

65	62	19	51	40	Nvidia 225 +0.2% Microsoft 517 +3.9% Alphabet 344 +0.4% Meta 750 -3.6% Amazon 250 +0.3% TSMC 452 +0.2% Semis (SOXX) 575 +1.5% Utilities (XLU) 39.33 -0.1%				
ITEMS READ	OUTLETS	SIGNALS	PAPERS	AI STORIES					

AI has industrialized financial scams, with recovery rates near zero

Scam losses hit roughly \$20 billion in 2025 while banks detect only 10% to 15% of losses and recover 2% to 5%, so prevention is the only viable defense. The report flags misaligned liability across banks, social platforms, and telecoms as the reason prevention stays underfunded.

Victim losses were estimated at roughly \$20 billion in 2025, up 26% year over year.

[Read the source](#)

EU regulators: the quantum deadline may beat commercial quantum

European financial regulators warn Q-Day could arrive before quantum computers are commercially useful and want post-quantum cryptography adopted by end of 2026. Harvest-now-decrypt-later attacks make encrypted data stolen today a future liability.

Regulators urge adoption of post-quantum cryptography by the end of 2026.

[Read the source](#)

The Clearing House picks Quant for a 25-bank tokenized deposit network

A settlement network for tokenized deposits across 25 major U.S. banks, connected to existing rails like RTP and CHIPS, is a concrete step toward regulated on-chain money. It reaches participating banks in the first half of 2027, setting a shared standard banks will have to plan around.

The Clearing House processes over \$2 trillion daily.

[Read the source](#)

Snorkel AI's \$3.5B raise shows training data is its own value chain

An eighteenfold revenue surge to \$375M annualized on a shift to hybrid synthetic-plus-expert datasets confirms frontier model development still runs on proprietary, verified data, not public web scraping. Data provenance and quality assurance are emerging as a distinct spend layer in AI infrastructure.

Snorkel raised \$350 million at a \$3.5 billion valuation with \$375 million in annualized revenue.

[Read the signal](#)

Alibaba claims China's most powerful AI chip and a 20GW data center bet

The Zhenwu V900, clusterable into units of 500,000, is a concrete non-US vendor capability claim backed by a 2032 global capacity target. The open question is whether Alibaba can sustain that capex without US supply chains.

Alibaba targets 20 gigawatts of data center capacity by 2032.

[Read the signal](#)

Banks are fighting over small-business bankers, and AI is the hiring weapon

Business banker openings at major competitors jumped 105% as banks chase business owners for cross-selling. BofA is pairing aggressive hiring with AI tools to speed underwriting, and targeting 30% client share, up from 27.6%.

Job openings at major competitors rose 105% over the past year.

[Read the source](#)

RESEARCH · 51 ANALYZED, TOP 8

■■■
IndicBankBench: Evaluating Safety and Reliability of Language Model Assistants in Indian Retail Banking

State-of-the-art LLM assistants complete less than 60% of grounded, multi-turn Indian retail banking tasks reliably (success on all three repeated trials), and at-least-once success overstates dependable performance b...

BEARS ON 1 · RIGOR HIGH

■■■
Where Does Exactly-Once Live? Model, Harness, and Tool-Contract Effects on Duplicate Side Effects in LLM Agents

Whether an LLM agent avoids duplicate side effects on failed writes is determined by the fault type: frontier models handle read-back-resolvable failures well (0.5% duplicates) but duplicate in 56-74% of ambiguous cas...

BEARS ON 2 · RIGOR HIGH

■ ■ ■

Same Bit Width, Different Outcomes: Post-Training Quantization of Text-to-Speech Across Architectures

Post-training quantization of TTS models to the same bit width produces wildly model-specific outcomes because the sensitive component differs unpredictably across architectures, but a staged ablation plus calibration...

BEARS ON 1 · RIGOR HIGH

■ □ □

Human-AI-Powered Hypothesis Testing: Cost-Aware Selective AI Scoring and Sequential Human Escalation

A sequential policy (SCALE) that selectively queries a noisy LLM judge and escalates only some AI-scored items to human review can achieve a statistically valid population-level hypothesis test at near-optimal first-o...

