Ari S. Morcos

SENIOR STAFF RESEARCH SCIENTIST, META AI RESEARCH, FAIR TEAM

☐ (+1) 818-481-9555 | ✉ arimorcos@gmail.com

🏠 www.arimorcos.com | ⬡ arimorcos | in arimorcos | 🐦 arimorcos

Research interests

Machine Learning and Deep Learning

My main research focus is to design and execute rigorous experiments to understand the solutions found by deep neural networks and most critically, their bottlenecks, so that we can intelligently design machine learning systems. I am also interested in developing methods to measure and induce abstract representations in neural networks.

Education

Harvard University *Cambridge, Massachusetts, USA*
PH.D. IN NEUROBIOLOGY *July 2011 - April 2016*
- Advisor: Christopher Harvey.
- Thesis: Population dynamics in parietal cortex during evidence accumulation for decision-making.

University of California, San Diego *La Jolla, California, USA*
B.S. IN PHYSIOLOGY AND NEUROSCIENCE *September 2008 - March 2011*

Work and Research Experience

Meta AI Research, FAIR Team *Menlo Park, CA*
SENIOR STAFF RESEARCH SCIENTIST *August 2023 - PRESENT*
STAFF RESEARCH SCIENTIST *March 2020 - August 2023*
SENIOR RESEARCH SCIENTIST *September 2018 - March 2020*
- Worked on varied ML topics including data pruning, lottery tickets, and model compression
- Research Manager

DeepMind *London, UK*
RESEARCH SCIENTIST *September 2016 - August 2018*
- Used neuroscience-inspired techniques to understand and improve deep neural networks.
- Investigated properties of learned network solutions which are predictive of generalization ability, both in individual networks and across groups of networks.
- Developed new behavioral and introspective methods to measure, understand, and improve abstraction abilities in deep neural networks.

Harvard University *Cambridge, Massachusetts, USA*
GRADUATE RESEARCH WITH CHRISTOPHER HARVEY *July 2011 - April 2016*
- Developed an evidence accumulation task for head-restrained mice in virtual reality and used calcium imaging of activity in large neuronal populations along with novel computational analyses to study the neuronal population activity dynamics underlying decision-making.
- Using machine learning approaches that had not been applied to neuroscience data sets previously, we identified multiple major features of the population activity, including a distributed code among heterogeneous individual neurons, history signals for past events lasting seconds, and structured trial-trial variability that was predictive of past and future activity patterns.
- Together, our results provide data that contradict the long-standing models of evidence accumulation and that propose a novel model of computation based on large-scale neuronal population dynamics.

Salk Institute for Biological Studies *La Jolla, California, USA*
UNDERGRADUATE RESEARCH WITH FRED H. GAGE *January 2009 - December 2010*
- Investigated the role of REST in adult neurogenesis.

Publications

Bordes F, Shekhar S, Ibrahim M, Bouchacourt D, Vincent P, and **Morcos AS**. "PUG: Photorealistic and Semantically Controllable Synthetic Data for Representation Learning." *arXiv preprint.*, 2023.

Tirumala K*, Simig D*, Aghajanyan A, and **Morcos AS**. "D4: Improving LLM Pretraining via Document De-Duplication and Diversification." *arXiv preprint.*, 2023.

Shekhar S, Bordes F, Vincent P, and **Morcos AS**. "Objectives Matter: Understanding the Impact of Self-Supervised Objectives on Vision Transformer Representations." *arXiv preprint.*, 2023.

Wortsman M*, Dettmers T*, Zettlemoyer L, **Morcos AS**, Farhadi A, and Schmidt L. "Stable and low-precision training for large-scale vision-language models." *arXiv preprint.*, 2023.

Ranadive O, Thakurdesai N, **Morcos AS**, Leavitt M, and Deny S. "On the special role of class-selective neurons in early training." *arXiv preprint.*, 2023.

Kira S, Safaii H, **Morcos AS**, Panzeri S, and Harvey CD. "A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions." *Nature Communications*, 2023; 14(2121).

