Slashdot

- UK Official Promises Statements 'Around VPNs' and Further Teen Restrictions on Chatbots and Social Media (2026/06/21 20:54)
  PC Gamer reports: The UK government is considering an Australia-style ban on social media for under-16s, with Prime Minister Keir Starmer saying that the ban could take effect as soon as spring next year. As for the much nearer future, Science and Technology Secretary Liz Kendall told BBC Breakfast earlier this week, "We will make further statements in July about VPNs and further restrictions." To be clear, no specific restrictions have yet been announced and Kendall sounded somewhat cautious about an outright ban during a parliament debate that took place the same day. "I have commissioned further research about their usage. There are really important issues to balance here," she says. "Many people want to use VPNs for privacy — that is important — but we know that some children use them to get around restrictions. I will come back to that in July in our response to the consultation." So, we'll have to wait until next month for anything definite, but it's hard not to feel like a full ban on VPNs is already on the table. If that does come to pass, more than the contents of my Bluesky inbox will be at stake. Utah in the US has already tried to implement a full VPN ban (though this was postponed until September after Aylo, the parent company of Pornhub, challenged the law in court)... [T]he UK could just be the next domino after Utah, potentially setting off a chain reaction that affects users around the world. The article also argues that age checks can also be a privacy nightmare "with the security breach that exposed the personal info of 70,000 Discord users last year being one case in point." Here's the complete statement from UK Technology Secretary Kendall. "I'll come back in July with a further statement around VPNs but also additional measures that we want to look at, further restrictions on AI chatbots that parents have found very worrying, more about overnight curfews or breaks in doomscrolling for 16- and 17-year-olds." Read more of this story at Slashdot.
- Cops Keep Getting Arrested for Using Flock's Cameras to Stalk People (2026/06/21 19:40)
  404 Media remembers how a Florida police office looked up his ex-girlfriend's license plate in the Flock automated license plate reader system at least 69 times in 2024 — even searching for her mom's license plate at least 24 times. The police office was charged with stalking and hacking-related offenses, serving one day in prison with five years of probation — but his case "was not a one-off." [Alternate link via Bruce Schneier] Local news reports from around the country repeatedly detail police abusing the Flock surveillance system in order to stalk their partners or ex-partners. The contours of each story are much the same, with the police officer in question using their access to the system to repeatedly track a specific person over the course of weeks or months. The cases highlight the fact that Flock can be used to track the whereabouts of individual people, that police do not get a warrant in order to use the system, and that, if they have access to the system, they have the technical ability to look up any license plate they want for any reason they want. An April study by the civil rights group Institute for Justice found that at least 18 police officers have been caught around the country using Flock to stalk a romantic interest in the last few years; another database, called the ALPR Abuse Library, has documented 20 specific cases of "stalking/targeting" around the country. The known cases of police stalking are almost certainly a vast underreporting of the overall abuse, because they largely include only cases in which the behavior was so egregious that it led to police officers being fired, arrested, or both. Flock told 404 Media that it is "aware of 15 incidents of abuse, each surfaced because of the transparency and accountability features deliberately built into our platform.... There are also 140,000 monthly active users of Flock, so the relatively rare instances of abuse, while obviously wrong and awful, are exactly that — rare," a Flock spokesperson told 404 Media. [One in

10,000.] "Humans are fallible; unlike most tools society provide law enforcement, Flock ensures that in the instances when our technology is misused, the evidence used to hold responsible parties accountable, is right there in our system. We also encourage all our customers to have a usage policy, regular training, and to implement our Audit Assistance tool, which proactively flags unintended use...." But it is also the case that Flock has strenuously fought against lawsuits and potential regulations that are seeking to require police to get a warrant to use the system. And many cases of abuse have not been detected by police departments themselves but by those private citizens, journalists, and stalking victims who have found patterns of abuse in public records files they have obtained from their local police departments. In most cases of Flock-related stalking reviewed by 404 Media, the abuse occurred over the course of months or years, and the victims were subjected to dozens or hundreds of lookups. Other abuse cases have been discovered using the website HaveIBeenFlocked.com, a website that compiles Flock searches released via public records requests and turns them into a searchable database. Flock has repeatedly tried to get that website taken down, as we have previously reported. Read more of this story at Slashdot.

