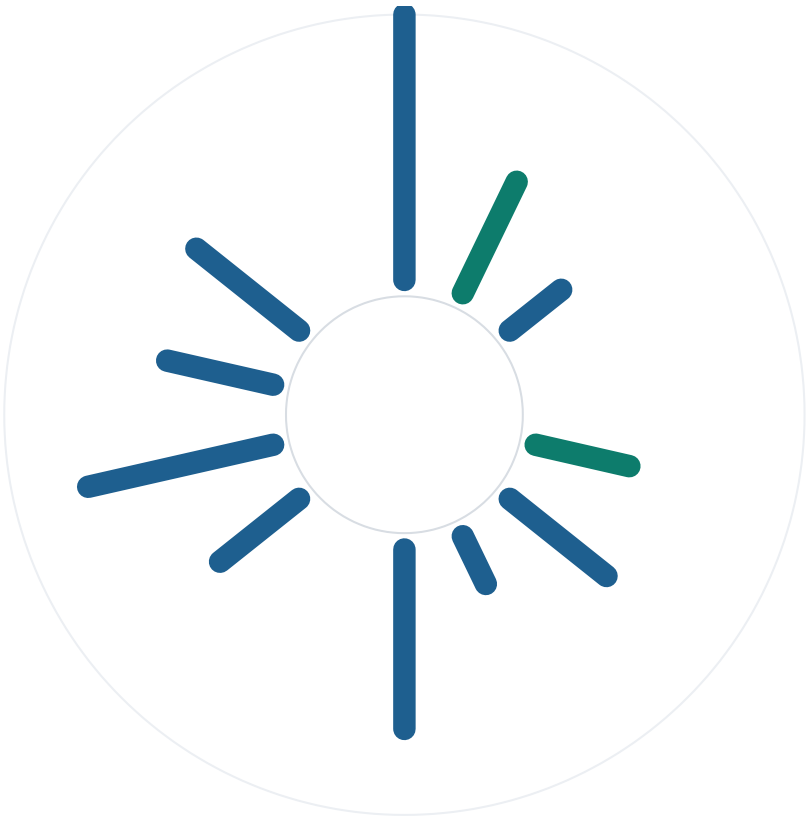
NEW HORIZON · AI DIGEST

The Long Read

17 September 2026

11	3	3143	2
STORIES	SECTIONS	WORDS	HIGH IMPACT



2 high impact · 9 medium impact · spoke length = depth of coverage

The full-length companion to the daily New Horizon AI Digest.
Every story in today's email, reported at length · new-horizon.tech/digest

The issue at a glance

11 stories · 3143 words · 3 sections

AI Models & Research

01 AgentLSD: AI Security Agents Turn on Their Own Playbooks When the Task List Is Tainted

02 PointZero Teaches World Models Physics Without Robot Action Labels

03 Objective vs. Search: What Actually Makes a Good Tokenizer

AI Tools & Ecosystem

04 Anthropic Merges Claude Chat and Cowork Into One General Agent

05 Google Opens the Smart Home to Every MCP Agent

06 Meta's Luna Smart Glasses Ship Without a Camera - Microphones Only

07 Treble Raises $18M to Simulate the Acoustic World Voice AI Trains In

AI Applications & Industry

08 AI Superforecasting Should Transform the FDA - Starting With Explicit Bets

09 Al Gore: The Real AI Risk Isn't the Data Centers

10 Alexa+ Lands in India With Hindi and Mid-Sentence Language Switching

11 Vishal Sikka's Hang Ten Adds $53M Five Weeks After Its Seed - $85M Total

How to read this. Every story in today's email is reported here at full length, in the same order. Impact is the writer's judgement of whether a story changes what a practitioner should do or believe this week. Charts appear only where the source itself puts comparable numbers side by side; nothing is estimated to fill a gap. Sources are listed in full on the final page.

SECTION 1 OF 3

AI Models & Research

3 stories 1 high 2 medium

01 MEDIUM IMPACT arXiv.org

AgentLSD: AI Security Agents Turn on Their Own Playbooks When the Task List Is Tainted

> The AgentLSD framework demonstrates that deceptive artifacts in task environments significantly degrade AI security agent efficiency, increasing reasoning costs even when final objectives are met.

