

Frederic Sala

Computer Sciences 5385
1210 W Dayton Street
Madison, WI 53706

email: fredsala@cs.wisc.edu
url: pages.cs.wisc.edu/~fredsala

Academic positions

- 2021-Present Assistant Professor
University of Wisconsin-Madison. Department of Computer Sciences
Affiliations: Data Science Institute (DSI), Institute for Foundations of Data Science (IFDS)
- 2017-2020 Postdoctoral Scholar
Stanford University. Computer Science Department

Education

- 2013-2016 Ph.D. in Electrical Engineering, University of California, Los Angeles (UCLA)
Dissertation: “Algorithms and coding techniques for reliable data management and storage”
(Outstanding Ph.D. Dissertation Award - Signals & Systems Track)
- 2011-2013 M.S. in Electrical Engineering, University of California, Los Angeles (UCLA)
Thesis: “Novel coding strategies for multi-level non-volatile memories”
(Outstanding M.S. Thesis Award - Signals & Systems Track)
- 2006-2010 B.S. in Electrical Engineering, University of Michigan, Ann Arbor

Honors & awards

- 2024 **DARPA Young Faculty Award**
- 2023 Best Paper Award Honorable Mention, NeurIPS R0-FoMo Workshop
- 2023 DataComp 2023 Competition Winner, Data Filtering Track Small
- 2022 Best Student Paper Runner-up, UAI 2022
- 2017 Outstanding Ph.D. Dissertation Award
UCLA Department of Electrical Engineering (Signals & Systems Track)
- 2015-2016 UCLA Dissertation Year Fellowship
- 2015 Qualcomm Innovation Fellowship Finalist
- 2013 Edward K. Rice Outstanding Masters Student Award
UCLA Henry Samueli School of Engineering & Applied Science
- 2013 Outstanding M.S. Thesis Award
UCLA Department of Electrical Engineering (Signals & Systems Track)
- 2012 National Science Foundation Graduate Research Fellowship (NSF GRFP)

Publications & talks

Preprints

1. N. Roberts, S. Guo, Z. Gao, S. Namburi, S. Crompt, C. Wu, C. Duan, and **F. Sala**, “Pretrained Hybrids with MAD Skills,” 2024.
2. D. Adila, C. Shin, Y. Zhang, and **F. Sala**, “Is Free Self-Alignment Possible?,” 2024.
3. T. Zhang, L. Cai, N. Roberts, J. Li, N. Guha, and **F. Sala**, “Stronger Than You Think: Benchmarking Weak Supervision on Realistic Tasks,” 2024.
4. T. Huang, C. Cao, V. Bhargava, and **F. Sala**, “The AlCHEmist: Automated Labeling 500x CHEaper than LLM Data Annotators,” 2024.
5. H. Vishwakarma, Y. Chen, S. Namburi, S. Tay, R. Vinayak, and **F. Sala**, “PabLO: Improving Semi-Supervised Learning with Pseudolabeling Optimization,” 2024.
6. C. Shin, J. Zhao, S. Crompt, H. Vishwakarma, and **F. Sala**, “OTTER: Improving Zero-shot Classification via Optimal Transport,” 2024.
7. H. Vishwakarma, Y. Chen, S. J. Tay, S. Namburi, **F. Sala**, and R. Vinayak, “Pearls from Pebbles: Improved Confidence Functions for Auto-labeling,” 2024.
8. D. McNeela, **F. Sala**, and A. Gitter, “Product Manifold Representations for Learning on Biological Pathways,” 2024.