Balestiero R, Ibrahim M, Sobal V*, **Morcos AS***, Shekhar S*, Goldstein T*, Bordes F*, Bardes A*, Mialon G*, Tian Y*, Schwarzschild A*, Gordon Wilson A*, Geiping J*, Garrido Q*, Fernandez P*, Bar A*, Pirsiavash H*, LeCun Y*, and Goldblum M*. "A Cookbook of Self-Supervised Learning." *arXiv preprint.*, 2023.

Abbas A, Tirumala K, Simig D, Ganguli S, and **Morcos AS**. "SemDeDup: Data-efficient learning at web-scale through semantic deduplication." *arXiv preprint.*, 2023. ***Best Paper Award*** at ICLR Multimodal Representation Learning Workshop.

Wijmans, E, Savva, M, Essa I, Lee S **Morcos AS**, and Batra, D. "Emergence of Maps in the Memories of Blind Navigation Agents." *International Conference on Learning Representations (ICLR)*, 2023. ***Notable - Top 5% (Oral), Outstanding Paper Award***.

Ibrahim M, Garrido Q, **Morcos AS**, and Bouchacourt D. "The Robustness Limits of SoTA Vision Models to Natural Variation." *Transactions on Machine Learning Research (TMLR)*, 2023.

Sorscher B*, Geirhos R*, Shekhar S, Ganguli S, and **Morcos AS**. "Beyond neural scaling laws: beating power law scaling via data pruning." *Neural Information Processing Systems (NeurIPS)*, 2022. ***Oral, Outstanding Paper Award***.

Wortsman M, Gururangan S, Li S, Farhadi A, Schmidt L, Rabbat M, **Morcos AS**. "lo-fi: distributed fine-tuning without communication." *Transactions on Machine Learning Research (TMLR)*, 2022.

Ibrahim M*, Bouchacourt D*, **Morcos AS**. "Robust Self-Supervised Learning with Lie Groups." *arXiv preprint.*, 2022.

Wortsman M, Ilharco G, Gadre SY, Roelofs R, Gontijo-Lopes R, **Morcos AS**, Namkoong H, Farhadi A, Carmon Y, Kornblith S, and Schmidt L. "Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time." *International Conference of Machine Learning (ICML)*, 2022.

Xie S, Zhu S, **Morcos AS**, and Vedantam R. "COAT: Measuring Object Compositionality in Emergent Representations." *International Conference of Machine Learning (ICML)*, 2022.

Kira S, Safaai H, **Morcos AS**, Harvey CD, and Panzeri S. "A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions." *biorXiv preprint*, 2022.

Gonzalez Ortiz JJ, Frankle J, Rabbat M, **Morcos AS**, and Ballas N. "Trade-offs of Local SGD at Scale: An Empirical Study." *arXiv preprint.*, 2021.

Chin TW, **Morcos AS**, and Marculescu D. "Joslim: Joint Widths and Weights Optimization for Slimmable Neural NetworksRepresentations." *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 2021.

Shen S, Baevski A, **Morcos AS**, Keutzer K, Auli M, and Kiela D. "Reservoir Transformers." *Association of Computational Linguistics (ACL)*, 2021.

Bouchacourt D, Ibrahim M, and **Morcos AS**. "Grounding inductive biases in natural images: invariance stems

from variations in data." *Neural Information Processing Systems (NeurIPS)*, 2021.

Valente M, Pica P, Bondanelli G, Moroni M, Runyan CA, **Morcos AS**, Harvey CD, and Panzeri S. "Correlations enhance the behavioral readout of neural population activity in association cortex." *Nature Neuroscience*, 24, 975–986 (2021).

d'Ascoli S, Biroli G, Sagun L, and **Morcos AS**. "Transformed CNNs: recasting pre-trained convolutional layers with self-attention." *arXiv*, 2021.

Ryali CK, Schwab DJ, and **Morcos AS**. "Characterizing and Improving the Robustness of Self-Supervised Learning through Background Augmentations." *arXiv*, 2021.

d'Ascoli S, Touvron H, Leavitt M, **Morcos AS**, Biroli G, and Sagun L. "ConViT: Improving Vision Transformers with Soft Convolutional Inductive Biases." *International Conference of Machine Learning (ICML)*, 2021.