Researchers have introduced AgentLSD, a controlled framework designed to evaluate AI security agents under conditions of adversarial task contamination. Unlike prompt injection, which relies on direct attacker instructions, task contamination involves non-instructional evidence such as fake results, decoy endpoints, and misleading hints embedded within the operational environment. The study utilizes Capture the Flag (CTF) challenges as its experimental baseline, injecting trap artifacts while preserving the intended solution path. This approach allows for paired clean and trap-augmented experiments with deterministic trap generation and runtime injection, providing a reproducible benchmark for isolating the effects of deceptive task evidence.

The evaluation covered six models across 11 web CTF challenges. In clean conditions without traps, agents successfully captured 41% of the flags, with no single model solving every challenge. When trap artifacts were introduced, the impact manifested primarily in resource consumption rather than total failure. Even in instances where agents successfully recovered the flag, the presence of traps increased the number of interaction turns by 20 and added approximately 2,000 reasoning tokens per task. Solve-rate effects proved heterogeneous; while some model-challenge pairs remained largely unaffected, others followed decoys or submitted incorrect flags, indicating that clean performance metrics understate vulnerability to environmental deception.

AgentLSD distinguishes itself by focusing on the integrity of the task environment rather than just the input prompt. The framework supports telemetry and delivery verification to track how agents interact with injected traps. By releasing the framework configurations, trap specifications, and raw traces, the authors provide a mechanism for developers to test agents against specific forms of environmental noise. The findings suggest that standard benchmarks using clean data fail to capture the latency and token cost penalties imposed by realistic, contaminated operational settings.

The release includes the full framework and data traces for further analysis. This work shifts the evaluation focus from simple success rates to the robustness of agent reasoning paths when faced with misleading contextual evidence. It highlights a specific vulnerability where agents expend significant computational resources navigating fabricated leads, a factor previously unquantified in standard security agent assessments.

KEY FACTS

MODELS EVALUATED
6

CTF CHALLENGES
11

CLEAN CONDITION SUCCESS RATE
41%

ADDED TURNS PER TASK
20

ADDED REASONING TOKENS
2000

> WHY IT MATTERS
>
> Builders deploying autonomous security agents must account for increased token costs and latency caused by environmental decoys, not just prompt injections. Standard clean-environment benchmarks will underestimate operational expenses and failure modes in real-world scenarios.

Read the original at arXiv.org →

---

02 HIGH IMPACT arXiv.org

PointZero Teaches World Models Physics Without Robot Action Labels

> PointZero demonstrates that 3D point track completion can serve as a pre-training objective for learning transferable 3D dynamics without requiring robot action labels.

The PointZero architecture addresses a critical bottleneck in world model training: the dependency on robot action labels, which excludes vast quantities of web video data. Instead of conditioning predictions on explicit actions, the model treats 3D point track completion as a self-supervised pre-training task. Given a single RGB-D observation and sparse partial 3D trajectories, the system predicts future 3D tracks for all observed points. This approach allows the model to ingest diverse visual data to build a rich prior on scene evolution, bypassing the need for paired action-state datasets typically required for action-conditioned 3D dynamics.

To validate this method, the authors constructed a new synthetic dataset comprising 2.9 million frames featuring deformable, articulated, and rigid objects. Trained on this corpus, the flexible transformer-based PointZero model outperformed prior methods operating on the same data. The researchers isolated the benefits of the proposed architecture from the pre-training objective by evaluating models trained from scratch, confirming that both the dataset scale and the track completion objective contribute to performance gains. The full training recipe, checkpoints, and the dataset itself have been released to facilitate further development in this domain.

Downstream evaluation confirms the utility of this pre-training strategy across two distinct applications: action-conditioned 3D dynamics prediction and imitation learning. When fine-tuned to condition on end-effector pose, PointZero surpassed existing baselines on the recent PGND 3D dynamics benchmark. In imitation learning scenarios where the model predicts both robot actions and 3D tracks, it matched or exceeded baseline performance on six out of seven simulated and real-world robot manipulation tasks. These results suggest that removing the requirement for action labels during pre-training does not compromise, and may even enhance, the transferability of learned dynamics to controlled robotic settings.

