Andrei Barbu

(315) 333-0621 andrei@0xab.com abarbu@csail.mit.edu January 22, 2025 http://0xab.com

Academic Employment and Education

Research Scientist *2016 – present*
Massachusetts Institute of Technology, CSAIL

Part of the MIT/Harvard Center for Brains, Minds, and Machines
Toyota Research Institute grant *Using vision and language to read minds*
MIT/IBM Laboratory for Brain-inspired Multimedia Machine Comprehension (BM^{3}C)
Research on grounding language in perception and robotics. Robots that understand commands and learn through linguistic interactions. Physics and planning integrated with visual understanding. Understanding the limits of human and machine vision.

Postdoctoral Associate *2013 – 2016*
Massachusetts Institute of Technology, CSAIL

Part of the MIT/Harvard Center for Brains, Minds, and Machines
Research on vision and language. How grounding language in perception allows us to understand human cognition and the limits of language understanding.

PhD, Artificial Intelligence *2008 – 2013*
Purdue University, School of Electrical and Computer Engineering

Title: Reasoning across language and vision in machines and humans
Advisor: Jeffrey Mark Siskind

Top performers in DARPA *Mind's Eye*

Bachelor of Computer Science *2004 – 2008*
University of Waterloo

Employment History

Software Development Engineer *May 2010 – Aug. 2010*
Amazon.com Seattle, USA

migrated from a monolithic database design to a distributed service-oriented system
designed & developed distributed services to collect metrics about internal systems
coordinated with other teams to develop internal tools and alarms

Kernel Developer *May 2008 – Aug. 2008*
Google Summer of Code 2008 – the GNU Hurd Purdue University, USA

designed and implemented an instrumentation framework for GNU Mach
implemented the DWARF2 debugging standard
debugger improvements
bug fixes and other new features

Software Developer and Knowledge Systems Architect *July 2005 – May 2006*
Dalin Software SA Laussane, Switzerland

ontology, temporal, and spatial logic frameworks
designed extensions to OWL, the Ontology Web Language
designed and developed tools for querying and updating OWL+SWRL ontologies

Publications

Population Transformer: Learning Population-level Representations of Intracranial Activity *April 2025*
G. Chau, C. Wang, S. Talukder, V. Subramaniam, S. Soedarmadji, Y. Yue, B. Katz, A. Barbu P1
International Conference on Learning Representations (ICLR)

Brain Treebank: Large-scale intracranial recordings from naturalistic language stimuli *December 2024*
V. Subramaniam, C. Wang, A. U. Yaari, A. K. Singh, D. Rosenfarb, J. DeWitt, P. Misra, J. R. Madsen, S. Stone, G. Kreiman, B. Katz, I. Cases, A. Barbu P2
Neural Information Processing Systems (NeurIPS)
oral, sub 2% acceptance rate
`https://braintreebank.dev`

BrainBits: How Much of the Brain are Generative Reconstruction Methods Using? *December 2024*
D. Mayo, C. Wang, A. Harbin, A. Alabdulkareem, A. E. Shaw, B. Katz, A. Barbu P3
Neural Information Processing Systems (NeurIPS)

Revealing Vision-Language Integration in the Brain with Multimodal Networks *July 2024*
V. Subramaniam, C. Conwell, C. Wang, G. Kreiman, B. Katz, I. Cases, A. Barbu P4
International Conference on Machine Learning (ICML)

Baba Is AI: Break the Rules to Beat the Benchmark *July 2024*
N. Cloos, M. Jens, M. Naim, Y. Kuo, I. Cases, A. Barbu, C. J. Cueva P5
ICML 2024 Workshop LLMs and Cognition

How hard are computer vision datasets? Calibrating dataset difficulty to viewing time *December 2023*
D. Mayo, J. Cummings, X. Lin, D. Gutfreund, B. Katz, A. Barbu P6
Neural Information Processing Systems (NeurIPS)
`https://objectnet.dev/mvt`

BrainBERT: Self-supervised representation learning for intracranial recordings *March 2023*
C. Wang, V. Subramaniam, A. Yaari, G. Kreiman, B. Katz, I. Cases, A. Barbu P7
International Conference on Learning Representations (ICLR)

Zero-shot linear combinations of grounded social interactions with Linear Social MDPs *January 2023*
R. Tejwani, Y. Kuo, T. Shu, B. Stankovits, D. Gutfreund, J. Tenenbaum, B. Katz, A. Barbu P8
Conference on Artificial Intelligence (AAAI)

