Kempner
INSTITUTE

For the Study of Natural
& Artificial Intelligence
at Harvard University
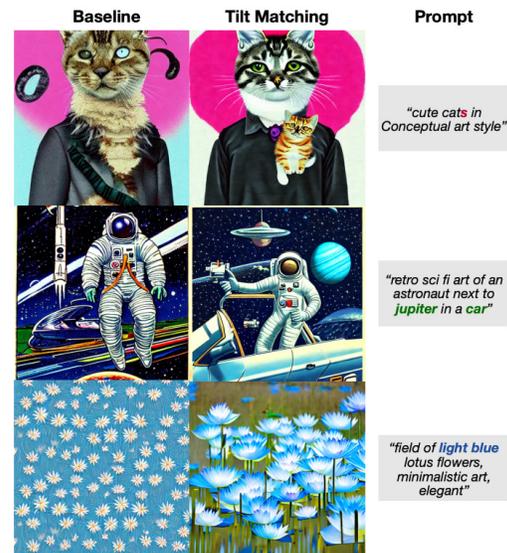
Research Roundup

January 2026

Discover our latest insights in natural and artificial intelligence research

Highlights

- A team led by Graduate Fellow Colton Casto published a paper in *Neuron* on the **cerebellar components of the human language network**.

- Associate Faculty Marinka Zitnik and her team introduced **Magneton**, an environment for developing substructure-aware protein models.

- A team including Associate Faculty Cengiz Pehlevan published a paper in *Nature* demonstrating experimental evidence of **predictive coding of reward in the hippocampus.**

Featured Figure

Baseline	Tilt Matching	Prompt
		"cute cats in Conceptual art style"
		"retro sci fi art of an astronaut next to jupiter in a car"
		"field of light blue lotus flowers, minimalistic art, elegant"

Examples of improvements to Stable Diffusion model outputs via **Tilt Matching**, a new method introduced by incoming Institute Investigator Michael Albergo and collaborators.

Featured Formula

$$\overbrace{h_{t+1}(\nu)}^{\substack{\text{Next Latent}\\\text{Prediction}}} = T_{a_t}^{-1} \underbrace{\overbrace{\psi_1(\nu)}^{\substack{\text{Internal Flow}\\\text{Transform}}}}_{\substack{\text{Self Action}\\\text{Transform}}} \cdot \mathrm{U}_\theta \overbrace{\underbrace{\left[h_t(\nu)\right.}_{\substack{\text{Update}\\\text{Memory}}}; \underbrace{\left. E_\theta[f_t, h_t](\nu)\right]}_{\text{Encoder Output}}}^{\substack{\text{Latent Defined}\\\text{over }\nu}}.$$

Self-Motion Flow Equivariant Recurrence Relation, from the preprint **Flow Equivariant World Models: Memory for Partially Observed Dynamic Environments.**

From the abstract: "By structuring world model representations with respect to internal and external motion, flow equivariance charts a scalable route to data efficient, symmetry-guided, embodied intelligence."

Featured Artifacts

Greater than the Sum of Its Parts: Building Substructure into Protein Encoding Models.
Robert Calef, Arthur Liang, Manolis Kellis, and Marinka Zitnik.

Get the Code ▶

Flow Equivariant World Models: Memory for Partially Observed Dynamic Environments.
Hansen Jin Lillemark, Benhao Huang, Fangneng Zhan, Yilun Du, and T. Anderson Keller.

Get the Code ▶

kempnerinstitute.harvard.edu

Kempner
INSTITUTE

For the Study of Natural
& Artificial Intelligence
at Harvard University

AI innovation

A note on the dynamics of extended-context disordered kinetic spin models.
Jacob A. Zavatone-Veth and Cengiz Pehlevan. Journal of Physics. A, Mathematical and Theoretical 59(4) :045001 (2026).

Investigating the Development of Task-Oriented Communication in Vision-Language Models.
Boaz Carmeli, Orr Paradise, Shafi Goldwasser, Yonatan Belinkov, and Ron Meir. arXiv:2601.20641v1 (2026).

Hyperparameter Transfer with Mixture-of-Expert Layers.
Tianze Jiang, Blake Bordelon, Cengiz Pehlevan, and Boris Hanin. arXiv:2601.20205v1 (2026).