RIGOR HIGH

■ □ □

No More Free Lunch: Corpus Task Complexity Matters as Corpora Grow

Long-context model performance on corpus tasks depends on the task's complexity class: tasks requiring quadratic or more oracle calls (high CTC) degrade far faster with corpus size than linear tasks, and reverse stand...

BEARS ON 1 · RIGOR MEDIUM

■ ■ ■

Sharp Limits for Honest Uncertainty in Hard-Budget Repeated Evaluation

Under a hard evaluation budget, honest confidence intervals for a benchmark score on a fixed grid need replication, and the optimal interval width follows sharp laws: $\lceil [M(t+1)]^{-1/2} \rceil$ when every task must be observed,...

BEARS ON 1 · RIGOR HIGH

■ □ □

When Fancy Eviction Fails: Rethinking Cache Replacement For LLM Prefix Reuse

Sophisticated cache-eviction algorithms provide little to no benefit over simple LRU for LLM prefix caching because session-paced reuse makes recency unusually predictive, and gains should instead come from targeted a...

RIGOR HIGH

■ □ □

AI-Moderated Interviews for Market Research and Digital Twins Calibration

AI-moderated interviews match human-moderated interviews in qualitative depth and cover more themes and customer needs per dollar, but the extra adaptivity of AI moderation adds no quantitative predictive benefit for...

RIGOR HIGH

THE COLUMN

The Week in Money Nobody Can Recover

Let us close the week with an accounting theme: everywhere you look, someone is spending money to defend against a bill that has not arrived yet.

The scam economy has an allocation problem. The sharpest number in the [Thomson Reuters report](#) is not the \$20 billion in losses, it is the 2% to 5% recovery rate. When recovery is functionally zero, fraud stops being a reimbursement expense and becomes an industry, which is what generative AI has made it. The report's diagnosis, misaligned liability among banks, platforms, and telecoms, is really a claim that nobody owns the externalized cost, so nobody invests in preventing it. What would strengthen that reading: any regulatory move assigning liability upstream, say to platforms that host the lures. What would weaken it: banks quietly improving detection well above that 10% to 15% range, which would suggest the market can fix this without a mandate.

Everyone is prepaying future liabilities, including cryptographers. European regulators warn in the [Q-Day advisory](#) that harvest-now-decrypt-later attacks mean data stolen today is already compromised in 2032, so they want post-quantum cryptography adopted by end of 2026. It is the same structure as the fraud story in reverse: pay now to avoid a bill later, with the payment due before anyone can prove the threat is real. Meanwhile the actual plumbing is being rebuilt too. [The Clearing House's tokenized deposit network](#), spanning 25 banks and bridging to RTP and CHIPS by mid-2027, is the rare infrastructure story where the standard arrives before the fight over the standard. That coordination is worth more than the tokenization itself.

The data layer keeps getting paid. [Snorkel's \\$3.5 billion raise](#) is the week's clearest evidence that verified training data is a distinct line item, not a byproduct. Eighteenfold revenue growth to \$375 million annualized tells you enterprises will pay for provenance, which supports the broader view that [training-data IP liability is now a quantifiable balance-sheet cost for AI vendors](#). If verification is a cost center buyers demand, the vendors who cannot document clean data get excluded from procurement, and the industry segments on paperwork. The counterpoint is that synthetic data could eventually make the whole verification layer cheap to substitute; Snorkel's own hybrid model hedges that, which is telling.

One honest wrinkle: all this prepaying assumes the spending produces revenue. The same week brings [Alibaba's Zhenwu V900 and a 20-gigawatt bet](#) against a backdrop where [announced aggregate capex materially exceeds what will be deployed](#). Betting on 2032 capacity is the ultimate pay-now-for-maybe-later trade.

The banks fighting over small-business bankers with AI underwriting tools at least bill someone this year.