Conference Publications

- 2024 D. Adila, C. Shin, L. Cai, and **F. Sala**, “Zero-Shot Robustification of Zero-Shot Models,” in *International Conference on Learning Representations (ICLR)*, 2024.
- 2024 Z. Shi, Y. Ming, Y. Fan, **F. Sala**, Y. Liang, “Domain Generalization via Nuclear Norm Regularization,” in *Conference on Parsimony and Learning (CPAL)*, 2024 (**oral**).
- 2023 N. Roberts, X. Li, D. Adila, S. Crompt, T. Huang, J. Zhao, and **F. Sala**, “Geometry-Aware Adaptation for Pretrained Models,” in *Neural Information Processing Systems (NeurIPS)*, 2023.
- 2023 H. Vishwakarma, **F. Sala**, and R. Vinayak, “Promises and Pitfalls of Threshold-based Auto-labeling,” in *Neural Information Processing Systems (NeurIPS)*, 2023 (**spotlight**).
- 2023 C. Shin, S. Crompt, D. Adila, and **F. Sala**, “Mitigating Source Bias for Fairer Weak Supervision,” in *Neural Information Processing Systems (NeurIPS)*, 2023.
- 2023 T. Huang, H. Vishwakarma, and **F. Sala**, “Train ‘n Trade: Foundations of Parameter Markets,” in *Neural Information Processing Systems (NeurIPS)*, 2023.
- 2023 M. Chen, N. Roberts, K. Bhatia, J. Wang, C. Zhang, **F. Sala**, and C. Ré, “Skill-it! A Data-Driven Skills Framework for Understanding and Training Language Models,” in *Neural Information Processing Systems (NeurIPS)*, 2023 (**spotlight**).
- 2023 N. Guha, M. Chen, K. Bhatia, A. Mirhoseini, **F. Sala**, and C. Ré, “Embroid: Unsupervised Prediction Smoothing Can Improve Few-Shot Classification,” in *Neural Information Processing Systems (NeurIPS)*, 2023.
- 2023 S. Namburi, M. N. Sreedhar, S. Srinivasan, and **F. Sala**, “The Cost of Compression: Investigating the Impact of Compression on Parametric Knowledge in Language Models,” in *Empirical Methods in Natural Language Processing (EMNLP) Findings*, 2023.
- 2023 S. Crompt, M. Bilinski, R. Gabrys, and **F. Sala**, “The Credential is Not Enough: Combining Honey-pots and Fake Credentials for Cyber-Defense,” in *Conference on Decision and Game Theory for Security (GameSec-23)*, 2023.
- 2023 B. Boecking, W. Neiswanger, N. Roberts, S. Ermon, **F. Sala**, and A. Dubrawski, “Generative Modeling Helps Weak Supervision (and Vice Versa),” in *International Conference on Learning Representations (ICLR)*, 2023.
- 2023 S. Arora, A. Narayan, M. Chen, L. Orr, N. Guha, K. Bhatia, I. Chami, **F. Sala**, and C. Ré, “Ask Me Any-