The Aligned Multimodal Movie Treebank: An audio, video, dependency-parse treebank *September 2022*
A. Yaari, J. DeWitt, H. Hu, B. Stankovits, S. Felshin, Y. Berzak, H. Aparicio, B. Katz, I. Cases, A. Barbu P9
Empirical Methods in Natural Language Processing (EMNLP)

Developing a Series of AI Challenges for the United States Department of the Air Force *September 2022*
V. Gadepally, G. Angelides, A. Barbu, A. Bowne, L. Brattain, T. Broderick, A. Cabrera, G. Carl, R. Carter, M. Cha, E. Cowen, J. Cummings, B. Freeman, J. Glass, S. Goldberg, M. Hamilton, T. Heldt, K. W. Huang, P. Isola, B. Katz, J. Koerner, Y. Lin, D. Mayo, K. McAlpin, T. Perron, J. Piou, H. M. Rao, H. Reynolds, K. Samuel, S. Samsi, M. Schmidt, L. Shing, O. Simek, B. Swenson, V. Sze, J. Taylor, P. Tylkin, M. Veillette, M. L Weiss, A. Wollaber, S. Yuditskaya, J. Kepner P10
IEEE Conference on High Performance Extreme Computing (HPEC)

Quantifying the Emergence of Symbolic Communication *August 2022*
E. Cheng, Y. Kuo, I Cases, B. Katz, A. Barbu P11
CogSci

Incorporating rich social interactions into MDPs *March 2022*
R. Tejwani, Y. Kuo, T. Shu, B. Stankovits, D. Gutfreund, J. Tenenbaum, B. Katz, A. Barbu P12
International Conference on Robotics and Automation (ICRA)
Excellent Paper award at IROS Cognitive and Social Aspects of Human Multi-Robot Interaction Workshop

Trajectory prediction with linguistic representations *March 2022*
Y. Kuo, X. Huang, A. Barbu, S. McGill, B. Katz, J. Leonard, G. Rosman P13
International Conference on Robotics and Automation (ICRA)

Neural Regression, Representational Similarity, Model Zoology & Neural Taskonomy at Scale in Rodent Visual Cortex *December 2021*
C. Conwell, D. Mayo, A. Barbu, M. Buice, G. Alvarez, B. Katz P14
Neural Information Processing Systems (NeurIPS)

Perceiving social events in a physical world *September 2021*
T. Shu, A. Netanyahu, M. Kryven, J. Muchovej, N. Shenoy,
B. Katz, A. Barbu, T. Ullman, J. Tenenbaum P15
Vision Sciences Society Annual Meeting

Social Interactions as Recursive MDPs *August 2021*
R. Tejwani, Y. Kuo, B. Katz, and A. Barbu P16
Conference on Robot Learning (CoRL)

Compositional Networks Enable Systematic Generalizationfor Grounded Language Understanding *August 2021*
Y. Kuo, B. Katz, and A. Barbu P17
Findings of Empirical Methods in Natural Language Processing (EMNLP)

Measuring Social Biases in Grounded Vision and Language Embeddings *June 2021*
C. Ross, B. Katz, and A. Barbu P18
Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

Multi-resolution modeling of a discrete stochastic process identifies causes of cancer *May 2021*
A. Yaari, M. Sherman, O. Clarke Priebe, P. Loh, B. Katz, A. Barbu, B. Berger P19
International Conference on Learning Representations (ICLR)

PHASE: PHysically-grounded Abstract Social Events for Machine Social Perception *January 2021*
A. Netanyahu, T. Shu, B. Katz, A. Barbu, J. Tenenbaum P20
Conference on Artificial Intelligence (AAAI)
Best paper award at the NeurIPS SVRHM workshop

Genetic and epigenetic factors interact to explain somatic mutation rates at multiple length scales *October 2020*
A. Yaari, M. Sherman, O. Clarke Priebe, P. Loh, A. Barbu, B. Katz, B. Berger P21
American Society of Human Genetics
Best paper award semi-finalist and reviewer's choice award

Learning a natural-language to LTL executable semantic parser for grounded robotics *November 2020*
C. Wang, C. Ross, Y. Kuo, B. Katz, and A. Barbu P22
Conference on Robot Learning (CoRL)