- After Six Years Of Work and Over 360 Patches, Linux 7.2 Finally Removes Bug-Prone strncpy (2026/06/21 18:12)
  Tech Times reports: Linux 7.2's merge window closed out a cleanup campaign on Friday that most kernel developers had stopped expecting to see end: the complete removal of strncpy(), a C string-copy function that the kernel's own documentation labels "actively dangerous," from every subsystem, driver, and architecture-specific file in the kernel source tree. The merge landed June 20, 2026. After around 362 commits spread across six years of incremental work, no call site using the function remained, and the function itself — including the last per-CPU-architecture optimized implementations — was struck from the source. The removal matters beyond housekeeping. strncpy() is a persistent source of a specific class of memory error: kernel buffers that contain sensitive data can leak bytes past an unterminated string boundary, a pattern that enables memory disclosure vulnerabilities. Eliminating the function from the tree removes that entire class from the kernel's attack surface — and, critically, makes strncpy() unavailable to any future contributor, turning a best-practice suggestion into an enforced policy. Phoronix notes it's replaced by five different functions: In place of strncpy, Linux kernel code should use strscpy() for NUL terminated destinations, strscpy_pad() for NUl-terminated destinations with zero-padding, strtomem_pad() for non-NUL-terminated fixed-width fields, memcpy_and_pad() for bounded copies with explicit padding, or memcpy() for known-length memory copies. "The reason five functions were needed," explains Tech Times, "is that different parts of the kernel were using strncpy() for five semantically distinct memory operations — each with a different intent, different termination requirement, and different padding behavior. " The original function obscured all of those differences under a single ambiguous name. The 362-commit campaign to replace it was, in effect, a codebase-wide audit that forced every call site to declare its actual intent in code That is an engineering outcome with lasting value: the kernel's string-handling semantics are now explicit where they were previously implicit, and future maintainers can read a function name and understand what a copy operation actually does. Read more of this story at Slashdot.
- US Bill Would Mandate AI Chip Location Tracking to Thwart China and Other Adversaries (2026/06/21 16:34)
  NBC News reports: A group of companies that specialize in tracking international shipments of sensitive technologies is backing a Capitol Hill bill that would require America's most powerful AI chips to incorporate stronger security mechanisms aimed at preventing the chips from reaching China and other adversaries. The letter, signed by six companies, says the Chip Security Act (CSA) would increase American chip companies' competitiveness and close key loopholes in the U.S. export control regime. The move clashes with claims from semiconductor lobbying groups that the requirements would constrain America's booming chip industry. Sent to congressional leadership Thursday morning and seen by NBC News, the dispatch instead argues that more robust security verification would assure chip customers and manufacturers that they are abiding

by sensitive restrictions on chip sales. The companies argue that the boosted confidence will "lead to increased sales, faster export approvals, larger transactions, greater access to new markets, and more expansive chip deals." Despite U.S. export control laws banning sales of advanced AI chips to certain countries, including China, loopholes in current requirements have allowed billions of dollars' worth of America's best AI chips to be sold to entities in third-party countries that can then forward them to China. In just one case in March, the Justice Department charged three people with conspiring to forward $2.5 billion of AI chips to China. The CSA aims to address those loopholes, mandating that chip exporters better track where advanced chips are sent, via either bespoke location-verification hardware or software that can run on existing hardware. That, bill proponents claim, would ensure that sensitive chips could be sold to countries like Malaysia or Indonesia without fear of further transfer to China... Experts say that because chips perform the advanced computations required for frontier AI systems, cutting off access to the chips is crucial to prevent geopolitical rivals from using AI systems for military or economic purposes. Read more of this story at Slashdot.