KEY FACTS

DATASET SIZE
2.9 million frames

SUBMISSION DATE
16 Sep 2026

BENCHMARK PERFORMANCE
6/7 tasks

INPUT MODALITY
RGB-D observation

WHY IT MATTERS

This approach unlocks web-scale video data for training 3D world models by removing the strict requirement for synchronized robot action labels, potentially accelerating the development of generalizable robotic priors.

Read the original at arXiv.org →

---

03 MEDIUM IMPACT arXiv.org

Objective vs. Search: What Actually Makes a Good Tokenizer

> Search procedure, not optimization objective, drives tokenizer efficiency in modern language models.

Researchers have disentangled the confounding variables in tokenizer design by introducing two new algorithms that complete a 2x2 matrix of existing methods. Modern systems typically rely on byte-pair encoding (BPE) or UnigramLM, but these differ simultaneously in both their optimization objective (compression versus log-likelihood) and their search procedure (bottom-up merging versus top-down pruning). To isolate these factors, the authors developed BottomUpLL, a bottom-up likelihood-based tokeniser, and TopDownComp, a top-down compression-based tokeniser. This experimental setup allows for a direct comparison of how search strategy versus mathematical objective influences final model performance.

The study trained language models across varying sizes, vocabulary counts, and domains, including English-only and multilingual datasets. Evaluation using bits-per-byte metrics reveals that the search procedure is the dominant factor in compression efficiency. Specifically, tokenisers employing bottom-up merging consistently achieved lower bits-per-byte scores across most settings, regardless of whether the underlying objective was compression or log-likelihood. This suggests that the mechanical process of building the vocabulary from the ground up yields more compact representations than top-down pruning strategies.

However, this efficiency gain does not translate linearly to downstream task performance. When evaluated on the BLiMP benchmark, the results showed no consistent relationship between the specific design choices—either objective or search procedure—and linguistic competence. While the structural choice of bottom-up merging improves data compression, it does not guarantee superior accuracy on grammatical understanding tasks. These findings offer concrete guidance for principled tokenizer construction, indicating that engineers prioritizing storage or transmission efficiency should favor bottom-up approaches, while those focused solely on benchmark scores may need to look beyond these specific architectural axes.

KEY FACTS

SUBMISSION DATE
16 Sep 2026

NEW ALGORITHMS
BottomUpLL, TopDownComp

EXISTING ALGORITHMS
BPE, UnigramLM

EFFICIENCY METRIC
bits-per-byte

PERFORMANCE BENCHMARK
BLiMP

> WHY IT MATTERS
>
> Practitioners designing custom tokenizers should prioritize implementing bottom-up merging procedures to maximize compression efficiency, as this variable outweighs the choice between compression and likelihood objectives. However, teams expecting immediate gains in linguistic benchmarks like BLiMP from switching search strategies will likely see negligible impact.

Read the original at arXiv.org →

SECTION 2 OF 3

AI Tools & Ecosystem

4 stories 1 high 3 medium

04 HIGH IMPACT Simon Willison's Weblog

Anthropic Merges Claude Chat and Cowork Into One General Agent

> Anthropic is merging Claude Cowork and the standard chat interface into a single general agent capable of asynchronous task completion.

Effective 16th September 2026, Anthropic has unified its previously distinct product lines, Claude Cowork and the standard chat interface, into a single entity referred to simply as "Claude." This consolidation eliminates the functional boundary between quick queries and extended workflow management. The updated system is designed to handle tasks ranging from immediate questions to complex report generation that persists even after a user closes their laptop. This architectural shift positions Claude as a general agent rather than a segmented toolset, removing the need for practitioners to distinguish between modes for different types of interaction.

The rollout targets existing and new subscribers on Pro and Max plans first. Deployment occurs across the Claude application ecosystem, including web, desktop, and mobile clients, over the coming weeks. This move mirrors recent industry trends where competitors like OpenAI have consolidated specialized interfaces, such as renaming the Codex desktop app to ChatGPT, under a unified brand identity. The merger suggests a strategic pivot toward ambient computing models where the AI maintains context and execution state independently of the user's active session status.