- thing: A simple strategy for prompting language models,” in *International Conference on Learning Representations (ICLR)*, 2023.
- 2022 H. Vishwakarma, N. Roberts, and **F. Sala**, “Lifting Weak Supervision To Structured Prediction,” in *Neural Information Processing Systems (NeurIPS)*, 2022.
- 2022 N. Roberts, X. Li, T. Huang, D. Adila, S. Schoenberg, C. Liu, L. Pick, H. Ma, A. Albarghouthi, and **F. Sala**, “AutoWS-Bench-101: Benchmarking Automated Weak Supervision with 100 Labels,” in *Neural Information Processing Systems (NeurIPS)*, 2022 (Datasets and Benchmarks track).
- 2022 R. Tu, N. Roberts, M. Khodak, J. Shen, **F. Sala**, and A. Talwalkar, “NAS-Bench-360: Benchmarking Neural Architecture Search on Diverse Tasks,” in *Neural Information Processing Systems (NeurIPS)*, 2022 (Datasets and Benchmarks track).
- 2022 C. Shin, W. Li, H. Vishwakarma, N. Roberts, and **F. Sala**, “Universalizing Weak Supervision,” in *International Conference on Learning Representations (ICLR)*, 2022.
- 2022 M. Chen, D. Fu, D. Adila, M. Zhang, **F. Sala**, K. Fatahalian, and C. Ré, “Shoring Up the Foundations: Fusing Model Embeddings and Weak Supervision,” in *Uncertainty in Artificial Intelligence (UAI)*, 2022, **(oral, best student paper runner-up)**.
- 2021 M. Chen, B. Cohen-Wang, S. Mussmann, **F. Sala**, and C. Ré, “Comparing the Value of Labeled and Unlabeled Data in Method-of-Moments Latent Variable Estimation,” in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2021.
- 2021 S. Hooper, M. Wornow, Y. H. Seah, P. Kellman, H. Xue, **F. Sala**, C. Langlotz, and C. Ré, “Cut Out The Annotator, Keep The Cutout: Better Segmentation With Weak Supervision,” in *International Conference on Learning Representations (ICLR)*, 2021.
- 2020 D. Fu, M. Chen, **F. Sala**, S. Hooper, K. Fatahalian, and C. Ré, “Fast and Three-rious: Speeding Up Weak Supervision with Triplet Methods,” in *International Conference on Machine Learning (ICML)*, 2020.
- 2020 I. Chami, A. Wolf, D. Juan, **F. Sala**, S. Ravi, and C. Ré, “Low-Dimensional Hyperbolic Knowledge Graph Embeddings,” in *Annual Conf. Association for Computational Linguistics (ACL)*, Seattle, Wa, 2020.
- 2020 Z. Kuang, **F. Sala**, N. Sohoni, S. Wu, A. Cordova-Palomera, J. Dunnmon, J. Priest, and C. Ré, “Ivy: Instrumental Variable Synthesis for Causal Inference,” in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Palermo, Italy, 2020.
- 2019 **F. Sala**^{*}, P. Varma^{*}, S. Sagawa, J. Fries, D. Fu, S. Khattar, A. Ramamoorthy, K. Xiao, K. Fatahalian, J. Priest, and C. Ré, “Multi-resolution weak supervision for sequential data,” in *Neural Information Processing Systems (NeurIPS)*, Vancouver, Canada, 2019.
- 2019 P. Varma^{*}, **F. Sala**^{*}, A. He, A. J. Ratner, and C. Ré, “Learning dependency structures for weak supervision models,” in *International Conference on Machine Learning (ICML)*, Long Beach, CA, 2019.
- 2019 A. Gu, **F. Sala**, B. Gunel, and C. Ré, “Learning mixed-curvature representations in product spaces,” in *International Conference on Learning Representations (ICLR)*, New Orleans, LA, 2019.
- 2019 A. J. Ratner, B. Hancock, J. Dunnmon, **F. Sala**, S. Pandey, and C. Ré, “Training complex models with multi-task weak supervision,” in *AAAI Conference on Artificial Intelligence*, Honolulu, HI, 2019.
- 2018 **F. Sala**, C. De Sa, A. Gu, and C. Ré, “Representation tradeoffs for hyperbolic embeddings,” in *International Conference on Machine Learning (ICML)*, Stockholm, Sweden, 2018.
- 2018 R. Gabrys and **F. Sala**, “Codes correcting two deletions,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Vail, CO, 2018.
- 2017 S. Kabir, **F. Sala**, G. V. d. Broeck, and L. Dolecek. “Coded machine learning: Joint informed replication and learning for linear regression,” in *Proc. IEEE 54th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2017.
- 2016 K. Mazooji, **F. Sala**, G. V. d. Broeck, and L. Dolecek, “Robust channel coding strategies for machine learning data,” in *Proc. IEEE 54th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2016.
- 2016 **F. Sala**, R. Gabrys, C. Schoeny, K. Mazooji, and L. Dolecek, “Exact sequence reconstruction for insertion-correcting codes,” in *Proc. IEEE International Symp. on Information Theory (ISIT)*, Barcelona,

Spain, 2016.

- 2015 **F. Sala**, C. Schoeny, D. Divsalar, and L. Dolecek, "Asymmetric error-correcting codes for Flash memories in high-radiation environments," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, 2015.
- 2015 C. Schoeny, **F. Sala**, and L. Dolecek, "Analysis and coding schemes for the Flash Normal-Laplace mixture channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, 2015.
- 2015 **F. Sala**, C. Schoeny, R. Gabrys, and L. Dolecek, "Three novel combinatorial theorems for the insertion/deletion channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, 2015.
- 2014 **F. Sala**, R. Gabrys, and L. Dolecek, "Deletions in multipermutations," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Honolulu, HI, 2014.
- 2014 R. Gabrys, E. Yaakobi, F. Farnoud, **F. Sala**, J. Bruck, and L. Dolecek, "Single-deletion-correcting codes over permutations," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Honolulu, HI, 2014.
- 2013 N. Bitouzé, **F. Sala**, S. M. S. Tabatabaei, and L. Dolecek, "A practical framework for efficient file synchronization," in *Proc. IEEE 51st Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2013.
- 2013 **F. Sala** and L. Dolecek, "Counting sequences obtained from the synchronization channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Istanbul, Turkey, 2013.