Meta Flow Maps enable scalable reward alignment.
Peter Potaptchik, Adhi Saravanan, Abbas Mammadov, Alvaro Prat, Michael S. Albergo, and Yee Whye Teh. arXiv:2601.14430v1 (2026).

Circuit Mechanisms for Spatial Relation Generation in Diffusion Transformers.
Binxu Wang, Jingxuan Fan, and Xu Pan. arXiv:2601.06338v1 (2026).

Bi-Orthogonal Factor Decomposition for Vision Transformers.
Fenil R. Doshi, Thomas Fel, Talia Konkle, and George Alvarez. arXiv:2601.05328v1 (2026).

Flow Equivariant World Models: Memory for Partially Observed Dynamic Environments.
Hansen Jin Lillemark, Benhao Huang, Fangneng Zhan, Yilun Du, and T. Anderson Keller. arXiv:2601.01075v1 (2026).

Max-Entropy Reinforcement Learning with Flow Matching and A Case Study on LQR.
Yuyang Zhang, Yang Hu, Bo Dai, and Na Li. arXiv:2512.23870v1 (2025).

GQ-VAE: A gated quantized VAE for learning variable length tokens.
Theo Datta, Kayla Huang, Sham Kakade, and David Brandfonbrener. arXiv:2512.21913v1 (2025).

Tilt Matching for Scalable Sampling and Fine-Tuning.
Peter Potaptchik, Cheuk-Kit Lee, and Michael S. Albergo. arXiv.:2512.21829v1 (2025).

Block-Recurrent Dynamics in Vision Transformers.
Mozes Jacobs, Thomas Fel, Richard Hakim, Alessandra Brondetta, Demba Ba, and T. Anderson Keller. arXiv:2512.19941v1 (2025

Applications of AI

Autonomous Knowledge Graph Exploration with Adaptive Breadth-Depth Retrieval.
Joaquín Polonuer, Lucas Vittor, Iñaki Arango, Ayush Noori, David A. Clifton, and Marinka Zitnik. arXiv:2601.13969v1 (2026).

Greater than the Sum of Its Parts: Building Substructure into Protein Encoding Models.
Robert Calef, Arthur Liang, Manolis Kellis, and Marinka Zitnik. arXiv: 2512.18114v1 (2025).

Interpretability and AI theory

Will it Merge? On The Causes of Model Mergeability.
Adir Rahamim, Asaf Yehudai, Boaz Carmeli, Leshem Choshen, Yosi Mass, and Yonatan Belinkov. arXiv:2601.06672v1 (2026).

Disordered Dynamics in High Dimensions: Connections to Random Matrices and Machine Learning.
Blake Bordelon and Cengiz Pehlevan. arXiv:2601.01010v2 (2026).

Demystifying LLM-as-a-Judge: Analytically Tractable Model for Inference-Time Scaling.
Indranil Halder and Cengiz Pehlevan. arXiv.:2512.19905v1 (2025).

Neuroscience & Cognitive Science

Emotion may indirectly link rendering and social reasoning.
Halely Balaban and Tomer D. Ullman. Trends in Cognitive Sciences S1364- 6613(25)00358-4 (2026).

The cerebellar components of the human language network.
Colton Casto, Moshe Poliak, Greta Tuckute, Hannah Small, Patrick Sherlock, Agata Wolna, Benjamin Lipkin, Anila M. D'Mello, and Evelina Fedorenko. Neuron:S0896- 6273(25)00986-9 (2026).

Predictive coding of reward in the hippocampus.
Mohammad Yaghoubi, M. Ganesh Kumar, Andres Nieto-Posadas, Coralie-Anne Mosser, Thomas Gisiger, Émmanuel Wilson, Cengiz Pehlevan, Sylvain Williams, and Mark P. Brandon. Nature (2026).

Credit Assignment via Neural Manifold Noise Correlation.
Byungwoo Kang, Maceo Richards, and Bernardo Sabatini. arXiv:2601.02636v1 (2026).

Note: This is a partial list of articles and preprints published by Kempner-affiliated researchers in the last month. Papers are listed by topic and publication/upload date, with the most recent first.

View this report and other Kempner Research-Round-ups:
kempnerinstitute.harvard.edu/kempner-community/research-roundup/