

◆ DISPATCH 018 · 2026-05-06

# Agents Buy Domains, Gemma Ships Drafters, and Local Catches Up to 65 Percent of the Job

👉 GSV GIVE THE AGENT A WALLET, NOT A CARD

2026-05-06 · 00:30:02

*“Agents can buy domains now without ever touching a credit card number.”*

— LENAR KESS, TODAY'S NARRATION

Agents can now sign up for Cloudflare and buy a domain through a tokenized payment protocol Cloudflare and Stripe co-designed. Google ships first-party multi-token prediction drafters for the entire Gemma 4 family the same week the LocalLLaMA community gets a 2.5x speedup on Qwen 3.6 27B from a hand-built llama.cpp branch. OpenAI swaps the ChatGPT default to GPT-5.5 Instant. NVIDIA, Microsoft, and OpenAI publish MRC, the multipath transport protocol behind Blackwell-era frontier training. And on the labor side, Dario Amodei trades his white-collar bloodbath line for the Jevons Paradox onstage with Jamie Dimon.

- [Cloudflare and Stripe ship the agent-as-customer protocol](#)
- [NVIDIA, Microsoft, and OpenAI publish MRC as an open OCP spec](#)
- [Google releases MTP drafters for the Gemma 4 family](#)
- [Community reports 2.5x and 100 tok/s on Qwen 3.6 27B with MTP](#)
- [GPT-5.5 Instant becomes the ChatGPT default](#)
- [Dario Amodei pivots from bloodbath to Jevons](#)
- [A clean indirect prompt-injection screenshot from r/ClaudeAI](#)
- [Trajectory-level fraud detection for LLM agents](#)
- [Agentic safety as interaction topology, not model alignment](#)
- [Agents as marginal token allocators](#)
- [ProgramBench: rebuilding 200 binaries from scratch](#)
- [A measured 65/20/15 cloud-vs-local routing rule](#)

- [François Fleuret on what LLMs are still missing](#)
- [Telus Digital alters call-centre agent accents in real time](#)
- [Write some software, give it away for free](#)

## CHAPTERS

- 00:00:04 Agents become Cloudflare customers
- 00:03:24 MRC: the network OpenAI was actually waiting for
- 00:06:18 Gemma 4 ships its drafters; Qwen runs at 100 tokens a second
- 00:10:05 The 65 / 20 / 15 routing rule
- 00:12:36 GPT-5.5 Instant becomes the default — and remembers where it heard things
- 00:14:53 Indirect prompt injection becomes a normal Tuesday
- 00:18:02 Two papers, one frame: the agent stack is a system
- 00:20:49 François Fleuret's three-item to-do list
- 00:23:03 Two ways to think about jobs
- 00:26:23 ProgramBench, Nonograph, and the sign-off

---

CANONICAL

<https://braid.opentangle.com/braid/episodes/2026-05-06.html>