Vedantam R, Szlam A, Nickel M, **Morcos AS**, and Lake B. "CURI: A Benchmark for Productive Concept Learning Under Uncertainty." *International Conference of Machine Learning (ICML)*, 2021.

Leavitt M and **Morcos AS**. "Selectivity considered harmful: evaluating the causal impact of class selectivity in DNNs." *International Conference on Learning Representations*, 2021.

Frankle J, Schwab DJ, and **Morcos AS**. "Training BatchNorm and Only BatchNorm: On the Expressive Power of Random Features in CNNs." *International Conference on Learning Representations*, 2021.

Cai T, Frankle J, Schwab DJ, and **Morcos AS**. "Are all negatives created equal in contrastive learning?" *Science Meets Engineering of Deep Learning @ ICLR*, 2021.

Leavitt M* and **Morcos AS***. "Towards falsifiable interpretability." *ML Retrospectives @ NeurIPS*, 2020.

Bartoldson BR, **Morcos AS**, Barbu A, and Erlebacher G. "The Generalization-Stability Tradeoff In Neural Network Pruning." *Neural Information Processing Systems (NeurIPS)*, 2020.

Chin TW, **Morcos AS**, and Marculescu D. "PareCO: Pareto-aware Channel Optimization for Slimmable Neural Networks." *CVPR Workshops*, 2021.

Litany O, **Morcos AS**, Sridhar S, Guibas L, and Hoffman J. "Representation learning through latent canonicalizations." *Winter Conference on Applications of Computer Vision (WACV)*, 2021.

Valente M, Pica G, Runyan CA, **Morcos AS**, Harvey CD, and Panzeri S. "Correlations enhance the behavioral readout of neural population activity in association cortex." *Nature Neuroscience 24, 975–986*, 2021.

Yang G, Zhang A, **Morcos AS**, Pineau J, Abbeel P, Parikh D, and Calandra R. "Plan2Vec: Unsupervised Representation Learning by Latent Plans." *Learning for Dynamics and Control (L4DC)*, 2020.

Yu H, Edunov S, Tian Y, and **Morcos AS**. "Playing the lottery with rewards and multiple languages: lottery tickets in RL and NLP." *International Conference on Learning Representations (ICLR)*, 2020.

Leavitt M and **Morcos AS**. "Linking average- and worst-case perturbation robustness via class selectivity and dimensionality." *Workshop on Uncertainty and Robustness in Deep Learning @ ICML*, 2020.

Frankle J, Schwab DJ, and **Morcos AS**. "The early phase of neural network training." *International Conference on Learning Representations (ICLR)*, 2020.

Wijmans E, Kadian A, **Morcos AS**, Lee S, Essa I, Parikh D, Savva M, and Batra D. "DD-PPO: Learning Near-Perfect PointGoal Navigators from 2.5 Billion Frames." *International Conference on Learning Representations (ICLR)*, 2020.

Wijmans E, Straub J, Batra D, Essa I, Hoffman J, **Morcos AS**. "Analyzing Visual Representations in Embodied Navigation Tasks." *arXiv preprint*.

2020 Caron M, **Morcos AS**, Bojanowski P, Mairal J, and Joulin A. "Pruning convolutional neural networks with self-supervision." *arXiv preprint*.

Morcos AS, Yu H, Paganini M, and Tian Y. "One ticket to win them all: generalizing lottery ticket initializations across datasets and optimizers." *Neural Information Processing Systems (NeurIPS)*, 2019.

Tian Y, Jiang T, Gong Q, and **Morcos AS**. "Luck matters: understanding training dynamics of deep ReLU

networks." *arXiv preprint*.

Hill F*, Santoro A*, Barrett DGT, **Morcos AS**, and Lillicrap T. "Learning to make analogies by contrasting abstract relational structure." *International Conference on Learning Representations (ICLR)*, 2019.

Barrett DGT*, **Morcos AS***, and Macke JH. "Analyzing biological and artificial neural networks: challenges with opportunities for synergy?" *Current Opinion in Neurobiology*, Vol. 55, pp 55-64, 2019. doi: https://doi.org/10.1016/j.conb.2019.01.007..

Morcos AS*, Raghu M, and Bengio S. "Insights on representational similarity in neural networks with canonical correlation." *Neural Information Processing Systems (NeurIPS)*, 2018.

