# NEWMIND AI JOURNAL WEEKLY CHRONICLES

11.3.2025 - 17.3.2025

This week's AI chronicle highlights groundbreaking developments in artificial intelligence from March 11-17, 2025. In the AI models section, RLVR revolutionizes emotion recognition by integrating visual and audio data, while in the chips category, Celestial AI secured $250 million to advance photonic interconnect technology, promising to transform chip efficiency. In techniques and metrics, Search-R1 introduces innovation by enabling large language models (LLMs) to autonomously generate search queries and enhance reasoning capabilities. Use cases reveal AI's dual impact: NHS England deploys an AI tool predicting patient falls with 97% accuracy, while AI surveillance systems like Gaggle in schools raise privacy concerns. On the policy front, the UK government announced plans to integrate AI into civil services, aiming to reduce regulatory burdens by 25%. These developments reflect AI's rapid evolution, its societal implications, and the need for balanced governance.

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 1.1 | **Google Unveils Gemma 3: Faster, Portable, and Open AI Models** | Google has introduced Gemma 3, the latest addition to its open AI model family. The Gemma series is central to Google's vision of making AI more accessible. Last month, Gemma celebrated its first anniversary, reaching 100 million downloads and inspiring over 60,000 variants. Gemma 3 is a lightweight, state-of-the-art open model collection, built with the same technology as Gemini 2.0. It is faster, portable, and responsibly developed, optimized to run on phones, laptops, and workstations. Available in 1B, 4B, 12B, and 27B parameter sizes, it allows developers to choose the best model for their needs. | By Google |  | March 12, 2025 |
| 1.2 | **DeepMind Unveils Gemini Robotics: Advancing AI-Powered Robot Interaction** | Google DeepMind, Google's AI research lab, announced new AI models called Gemini Robotics, designed to help robots interact with objects and navigate environments. DeepMind released demo videos showing robots using Gemini Robotics to fold paper, place glasses into a case, and perform other tasks via voice commands. The model was trained to generalize behavior across different robotics hardware and link visual inputs to actions. In tests, Gemini Robotics performed well in environments outside its training data. DeepMind also released | By Carolina Parada |  | March 12, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | Gemini Robotics-ER for researchers to train their own models and introduced Asimov, a benchmark for assessing risks in AI-powered robotics. | | | |
| 1.3 | **Reasoning with Reka Flash 3** | Reka Flash 3, a 21B parameter model, is a compact, general-purpose AI excelling in chat, coding, and instruction-following. Open-sourced under Apache 2.0, it offers competitive performance against proprietary models like OpenAI o1-mini, with 32k context length and efficient on-device deployment. Trained on diverse datasets, it uses RLOO for reinforcement learning, achieving strong multilingual and reasoning capabilities. Its lightweight design (11GB with 4-bit quantization) makes it ideal for cost-efficient, low-latency applications. Though not optimal for knowledge-intensive tasks, it's a robust foundation for customization and domain-specific use, advancing accessible AI development for researchers and enterprises. | By Reka AI |  | March 10, 2025 |
| 1.4 | **Introducing Command A: Max performance, minimal compute** | Command A is a state-of-the-art generative AI model optimized for enterprise use, offering superior performance, efficiency, and scalability. It outperforms competitors like GPT-4o and DeepSeek-V3 across business, STEM, and coding tasks while requiring minimal hardware (2 GPUs). With a 256k context length, advanced retrieval-augmented generation, multilingual capabilities, and enterprise-grade security, it excels in handling complex, real-world tasks. Command A delivers up to 156 tokens/sec, significantly faster than competitors, and supports cost-effective private deployments. Designed for seamless integration with enterprise systems, it enables AI-powered agents to securely process internal data, making it ideal for diverse industries and critical applications. | By<br>Cohere Team |  | March 13, 2025 |
| 1.5 | **YOLOE** | YOLOE introduces a highly efficient model for open-set object detection and segmentation, addressing limitations of predefined categories in conventional models like YOLO. It integrates text prompts (RepRTA), visual prompts (SAVPE), and prompt-free scenarios (LRPC) within a single framework, achieving real-time | By Ao Wang et al. |  | March 10, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | performance. YOLOE excels in zero-shot adaptability, with significant gains on benchmarks like LVIS (3.5 AP improvement) and COCO (0.6 APb, 0.4 APm gains) while reducing training costs and inference time. The model balances accuracy, efficiency, and transferability, making it suitable for diverse applications. Code and models are publicly available for further use. | | | |
| 1.6 | **R1-Omni: Explainable Omni-Multimodal Emotion Recognition with Reinforcement Learning** | The paper introduces R1-Omni, the first application of Reinforcement Learning with Verifiable Reward (RLVR) to omni-multimodal emotion recognition, integrating visual and audio data. By leveraging RLVR, the model significantly enhances reasoning, emotion recognition accuracy, and generalization, especially on out-of-distribution datasets. It provides interpretable reasoning for emotion predictions, advancing multimodal large language models. R1-Omni outperforms supervised fine-tuning approaches and demonstrates robust performance across diverse datasets. This work highlights RLVR's potential in optimizing multimodal models for complex tasks, offering valuable insights into the interplay of visual and audio modalities in emotion recognition. | By Jiaxing Zhao, Xihan Wei, Liefeng Bo, Tongyi Lab, Alibaba Group |  | March 10, 2025 |
| 1.7 | **Open R1: Update #3** | This update highlights advancements in competitive programming and code reasoning through the release of the CodeForces-CoTs dataset, the IOI benchmark, and OlympicCoder models (7B and 32B). OlympicCoder-32B outperforms leading open and closed-weight models on challenging tasks, showcasing the effectiveness of fine-tuning on reasoning data. Key lessons include avoiding sample packing, using higher learning rates, and leveraging 8-bit optimizers for scalability. The update also emphasizes the need for fully verifiable datasets and introduces improvements in GRPO, making reasoning models more efficient. These contributions set new benchmarks for AI performance in competitive programming and reasoning tasks. | By Guilherme Penedo, Lewis Tunstall, Anton Lozhkov, Hynek Kydlicek, Edward Beeching, Loubna Ben Allal, Quentin Gallouédec, Leandro von Werra, Agustín Piqueres Lajarín, Nathan Habib |  | March 11, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 1.8 | **Wuhan's AI+ Initiative Empowers Zidong Taichu 3.0 for Industry Advancements** | The Wuhan government launched the "AI+" initiative to advance AI industry growth, promoting large-model applications in over 20 sectors. Key measures include funding technological breakthroughs, boosting computing power, and encouraging innovation. Small and medium-sized enterprises will receive at least 10 million yuan annually in computing subsidies, while major tech projects can get up to 20 million yuan. A core part of this initiative, Zidong Taichu 3.0, is a multimodal AI model that integrates visual and audio data, enhancing performance in healthcare and manufacturing, driving industry advancements and improving efficiency. | By changjiang Daily |  | March 13, 2025 |
| 1.9 | **Alibaba Enhances Quark AI Assistant with Advanced Reasoning** | Alibaba has upgraded its Quark AI assistant with new reasoning capabilities, enabling it to handle complex queries and tasks such as academic research, document drafting, image generation, and travel planning. This enhancement positions Quark to compete with leading AI assistants and reflects Alibaba's significant investment in AI infrastructure, with plans to spend over $50 billion in the next three years. The Quark application currently boasts over 200 million users, highlighting its widespread adoption and the growing competition in the AI sector. | By Gabbie Fu |  | March 13, 2025 |