- The Rust Ecosystem Gets an AI Security Engineer in Residence (2026/06/21 15:34)
  While the Rust Foundation has a Security Initiative to protect its ecosystem, "the threats have expanded," they announced this week, "and so has the kind of help maintainers need." Much of this comes back to a single shift: Automated tooling (much of it now built on large language models) has gotten good enough to surface real vulnerabilities in open source code quickly and at scale. That is useful, and several large Rust projects have already received and fixed credible issues found this way. The same tooling has also made it trivial to generate vulnerability reports that look plausible and are worthless. Maintainers across the ecosystem are losing real hours sorting these from the reports that matter, and the noise tends to bury the signal. So, with funding from the Alpha-Omega Project, the Rust Foundation is bringing on a full-time AI Security Engineer in Residence dedicated to the Rust ecosystem. This position is being funded with part of the $12.5M in open source security funding that the Linux Foundation announced in March. The role exists to take pressure off maintainers. The person in this position will use a mix of human-led and AI-assisted methods to proactively review Rust itself and the crates the ecosystem leans on most and help us separate real, exploitable issues from false positives and low-signal noise before anything reaches a maintainer... This role will run full-time for six months to start, with room to extend depending on what we learn and the funding available. Methods, playbooks, and prompts will be documented so the work doesn't end with the contract. We are grateful that Rust is not embarking on this work in isolation. Several other ecosystems have received parallel Alpha-Omega grants for the same kind of work (e.g., the PHP Foundation and the Drupal Association) and we plan to share tooling, triage practices, and what we learn rather than duplicating work A statement from Rust's new AI Security Engineer in Residence acknowledges that "One of our next challenges is the wave of bugs discovered by the next generation of AI-powered developer tools." Read more of this story at Slashdot.
- Canonical's Upcoming AI Tool: Talk to Ubuntu Instead of Typing (2026/06/21 14:34)
  This week the Ubuntu desktop's director of engineering announced they're bringing speech-to-text dictation to Ubuntu Desktop, aiming for an experience "that feels like a natural part of the desktop while respecting user privacy and running entirely on local hardware." "Speech recognition has become a common feature on modern platforms, and we think it should be a first-class experience on Ubuntu Desktop as well." More details from the blog It's FOSS: For Ubuntu 26.10, the initial version of Myna is expected to be a desktop dictation tool built around GNOME on Wayland with a push-to-talk mechanism gatekeeping when your microphone accepts input. Using it means holding a hotkey, speaking, and letting go. A small activity indicator shows while it is listening, and the transcribed text lands wherever the cursor was sitting when dictation started. Recognition itself happens inside a sandboxed component called the Canonical Inference Snap, while a Speech Orchestrator manages the session and an Audio Adapter handles whatever the microphone picks up, denoising and chunking it before it ever reaches the model...

Speech recognition will happen locally, and an internet connection is not needed once the appropriate model is installed... The audio data won't be sticking around either, being stored in a small in-memory buffer that gets discarded the moment the session ends. Features like dictation into password fields, wake words, continuous listening, voice assistants, voice commands, translation, speaker identification, and automatic language detection are all off the table... You should also know that Canonical is looking for feedback before the specs for Myna are finalized, especially from people who already rely on dictation or assistive tools on Linux. Read more of this story at Slashdot.