Jaderberg M*, Czarnecki WM*, Dunning I*, Marris L, Lever G, Castaneda AG, Beattie C, Rabinowitz NC, **Morcos AS**, Ruderman A, Sonnerat N, Green T, Deason L, Leibo JZ, Silver D, Hassabis D, Kavukcuoglu K, and Graepel T. "Human-level performance in first-person multiplayer games with population-based deep reinforcement learning." *Science*, 31 May 2019; Vol. 364, Issue 6443, pp. 859-865.

Eslami SMA*, Rezende DJ*, Besse F, Viola F, **Morcos AS**, Garnelo M, Ruderman A, Rusu AA, Danihelka I, Gregor K, Reichert DP, Beusing L, Weber T, Vinyals O, Rosenbaum D, Rabinowitz NC, King H, Hillier C, Botvinick M, Wierstra D, Kavukcuoglu K, and Hassabis D. "Neural scene representation and rendering." *Science*, 15 Jun 2018; Vol. 360, Issue 6394, pp. 1204-1210.
    Author contribution: A.S.M. designed and performed analysis experiments.

Ruderman A, Rabinowitz NC, **Morcos AS**, and Zoran D. "Pooling is neither necessary nor sufficient for appropriate deformation stability in CNNs." *arXiv preprint*.

Barrett DGT*, Hill F*, Santoro A*, **Morcos AS**, and Lillicrap T. "Measuring abstract reasoning in neural networks." *International Conference on Machine Learning (ICML)*, 2018. ***Selected for a long talk***.

Morcos AS, Barrett DGT, Rabinowitz NC, and Botvinick M. "On the importance of single directions for generalization." *International Conference on Learning Representations (ICLR)*, 2018.

Morcos AS and Harvey CD. "History-dependent variability in population dynamics during evidence accumulation in cortex." *Nature Neuroscience*, 2016; 19(12):1672-1681.

Morcos AS. "Population dynamics in parietal cortex during evidence accumulation for decision-making." *PhD Thesis, Harvard University*, April, 2016. https://dash.harvard.edu/handle/1/33493459.

Kim HJ, Denli AM, Wright R, Baul TD, Clemenson GD, **Morcos AS**, Zhao C, Schafer ST, Gage FH, and Kagalwala MN. "REST Regulates Non-Cell-Autonomous Neuronal Differentiation and Maturation of Neural Progenitor Cells via Secretogranin II." *Journal of Neuroscience*, 2015 Nov 4;35(44):14872-84.

Morcos AS. "Mechanisms and applications of adult neurogenesis." *Saltman Quarterly*, 2009 (6):35-36. .

*indicates equal contribution

Awards, Honors, & Fellowships

2023	**Multimodal Representation Learning Workshop @ ICLR Best Paper Award**
	SemDeDup: Data-efficient learning at web-scale through semantic deduplication
2023	**ICLR Outstanding Paper Award**
	Emergence of Maps in the Memories of Blind Navigation Agents *Top 4/ 4900*
2022	**NeurIPS Outstanding Paper Award**
	Beyond neural scaling laws: beating power law scaling via data pruning *Top 15/ 10,700*
2013-2016	**Stuart and Victoria Quan Pre-Doctoral Fellowship**
	Harvard Medical School

Community

Reviewer
Neural Information Processing Systems (NeurIPS), 2018, 2019
International Conference on Learning Representations (ICLR), 2019
International Conference on Machine Learning (ICML), 2019, 2020 (top reviewer), 2021
Transactions on Pattern Analysis and Machine Intelligence
Nature
Nature Communications
Nature Machine Intelligence

Area chair
International Conference on Artificial Intelligence and Statistics (AISTATS), 2020
Neural Information Processing Systems (NeurIPS), 2020, 2021, 2022, 2023
International Conference on Machine Learning (ICML), 2023
International Conference on Learning Representations (ICLR), 2020, 2021, 2022, 2023

Workshop organization
DataComp, ICCV 2023
Sparsity in Neural Networks: Advancing Understanding and Practice, 2021, 2022
Identifying and Evaluating Deep Learning Phenomena, ICML 2019