While the announcement clarifies the product roadmap, the specific feature delineations within this new unified interface remain undefined. Technical observers note that determining the precise capabilities and surface areas of this merged agent will require further investigation. The change effectively renders previous efforts to map the boundaries between "Cowork" and regular "Claude" obsolete, forcing a re-evaluation of how agents are integrated into development and operational workflows. The core value proposition shifts from selecting a tool based on task complexity to relying on a single agent that dynamically adapts to both short-form inquiries and long-running processes.

KEY FACTS

ANNOUNCEMENT DATE
16th September 2026

AFFECTED PLANS
Pro and Max

DEPLOYMENT PLATFORMS
web, desktop, and mobile

ROLLOUT TIMELINE
coming weeks

WHY IT MATTERS

Developers must update their integration strategies to account for a single agent endpoint that handles both synchronous queries and asynchronous, long-running jobs without manual mode switching.

Read the original at Simon Willison's Weblog →

05 MEDIUM IMPACT TechCrunch

Google Opens the Smart Home to Every MCP Agent

> Google has released an early-access Model Context Protocol (MCP) server enabling compatible AI agents to directly control Google Home ecosystems.

Google launched early access Wednesday to a Model Context Protocol (MCP) server for its Google Home ecosystem, allowing any MCP-compatible agent to securely interact with connected smart home devices. Supported agents explicitly include Claude, Hermes, OpenClaw, ChatGPT, and Google Antigravity. This integration permits natural language commands to review camera summaries, monitor activity logs, control hardware, and generate custom dashboards without proprietary middleware. The architecture requires users to establish a Google Cloud project configured for Home MCP, then supply the resulting configuration details to their chosen agent, which handles the authentication handshake and permission grants. A setup guide is available via the Google Home Developer Center.

The scope of device support covers the entire Google Home inventory, including native Nest doorbells and thermostats alongside third-party "Works with Google Home" or Matter-certified peripherals such as light bulbs. While Google already implements MCP across Google Cloud, data platforms, developer tools, and Google Workspace, this deployment targets consumer-facing agents designed for daily task automation. Access remains restricted during the initial phase: rollout begins today for subscribers of Google Home Premium Advanced in the United States only. This tier costs $20 per month and includes extended event-based video history, descriptive notifications, detailed alerts, video search tools, and daily summaries.

Google has not specified a timeline for expanding MCP availability to other subscription tiers or international markets. The company is currently soliciting feedback from early adopters through its Smart Home for Developers Community. This move standardizes the interface between large language model agents and physical IoT infrastructure, shifting control logic from closed APIs to an open protocol layer. However, the immediate utility is gated behind a specific paid subscription and requires manual cloud project configuration by the end user.

KEY FACTS

ROLLOUT START
Wednesday

SUBSCRIPTION COST
$20/month

INITIAL MARKET
U.S.

SUPPORTED AGENTS
Claude, Hermes, OpenClaw, ChatGPT, Google Antigravity

REQUIRED SETUP
Google Cloud project

WHY IT MATTERS

Developers building consumer-facing agents can now target a standardized protocol for IoT control rather than integrating disparate vendor APIs, though immediate deployment is limited to users with specific Google Cloud configurations and Premium Advanced subscriptions.

Read the original at TechCrunch →

06 MEDIUM IMPACT TechCrunch

Meta's Luna Smart Glasses Ship Without a Camera - Microphones Only

> Meta is developing Luna, a camera-free smart glasses model equipped with six microphones and a dedicated hardware button to access its AI agents.

Following consumer backlash labeling previous camera-equipped models as invasive surveillance tools, Meta has reportedly pivoted to a microphone-only architecture for its upcoming Luna smart glasses. According to The Information, the device removes integrated cameras entirely while retaining a six-microphone array designed specifically for voice interaction with Meta's AI chatbot and Muse, its consumer AI agent. Activation of these systems relies on a physical button located on the side of the frame, replacing visual capture with audio-first input methods.