WHAT BUILDERS ARE READING

TOOLING & MODELS

Jevmem – automatic project memory for Claude Code, built on Jev

Automatic project memory for Claude Code; pattern for persistent coding-agent context worth piloting internally

HN · 24 points · 9 comments · repo

Classified Estimates Show the NSA Is Paying Billions to Test AI Models

NSA is spending billions testing AI models; signals government model-security bar regulated firms may be measured against
HN · 76 points · 30 comments

U.S. appeals court upholds designation of Anthropic as supply chain risk

Court upholds Anthropic supply-chain-risk designation; may constrain vendor approvals if you build on Claude
HN · 54 points · 11 comments

THINGS HAPPEN · 23

POLICY & REGULATION · 7

OpenAI proposes global AI safety standards and incident tracking to US-China dialogue partners

axios.com · in Atlas

China's Supreme People's Court issues first adjudication guidance for AI disputes, requiring training-data evidence from developers

iam-media.com · in Atlas

OpenAI and Anthropic CEOs to address UN Security Council on AI safety amid model autonomy incidents

cnbc.com · in Atlas

California Executive Order N-9-26 directs study of mandatory AI kill switch and onsite lab auditors

ppc.land · in Atlas

NIST CAISI finds Z.ai's GLM-5.3 is most cyber-capable open-weight model to date, but lags US frontier by about four months

nist.gov · in Atlas

DOJ Statement of Interest Argues Against Mandatory Licensing for AI Training Data

news.bloomberglaw.com · in Atlas

EU Watchdog Makes Digitization A Priority For Regulators - Law360

law360.com · T2

DEALS & MARKETS · 2

Lovable's annualized revenue crosses \$600M as vibe coding takes off

techcrunch.com · T2

ElevenLabs' CEO on margins, IPO timing, and telling customers they're talking to a bot

techcrunch.com · T2

RESEARCH & EVIDENCE · 4

Peer-reviewed review frames AI data center grid integration as requiring multi-layer energy storage coordination

sciencedirect.com · in Atlas

St. Louis Fed study links industry-level AI use to US post-pandemic productivity recovery

frbsf.org · in Atlas

ACM study finds GitHub Copilot efficiency gains do not translate to throughput increases

cacm.acm.org · in Atlas

CFR report details China's persistent critical-mineral leverage post-Busan and gaps in Western de-risking

cfr.org · in Atlas

CHIPS & INFRASTRUCTURE · 2

Korea HBM exports to Malaysia surge fivefold as advanced packaging capacity diversifies

biz.chosun.com · in Atlas

IC-Link joins TSMC 3DFabric Alliance, securing European co-design access to CoWoS capacity

techtimes.com · in Atlas

BANKS & PAYMENTS · 1

Bank of America plans to double AI budget next year

cidive.com · T2

LABS & MODELS · 3

OpenAI initiates early-phase third-party safety evaluations for AI models

bloomberg.com · in Atlas

Anthropic releases Claude Opus 5.5 with 40% lower inference cost and higher agentic reliability

anthropic.com · in Atlas

OpenAI reduces GPT-6 Sol and Luna API costs by 50 percent with improved factuality

techcrunch.com · in Atlas

WORK & LABOR · 3

Bank of America to hire 1,000 more apprentices, invest \$150M in workforce development

foxbusiness.com · T3

AI adoption in Chinese enterprises boosts productivity but raises ...

ilo.org · in Atlas

Blizzard union contract establishes AI usage rules and layoff terms

hrexecutive.com · in Atlas

ELSEWHERE · 1

Decoding the New Premium Traveler - Skift

skift.com · T2

SOURCES · 62 OUTLETS ACROSS 65 ITEMS

finance.yahoo.com 6 · ffnews.com 5 · techcrunch.com 5 · bankingdive.com 4 · federalreserve.gov 2 · americanbanker.com 2 · paymentsdive.com 2 · federalregister.gov 2 · cnbc.com 2 · bloomberg.com 2 · genfinity.io 1 · financefeeds.com 1

Generated Sep 25, 2026 from 65 items the engines stored; every link resolves to a stored record or its source. Market prices as of 16:45 UTC, unofficial feed.