- New Super PAC Aims to Rally Tech Workers to Help Limit AI: 'the Guardrails Alliance' (2026/06/21 11:34)
  "A grassroots movement is forming among everyday tech workers who are demanding their companies develop and deploy AI responsibly," reports TechCrunch. Hoping to leverage that discontent is a new super PAC called the Guardrails Alliance. The New York Times reports that it launched Thursday with backers that included tech employees and labor unions: Guardrails positions itself as a populist political movement that runs on small donations from people in the trenches of the AI boom. The PAC has about $5 million at its disposal today and planGuardrails will buy ads to support Alex Bores, a New York congressional candidate who became Leading the Future's first target and is running in the primaries next week. s to raise $15 million this cycle — small potatoes compared to deep-pocketed adversaries like Leading the Future, which has more than $100 million from tech leaders like OpenAI president Greg Brockman... "This is not about matching [Leading the Future] dollar for dollar," [said the super PAC's co-founder, political operative Shaunna Thomas]. "What this vehicle is meant to do is be a political home for people who are concerned about the way the anti-regulation AI tech sector is trying to manipulate elections." Meanwhile a former Netflix and Warner Bros. executive has launched the Alliance for Responsible Innovation in the Arts & Media, reports Variety, calling it an AI-focused content coalition that says it's dedicated to supporting "responsible and sustainable AI innovation and the importance of human creativity." The initial members of the coalition, announced Monday, include Disney, the New York Times, Adobe, Condé Nast, the Financial Times, ITV, Advance, BBC, Cambridge University Press & Assessment, U.K. publisher Reach and Wiley. Many of the coalition's members have either struck deals with AI companies or are developing their own AI tools... The group plans to argue for legal and policy guardrails around AI's usage, with its funding directed towards analyses, tools and services focused on advancing those initiatives... One of the group's launch advisers is Damian Collins, OBE, who previously served as the U.K. Parliamentary Under-Secretary of State in the Department for Science, Innovation and Technology under prime ministers Boris Johnson and Liz Truss. "Using AI to break the law can never be an acceptable excuse," he said in a statement. "Laws around personal safety, intellectual property and financial crime still apply in the age of AI. This is why ARIAM has been created and why I'm proud to working with this necessary initiative." Read more of this story at Slashdot.
- Facial Recognition on Public Buses? Kansas City Says Yes (2026/06/21 07:34)
  An anonymous reader shared this report from the Associated Press: Officials in Kansas City, Missouri, are preparing to equip cameras on some public buses with facial recognition software capable of identifying passengers who appear on a list of banned riders or missing persons. Supporters and opponents alike view the effort as a major litmus test for tapping the AI-powered software on a U.S. public transportation system, positioning Kansas City as the latest epicenter of a fierce debate over whether the safety benefits of artificial intelligence are worth the privacy costs. "The idea of running face recognition on a camera that is pointed on live spaces in public is a line that until recently has never really been crossed in the last 25 years," said Jay Stanley, senior policy analyst for the Project on Speech, Privacy and Technology at the American Civil Liberties Union. The state of Missouri declined to help fund the project as expected due to concerns with the facial recognition component. Still, the city is pushing ahead with local and federal money, said Tyler Means, chief mobility and strategy officer at the Kansas City Transportation

Authority. "Privacy is always a tricky thing," Means said. "We've always had cameras on our buses. It's just new technology. I think in time it'll smooth over and people will realize, 'Well, it didn't really feel any different'...." Images captured by cameras aboard the buses would immediately be checked against any active alerts, generated when a missing person, banned rider or someone on a law enforcement watch list designated by the transportation authority is identified... After the buses return to the depot, the transportation authority would archive the regular video footage on a local server for up to five years. The company partnering with Kansas City to run the cameras "started using live facial recognition years ago to alert nursing homes when residents left the building," according to the article, and then "brought the technology to correctional institutions and schools." But this is its first attempt at bringing its cameras onto public transportation. The article also includes this quote from Will Owen, communications director for the Surveillance Technology Oversight Project. "City residents should not be guinea pigs for transit systems to test Silicon Valley's latest unproven, biased surveillance tech." Read more of this story at Slashdot.