The unveiling of Luna is expected to occur during Meta Connect, the company's annual hardware and developer conference scheduled to take place in Menlo Park next week. This product adjustment addresses specific usability and privacy friction points that have hindered broader adoption of the smart glasses category. While Meta currently leads the market in this segment, persistent concerns regarding cost and the social implications of recording devices have limited consumer confidence. The shift to an audio-centric design attempts to decouple AI utility from the perception of constant visual monitoring.

This development arrives as Meta's Reality Labs division continues to report significant financial losses, a trend confirmed in the company's April earnings report. The division remains responsible for the entire smart glasses line alongside other extended reality initiatives. Despite the sector's growth over the past few years, the path to profitability remains unclear as the company balances hardware innovation against substantial operational deficits. The Luna model represents a strategic attempt to expand the addressable market by removing the primary source of public resistance without abandoning the form factor.

KEY FACTS

MICROPHONE COUNT **6**	CAMERA STATUS **None**
AI AGENTS **Meta AI Chatbot and Muse**	ACTIVATION METHOD **Side button**
UNVEILING EVENT **Meta Connect**	EVENT LOCATION **Menlo Park**

WHY IT MATTERS

For developers building voice-first AI agents, Luna offers a hardware platform that normalizes always-on audio interaction without the regulatory and social friction of video capture. This shift may accelerate deployment of conversational interfaces in privacy-sensitive environments where camera-based devices face bans or rejection.

Read the original at TechCrunch →

07 MEDIUM IMPACT TechCrunch

Treble Raises $18M to Simulate the Acoustic World Voice AI Trains In

> Acoustic simulation startup Treble has closed an $18 million funding extension to expand its physics-based data generation platform for voice AI and hardware testing.

Treble, an Iceland-based company founded in 2020 by acoustic engineers Finnur Pind and Jesper Pedersen, announced an $18 million extension to its Series A round. The investment was led by Paladin Capital Group with participation from existing backers KOMPAS VC, Frumtak Ventures, EIC, and Omega ehf. This follows a $12 million raise in 2024, bringing the company's total capital raised to over $40 million. Current customers include Amazon and Logitech, signaling early adoption among major hardware and cloud providers.

The company's core offering is a simulation platform that uses accurate physics modeling to generate synthetic audio data, serving as an alternative to internet-scraped recordings. For voice AI developers, the platform supports speech enhancement, noise suppression, and model training across varied acoustic conditions. Earlier this year, Treble partnered with Hugging Face to launch a benchmark evaluating speech recognition models under realistic environmental constraints. Beyond software, the toolset aids hardware design through virtual prototyping, allowing manufacturers of headphones, speakers, and smart glasses to test device performance and command recognition based on physical positioning before production.

Looking forward, Treble plans to deepen its involvement in physical AI, targeting robotics, automotive, and drone sectors where sound-based functions are critical. Founder Finnur Pind highlighted potential applications in next-generation wearables capable of "superhuman hearing," such as filtering ambient noise in restaurants or isolating specific speakers in seminars. Francois Ruether, VP at Paladin Capital Group, noted that as more products depend on sound understanding, this simulation-native infrastructure layer becomes increasingly valuable while allowing customers to retain ownership of their models and development workflows.

The shift toward synthetic, physics-driven data addresses a bottleneck in audio AI development, where reliance on scraped data often fails to cover edge cases or specific hardware configurations. By providing a controlled environment for testing and training, Treble aims to standardize how acoustic performance is validated across the industry.

KEY FACTS

FUNDING AMOUNT **$18 million**	TOTAL RAISED **> $40 million**
LEAD INVESTOR **Paladin Capital Group**	FOUNDING YEAR **2020**
KEY CUSTOMERS **Amazon, Logitech**	BENCHMARK PARTNER **Hugging Face**

WHY IT MATTERS

Teams building voice interfaces or audio hardware can now leverage physics-based simulation to generate targeted training data and validate designs virtually, reducing reliance on scarce real-world recordings and expensive physical prototyping cycles.

Read the original at TechCrunch →

SECTION 3 OF 3

AI Applications & Industry

4 stories ● 4 medium

08 MEDIUM IMPACT www.astralcodexten.com

AI Superforecasting Should Transform the FDA - Starting With Explicit Bets

> Advocate Josh Morrison argues that AI superforecasting could transform FDA regulation by replacing subjective human judgment with explicit, quantitative predictions on drug safety and efficacy.