Encoding formulas as deep networks: Reinforcement learning for zero-shot execution of LTL formulas *October 2020*
Y. Kuo, B. Katz, and A. Barbu P23
International Conference on Intelligent Robots (IROS)

Deep compositional robotic planners that follow natural language commands *May 2020*
Y. Kuo, B. Katz, and A. Barbu P24
International Conference on Robotics and Automation (ICRA)

ObjectNet: A large-scale bias-controlled dataset for pushing the limits of object recognition models *December 2019*
A. Barbu, D. Mayo, J. Alverio, W. Luo, C. Wang, D. Gutfreund, J. Tenenbaum, and B. Katz P25
Neural Information Processing Systems (NeurIPS)
`https://objectnet.dev`

Deep video-to-video transformations for accessibility with an application to photosensitivity *June 2019*
D. Banda, A. Barbu, and B. Katz P26
Pattern Recognition Letters (PatRec)

Partially Occluded Hands: A new dataset for single-image hand pose estimation *December 2018*
B. Myanganbayar, C. Mata, G. Dekel, B. Katz, G. Ben-Yosef, A. Barbu P27
Asian Conference on Computer Vision (ACCV)

First steps to understanding how symbolic communication arises *December 2018*
J. Correa, Y. Kuo, B. Katz, and A. Barbu P28
Workshop on Emergent Communication at NeurIPS

Grounding semantic parsing using caption videos *November 2018*
C. Ross, A. Barbu, Yevgeni Berzak, Battushig Myanganbayar, and B. Katz P29
Empirical Methods in Natural Language Processing (EMNLP)
`http://aclweb.org/anthology/D18-1285`

Deep sequential models for sampling-based planning *October 2018*
Y. Kuo, A. Barbu, and B. Katz P30
International Conference on Intelligent Robots (IROS)
`http://0xab.com/papers/iros2018.pdf`

Deep compositional models for robotic planning and language *October 2018*
Y. Kuo, A. Barbu, and B. Katz P31
Workshop on Language and Robotics at IROS

Temporal Grounding Graphs for Language Understanding with Accrued Visual-Linguistic Context *August 2017*
R. Paul, A. Barbu, S. Felshin, B. Katz, and N. Roy P32
International Joint Conference on Artificial Intelligence (IJCAI)
`http://0xab.com/papers/ijcai17.pdf`

Saying What You're Looking For: Linguistics Meets Video Search *October 2016*
D. Barrett, A. Barbu, N. Siddharth, and J. M. Siskind P33
IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)

Anchoring and Agreement in Syntactic Annotations *September 2016*
Y. Berzak, Y. Huang, A. Barbu, A. Korhonen, and B. Katz P34
Empirical Methods in Natural Language Processing (EMNLP)

Do You See What I Mean? Visual Resolution of Linguistic Ambiguities *September 2015*
Y. Berzak, A. Barbu, D. Harari, B. Katz, and S. Ullman P35
Empirical Methods in Natural Language Processing (EMNLP)

A Compositional Framework for Grounding Language Inference, Generation, and Acquisition in Video *April 2015*
H. Yu, N. Siddharth, A. Barbu, and J. M. Siskind P36
Journal of Artificial Intelligence Research (JAIR)

Seeing is *Worse* than Believing: Reading People's Minds Better than Computer-Vision Methods Recognize Actions *September 2014*
A. Barbu, D. P. Barrett, W. Chen, N. Siddharth, C. Xiong, J. J. Corso, C. D. Fellbaum, C. Hanson, S. J. Hanson, S. Hélie, E. Malaia, B. A. Pearlmutter, J. M. Siskind, T. M. Talavage, and R. B. Wilbur P37
European Conference on Computer Vision (ECCV)

Seeing What You're Told: Sentence-Guided Activity Recognition In Video *September 2014*
N. Siddharth, A. Barbu, J. M. Siskind P38
Human-Machine Communication for Visual Recognition and Search Workshop at ECCV

Language-driven video retrieval *June 2014*
A. Barbu, N. Siddharth, J. M. Siskind P39
Vision Meets Cognition at CVPR

Seeing what you're told: sentence-guided activity recognition in video *June 2014*
N. Siddharth, A. Barbu, and J. M. Siskind P40
IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

Recognizing human activities from partially observed videos *June 2013*
Y. Cao, D. Barrett, A. Barbu, N. Siddharth, H. Yu, A. Michaux, Y. Lin, S. Dickinson, J. M. Siskind, and S. Wang P41
IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
poster 472/1870 (25.2%)