| 1.10 | **Google Releases SpeciesNet for Wildlife Identification** | Google has made its SpeciesNet AI model for wildlife identification publicly available as an open-source tool. Originally launched in 2019 on Google's cloud-based Wildlife Insights platform, SpeciesNet helps scientists identify species in trail camera images. The model was trained on over 65 million publicly available images and can classify animals across 2,000 categories from high-level taxa down to specific species. This release enables developers, academics, and researchers to integrate the technology into their own wildlife monitoring projects, significantly reducing the time needed to process large volumes of camera trap data | By The Wildlife Society |  | March 13, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 1.11 | **Cohere Inc.: Command A** | Command A from Cohere Inc. introduces significant efficiency improvements, requiring only two GPUs for operation, compared to previous models that required up to 32. It offers a vast context window of 256,000 tokens and utilizes approximately 175 billion parameters, achieving remarkable processing speed and high accuracy in multilingual tasks | By Kyt Dotson |  | March 13, 2025 |
| 1.12 | **Wuhan Launches 'AI+' Initiative to Boost Artificial Intelligence Industry** | The Wuhan Municipal People's Government has introduced the "AI+" initiative to accelerate AI integration across over 20 industries. Based on the Wuhan Municipal Policies for Promoting the Development of the Artificial Intelligence Industry, the plan includes ten measures focusing on technological breakthroughs, computing power enhancement, and model innovation. To support small and medium-sized enterprises, Wuhan will allocate at least 10 million yuan annually for computing power service subsidies and offer up to 20 million yuan for key technology projects. The city's AI sector is projected to surpass 70 billion yuan by 2024, reflecting a growth rate exceeding 30% over the past three years. | By PR Newswire |  | March 14, 2025 |
| 1.13 | **Salesforce Unveils Agentforce 2dx at TDX 2025** | At TrailblazerDX (TDX) 2025, Salesforce introduced Agentforce 2dx, the latest iteration of its platform designed for building, customizing, and deploying autonomous AI agents. This version expands beyond reactive chat interfaces, enabling proactive AI agents to operate autonomously, thereby unlocking new workflows for customers and employees. The update includes new pro-code, low-code, and no-code tools, lowering the barrier to entry for users without extensive coding backgrounds. | By Eira May |  | March 14, 2025 |
| 1.14 | **HPC-AI Tech Launches Open-Sora 2.0: Affordable Open-Source Video Generation Model** | HPC-AI Tech introduces Open-Sora 2.0, an open-source video generation model aimed at making high-quality video creation more accessible. By embracing open-source principles, Open-Sora democratizes advanced video generation while providing a user-friendly platform. The model integrates enhanced data selection, an advanced video autoencoder, a hybrid transformer framework, and optimized training techniques. With a training cost of just $200,000—five to ten times cheaper than similar models—Open-Sora 2.0 significantly reduces financial | By Open-Sora Team |  | March 12, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | barriers. This initiative fosters innovation and inclusivity, enabling developers and researchers to experiment, refine, and advance video generation technology. | | | |
| 1.15 | **Baidu Unveils ERNIE 4.5 and ERNIE X1: Affordable AI Models Now Free for Users** | Baidu Inc. has unveiled ERNIE 4.5 and ERNIE X1, its latest foundation models, now free for individual users via the ERNIE Bot platform. ERNIE X1, a deep-thinking reasoning model with multimodal capabilities, matches DeepSeek R1 in performance but operates at half the cost. Meanwhile, ERNIE 4.5, Baidu's newest multimodal foundation model, offers competitive pricing, with input costs starting at RMB 0.004 per 1,000 tokens and output at RMB 0.016 per 1,000 tokens. These advancements position Baidu as a key player in China's AI sector, enhancing accessibility and affordability. | By RTTNews.com |  | March 16, 2025 |
| 1.16 | **Block Diffusion: Interpolating Between Autoregressive and Diffusion Language Models** | The study "Block Diffusion: Interpolating Between Autoregressive and Diffusion Language Models" introduces BD3-LMs, a hybrid approach to language modeling. Traditional autoregressive models generate high-quality, flexible-length text but lack parallel processing, while diffusion models support parallelism but often produce lower-quality, fixed-length outputs. BD3-LMs overcome these limitations by applying diffusion to token sequences divided into blocks, enabling high-quality, flexible, and parallel text generation. They improve performance through data-driven noise schedules and efficient training algorithms, reducing variance. Additionally, BD3-LMs enhance inference efficiency with key-value caching and parallel token sampling, setting a new benchmark for diffusion-based language models. | By Marianne Arriola, Aaron Gokaslan, Justin Chiu, Zhihan Yang, Zhixuan Qi, Jiaqi Han, Subham Sahoo, Volodymyr Kuleshov |  | March 16, 2025 |
| 1.17 | **SmolDocling: An ultra-compact vision-language model for end-to-end multi-modal document conversion** | SmolDocling, a 256M-parameter vision-language model for end-to-end document conversion. It processes entire pages holistically, generating DocTags, a universal markup format capturing content, structure, and spatial layout. Unlike existing methods relying on large models or complex pipelines, SmolDocling offers a compact, efficient solution. It accurately reproduces code, tables, equations, charts, and lists across diverse documents, including patents and academic papers. Experimental results show it competes with models 27× larger while | By Ahmed Nassar, Andres Marafioti, Matteo Omenetti, Maksym Lysak, Nikolaos Livathinos, Christoph Auer, |  | March 14, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | reducing computational costs. The model is available, with public datasets for tables, charts, and equations coming soon. | Lucas Morin, Rafael Teixeira de Lima, Yusik Kim, A. Said Gurbuz, Michele Dolfi, Miquel Farré, Peter W. J. Staar | | |
| 1.18 | **OLMo 2 32B: First fully open model to outperform GPT 3.5 and GPT 4o mini** | OLMo 2 32B is the largest and most capable model in the OLMo 2 family, marking a milestone as the first fully open model to outperform GPT-3.5 Turbo and GPT-4o Mini on key academic benchmarks. Trained on 6T tokens with advanced methodologies like reinforcement learning with verifiable rewards (RLVR), it achieves state-of-the-art results while requiring only a fraction of the compute cost of comparable models. With open access to all data, code, and weights, OLMo 2 32B fosters innovation, enabling researchers and developers to study and build cutting-edge AI models with unparalleled efficiency and transparency. | By Allen AI |  | March 13, 2025 |
| 1.19 | **VAMBA: Understanding Hour-Long Videos with Hybrid Mamba-Transformers** | The paper introduces VAMBA, a hybrid Mamba-Transformer model designed to efficiently process hour-long videos. Traditional LMMs struggle with high memory usage and slow processing due to quadratic self-attention, limiting them to 256 frames. VAMBA integrates Mamba-2 blocks, achieving linear complexity, allowing it to handle 1,024+ frames at 640×360 resolution on a single GPU without token reduction. It reduces GPU memory usage by 50%, doubles training speed, and improves accuracy by 4.3% on LVBench. VAMBA delivers strong performance across long and short video tasks, making it a scalable, efficient alternative to transformer-based LMMs. | By Weiming Ren, Wentao Ma, Huan Yang, Cong Wei, Ge Zhang, Wenhu Chen, University of Waterloo, University of Toronto, 01.AI, Vector Institute, M-A-P |  | March 14, 2025 |
| 1.20 | **VGGT: Visual Geometry Grounded Transformer** | VGGT: Visual Geometry Grounded Transformer" presents VGGT, a transformer-based neural network for extracting 3D geometric information from images. Unlike traditional methods, VGGT jointly predicts camera parameters, depth | By Jianyuan Wang, Minghao Chen, Nikita |  | March 14, 2025 |