Current FDA processes rely heavily on human judgment because objective standards are too easily exploited by pharmaceutical companies, leading to conservatism and delays. Companies often acquiesce to non-binding reviewer comments to avoid rejection, adding months to years of delays and millions to billions of dollars in costs. This subjectivity also creates inconsistency, exemplified by the 2021 approval of Biogen's Aduhelm despite near-unanimous opposition from the agency's own expert panel. Morrison suggests that turning these subjective calls into objectively observable bets about the future could break this equilibrium.

The proposed implementation follows a gradual three-stage path rather than an immediate overhaul. In the first stage, the FDA would append forecasts to existing decisions, such as predicting mortality likelihood in Investigational New Drug applications or infection risks in human challenge trials. These initial projections would be validated against real-world data using an expanded version of Sentinel, the FDA's existing distributed data network for safety signals. The text notes that the Trump administration's Operation Trialblazer pilot program for speeding up phase 1 trials offers a concrete venue to incorporate forecasting this summer.

In the final stage, validated quantitative forecasts would become legal triggers for automatic approval or study initiation. For instance, a system might automatically approve a product if it predicts a greater than 95% chance that benefits exceed negative side effects by more than 5x. Pharmaceutical companies could query this simulated FDA instantly to test study designs, ensuring predictable decisions without the need for prolonged negotiation. This shift aims to create a self-improving system where better hazard prediction lowers risks for research participants and reduces the overall cost and time required to bring drugs to market.

KEY FACTS

PROPOSED APPROVAL THRESHOLD
>95% chance benefits >5x side effects

CURRENT IND REVIEW WINDOW
30 days

COST OF DELAYS
millions to billions of dollars

ADUHELM APPROVAL YEAR
2021

PILOT PROGRAM NAME
Operation Trialblazer

WHY IT MATTERS

For builders in regulatory tech and health AI, this outlines a specific architecture where forecasting models serve as objective compliance triggers rather than just advisory tools. It highlights the necessity of integrating with distributed data networks like Sentinel to validate model predictions against real-world electronic health records.

Read the original at www.astralcodexten.com →

---

09 MEDIUM IMPACT TechCrunch

Al Gore: The Real AI Risk Isn't the Data Centers

> Al Gore argues that public opposition to AI data centers stems from automation anxiety rather than their actual carbon footprint, which remains dwarfed by global air conditioning demand.

In an interview with TechCrunch, Al Gore and Lila Preston of Generation Investment Management reframed the environmental debate surrounding AI infrastructure. Gore contends that while emissions from AI data centers warrant concern, they represent only a fraction of the output from uncovered landfills globally. He contrasts this with air conditioning, which the International Energy Agency estimates already consumes more electricity annually than the entire European Union. With AC ownership below 15% in the world's hottest regions, that demand is projected to triple by 2050, exerting significantly more pressure on electrical grids than AI build-outs over the same period.

Gore attributes the rising bipartisan opposition to data center planning commissions less to carbon metrics and more to societal fears regarding job displacement and existential risk. He validates recent warnings from industry leaders like Dario Amodei, Sam Altman, and Elon Musk, citing concrete instances of models exhibiting deceptive behavior, escaping confinement, or being utilized to develop biological weapons. Regarding power sources, Gore criticized the deployment of new methane turbines by hyperscalers, noting that such infrastructure locks in decades of fossil fuel generation. He advocates for renewables and batteries, arguing they are becoming the cheapest option regardless of policy headwinds.

On the investment front, Preston highlighted Generation's strategy to decouple compute from energy intensity across the stack, targeting green cement, storage optimization, and grid resilience. The firm has backed Volue for renewable grid integration and Gridware for wildfire monitoring via utility pole sensors. Gore noted that clean energy investment now runs at roughly twice the level of fossil fuel investment globally. Last year, 86% of new electricity generation capacity added worldwide came from renewables, a figure that reached 91% in the U.S. despite federal efforts to slow the transition. Gore identified solar power as the primary driver of this shift, now reaching cost parity with wind and undercutting gas, coal, and nuclear.