- Polymarket Paid Dozens to Post Videos of Themselves 'Winning' With Fake Bets (2026/06/21 04:34)
  In January a college student posted a video showing him winning $100,000 on Polymarket — one of 145 that appeared to show bets adding up to almost $410,000, reports the Wall Street Journal. "But none of those bets were real." Instead its creator was "one of dozens of mostly college-age creators Polymarket paid to film themselves making fake trades and sometimes scoring fake wins," the Journal reports, citing interviews with the creators an an analysis of more than 1,100 of their videos: Polymarket built near-perfect copies of its website, then instructed creators to make simulated trades on those dummy sites and hide that they were being paid by Polymarket. To get the videos to go viral, Polymarket has recruited a social-media army to copy and re-post creators' footage. Though the New York-based company has been banned from offering its primary crypto platform in the U.S. since 2022, the social-media creators are paid to specifically target U.S. users, who can still access the site with a virtual private network... Polymarket hired and worked closely with a marketing contractor to promote the site. In a message reviewed by the Journal, that contractor told its social-media army to repost content made by 10 Polymarket creators in particular... These creators didn't initially identify themselves as paid by Polymarket, although one offered a $20 bonus code in his social-media bio... The company instructed creators not to disclose they are paid, according to creators who have worked with the company. They said the pay often added up to $2,000 to $3,000 a month... A handful of videos the Journal reviewed also contained short glimpses of URLs indicating the sites were test environments for Polymarket engineers... Creators said they send the finished videos to Polymarket for review. If a video isn't engaging enough, or if it bears obvious signs of being faked, Polymarket will ask for the videos to be reshot, the creators said... Polymarket sends creators bullet-point guidance on what to say, according to creators who have worked with the company and a recruiting website... Polymarket's viral clipping campaign racked up more than 140 million views on TikTok, YouTube and Instagram, according to the analytics provider Tubular... Internal materials show that Polymarket and Virality promote videos showing how easy it is to conduct insider trades on the platform. Polymarket has paid clippers to promote at least 19 videos discussing opportunities to use inside information or other tactics to manipulate markets. America's advertising laws "require people who are paid to endorse a product to disclose their ties," the article notes, "although there is some gray area about what's permitted." (After the Journal's investigation, the creators started adding "@polymarket partner" to their bios, the article points out._ And when asked for a comment, Polymarket "said it plans to conduct a comprehensive audit of active promotional content." Read more of this story at Slashdot.
- Gamers Sue PlayStation: It's Not Clear They're Selling Licenses Rather Than Ownership of Games (2026/06/21 01:34)
  The gaming news site Aftermath reports: Four gamers are suing Sony Interactive Entertainment for allegedly breaking a California law that requires digital storefronts selling games to make it clear people are buying licenses, not actually owning the games. Sony Interactive

Entertainment's PlayStation store uses language like "Buy Now" and "Confirm Purchase," lawyers wrote in a complaint filed on Thursday... "In reality, consumers who 'purchase' digital games through PlayStation do not obtain ownership of those products," lawyers wrote. "Instead, PlayStation grants only a limited, revocable license to access the software, subject to multiple restrictions contained in a separate Software Product License Agreement".... [T]he PlayStation store does have a disclosure. Above the "Confirm Purchase" button, there's a note: "By selecting [Confirm Purchase], you agree to complete the purchase in accordance with the PlayStation Terms of Service before using this content. You further acknowledge that your purchase of this digital product amounts to a license subject to the Software Product License Agreement." These four gamers aren't satisfied with that; they said in the complaint that it's too small, and that "a reasonable customer completing a purchase would not necessarily notice this disclosure." "It's a proposed class action complaint, meaning the group of four gamers is asking a judge to grant them class action status." Read more of this story at Slashdot.

- How Millions of Digital Home Devices Are Secretly Powering Cyberattacks (2026/06/20 23:23)
  The Wall Street Journal reports on internet-connected devices — and how every year millions of them "can contain a secret digital backdoor that opens up access to your home internet, so that anyone... can surf the web as if they were you." (And this is especially true for "knockoffs that you buy online"...) In a video report this week they tested two digital picture frames from Amazon and three streaming devices from Walmart "because we heard that they often ship with backdoor software used in cyberattacks. Security experts believe manufacturers are being paid to add this malware, but many people also get tricked into downloading the software onto their phones or computers... Within minutes of turning the devices on, there was a surge of internet traffic... Visits to gambling, porn, cryptocurrency and loads of other sketchy web sites started pouring in from users around the world." (And remote visitors also tried to access Outlook and Gmail accounts...) Residential proxy companies even rent out access to "tens of millions of home networks around the world," according to the report. "But the problem is actually worse than that. Hackers figured out a way to seize control of these backdoors, and they started taking over these residential networks. Last month authorities arrested a 23-year-old Ottawa man, saying he'd taken control of more than a million devices to launch some of the largest cyberattacks anyone had ever seen.." After a couple months the Journal's reporter collected logs of all the traffic, and sent it to an investigator at Comcast, who said both were conducting DDoS attacks. But estimate for the number of infected devices are as low as tens of millions or as high 500 million-plus. "We've seen nation state attacks launched through these kind of endpoints, which means your device sitting in your house is part of a nation state attack against another nation state... We've seen ad fraud, we've seen ticket scalping, we've seen financial fraud." But more importantly, "We have seen some of the largest computer attacks — meaning computers attacking other computers at human request — ever recorded in our digital history in the last several months." At cybersecurity conferences, some are warning "there are much larger ones on the horizon if we don't get a hold of this problem." The company making the picture frame "couldn't be reached for comment," while Amazon said it's been out of stock since last year. Both Amazon and Walmart said they take action when they confirm malware on a third-party product. Read more of this story at Slashdot.
- OpenAI Announces Benchmarks for AI Life Sciences Research. Its Best Model Failed 63.9% of the Test (2026/06/20 21:34)
  This week OpenAI announced a 750-task test to to measure "whether AI systems can support realistic life science research tasks, not just answer biology questions." But while OpenAI's top-performing GPT-Rosalind model led the rankings, Slashdot reader BrianFagioli notes that "it achieved a pass rate of just 36.1 percent, failing nearly two-thirds of benchmark tasks." Nerds.xyz points out that means "the best-performing model failed nearly two-thirds of the benchmark's tasks." The benchmark also revealed a familiar weakness. AI systems generally perform better when