*: Equal Contribution

Journal Articles

14. M. Chen, B. Nachman, **F. Sala**, "Resonant anomaly detection with multiple reference datasets," *Journal of High Energy Physics*, July 2023.
13. R. Yang, **F. Sala**, P. Bogdan, "Hidden network generating rules from partially observed complex networks," *Nature Communications Physics*, Aug. 2021.
12. C. Schoeny, **F. Sala**, M. Gottscho, I. Alam, P. Gupta, and L. Dolecek, "Context-aware resiliency: Unequal message protection for random-access memories," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6146-6159, Oct. 2019.
11. R. Gabrys and **F. Sala**, "Codes correcting two deletions," *IEEE Transactions on Information Theory*, vol. 65, no. 2, pp. 965-974, Feb. 2019.
10. S. S. Garani, L. Dolecek, J. Barry, **F. Sala**, and B. Vasic, "Signal processing and coding techniques for 2-D magnetic recording: An overview," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 286-318, Feb. 2018.
9. **F. Sala**, C. Schoeny, R. Gabrys, and L. Dolecek, "Exact reconstruction from insertions in synchronization codes," *IEEE Transactions on Information Theory*, vol. 63, no. 4, pp. 2428-2445, Apr. 2017.
8. **F. Sala**, C. Schoeny, S. Kabir, D. Divsalar, and L. Dolecek "On nonuniform noisy decoding for LDPC codes with application to radiation-induced errors," *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1438-1450, Apr. 2017.
7. **F. Sala**, N. Bitouzé, C. Schoeny, and L. Dolecek, "Synchronizing files under a large number of edits," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2258-2273, Jun. 2016.
6. R. Gabrys, E. Yaakobi, F. Farnoud, **F. Sala**, S. Bruck, and L. Dolecek, "Codes correcting erasures and deletions for rank modulation," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 136-150, Jan. 2016.
5. L.F. Wanner **et al.**, "NSF expedition on variability-aware software: Recent results and contributions," *Information Technology*, vol. 57, no. 3, pp. 181-198, Jun. 2015.
4. **F. Sala**, K.A.S. Immink, and L. Dolecek, "Error control schemes for modern Flash memories: Solutions for Flash deficiencies," *IEEE Consumer Electronics*, vol. 4, no. 1, pp. 66-73, Jan. 2015.

3. R. Gabrys, **F. Sala**, and L. Dolecek, "Coding for unreliable Flash memory cells," *IEEE Communication Letters*, vol. 18, no. 9, pp. 1491-1494, Jul. 2014.
2. **F. Sala**, R. Gabrys, and L. Dolecek, "Dynamic threshold schemes for multi-level non-volatile memories," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2624-2634, Jul. 2013.
1. E. Cheng, L. Liptak, and **F. Sala**, "Linearly many faults in 2-tree generated networks," *Networks*, vol. 55, no. 2, pp. 90-98, Mar. 2010.