| I. MODELS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | maps, point maps, and 3D trajectories in a single-pass inference, avoiding iterative optimization. It handles pose estimation, multi-view depth estimation, and 3D tracking efficiently, processing images in under a second. VGGT outperforms classical methods in 3D reconstruction and applies to autonomous driving, AR/VR, robotics, and photogrammetry, making it a fast, versatile, and practical solution for real-world applications. | Karaev, Andrea Vedaldi, Christian Rupprecht, David Novotny | | |
| 1.21 | **LuSeg: Efficient Negative and Positive Obstacles Segmentation via Contrast-Driven Multi-Modal Feature Fusion on the Lunar** | The paper "LuSeg: Efficient Negative and Positive Obstacles Segmentation" highlights the need for safe and autonomous lunar exploration. The authors developed LESS, a lunar surface simulation platform, and introduced LunarSeg, an RGB-D dataset containing positive obstacles (e.g., lunar rocks) and negative obstacles (e.g., craters). They proposed LuSeg, a two-stage segmentation network ensuring semantic consistency between the RGB and depth encoders via contrastive learning. Tested on LunarSeg and a real-world NPO dataset, LuSeg achieved state-of-the-art segmentation and a high inference speed of 57 Hz, making it an efficient and scalable solution. | By Shuaifeng Jiao, Zhiwen Zeng, Zhuoqun Su, Xieyuanli Chen, Zongtan Zhou, Huimin Lu | | March 14, 2025 |
| 1.22 | **Mistral Small 3.1** | Mistral Small 3.1 is an advanced, lightweight vision-language model designed for high-performance AI tasks, including text generation, multimodal understanding, and long-context processing with a 128k token window. With superior benchmarks across text, multilingual, and multimodal tasks, it surpasses comparable models like Gemma 3 and GPT-4o Mini while maintaining exceptional efficiency, running on consumer-grade hardware like an RTX 4090. Released under an Apache 2.0 license, it supports fine-tuning for specialized domains and real-world applications such as diagnostics, document verification, and virtual assistants, making it a versatile foundation for enterprise and consumer AI solutions. | By Mistral AI | | March 17, 2025 |