Seeing unseeability to see the unseeable *December 2012*
N. Siddharth, A. Barbu, and J. M. Siskind P42
Advances in Cognitive Systems, 2:77–94
oral 14/38 (37%), conference presentation associated with

Simultaneous object detection, tracking, and event recognition *December 2012*
A. Barbu, N. Siddharth, A. Michaux, and J. M. Siskind P43
Advances in Cognitive Systems, 2:203–20
oral 14/38 (37%), conference presentation associated with

Video in sentences out *August 2012*
A. Barbu, A. Bridge, Z. Burchill, D. Coroian, S. Dickinson, S. Fidler, A. Michaux, S. Mussman, N. Siddharth, D. Salvi, L. Schmidt, J. Shangguan, J. M. Siskind, J. Waggoner, S. Wang, J. Wei, Y. Yin, and Z. Zhang P44
Conference on Uncertainty In Artificial Intelligence (UAI), pp. 102–12
oral 24/304 (8%)

A visual language model for estimating object pose and structure in a generative visual domain *May 2011*
N. Siddharth, A. Barbu, and J. M. Siskind P45
IEEE International Conference on Robotics and Automation (ICRA)
oral 982/2004 (49%)

Learning physically-instantiated robotic game play through visual observation *May 2010*
A. Barbu, N. Siddharth, and J. M. Siskind P46
IEEE International Conference on Robotics and Automation (ICRA)
oral 856/2062 (42%)

Patents

Secure Large Language Models *in review*
C. Arnold, A. Alabdulkareem, B. Katz, A. Barbu A1
PCT/US24/58217

Deep Compositional Robotic Planners That Follow Natural Language Commands *in review*
Y. Kuo, B. Katz, A. Barbu A2
Provisional 62/944,924 and 62/944,932

Computer method and apparatus making screens safe for those with photosensitivity *July 2022*
A. Barbu, D. Banda, B. Katz A3
US patent US11381715

Correlating videos and sentences *November 2015*
J. M. Siskind, A. Barbu, N. Siddharth, and H. Yu A4
US patent US9183466

Technical Reports

Bias and Agreement in Syntactic Annotations *May 2016*
Yevgeni Berzak, Yan Huang, A. Barbu, Anna Korhonen, Boris Katz T1
arXiv:1605.04481
`http://arxiv.org/abs/1605.04481`

The Compositional Nature of Event Representations in the Human Brain *June 2013*
A. Barbu, N. Siddharth, C. Xiong, J. J. Corso, C. D. Fellbaum, C. Hanson, S. J. Hanson, S. Hélie, E. Malaia, B. A. Pearlmutter, J. M. Siskind, T. M. Talavage, and R. B. Wilbur T2
arXiv:1306.2293
`http://arxiv.org/abs/1306.2293`

Large-scale automatic labeling of video events with verbs based on event-participant interaction *April 2012*
A. Barbu, A. Bridge, D. Coroian, S. Dickinson, S. Mussman, S. Narayanaswamy, D. Salvi, L. Schmidt, J. Shangguan, J. M. Siskind, J. Waggoner, S. Wang, J. Wei, Y. Yin, and Z. Zhang T3
arXiv:1204.3616
`http://arxiv.org/abs/1204.3616`

Workshops organized

Human-inspired Computer Vision at ECCV 2024 *September 2024*
Lucia Schiatti, Mengmi Zhang, Yen-Ling Kuo, Vittorio Cuculo, Martin Schrimpf, A. Barbu W1
30 accepted papers, 4 invited speakers, over 200 attendees
`https://sites.google.com/view/hcvworkshop2024`

Linguistics Meets Image and Video Retrieval at ICCV 2019 *October 2019*
Amrita Saha, Hui Wu, Adriana I. Kovashka, A. Barbu, Xiaoxiao Guo, Karthik Sankaranarayanan, Samarth Bharadwaj, and Yupeng Gao W2
3 invited speakers, hosted the Fashion IQ challenge
`https://sites.google.com/view/lingir`

Vision and Language at CVPR 2019 *June 2019*
A. Barbu, Siddharth N., D. Gutfreund W3
22 accepted papers, 5 invited speakers, over 200 attendees
`languageandvision.com`