Workshop Publications

- 2024 C. Shin, J. Cooper, D. Adila, and **F. Sala**, "Weak-to-Strong Generalization Through the Data-Centric Lens," *ICML 2024 Workshop on Data-Centric Machine Learning Research (DMLR)*, Vienna, Austria, 2024.
- 2024 T. Zhang, L. Cai, N. Roberts, J. Li, N. Guha, and **F. Sala**, "Stronger Than You Think: Benchmarking Weak Supervision on Realistic Tasks," *ICML 2024 Workshop on Data-Centric Machine Learning Research (DMLR)*, Vienna, Austria, 2024.
- 2024 H. Vishwakarma, Y. Chen, S. Tay, S. Namburi, R. Vinayak, and **F. Sala**, "Pearls from Pebbles: Improved Confidence Functions for Auto-labeling," *ICML 2024 Workshop on Data-Centric Machine Learning Research (DMLR)*, Vienna, Austria, 2024.
- 2024 H. Vishwakarma, Y. Chen, S. Namburi, S. Tay, R. Vinayak, and **F. Sala**, "PabLO: Improving Semi-Supervised Learning with Pseudolabeling Optimization," *ICML 2024 Workshop on Data-Centric Machine Learning Research (DMLR)*, Vienna, Austria, 2024.
- 2024 D. Adila, C. Shin, L. Cai, and **F. Sala**, "Foundation Models Can Robustify Themselves, For Free," *NeurIPS 2023 Workshop on Robustness of Few-shot and Zero-shot Learning in Foundation Models (R0-FoMo)*, New Orleans, LA, 2023, (**oral, best paper honorable mention**).
- 2023 D. Adila, C. Shin, L. Cai, and **F. Sala**, "Foundation Models Can Robustify Themselves, For Free," *NeurIPS 2023 Workshop on Robustness of Few-shot and Zero-shot Learning in Foundation Models (R0-FoMo)*, New Orleans, LA, 2023, (**oral, best paper honorable mention**).
- 2023 H. Vishwakarma, **F. Sala**, and R. Vinayak, "Understanding Threshold-based Auto-labeling: The Good, the Bad, and the Terra Incognita," *NeurIPS 2023 Workshop on Adaptive Experimental Design and Active Learning in the Real World (RealML-2023)*, New Orleans, LA, 2023.
- 2023 T. Huang, C. Shin, S. J. Tay, D. Adila, and **F. Sala**, "Multimodal Data Curation Via Object Detection And Filter Ensembles," *ICCV 2023 Workshop: Towards the Next Generation of Computer Vision Datasets (TNGCV)*, Paris, France, 2023.
- 2023 D. McNeela, **F. Sala**, and A. Gitter, "Mixed Curvature Representation Learning for Biological Pathway Graphs," *ICML 2023 Workshop for Computational Biology*, Honolulu, HI, 2023.
- 2023 H. Vishwakarma, **F. Sala**, and R. Vinayak, "Promises and Pitfalls of Threshold-based Auto-labeling," *ICML 2023 Workshop on Data-centric Machine Learning Research*, Honolulu, HI, 2023.
- 2023 M. Chen, N. Roberts, K. Bhatia, J. Wang, C. Zhang, **F. Sala**, and C. Ré, "Skill-it! A Data-Driven Skills Framework for Understanding and Training Language Models," *ICML 2023 Workshop on Data-centric Machine Learning Research*, Honolulu, HI, 2023.
- 2023 T. Huang, C. Cao, S. Schoenberg, H. Vishwakarma, N. Roberts, and **F. Sala**, "ScriptoriumWS: A Code Generation Assistant for Weak Supervision," *ICLR 2023 Workshop on Deep Learning for Code (DL4C)*, Kigali, Rwanda, 2023.
- 2022 M. Chen, B. Nachman, and **F. Sala**, "Anomaly Detection with Multiple Reference Datasets in High Energy Physics," *NeurIPS 2022 Workshop on Machine Learning and the Physical Sciences*, New Orleans, LA, 2022.
- 2022 Z. Shi, Y. Ming, Y. Fan, **F. Sala**, and Y. Liang, "Domain Generalization with Nuclear Norm Regularization," *NeurIPS Workshop on Workshop on Distribution Shifts (DistShift)*, New Orleans, LA, 2022.
- 2022 R. Tu, N. Roberts, V. Prasad, S. Nayak, P. Jain, **F. Sala**, G. Ramakrishnan, A. Talwalkar, W. Neiswanger, and C. White, "AutoML for Climate Change: A Call to Action," *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*, New Orleans, LA, 2022.
- 2022 D. Adila, S. Crompt, S. Mo, **F. Sala**, "Causal Omnivore: Fusing Noisy Estimates of Spurious Correla-

- tions,” *ICML 2022 Workshop on Spurious Correlations, Invariance, and Stability*, Baltimore, MD, 2022.
- 2019 I. Chami, A. Wolf, **F. Sala**, and C. Ré, “Low-dimensional knowledge graph embeddings via hyperbolic rotations,” *NeurIPS 2019 Workshop on Graph Representation Learning*, Vancouver, Canada, 2019.
- 2017 **F. Sala**, S. Kabir, L. Dolecek, and G. V. d. Broeck. “Don’t fear the bit flips: Robust linear prediction through informed channel coding,” *Proc. ICML 2017 Workshop on Reliable Machine Learning in the Wild*, Sydney, Australia, 2017.
- 2017 C. Schoeny, **F. Sala**, M. Gottscho, I. Alam, P. Gupta, and L. Dolecek, “Context-aware resiliency: Unequal message protection for random-access memories,” in *Proc. IEEE Information Theory Workshop (ITW)*, Kaohsiung, Taiwan, 2017.
- 2016 **F. Sala**, H. Duwe, L. Dolecek, and R. Kumar, “A unified framework for error correction techniques in on-chip memories,” presented at *Workshop on Silicon Errors in Logic - System Effects (SELSE-12)*, Austin, TX, 2016.
- Best of SELSE-12 paper, presented in special session at IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)**, Toulouse, France, 2016.