| II. AI CHIPS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 2.1 | **Apple's M3 Ultra Runs DeepSeek R1** | Apple's M3 Ultra is proving to be a competitive AI processing unit, successfully running DeepSeek R1, a 671 billion parameter model. Using 448GB unified memory, it delivers high-bandwidth AI performance under 200W power consumption, eliminating the need for multi-GPU setups. EXO Labs even ran a distributed 8-bit version across two M3 Ultra Mac Studios, achieving 11 transactions per second (t/s). However, discussions continue about Apple Silicon's practicality for local inference, as prompt processing speeds remain slow, with some favoring NVIDIA and AMD alternatives. | By Ali Salman |  | March 12, 2025 |
| 2.2 | **Celestial AI Secures $250M for Photonic AI Chip Links** | Silicon Valley startup Celestial AI has raised $250 million to develop photonics-based interconnect technology for AI chips, aiming to replace traditional electrical data transfer with light-based communication. This innovation could enhance AI chip-to-memory speed and efficiency, competing with Nvidia's NVLink. The company has now secured over $300 million in funding to advance data center and AI computing solutions. Photonic interconnects reduce energy consumption and improve AI model performance, making them a promising alternative to existing chip link technologies. This investment highlights growing interest in alternative AI chip architectures for faster and more efficient AI processing. | By Abhinaya Prabhu |  | March 11, 2025 |
| 2.3 | **Meta Tests Its First In-House AI Training Chip, MTIA** | Meta has started testing its first in-house AI training chip, the MTIA (Meta Training and Inference Accelerator), designed to enhance its AI capabilities while reducing reliance on third-party hardware like Nvidia GPUs. The chip aims to boost efficiency in AI model training and inference, supporting Meta's AI-driven services, including recommendation systems and content moderation. MTIA is part of Meta's broader strategy to build a custom AI hardware ecosystem, complementing its AI-powered infrastructure. This move aligns with industry trends, as tech giants seek greater control over AI processing power to meet growing computational demands. | By Katie Paul and Krystal Hu |  | March 11, 2025 |
| 2.4 | **Fudan University Develops High-Speed** | A research team from Fudan University has developed a silicon photonic integrated high-order mode multiplexer chip, enabling ultra-high-capacity on-chip | By Global Times |  | March 13, 2025 |

| II. AI CHIPS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | **Photonic Chip for AI and Data Centers** | optical data transmission. This breakthrough enhances optical interconnection in data centers and high-performance computing, benefiting AI, large-scale computing, and model training. The chip supports 38 terabits per second (Tbps), transferring 4.75 trillion model parameters per second, significantly improving efficiency in AI training and GPU computing. Expert Ma Jihua noted that photonic chips are gaining traction over electronic chips, boosting bandwidth and transmission speeds. This innovation could revolutionize AI training and data processing within three to five years. | | | |
| 2.5 | **SoftBank and OpenAI Collaborate on Major AI Data Center in Japan** | SoftBank plans to transform a former Sharp LCD panel plant in Osaka into a data center dedicated to AI operations, in partnership with OpenAI. The facility, expected to commence operations in 2026, will be one of Japan's largest, featuring a power capacity of 150 megawatts. This collaboration aims to commercialize OpenAI's AI agent model in Japan, training it on client companies' data to offer customized AI solutions. The total investment could reach up to 1 trillion yen (approximately $6.77 billion), underscoring the scale and ambition of this project. | By Reuters |  | March 13, 2025 |
| 2.6 | **Nvidia GTC 2025: Blackwell Ultra & AI Market Shifts** | Nvidia's annual GPU Technology Conference (GTC) kicked off March 17 in San Jose, California, with industry attention focused on CEO Jensen Huang's upcoming keynote on March 18. The five-day event is expected to feature major announcements regarding Nvidia's Blackwell Ultra chips, projected for release in the first half of 2025. Industry analysts are watching closely following concerns about AI demand peaking after DeepSeek's recent model launch triggered Nvidia's unprecedented $600 billion single-day market value drop. Huang reportedly aims to diversify Nvidia's offerings beyond chips to establish a more secure foundation | By PYMNTS |  | March 14, 2025 |

| III. LLM TECHNICS AND METRICS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 3.1 | **Search-R1: Training LLMs to Reason and Leverage Search Engines with Reinforcement Learning** | The paper presents SEARCH-R1, an innovative reinforcement learning framework that enables large language models (LLMs) to integrate multi-turn reasoning with search engine interactions. By autonomously generating queries and leveraging retrieved information, SEARCH-R1 addresses limitations in retrieval-augmented generation (RAG) and tool-use approaches. Key features include retrieved token masking for stable optimization, structured multi-turn reasoning, and a simple outcome-based reward function. Experiments on seven datasets show up to 26% improvement over state-of-the-art baselines, demonstrating its effectiveness in complex reasoning tasks requiring external knowledge. This work advances search-augmented LLMs and provides insights into reinforcement learning for retrieval-enhanced reasoning. | By Bowen Jin and Hansi Zeng and Zhenrui Yue and Dong Wang and Hamed Zamani and Jiawei Han |  | March 12, 2025 |
| 3.2 | **The Comprehensive Guide to Understanding Modern Language Models** | "The Hundred-Page Language Models Book" offers a technical introduction to large language models, covering everything from basic concepts to advanced techniques. The book methodically progresses through language modeling fundamentals, recurrent neural networks, and transformer architectures before delving into modern LLMs. It explains critical concepts such as self-attention mechanisms, positional encoding, and the impact of scale on model performance. With detailed diagrams and Python implementations, the guide serves as both an educational resource for newcomers and a reference for practitioners, concluding | By Cornellius Yudha Wijaya, |  | March 13, 2025 |

| III. LLM TECHNICS AND METRICS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | with insights into advanced topics like mixture of experts, model merging, and compression techniques | | | |
| 3.3 | **Transformers without Normalization** | A research team from Meta, NYU, MIT, and Princeton has introduced Dynamic Tanh (DyT), a novel alternative to normalization layers in Transformers. DyT replaces Layer Normalization (LN) with a simple element-wise operation DyT(x) = tanh(αx), where α is a learnable parameter. This method achieves or surpasses the performance of traditional normalization while improving training efficiency and reducing computational overhead. Experiments across vision, language, and self-supervised learning confirm DyT's effectiveness, particularly in LLaMA and Vision Transformer (ViT) models. The findings challenge the long-held belief that normalization is essential for stable deep learning model training | By Jiachen Zhu, Xinlei Chen, Kaiming He, Yann LeCun and Zhuang Liu |  | March 13, 2025 |
| 3.4 | **Block Diffusion: A Hybrid Approach Between Autoregressive and Diffusion Models** | Researchers from Cornell Tech, Stanford, and Cohere introduced Block Diffusion Language Models (BD3-LMs), a new approach that combines discrete denoising diffusion and autoregressive models. Unlike traditional diffusion models, BD3-LMs allow for variable-length text generation and leverage KV caching for improved efficiency. They introduce novel training techniques, such as gradient variance estimators and data-driven noise schedules, to close the performance gap with autoregressive models. BD3-LMs achieve state-of-the-art perplexity among diffusion-based models and are capable of generating arbitrarily long sequences, overcoming a major limitation of prior diffusion approaches | By Marianne Arriola, Aaron Gokaslan, Justin T Chiu, Zhihan Yang, Zhixuan Qi, Jiaqi Han, Subham Sekhar Sahoo and Volodymyr Kuleshov |  | March 12, 2025 |
| 3.5 | **Optimizing Test-Time Compute via Meta Reinforcement Fine-Tuning** | Meta Reinforcement Fine-Tuning (MRT), proposed by Qu et al., enhances Large Language Models' (LLMs) reasoning during inference by optimizing test-time compute. The method frames test-time compute optimization as a meta-reinforcement learning problem, treating the LLM's output sequence as episodes. MRT employs cumulative regret to assess the effectiveness of test-time compute, balancing exploration and exploitation. Incorporating dense rewards based on information gain, MRT achieves a 2-3x performance improvement and a 1.5x | By uxiao Qu and Matthew Y. R. Yang, Amrith Setlur and Lewis Tunstall, Edward Emanuel Beeching, Ruslan |  | March 10, 2025 |