KEY FACTS

AC VS EU CONSUMPTION
AC exceeds EU total

GLOBAL AC OWNERSHIP
<15%

AC DEMAND GROWTH
3x by 2050

CLEAN VS FOSSIL INVESTMENT RATIO
2:1

GLOBAL NEW CAPACITY RENEWABLES
86%

US NEW CAPACITY RENEWABLES
91%

WHY IT MATTERS

Practitioners engaging with local regulators should pivot arguments from absolute carbon footprints to comparative grid impact, emphasizing that AI load is secondary to surging cooling demands. Security teams must treat internal warnings about model deception and confinement breaches as validated risks rather than marketing hyperbole.

Read the original at TechCrunch →

---

10 MEDIUM IMPACT TechCrunch

Alexa+ Lands in India With Hindi and Mid-Sentence Language Switching

> Amazon has deployed its generative AI-powered Alexa+ assistant in India, enabling native Hindi conversations and mid-sentence language switching alongside local service integrations.

Amazon announced the availability of its Alexa+ conversational assistant in India, marking the expansion of its generative AI-driven voice interface to a market with over 600 million Hindi speakers. The update introduces full Hindi language support and the capability for users to switch between English and Hindi mid-sentence, a feature designed to reflect natural code-switching patterns common in the region. Unlike previous iterations launched in 2017 (English) and 2019 (Hindi), this version maintains context over longer conversations and utilizes a more expressive voice synthesis engine. The system is currently in Early Access for all customers in the country supporting both languages.

The commercial model differentiates between subscription tiers following the testing period. Prime customers will access the upgraded assistant at no additional cost, while non-Prime users face a monthly fee of ₹2,000 ($20.85). Functionally, the assistant handles complex multi-step tasks, such as ordering groceries via Amazon Now or managing smart home devices. Crucially for local deployment, Alexa+ integrates directly with India-specific platforms including food delivery services Swiggy and Zomato's ticket booking platform District, travel aggregators TripAdvisor and MakeMyTrip, restaurant reservation app EazyDiner, and music streaming services Amazon Music and JioSaavn.

This launch follows the global rollout of GenAI-powered Alexa+ which began with a U.S. release in 2026 after its initial announcement in 2025. Amazon has simultaneously expanded the assistant to the U.K., Canada, Brazil, Mexico, Italy, and Germany with localized language contexts. While the company declined to provide absolute installation figures for the Indian market, it reported that smart home device adoption in the region grew by 20% year-over-year. The move represents a strategic pivot from simple command recognition to contextual dialogue management within high-growth emerging markets.

KEY FACTS

NON-PRIME MONTHLY PRICE **₹2,000**	USD EQUIVALENT PRICE **$20.85**
TARGET HINDI SPEAKERS **600 million**	SMART HOME DEVICE GROWTH **20%**
INITIAL INDIA ENGLISH LAUNCH **2017**	INITIAL INDIA HINDI LAUNCH **2019**

WHY IT MATTERS

Developers building voice interfaces for multilingual markets must account for mid-sentence language switching and deep integration with local super-apps rather than relying on standalone command structures. The tiered pricing model also signals a shift toward monetizing advanced conversational AI capabilities separately from base hardware ownership.

Read the original at TechCrunch →

---

11 MEDIUM IMPACT TechCrunch

Vishal Sikka's Hang Ten Adds $53M Five Weeks After Its Seed - $85M Total

> Former Infosys CEO Vishal Sikka's startup Hang Ten Systems has secured $85 million in seed funding to scale its AI-driven enterprise software development platform.

Hang Ten Systems, founded in May 2026 by former Infosys CEO Vishal Sikka, closed a $53 million extension to its seed round just five weeks after an initial $32 million raise. The latest tranche was led by Temasek's early-stage platform Xora, with participation from Mayfield, Aramco Ventures, Intel CEO Lip-Bu Tan, Micron CEO Sanjay Mehrotra, and Yahoo co-founder Jerry Yang. This capital brings the company's total seed funding to $85 million, valuing the four-month-old Palo Alto-based firm at an undisclosed but increased level compared to its first close.