Monographs & Book Chapters

- 2016 L. Dolecek and **F. Sala**, “Channel coding methods for non-volatile memories,” *Foundations and Trends in Communications and Information Theory*, vol. 13, no. 1, pp. 1-136, Feb. 2016.
- 2016 **F. Sala**, C. Schoeny, and L. Dolecek, “Advanced algebraic and graph-based ECC schemes for Flash memories,” in *3D Flash Memories*, Rino Micheloni, Ed. Springer, 2016, pp. 321-348.

Theses

- 2016 **F. Sala**, “Algorithms and coding techniques for reliable data management and storage.” Ph.D. Dissertation, University of California, Los Angeles, December 2016. Outstanding Ph.D. Dissertation Award - Signals & Systems Track
- 2013 **F. Sala**, “Novel coding strategies for multi-level non-volatile memories.” M.S. Thesis, University of California, Los Angeles, May 2013. Distinguished M.S. Thesis Award - Signals & Systems Track

Patents

- 2019 **F. Sala**, C. Hu, H. Zheng, D. Niu, and M. Chang, Samsung Electronics Co Ltd, “Virtual bucket multiple hash tables for efficient memory in-line deduplication application,” Patent US10496543B2, Granted: Dec. 2019.
- 2018 **F. Sala**, C. Hu, H. Zheng, D. Niu, and M. Chang, Samsung Electronics Co Ltd, “Optimized hopscotch multiple hash tables for efficient memory in-line deduplication application,” Patent US9983821B2, Granted: May 2018.

Workshops, Tutorials, and Competitions Organized

- 2024 Spring Symposium on Clinical Foundation Models, AAAI 2024.
- 2023 AutoML Cup Competition, AutoML Conference 2023 (Chair).
- 2022 Theory and Practice of Efficient and Accurate Dataset Construction tutorial, NeurIPS 2022.
- 2022 AutoML Decathlon Competition, NeurIPS 2022.
- 2020 Differential geometry meets deep learning (DiffGeo4DL) workshop, NeurIPS 2020.
- 2019 Graph representation learning workshop, ICLR 2019.
- 2018 Relational representation learning (R2L) workshop, NeurIPS 2018.

Talks

- 2024 **F. Sala**, “Data-Efficient Adaptation for Pretrained Decision-Making Models”, Midwest Machine Learning Symposium (MMLS), May 20, 2024.
- 2024 **F. Sala**, “Data-Efficient Adaptation for Pretrained Machine Learning Models”, Naval Information Warfare Center (NIWC) Pacific Cyber Chats, March 28, 2024.
- 2023 **F. Sala**, “Getting More From Structured Data: Weak Supervision and Non-Euclidean Machine Learning”, UCB Think Tank Talks, November 14th, 2023.
- 2023 **F. Sala**, “How to Aggregate Your Large Language Model Objects,” Systems, Information, Learning and Optimization (SILO), Madison, WI, July 28th, 2023.
- 2023 **F. Sala**, “Moment Techniques for Weakly-Supervised Learning,” University of Wisconsin Applied Algebra Seminar, March 30, 2023.
- 2023 **F. Sala**, “Weak Supervision: A Machine Learning Paradigm for All Seasons,” Middle East Technical University (METU) EEE Seminar, March 6, 2023.
- 2023 **F. Sala**, “Weak Supervision: A Machine Learning Paradigm for All Seasons,” Information Theory and Applications Workshop (ITA), February 14, 2023.
- 2022 **F. Sala**, “Efficiently Constructing Datasets for Diverse Datatypes,” Stanford MLSys Seminar Series, Stanford, CA, January 20, 2022.
- 2020 **F. Sala**, “The World Isn’t Flat: Towards Non-Euclidean Machine Learning,” Systems, Information, Learning and Optimization (SILO), Madison, WI, November 18th, 2020.
- 2019 **F. Sala**, “Putting non-Euclidean geometry to work in ML: Hyperbolic and product manifold embeddings,” Physics in ML Workshop in ml4science, Berkeley, CA, May 29, 2019.
- 2018 **F. Sala**, “Codes, embeddings, & non-Euclidean geometry: Unexpected allies in the fight to clean data,” CROSS Research Symposium at UCSC, Santa Cruz, CA, October 3, 2018.
- 2016 **F. Sala**, “Error-correcting codes for radiation-induced error patterns in flash memories,” Non-Volatile Memories Workshop (NVMW), San Diego, CA, March 7, 2016.
- 2016 **F. Sala**, “Exact reconstruction of coded data from traces.” Graduation Day Talk at Information Theory and Applications Workshop (ITA), La Jolla, CA, February 3, 2016.
- 2014 **F. Sala**, “Constrained rank modulation for flash memories,” Non-Volatile Memories Workshop (NVMW), San Diego, CA, March 11, 2014.
- 2013 **F. Sala**, “Reducing energy usage through a novel file synchronization algorithm,” DIMACS Workshop on Algorithms for Green Data Storage, Rutgers University, New Brunswick, NJ, December 18, 2013.
- 2013 **F. Sala**, “Dynamic threshold schemes for multi-level non-volatile memories,” Non-Volatile Memories Workshop (NVMW), San Diego, CA, March 5, 2013.