everything is presented as text. Once they are forced to work with supporting documents, figures, or complex datasets, performance drops noticeably. GPT-Rosalind's pass rate fell from 45.1 percent on text-only tasks to 28.1 percent on tasks involving artifacts or URLs. To be fair, the benchmark is not intended to suggest AI is useless in research. Quite the opposite. OpenAI found that models are becoming increasingly capable of scientific communication, evidence synthesis, and translating research findings into practical explanations. Those are valuable skills, particularly for researchers drowning in information. But LifeSciBench serves as a useful reminder that today's AI systems are still far from autonomous scientists. They can help. They can assist. They can sometimes provide surprisingly useful insights. What they cannot reliably do, however, is replace the expertise, judgment, and skepticism that real scientific research requires. Read more of this story at Slashdot.

- Remembering When Alan Turing Developed a Portable Voice Encryption Device (2026/06/20 20:34)
  Long-time Slashdot reader smooth wombat writes: Alan Turing, one of the more famous people who worked at Bletchley Park to decipher the German Enigma coding machine, was also working on a separate project. His private papers, known as the Bayley papers for his assistant Donald Bayley who held onto the papers until his death in 2020, reveal Turning had produced a working model of a portable voice encryption device. He even demonstrated it by using a Winston Churchill speech recording. "Weighing just 39 kg, including its power pack," Jack Copeland wrote in an article for IEEE Spectrum, "Delilah would be at home in a truck, a trench, or a large backpack." More from Popular Mechanics: Turingâ(TM)s work at Bletchley Park actually informed the Delilah experimentation he was doing at Hanslope Park, and not just because he used Red Forms, the Army-issue sheets Hanslope staffers were meant to use to alert Bletchley staffers to enemy signals, as his personal scrap paper for Delilah experiments. He drew inspiration from one of the German cipher machines they had decoded at Bletchley; not the famed Enigma machine, but rather the SZ42. While the former relied on Morse Code, the latter utilized a 5-bit telegraph code, which Copeland notes âoewas a forerunner of ASCII and Unicode and is still used by some ham radio operators.â The SZ42 produced an obscuring key of telegraph characters, with an identical key produced to both the sender and receiver. If it could be done for text, Turing reasoned it could be done for sound as well... [T]he reason Delilah fell to the wayside of history isnâ(TM)t because it was a failure, but rather because it simply wasnâ(TM)t needed anymore. By the time Turing had built and demonstrated his device, the war was over. What good was a portable voice encryptor if you had no major enemies trying to intercept your calls, the government reasoned. So funding for the project stopped, and Turingâ(TM)s two-year experiment ended with a whimper. Turingâ(TM)s time as an electrical engineer at Hanslope Park became a footnote in his story, if even that. Read more of this story at Slashdot.
- Tech Pundit Cringely Co-Founds Startup '2Brains Inc' to Solve LLM Hallucinations (2026/06/20 19:34)
  Long-time tech pundit Robert Cringely started his career at the Stanford Artificial Intelligence Lab back in 1978. Last month 73-year-old Cringely explained why his site went on a two-year hiatus — and it's not just because of a heart attack and a stroke last July: Just like everyone else, I've been busy all this time on Artificial Intelligence, founding with two partners a company called 2Brains... The work we were doing together is unfinished, but it's not stopped. The patents are filed, the architecture is documented, and the small team continuing the work includes me. Cringely's first piece made the cast that "the trillion-dollar bet the AI industry is making right now may be wrong, and that there's an architectural alternative we've patented and built." In Machines of Loving Grace, Amodei made the case that scaling compute would eventually solve essentially every hard problem in artificial intelligence. Buried in that optimism — or maybe not buried, maybe right out in the open — was a quiet absolution. Hallucinations, the embarrassing tendency of these systems to state falsehoods with total confidence, would take care of themselves. Make the models big enough, train them long enough, and the problem dissolves. You don't have to solve it. You just have to wait, and spend. And so the entire AI industry breathed a sigh of relief. I have spent forty years watching this industry, and I know a permission slip