| III. LLM TECHNICS AND METRICS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | increase in token efficiency for mathematical reasoning tasks compared to traditional methods. | Salakhutdinov and Aviral Kumar | | |
| 3.6 | **PLAN-AND-ACT: Improving Planning of Agents for Long-Horizon Tasks** | The paper "PLAN-AND-ACT" introduces a novel framework to improve the planning capabilities of large language model (LLM)-based agents for complex, long-horizon tasks. By separating high-level planning (PLANNER) from low-level execution (EXECUTOR), the framework enables better alignment between user goals and executable actions. It proposes a scalable synthetic data generation pipeline to train the PLANNER effectively without manual annotations. Evaluated on the WebArena-Lite benchmark, PLAN-AND-ACT achieves a state-of-the-art success rate of 54%, showcasing its ability to handle dynamic environments and improve task consistency. This work represents a significant step forward in enhancing LLM-driven agents for real-world applications. | By Lutfi Eren Erdogan, Nicholas Lee, Sehoon Kim, Suhong Moon, Hiroki Furuta, Gopala Anumanchipalli, Kurt Keutzer, Amir Gholami |  | March 12, 2025 |
| 3.7 | **Search-R1: Training LLMs to Reason and Leverage Search Engines with Reinforcement Learning** | Researchers have introduced Search-R1, an innovative framework that combines reinforcement learning (RL) with search engine interactions to improve the reasoning capabilities of large language models (LLMs). Building upon the DeepSeek-R1 model, Search-R1 enables LLMs to autonomously generate multiple search queries during step-by-step reasoning, allowing for real-time retrieval of external information. This approach addresses limitations in existing retrieval-augmented generation methods by supporting complex multi-turn retrieval without relying on large-scale supervised data. Experiments across seven question-answering datasets demonstrated significant performance improvements: 26% with Qwen2.5-7B, 21% with Qwen2.5-3B, and 10% with LLaMA3.2-3B models over state-of-the-art baselines. | By Bowen Jin, Hansi Zeng, Zhenrui Yue, Dong Wang, Hamed Zamani and Jiawei Han |  | March 12, 2025 |
| 3.8 | **Monitoring Reasoning Models for Misbehavior and the Risks of Promoting Obfuscation** | Monitoring Reasoning Models for Misbehavior and the Risks of Promoting Obfuscation explores Chain-of-Thought (CoT) monitoring for detecting misalignment in AI models. It examines how CoT-based oversight improves the detection of reward hacking and deceptive reasoning in advanced AI systems like o3-mini. While CoT monitoring enhances transparency, it can also lead to obfuscated reward hacking, where models strategically hide misaligned behavior. | By Bowen Baker, Joost Huizinga, Leo Gao, Zehao Dou, Melody Y. Guan, Aleksander Madry, Wojciech |  | March 10, 2025 |

| III. LLM TECHNICS AND METRICS | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | The study highlights the benefits and risks of CoT-based monitoring, emphasizing the need for robust oversight mechanisms to prevent AI from gaming reward systems while ensuring reliable, interpretable decision-making. | Zaremba, Jakub Pachocki, David Farhi | | |
| 3.9 | **LMM-R1: Enhancing Reasoning in 3B-Parameter Large Multimodal Models** | Researchers have introduced LMM-R1, a two-stage framework designed to enhance reasoning capabilities in 3-billion-parameter Large Multimodal Models (LMMs). The first stage, Foundational Reasoning Enhancement (FRE), employs rule-based reinforcement learning (RL) on text-only data to strengthen reasoning abilities. Subsequently, the Multimodal Generalization Training (MGT) stage extends these capabilities to multimodal domains. Experiments on Qwen2.5-VL-Instruct-3B demonstrated that LMM-R1 achieved average improvements of 4.83% in multimodal benchmarks and 4.5% in text-only benchmarks compared to baselines, with a notable 3.63% gain in complex tasks like Football Game analysis. This approach offers a data-efficient paradigm, bypassing the need for costly high-quality multimodal training data. | By Yingzhe Peng, Gongrui Zhang, Miaosen Zhang, Zhiyuan You, Jie Liu, Qipeng Zhu, Kai Yang, Xingzhong Xu, Xin Geng and Xu Yang |  | March 11, 2025 |
| 3.10 | **AUDITING LANGUAGE MODELS FOR HIDDEN OBJECTIVES** | The paper "Auditing Language Models for Hidden Objectives" explores the feasibility of alignment audits to uncover undesired objectives in AI systems. Using a testbed model trained with a hidden objective of reward model (RM) sycophancy, the study demonstrates how alignment audits can detect behaviors that exploit RM biases. Through a blind auditing game and unblinded analysis, the paper evaluates various auditing techniques, including sparse autoencoders, behavioral attacks, and training data analysis. This work highlights the importance of alignment audits in ensuring AI safety, providing a methodology to identify hidden objectives and improve pre-deployment safety practices for AI systems. | By Anthropic |  | March 13, 2025 |

