



Research Roundup

May 2026

Discover our latest insights
in natural and artificial
intelligence research

- ⬡ Kempner Graduate Fellow Shuze Liu, Associate Faculty Samuel Gershman and Ph.D. student Yang Xiang published a paper in *Psychological Review* entitled "**Probabilistic Forecasting Guides Dynamic Decisions.**"
- ⬡ Kempner Associate Faculty Patrick Slade and his team published a paper entitled "**Estimating and Interpreting How Humans Prioritize Multiple Movement Goals During Walking**" in the *Journal of NeuroEngineering and Rehabilitation*.
- ⬡ In a new preprint, Kempner researchers introduced the "**Recurrent Transformer,**" a modified transformer architecture that "improve cross-entropy over a parameter-matched Transformer baseline and achieve the improvement with fewer layers (fixed parameters)."

Featured Figure



Figure from the preprint "**ArchSym: Detecting 3D-Grounded Architectural Symmetries in the Wild.**"

From a single RGB image (left in each pair), the researchers' model recovers dominant 3D symmetry planes (right) even when they are partially occluded or not directly visible.

Featured Formula

$$d^L = \frac{1}{N} \sum_i \mathbb{E} \left(\left| \sum_j \left(\frac{\partial z_i^L}{\partial z_j^0} z_j^0 - \frac{\partial z_i^L}{\partial z_j^{0\dagger}} z_j^{0\dagger} \right) \right|^2 \right)$$

Formula characterizing an order parameter that arises from considering how an infinitesimal symmetry transformation is propagated through a type of equivariant deep neural network. From the preprint "**Spontaneous Symmetry Breaking and Goldstone Modes for Deep Information Propagation.**"

Featured Artifacts

ArchSym: Detecting 3D-Grounded Architectural Symmetries in the Wild.

Hanyu Chen, Ruojin Cai, Steve Marschner, and Noah Snavely. (2026).

[Get the Code](#) ▶

OptimusKG: Unifying Biomedical Knowledge in a Modern Multimodal Graph.

Lucas Vittor, Ayush Noori, Iñaki Arango, Joaquín Polonuer, Sam Rodrigues, Andrew White, David A. Clifton, and Marinka Zitnik (2026).

[Get the Code](#) ▶

AI Innovation

Less Data, Faster Training: Repeating Smaller Datasets Speeds Up Learning via Sampling Biases.

Jingwen Liu, Ezra Edelman, Surbhi Goel, and Bingbin Liu. arXiv:2605.20314 (2026).

The Two Clocks and the Innovation Window: When and How Generative Models Learn Rules.

Binxu Wang, Emma Lucia Byrnes Finn, and Bingbin Liu. arXiv:2605.10019v1 (2026).

Decentralized Diffusion Policy Learning for Enhanced Exploration in Cooperative Multi-Agent Reinforcement Learning.

Yuyang Zhang, Haldun Balim, and Na Li. arXiv:2605.07101v1 (2026).

Dynamic Vine Copulas: Detecting and Quantifying Time-Varying Higher-Order Interactions.

Houman Safaai and Alessandro Marin Vargas. arXiv:2605.03061v2 (2026).

A Unified Perspective on Fine-Tuning and Sampling with Diffusion and Flow Models.

Carles Domingo-Enrich, Yuanqi Du, and Michael S. Albergo. arXiv:2605.00229v1 (2026).

Long-Tail Internet Photo Reconstruction.

Yuan Li, Yuanbo Xiangli, Hadar Averbuch-Elor, Noah Snavely, and Ruojin Cai. arXiv:2604.22714v1 (2026).

The Recurrent Transformer: Greater Effective Depth and Efficient Decoding.

Costin-Andrei Oncescu, Depen Morwani, Samy Jelassi, Alexandru Meterez, Mujin Kwun, and Sham Kakade. arXiv:2604.21215v1 (2026).

Amortized Vine Copulas for High-Dimensional Density and Information Estimation.

Houman Safaai. arXiv:2604.20568v2 (2026).

Interpretability and AI Theory

Artificial Intelligence for Science: The Easy and Hard Problems.

Ruairidh Battleday and Sam Gershman. Philosophical Transactions of the Royal Society A 384(2320):20240530. (2026).

Spontaneous Symmetry Breaking and Goldstone Modes for Deep Information Propagation.

Nabil Iqbal, T. Anderson Keller, Yue Song, Takeru Miyato, and Max Welling. arXiv:2605.14685v1 (2026).

Interpretability Can Be Actionable.

Hadas Orgad, Fazl Barez, Tal Haklay, Isabelle Lee, Marius Mosbach, Anja Reusch, Naomi Saphra, Byron Wallace, Sarah Wiegrefe, Eric Wong, Ian Tenney, and Mor Geva. arXiv:2605.11161v1 (2026).