when I see one. Because that is what the essay became, whatever Amodei intended. It gave every other person writing nine- and ten-figure checks a reason not to worry about the one thing that should worry them most. The hallucination problem is the difference between a clever toy and a system a hospital or a bank or a court can actually rely on. It is the whole ballgame for enterprise AI. And the prevailing wisdom, blessed from the top, is that you needn't address it directly. Scale will provide... A small company I helped start, 2Brains Inc., set out in 2022 to solve hallucinations — before ChatGPT, before the scaling consensus hardened into received truth, back when the polite assumption was that the problem was simply insurmountable. We did not solve it by waiting for bigger models. We solved it architecturally, by separating the part of the system that generates language from the part that retrieves and verifies facts, and reconciling the two before anything reaches the user. It runs on ordinary processors. It is cheap. And on the industry's own benchmark for this kind of faithfulness, it more than doubles the published baseline, with no fabricated facts in the verified case at all. The article asks whether scaling will, at tremendous cost, eventually reduce hallucinations — or even worse, if the largest companies in the world "are spending a fortune chasing a cure that is not coming." And last week Cringely pitched more advantages for their solution, noting that most prompts aren't even chatbot-level creative prompts — but just requests to retrieve simple data: The reason 2Brains doesn't lie and the reason it's cheap are the same reason. It looks the fact up instead of guessing it — so it cannot fabricate, and the lookup runs on a processor that sips power instead of a chip that gulps it. Trust and thrift are not a trade-off you balance against each other. They fall out of a single design decision. You do not pay extra for the honest version. The honest version is the cheap version. That sentence is the whole company. Read more of this story at Slashdot.

- Waymo Recalls About 3,900 Robotaxis After Some Drove Into 'Freeway Construction Zones' (2026/06/20 18:34)
  CNBC reports: Waymo is recalling almost 3,900 robotaxis in the U.S. to fix software issues after some cars drove into freeway construction zones, according to notices filed with the National Highway Traffic Safety Administration. The voluntary recall, the Alphabet-owned company's second in just over a month, followed 13 known incidents where Waymo robotaxis drove into construction zones on freeways in Phoenix, or entered freeway lanes with active construction in the San Francisco area, the filings published Thursday said... A letter posted to the regulator's website... noted that, "Driving through a closed construction zone increases the risk of a crash..." [Waymo said in a statement emailed to CNBC] "We voluntarily restricted freeway operations last month while making improvements, proactively notified state and federal regulators, and decided to file a voluntary software recall with NHTSA. We continue to safely serve riders on surface streets in all the cities where we operate...." The company implemented another voluntary recall in May after some of its robotaxis had driven into flooded zones or standing water. The NHTSA Safety Board also initiated a probe of Waymo after a January incident in which a robotaxi illegally passed a stopped school bus. Read more of this story at Slashdot.