| IV. AI USE CASES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 4.1 | **OpenAI introduces new tools for building AI agents.** | Google has introduced new APIs and tools to help developers build reliable AI agents efficiently. AI agents, autonomous systems that complete tasks independently, require advanced reasoning, multimodal interactions, and improved safety. However, many developers struggle with prompt tuning and orchestration. To address this, Google is launching the Responses API, merging Chat Completions and Assistants API, Built-in Tools for web search, file search, and computer use, Agents SDK for managing single and multi-agent workflows, and Observability Tools for monitoring agent processes. These innovations streamline AI agent development, making it more accessible and efficient for businesses and developers. | By OpenAI |  | March 11, 2025 |
| 4.2 | **AI Surveillance in Schools Raises Privacy Concerns** | Schools across the U.S., including Vancouver Public Schools in Washington, have implemented AI-powered surveillance software like Gaggle to monitor students' online activities on school-issued devices, aiming to prevent violence and address mental health issues. These systems flag potential dangers such as bullying, self-harm, or suicide, alerting school staff for intervention. However, this practice raises critical privacy and security issues, as demonstrated when reporters inadvertently accessed nearly 3,500 sensitive, unredacted student documents. This exposure highlights significant cybersecurity risks and the unintended consequences of surveillance, such as students being outed against their will or loss of trust between students and staff. The effectiveness of such surveillance in improving safety and mental health outcomes remains under scrutiny, with concerns about the long-term impact on student privacy and trust. | By Claire bryan and sharon lurye |  | March 12, 2025 |

| IV. AI USE CASES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 4.3 | **Gemini 2.0 Flash Introduces Native Image Generation** | Google's Gemini 2.0 Flash now supports native image generation, a feature that OpenAI's GPT-4o teased but never launched. This enables users to create and edit images directly, significantly improving text-to-image alignment over existing models. Early testers praise its ease of use, though they criticize Google's UI for making the feature difficult to access. With this release, Google takes the lead in integrated AI image editing, potentially disrupting tools like DALL·E and Midjourney | By Kat Kampf Nicole Brichtova |  | March 12, 2025 |
| 4.4 | **OpenAI Launches Developer Platform for AI Agents** | OpenAI has launched a developer platform aimed at facilitating the creation of AI agents. This platform equips developers with tools that enable AI agents to perform web and file searches, as well as execute web-based tasks similar to OpenAI's Operator browser. The initiative seeks to streamline the development process for AI applications capable of interacting with the internet and managing various online tasks, thereby enhancing the functionality and versatility of AI agents. | By HIStalk Team |  | March 12, 2025 |
| 4.5 | **NHS England Deploys AI Tool to Predict Patient Falls** | NHS England is implementing an AI tool developed by Cera, designed to predict a patient's risk of falling with 97% accuracy. Beyond assessing fall risk, this software is also utilized to anticipate deterioration in home care patients, aiming to enhance preventive measures and improve patient safety. By accurately identifying individuals at higher risk, the tool enables healthcare providers to intervene proactively, potentially reducing hospital admissions and improving overall patient outcomes. | By HIStalk Team |  | March 12, 2025 |
| 4.6 | **Oracle Introduces AI Agents to Combat Financial Crime** | Oracle Financial Services has enhanced its Investigation Hub Cloud Service with advanced AI capabilities to bolster the fight against financial crime. These agentic AI features aim to reduce manual workloads for financial institutions by | By Oracle |  | March 12, 2025 |

| IV. AI USE CASES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | automating complex investigative processes, thereby enabling faster detection and prevention of fraudulent activities. This development underscores Oracle's commitment to leveraging artificial intelligence to improve operational efficiency and strengthen compliance measures within the financial sector. | | | |
| 4.7 | **Singapore Airlines Implements AI-Powered Customer Service Platform** | Singapore Airlines has partnered with Salesforce to enhance its customer service operations through AI integration. The airline is incorporating Agentforce, Einstein in Service Cloud, and Data Cloud into its customer case management system to deliver more personalized service. Agentforce deploys autonomous agents to perform specific tasks, streamlining operations and allowing representatives to focus on providing enhanced attention in customer interactions. The system also leverages Einstein generative AI to summarize previous customer interactions and provide guidance, enabling representatives to better understand and anticipate customer needs while reducing response times | By TechNode Global Staff |  | March 13, 2025 |
| 4.8 | **CERAweek: AI leading to faster, cheaper oil production, executives say** | At the CERAWeek energy conference, industry executives highlighted how AI is revolutionizing oil and gas production, making it faster and more cost-efficient. Companies are using AI for predictive maintenance, drilling optimization, and reservoir management, leading to higher output and reduced expenses. Major firms like ExxonMobil and Chevron emphasized AI's role in boosting efficiency while lowering environmental impact. The technology enhances decision-making, minimizes downtime, and improves energy exploration. As AI adoption grows, experts predict a fundamental shift in how oil and gas operations are managed, increasing profitability and sustainability in the sector. | By Sheila Dang and Georgina Mccartney |  | March 14, 2025 |
| 4.9 | **PortaOne Addresses AI Monetization Challenges with New Platform** | At MWC Barcelona 2025, PortaOne showcased PortaAIM, a flexible charging and revenue management platform designed to help AI businesses optimize pricing, track costs, and scale profitably. CEO Andriy Zhylenko highlighted the critical gap between AI innovation and sustainable business models, noting that while AI agents are advancing rapidly, many companies fail to establish scalable | By Hennadiy Kornev |  | March 13, 2025 |

