The signal: early signal (a first release or announcement), growing (several players or wide use), breaking through (widely adopted, becoming a standard).

Open models

Kimi K3.1 Pro

Moonshot AI · 23 October 2026 · growing · [source](#)

A new AI model available via API with full model weights scheduled for release by November 4, 2026.

Why it may matter: Benchmarks show it rivals top proprietary models at lower pricing and will have open weights.

Usable today by: developers

Beam

Reflection · 16 October 2026 · early signal · [source](#)

A 501B parameter open-weight model designed for software engineering, reasoning, and long-running agentic tasks.

Why it may matter: It brings massive open-weight capabilities to independent developers and agent systems via platforms like Ollama Cloud.

Usable today by: developers

Timestral

Mistral AI · 19 October 2026 · early signal · [source](#)

A series of foundation models from Mistral AI dedicated to time-series data.

Why it may matter: It introduces specialized foundation models for forecasting and data analysis beyond text.

Usable today by: developers

Clef

Cloudflare · 19 October 2026 · early signal · [source](#)

A 27B parameter multimodal decision model available via Ollama for local and cloud execution.

Why it may matter: It brings decision-focused multimodal models from an infrastructure provider into local and hybrid workflows.

Usable today by: developers

Agents and frameworks

Fly.io Sprites with Ollama

Fly.io & Ollama · 13 October 2026 · early signal · [source](#)

An integration allowing AI agents to safely use Ollama's cloud models inside isolated Fly.io cloud computers.

Why it may matter: It provides a secure sandboxed environment to isolate autonomous agents running code from uncontrolled actions.

Usable today by: developers

AI Agent Payment Suite for Stablecoins

Fireblocks · 22 October 2026 · early signal · [source](#)

A payment suite allowing AI agents to independently execute transactions using stablecoins.

Why it may matter: Supports emerging infrastructure for direct financial settlement between autonomous systems.

Usable today by: developers and early adopters

Protocols and agent payments

Open Wallet Standard

MoonPay · 23 October 2026 · early signal · [source](#)

An open wallet standard designed for payments executed by autonomous AI agents.

Why it may matter: It standardizes how AI agents can independently manage transactions and wallets.

Usable today by: developers

x402 SDK

x402 · 23 October 2026 · early signal · [source](#)

A critical vulnerability in the x402 SDK that allows malicious APIs to drain AI agents' token balances.

Why it may matter: It highlights practical security risks when deploying automated payment protocols for agents.

Usable today by: developers

Browser and computer use

Pure Vision-Powered Computer Use Agent

Yancheng Assistant · 22 October 2026 · early signal · [source](#)

A computer use agent operating purely on vision for tasks like bulk job applications and content migration.

Why it may matter: It illustrates the shift toward visual OS interaction without relying on underlying API integrations.

Usable today by: early adopters and developers

Local and on-device

LFM2.5-24B-A2B

Liquid AI · 13 October 2026 · growing · [source](#)

A compact 24B parameter model designed for running fast, private agents directly on-device.

Why it may matter: It enables complex agent workflows on local hardware without reliance on external cloud APIs.

Usable today by: developers

Voice

Dessa

Block · 23 October 2026 · early signal · [source](#)

A full-duplex speech model trained entirely on conversations without text intermediate processing.

Why it may matter: The model aims for real-time, human-like voice interaction without relying on traditional speech-to-text pipelines.

Usable today by: developers and researchers

Granite Speech 5.0

IBM · 21 October 2026 · growing · [source](#)

A speech model designed to transcribe, translate, and attribute faster and more accurately.

Why it may matter: It provides enterprise-grade open alternatives for advanced transcription and multilingual voice processing.

Usable today by: developers

Voxtral 3

Mistral AI · 20 October 2026 · early signal · [source](#)

A new speech model from Mistral AI.

Why it may matter: It expands the options for alternative European models in voice and audio processing.

Usable today by: developers

Developer tools

DeepInfer

DeepSeek · 20 October 2026 · early signal · [source](#)

A high-performance inference framework designed specifically for DeepSeek models.

Why it may matter: It optimizes local and large-scale execution of DeepSeek models for faster inference.

Usable today by: developers

Rules and regulation

DeepSeek Federal Civilian Agency Ban

U.S. Department of Homeland Security (DHS) · 19 October 2026 · breaking through · [source](#)

An official DHS order prohibiting the use of DeepSeek inside U.S. federal civilian agencies.

Why it may matter: Government restrictions on foreign AI models force organizations to reconsider their choice of model providers.

Usable today by: researchers

Sources read up to 24 October 2026: Hacker News (newest item 23 October 2026); Simon Willison's blog (newest item 23 October 2026); Hugging Face blog (newest item 23 October 2026); Ollama blog (newest item 19 October 2026); Model Context Protocol blog (newest item 22 August 2026); news: open models (newest item 24 October 2026); news: Mistral, Cohere, Perplexity, Hugging Face, ElevenLabs (newest item 24 October 2026); news: MCP, A2A, x402, agent

payments (newest item 23 October 2026); news: browser and computer use (newest item 24 October 2026); news: local AI and voice (newest item 24 October 2026); Dutch news: AI start-ups and the AI Act (newest item 23 October 2026).

Every item comes from one of these fetched pieces, with the link to that piece; nothing comes from the language model's own knowledge. An inexpensive language model summarises and selects; renewed every week. This is a draft: research, not advice.

Floris Knol

Research into AI technology you can use every day

florisknol.com