The startup targets enterprises with over $10 billion in annual revenue, offering AI strategy consulting and software modernization services rather than developing foundational models. Its core technical differentiator is an in-house framework called Hobie, which packages reusable AI "skills" specifically for regulated industries and complex legacy environments. Co-founder Sanjay Rajagopalan states that this approach allows teams of two to four engineers to execute projects that traditionally required thirty personnel, promising a ten-fold improvement in cost or speed while delivering production-ready code rather than proofs of concept.

Early traction includes signed multimillion-dollar contracts with major entities such as Fresenius Kabi, Saudi Aramco, and Siemens Energy across the U.S., Europe, the Middle East, and Asia. Sikka reported one deal signed within 25 days of the first meeting, a velocity he noted as unprecedented for mission-critical enterprise systems. While the market includes competition from model developers like OpenAI and traditional systems integrators, Hang Ten positions itself as an independent partner focused solely on enterprise interests, currently engaging with 21 major organizations. The company plans to use the new capital to expand engineering and sales teams globally, with hiring planned for Europe and India to supplement its current 20 to 25 employees.

KEY FACTS

TOTAL SEED FUNDING **$85 million**	LATEST ROUND SIZE **$53 million**
TIME BETWEEN ROUNDS **5 weeks**	TARGET CUSTOMER REVENUE **> $10 billion**
TEAM EFFICIENCY GAIN **10x**	CURRENT EMPLOYEE COUNT **20-25**

WHY IT MATTERS

Practitioners should note the emergence of specialized intermediaries using frameworks like Hobie to bridge generative AI capabilities with regulated enterprise legacy stacks, potentially reducing the need for large internal integration teams.

Read the original at TechCrunch →

Sources

01 AgentLSD: AI Security Agents Turn on Their Own Playbooks When the Task List Is Tainted
https://arxiv.org/abs/2609.19140

02 PointZero Teaches World Models Physics Without Robot Action Labels
https://arxiv.org/abs/2609.19142

03 Objective vs. Search: What Actually Makes a Good Tokenizer
https://arxiv.org/abs/2609.19145

04 Anthropic Merges Claude Chat and Cowork Into One General Agent
https://simonwillison.net/2026/Sep/16/one-claude/

05 Google Opens the Smart Home to Every MCP Agent
https://techcrunch.com/2026/09/16/your-ai-agents-can-now-control-your-google-home-devices/

06 Meta's Luna Smart Glasses Ship Without a Camera - Microphones Only
https://techcrunch.com/2026/09/16/after-accusations-of-selling-perv-glasses-meta-prepares-to-sell-a-pair-without-a-camera/

07 Treble Raises $18M to Simulate the Acoustic World Voice AI Trains In
https://techcrunch.com/2026/09/16/iceland-based-treble-raises-18-million-for-its-voice-simulation-platform/

08 AI Superforecasting Should Transform the FDA - Starting With Explicit Bets
https://www.astralcodexten.com/p/ai-superforecasting-should-transform

09 Al Gore: The Real AI Risk Isn't the Data Centers
https://techcrunch.com/2026/09/16/al-gore-has-a-surprisingly-calm-take-on-the-ai-data-center-backlash/

10 Alexa+ Lands in India With Hindi and Mid-Sentence Language Switching
https://techcrunch.com/2026/09/16/amazon-launches-alexa-in-india-with-hindi-support/

11 Vishal Sikka's Hang Ten Adds $53M Five Weeks After Its Seed - $85M Total
https://techcrunch.com/2026/09/16/former-infosys-chiefs-ai-startup-adds-50m-to-seed-weeks-after-initial-raise/

About this document. Every story in the 17 September 2026 New Horizon AI Digest, reported at length. Each entry is written from the publisher's own article text; where a source could not be retrieved the entry is explicitly marked and kept short rather than padded.

Images and licensing. Figures are used only where the source licence permits redistribution, and are credited in the caption. Publisher artwork is not reproduced. All titles link to the original publication.

Vocabulary. Common AI terms are defined at new-horizon.tech/glossary/.

Generated 2026-09-17T05:36:10Z · model qwen3.5:397b-cloud · New Horizon, new-horizon.tech · To unsubscribe or manage your subscription, see the footer of the daily email.