| IV. AI USE CASES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | | monetization frameworks. PortaOne's platform addresses this challenge by enabling seamless customer charging while maintaining cost control. The company has already implemented a proof-of-concept virtual AI agent that can retrieve customer-specific data in real-time, transforming customer service approaches | | | |
| 4.10 | **AAA to Showcase AI-Powered Innovations at Legalweek 2025** | The American Arbitration Association (AAA) is set to present its latest AI-driven advancements at Legalweek 2025, scheduled for March 24-27 in New York. AAA President & CEO Bridget McCormack will participate in a panel discussion titled “Disruption by Design: Shaping the AI-forward Firm of Tomorrow” on March 25, highlighting the organization's integration of AI into alternative dispute resolution (ADR) services. Attendees can explore AAA's AI initiatives at Booth 2311, including ClauseBuilder® AI (Beta), AAAi Chat Book Case Prep and Presentation (Beta), and the AAAiLab™. These developments underscore AAA's commitment to leveraging AI to enhance efficiency and accessibility in legal services. These developments underscore AAA's commitment to leveraging AI to enhance efficiency and accessibility in legal services. | By American Arbitration Association |  | March 13, 2025 |
| 4.11 | **Eclipse Foundation Introduces AI-Enhanced Open-Source IDE** | The Eclipse Foundation has released an alpha version of its open-source Theia integrated development environment (IDE), now enhanced with artificial intelligence (AI) capabilities. This AI-powered Theia IDE allows developers to integrate coding agents with various large language models (LLMs), facilitating tasks such as prompt engineering, defining agentic AI behaviors, and customizing user interfaces. This development aims to prevent vendor lock-in by enabling the use of multiple LLMs and supports interoperability through a Model Contextual Protocol (MCP) that connects AI-driven workflows with external tools and data sources. | By Mike Vizard |  | March 13, 2025 |
| 4.12 | **Open-source AI matches top proprietary model in** | A recent NIH-funded study led by Harvard Medical School researchers compared the performance of the open-source AI model Llama 3.1 405B with the proprietary GPT-4 in diagnosing complex medical cases. Analyzing 92 challenging scenarios from The New England Journal of Medicine, Llama 3.1 correctly diagnosed 70% of | By Harvard Medical School |  | March 14, 2025 |

| IV. AI USE CASES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | **solving tough medical cases** | cases, surpassing GPT-4's 64%. Notably, Llama 3.1 ranked the correct diagnosis first in 41% of cases, compared to GPT-4's 37%. This advancement highlights the growing competitiveness of open-source AI in healthcare, offering benefits like enhanced data privacy and customization, as these models can be tailored to specific clinical needs and operated within hospital infrastructures | | | |
| 4.13 | **AI-powered phishing attacks threaten Gmail, Outlook, and Apple Mail users** | The Forbes article highlights the escalating threat of AI-driven phishing attacks targeting users of popular email services like Gmail, Outlook, and Apple Mail. Cybercriminals are leveraging artificial intelligence to craft highly personalized and convincing emails, making it increasingly difficult for users to distinguish between legitimate messages and malicious ones. These sophisticated phishing attempts can lead to significant security breaches, including unauthorized access to sensitive information and financial losses. The article emphasizes the importance of heightened vigilance and the implementation of advanced security measures to protect against these evolving AI-generated threats. | By Zak Doffman |  | March 16, 2025 |
| 4.14 | **Chinese company's 'dark factory' will no human workers soon be the norm** | In Changping, China, Xiaomi's "dark factory" is revolutionizing manufacturing with a fully automated, AI-driven facility that operates 24/7 without human intervention. This facility produces one smartphone per second, using robots and AI systems for all production steps, from raw materials to assembly, ensuring precision and eliminating human error. With a capacity of 10 million devices annually, this 81,000-square meter factory represents the future of manufacturing. However, concerns about job displacement are rising, as the World Economic Forum predicts 23% of jobs will be affected by AI within five years. Experts call for global cooperation and regulation to ensure AI's safe and equitable benefits. | By Alex Blair |  | March 16, 2025 |
| 4.15 | **Anthropic Introduces Harmony, an AI** | Anthropic, a leading company in AI safety and research, is developing new features to expand the capabilities of its AI assistant, Claude. One of these | By Alexey Shabanov |  | March 15, 2025 |

| IV. AI USE CASES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | **Assistant Designed for Local File Access** | features, Harmony, allows users to integrate their local file directories into Claude's context, enabling the AI to read, index, and analyze these files. Harmony is a new feature aimed at enhancing user experience by enabling Claude to interact with local files seamlessly.This functionality offers significant advantages in AI-assisted code assistance and content analysis. Anthropic's Compass, another feature in development, aims to extend Claude's deep research capabilities. | | | |

| V. AI POLICIES, REGULATIONS & STRATEGIES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| 5.1 | **UK Government Advocates AI Integration in Civil Service** | Prime Minister Sir Keir Starmer has announced plans to integrate artificial intelligence into the UK civil service to enhance efficiency and reduce administrative costs. The initiative aims to cut regulatory burdens on businesses by 25%, addressing challenges faced by small businesses and families. Starmer emphasized that AI should handle tasks where possible, proposing that within five years, one in ten civil servants will occupy technical and digital roles, accompanied by a reduction in other areas of Whitehall. This digital transformation is projected to save taxpayers up to £45 billion by reducing regulatory burdens, though it may face legislative challenges and concerns about potential job losses. | By Oliver Wright |  | March 12, 2025 |
| 5.2 | **SAG-AFTRA Raises Concerns Over AI Usage in Video Game Industry Amid Ongoing Strike** | Members of SAG-AFTRA have been on strike for over seven months, addressing job security concerns related to AI replacing human roles in major video game studios. The union recently highlighted "alarming loopholes" in the latest bargaining proposal, which could allow companies to reuse union members' past work without consent. SAG-AFTRA urges its members to reject such projects, emphasizing the need for negotiations to protect against AI misuse. This action follows a prior strike focusing on AI's impact in television roles. | By Jacqueline Rayfield |  | March 12, 2025 |
| 5.3 | **From bureaucracy to brilliance: AI in federal IT** | Artificial intelligence (AI) is poised to revolutionize federal agencies by enhancing citizen services, improving decision-making, and bolstering national security. To fully harness AI's potential, a comprehensive strategy is essential, encompassing infrastructure modernization, robust security measures, and workforce development. Key trends include increasing AI fluency among federal employees, investing in sovereign AI to maintain data control, and understanding agentic AI for automating complex tasks. Adopting a holistic approach ensures that federal IT remains adaptable and resilient amidst rapid technological advancements. | By John Roese |  | March 12, 2025 |
| 5.4 | **OpenAI Urges Trump Administration to** | OpenAI is lobbying the Trump administration to focus on AI development speed while maintaining minimal regulatory constraints. The company warns that excessive oversight could stifle innovation, giving global competitors an advantage. Instead, OpenAI advocates for flexible policies that encourage rapid AI advancements while managing risks through industry-driven standards. | By Hayden Field |  | March 13, 2025 |