Industry Experience

- 2020-present Snorkel AI, Scientist/Consultant
- 2015 Samsung Electronics, Research Engineering Intern
Memory Systems Laboratory (MSL) - New Memory Technologies Group
- 2014 Jet Propulsion Laboratory (JPL), Research Engineering Intern
Information Processing Group
- 2010-2011 Microsoft Corporation, Software Development Engineer
Windows Phone Group

Funding Awarded

- 2024 **DARPA:** Young Faculty Award. “Active Adaptation for Decentralized Foundation Models”. \$500,000.
- 2024 **Air Force:** STTR X23.E Phase 1. “Rapid Labeling and Weak Supervision in USAF Data Science Pipelines”. \$35,000 (share).
- 2023-2024 **Army Research Lab (ARL):** Strengthening Teamwork for Robust Operations in Novel Groups (STRONG) Program. “Leveling Up Human-Machine Collaborations by Leveraging Distributed Causal Insights”. \$100,000.
- 2023-2024 **American Family Funding Initiative** Award. “Data-Efficient Customization for Large Pretrained Models”. \$100,000.
- 2023-2024 **American Family Funding Initiative** Award. “Improving Auto-labeling with Confidence Functions”. (Co-PI; PI: Ramya Vinayak). \$100,000.
- 2022-2023 **UW Madison OVCRGE** Fall Research Competition Award. “Foundations of Autolabeling Systems” (PI; Co-PI: Ramya Vinayak). \$53,500.
- 2022-2023 **American Family Funding Initiative** Award. “Auto-labeling Foundations” (Co-PI; PI: Ramya Vinayak). \$100,000.
- 2021-2025 **NSF SHF:** Medium. “Program Synthesis for Weak Supervision” (Co-PI; PI: Aws Albarghouthi). \$916,000.

Service

Area Chair/Editor

Transactions on Machine Learning Research (TMLR): Action Editor
Conference on Neural Information Processing Systems (NeurIPS)
International Conference on Learning Representations (ICLR)
International Conference on Machine Learning (ICML)
AAAI Conference on Artificial Intelligence
International Joint Conferences on Artificial Intelligence (IJCAI)

Conference Reviewing

Conference on Neural Information Processing Systems (NeurIPS)
Top Reviewer: 2018, 2019
International Conference on Machine Learning (ICML)
Top Reviewer: 2021
International Conference on Artificial Intelligence and Statistics (AISTATS)
International Conference on Learning Representations (ICLR)
Conference on Uncertainty in Artificial Intelligence (UAI)
International Joint Conference on Artificial Intelligence (IJCAI)
AAAI Conference on Artificial Intelligence
SysML Conference
International Symposium on Information Theory (ISIT)

Journal Reviewing

Transactions on Machine Learning Research (Action Editor)
IEEE Transactions on Information Theory

IEEE Transactions on Signal Processing
IEEE Journal on Selected Areas in Communications
IEEE Transactions on Communications
IEEE Communications Letters
IEEE Signal Processing Letters
Designs, Codes and Cryptography
International Journal of Electronics and Communications

Students

Graduate Students:

- Harit Vishwakarma (co-advised with Ramya Vinayak).
- Nicholas Roberts. **Awards:** Jacquelin Perry Prove AI Fellowship 2022, MLCommons Rising Star 2023.
- Changho Shin. **Awards:** Qualcomm Innovation Fellowship Finalist 2024.
- Dyah Adila. **Awards:** Qualcomm Innovation Fellowship Finalist 2024.
- Sonia Crompt. **Awards:** UW-Madison Computer Sciences Summer Fellowship 2022.
- Brian Huang.
- Kendall Park.