Combining Vision and Language ACCV 2018 *June 2018*
Qi Wu, Peng Wang, Chuang Gan, Fumin Shen, Xiaodong He, A. Barbu, Anton van den Hengel W4
A workshop introducing new researchers to problems in vision and language, 7 invited speakers
`http://qi-wu.me/accv_v2l/`

Vision and Language at CVPR 2018 *June 2018*
A. Barbu, Siddharth N., Tao Mei, Y. Berzak, N. Shukla, J. Luo, R. Sukthankar, D. Gutfreund W5
16 accepted papers, 7 invited speakers, over 200 attendees
`languageandvision.com`

Vision and Language at CVPR 2017 *July 2017*
A. Barbu, Tao Mei, Siddharth Narayanaswamy, Puneet Kumar Dokania, Quanshi Zhang, Nishant Shukla, Jiebo Luo, Rahul Sukthankar W6
17 accepted papers, 7 invited speakers, over 200 attendees
`languageandvision.com/2017.html`

Vision and Language at CVPR 2015 *June 2015*
A. Barbu, Georgios Evangelopoulus, Daniel Harari, Krystian Mikolajczyk, Siddharth Narayanaswamy, Caiming Xiong, Yibiao Zhang W7
14 accepted papers, 9 invited speakers, over 100 attendees
`languageandvision.com/2015.html`

PhD students supervised

Christopher Arnold *2023-present*
Multimodal story understanding S1
Co-supervised with Boris Katz.

Christopher Z Wang *2021-present*
Object recognition S2
Co-supervised with Boris Katz.

Vighnesh Subramaniam *2023-present*
Social perception S3
Co-supervised with Boris Katz.

David Mayo *2020-present*
Object recognition S4
Co-supervised with Boris Katz.

Adam Yaari *2018-2022*
Language and neuroscience S5
CEO and founder of Serinus Biosciences. Co-supervised with Boris Katz.

Yen-Ling Kuo *2017-2022*
Grounding language in planning S6
Faculty at University of Virginia. Co-supervised with Boris Katz.

Candace Ross *2016-2022*
Grounded language acquisition S7
Scientist at FAIR. Co-supervised with Boris Katz.

Master's students supervised

Abdulrahman Alabdulkareem *2024*
Understanding modalities in detail M1
Co-supervised with Boris Katz.

Anastasiia Uvarova *current*
Modeling the development of language M2
Co-supervised with Boris Katz.

Laura Queipo *2023*
Causal summarization M3
Co-supervised with Boris Katz.

Vighnesh Subramaniam *2023*
Understanding multimodal integration in the brain M4
Co-supervised with Boris Katz.

Meagan R. Jens *2023*
Machines that change the rules to win M5
Co-supervised with Boris Katz.

Dana Rosenfarb *2022*
Creating a large-scale benchmark for language in the brain M6
Co-supervised with Boris Katz.

Felipe Monsalve *2022*
Developing a secure ML platform for healthcare applications M7
Co-supervised with Boris Katz.

Emily Cheng *2022*
Understanding language evolution M8
Now a PhD student. Co-supervised with Boris Katz.

Aaditya K. Singh *2021*
Deep Attentional Modulation for Zero-shot Learning in Object Recognition M9
Now at PhD student at UCL. Co-supervised with Boris Katz.

Christopher Z. Wang *2021*
Modeling child language acquisition as grounded semantic parsing M10
Now at PhD student at MIT. Co-supervised with Boris Katz.

Allison Fu *2021*
Cognitively-plausible translation without aligned corpora M11
Co-supervised with Boris Katz.

Julian Alverio *2021*
Robots that understand language M12
Co-supervised with Boris Katz.

David Mayo *2018*
Understanding human and machine object recognition at scale M13
Co-supervised with Boris Katz.

Dalitso Banda *2018*
Removing dangerous visual stimuli from videos M14
Co-supervised with Boris Katz.

Battushig Myanganbayar *2018*
Inverse graphics for hand-pose reconstruction M15
Co-supervised with Boris Katz now at Apple.

Maria Ryskina *2016*
Imaginative Paraphrase Recognition via Join Vision-Language Video Analysis M16
Co-supervised with Boris Katz. Now a PhD student at CMU.

Sergey Voronov *2016*
Common Sense Reasoning Through Imagination M17
Co-supervised with Boris Katz. Now a PhD student at University of Ohio.

Nicolas Rakover *2016*
A uniform Representation for Visual Concepts M18
Co-supervised with Boris Katz. Now at Google.