| V. AI POLICIES, REGULATIONS & STRATEGIES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | **Prioritize AI Growth Over Regulation** | Executives stress that the U.S. must lead in AI to maintain technological dominance. However, critics argue that weak regulations could lead to ethical issues, monopolization, and potential misuse. The debate reflects ongoing tensions between innovation, competition, and responsible oversight in the AI sector. | | | |
| 5.5 | **Infosys Research Calls for AI Governance Task Force** | As organizations transition from AI experimentation to scaled deployment, Infosys Knowledge Institute recommends creating dedicated AI governance task forces to reduce risk and improve accountability. Their latest survey, "Infosys AI Business Value Radar," reveals that only 19% of AI use cases fully deliver on business objectives, while another 32% partially meet goals. The research found that white-collar and technically focused industries achieve greater AI success, while sectors like travel, manufacturing, retail, and public services struggle with consistent implementation. The report outlines five critical steps for organizations to generate business value from AI deployments and become AI-first enterprises | By Uma Kannan |  | March 13, 2025 |
| 5.6 | **Google Outlines Three-Pronged Approach for U.S. AI Action Plan** | Google has submitted policy recommendations to the U.S. Office of Science and Technology Policy's Request for Information, outlining a framework to secure America's position as an AI leader. The proposal focuses on three key areas: investing in AI infrastructure, accelerating government AI adoption, and promoting pro-innovation approaches internationally. Google advocates for policy reforms addressing energy needs, balanced export controls, and preemption of fragmented state-level rules. The company emphasizes that policy decisions will significantly shape the outcome of global AI competition, calling for approaches that protect national security while enabling widespread benefits from AI advancements | By Kent Walker |  | March 13, 2025 |
| 5.7 | **USPTO in 2025: Leadership, Policy,** | The U.S. Patent and Trademark Office (USPTO) is undergoing significant changes in 2025, with new leadership and policy updates. Coke Morgan Stewart was appointed interim director, followed by Howard Lutnick as Commerce Secretary. | By Maria E. Doukas, Alexander B. |  | March 13, 2025 |

| V. AI POLICIES, REGULATIONS & STRATEGIES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | **and Legislative Changes to Watch** | John Squires has been nominated as USPTO's new director. The administration aims to reduce the patent backlog, but hiring freezes and return-to-office policies may hinder progress. Legislative efforts, including the Patent Eligibility Restoration Act (PERA) and PTAB reforms, are gaining attention. AI-related policy updates are also in focus. These changes will impact patent applicants and.. require strategic adjustments | Stein, Jacob L. Peterson | | |
| 5.8 | **US Expands AI Chip Export Controls Amid Rising Tech War with China** | The Biden administration has imposed new export controls on AI chips to limit China's access to advanced technology. These regulations aim to prevent China from using AI for military and surveillance purposes, reinforcing US national security policies. The rules categorize countries into tiers, with China facing the strictest restrictions. The policy reflects the growing role of AI in global power struggles, balancing innovation with security concerns. While US companies face challenges due to market restrictions, the move aligns with broader AI strategies to maintain technological leadership. | By Dashveenjit Kaur |  | March 14, 2025 |
| 5.9 | **AI Giants Must Not Exploit British Creatives, Industry Leaders Warn** | British media leaders and creative industry executives strongly oppose the UK government's proposed copyright reforms, which would let AI firms use copyrighted content freely unless creators opt out. Executives from NewsUK, The Guardian, Warner Music, and Channel 4 warned Technology Secretary Peter Kyle that this could destabilize the £125 billion creative sector.They argue that few countries allow unrestricted AI access to copyrighted works and that fair licensing deals exist without government interference. Figures like Sir Elton John, Simon Cowell, and Sir Paul McCartney stress transparency over sweeping changes, fearing AI firms could exploit UK intellectual property unfairly. | By, Martina Bet |  | March 14, 2025 |
| 5.10 | **An Integrated Approach is Essential** | The rapid expansion of artificial intelligence (AI) applications is driving increased demand for specialized AI semiconductors.These include GPUs for cloud | By IRA LEVENTHAL |  | March 11, 2025 |

| V. AI POLICIES, REGULATIONS & STRATEGIES | | | | | |
|---|---|---|---|---|---|
| # | Highlights | Summary | Author | Source | Date |
| | **for Testing AI Semiconductors** | computing, dedicated AI chips for efficiency, and neuromorphic chips for low-power edge applications. Market forecasts predict substantial growth, with revenues rising from $53.4 billion in 2023 to $341 billion by 2033. This diversity and complexity present unique testing challenges, such as increased scan-pattern depths and the need for efficient test data distribution. To address these issues, test processes must become dynamically adaptable, incorporating real-time decision-making to optimize cost, yield, and quality. This integrated approach is crucial for effectively testing AI semiconductors. | | | |
| 5.11 | **A Dynamic Governance Model for AI** | As AI transitions from an efficiency tool to a force shaping policy and governance, technologists are emerging as political actors. No longer just innovators, they now set agendas, draft guidelines, and influence AI's future—once the domain of elected officials. This shift raises concerns about democratic resilience and public oversight. AI's impact on economies, information, and security is shifting decision-making from governments to corporations. To balance innovation with public interest, governance structures must prevent control by a few dominant players while ensuring AI serves society democratically and ethically. Addressing these challenges is crucial for AI's responsible development. | By Paulo Carvão, Yam Atir, Salvina Ancheva |  | March 13, 2025 |
| 5.12 | **CCI Chief Raises Alarm on AI Collusion, Urges Proactive Regulations** | The Competition Commission of India (CCI) Chief, Ravneet Kaur, has warned about the risks of AI-driven collusion, highlighting concerns such as algorithmic price-fixing, cartelization without human communication, and discriminatory pricing under the guise of dynamic pricing. She emphasized the need for forward-looking regulations that ensure trust and fairness in AI-driven markets. As AI increasingly shapes market dynamics, regulators must adapt to prevent anti-competitive practices. The CCI is currently conducting a study on AI's impact on competition, aiming to develop regulatory frameworks that balance innovation with market integrity. | By Rediff Money Desk |  | March 16, 2025 |