Interpretability and AI Theory

Spectral Dynamics in Deep Networks: Feature Learning, Outlier Escape, and Learning Rate Transfer.

Clarissa Lauditi, Cengiz Pehlevan, and Blake Bordelon. arXiv:2605.07870v1 (2026).

Task Relevance Is Not Local Replaceability: A Two-Axis View of Channel Information.

Houman Safaai, Andrew T. Landau, Celia C. Beron, Yasin Mazloumi, and Bernardo L. Sabatini. arXiv:2605.07086v1. (2026).

Differentiable Faithfulness Alignment for Cross-Model Circuit Transfer.

Shun Shao, Binxu Wang, Shay B. Cohen, Anna Korhonen, and Yonatan Belinkov. arXiv:2604.24302v1 (2026).

A Dimensional R2 Regression Metric.

Jaesung Yoo, Stefan Lemke, Jian Zhong Guo, Kanaka Rajan, and Adam Hantman. arXiv:2605.01066v1. (2026).

Supernodes and Halos: Loss-Critical Hubs in LLM Feed-Forward Layers.

Audrey Cherilyn and Houman Safaai. arXiv:2604.23475 (2026).

Linear Equivalence of Nonlinear Recurrent Neural Networks.

David G. Clark. arXiv:2604.23489v2 (2026).

Reasoning Models Know What's Important, and Encode It in Their Activations.

Yaniv Nikankin, Martin Tutek, Tomer Ashuach, Jonathan Rosenfeld, and Yonatan Belinkov. arXiv:2604.18307v1 (2026).

Applications of AI

DT-Transformer: A Foundation Model for Disease Trajectory Prediction on a Real-World Health System.

Yunying Zhu, Andrew R. Weckstein, Kueiyu Joshua Lin, and Jie Yang. arXiv:2605.14227v1 (2026).

ArchSym: Detecting 3D-Grounded Architectural Symmetries in the Wild.

Hanyu Chen, Ruojin Cai, Steve Marschner, and Noah Snavely. arXiv:2604.22202v1 (2026).

OptimusKG: Unifying Biomedical Knowledge in a Modern Multimodal Graph.

Lucas Vittor, Ayush Noori, Iñaki Arango, Joaquín Polonuer, Sam Rodrigues, Andrew White, David A. Clifton, and Marinka Zitnik. arXiv:2604.27269v1 (2026).

Towards a General-Purpose Foundation Model for Functional MRI Analysis.

Cheng Wang, Yu Jiang, Zhihao Peng, Chenxin Li, Chang-Bae Bang, Lin Zhao, Wanyi Fu, et al. Nature Biomedical Engineering (2026).



Neuroscience & NeuroAI

Multilingual Computational Models Capture a Shared Meaning Component in Brain Responses Across 21 Languages.

Andrea Gregor de Varda, Saima Malik-Moraleda, Greta Tuckute, and Evelina Fedorenko. bioRxiv:2025.02.01.636044 (2026).

Decoding Alignment Without Encoding Alignment: A Critique of Similarity Analysis in Neuroscience.

Johannes Bertram, Luciano Dyballa, T. Anderson Keller, Savik Kinger, and Steven W. Zucker. arXiv:2605.05907v1 (2026).

Probabilistic Forecasting Guides Dynamic Decisions.

Shuze Liu, Yang Xiang, and Samuel J. Gershman. Psychological Review (2026).

NeuroClaw Technical Report.

Cheng Wang, Zhibin He, Zhihao Peng, Shengyuan Liu, Yufan Hu, Yang Carl, He Lifang, Lichao Sun, Xiang Li, and Yixuan Yuan. arXiv:2604.24696v2 (2026).

Perception and Neural Representation of Intermittent Odor Stimuli in Mice.

Luis E. Boero, Hao Wu, Joseph D. Zak, Paul Masset, Farhad Pashakhanloo, Siddharth Jayakumar, Bahareh Tolooshams, Demba Ba, and Venkatesh N. Murthy. Nature Communications (2026).

Centralized Brain Networks Controlling Antennal Grooming Coordination.

Pembe Gizem Özdil, Jonathan Arreguit, Clara Scherrer, Femke Hurtak, Auke Ijspeert, and Pavan Ramdya. Nature Communications (2026).

Estimating and Interpreting How Humans Prioritize Multiple Movement Goals During Walking.

Jordan N. Feldman, David Morales Loro, Andrew Chin, Sarah Erickson, Sarah Li, Eliana Bonbrest, and Patrick Slade. Journal of NeuroEngineering and Rehabilitation (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/