

DYAH ADILA

CONTACT INFORMATION	<i>E-mail:</i> adila@wisc.edu <i>Website:</i> dyahadila.github.io	
EDUCATION	Ph.D. in Computer Science University of Wisconsin - Madison • Advisor: Fred Sala M.S. in Computer Science University of Minnesota - Twin Cities B.Eng. Nanyang Technological University (NTU), Singapore	Sep 2021 - Present Sep 2019 - May 2021 Aug 2013 - Jun 2017
SELECTED CONFERENCE & JOURNAL PUBLICATIONS	Y. Zhang, D. Adila , C. Shin, and F. Sala, <i>Personalize Your LLM: Fake it then Align It</i> in North American Chapter of the Association for Computational Linguistics (NAACL) Findings, 2025. D. Adila , S. Zhang, B. Han, Y. Wang, <i>Discovering Bias in Latent Space: An Unsupervised Debiasing Approach</i> , in International Conference on Machine Learning (ICML), 2024. [Paper] [Code] D. Adila* , C. Shin*, L. Cai, F. Sala, <i>Zero-Shot Robustification of Zero-Shot Models</i> , in International Conference on Learning Representations (ICLR) 2024. [Paper] [Code] [Blog] N. Roberts*, X. Li*, D. Adila , S. Crompt, B. Huang, J. Zhao, F. Sala, <i>Geometry-Aware Adaptation for Pretrained Models</i> , in Neural Information Processing Systems (NeurIPS), 2023. [Paper] C. Shin, S. Crompt, D. Adila , S., F. Sala, <i>Mitigating Source Bias for Fairer Weak Supervision</i> , in Neural Information Processing Systems (NeurIPS), 2023. [Paper] M. Chen*, D. Fu*, D. Adila , M. Zhang, F. Sala, K. Fatahalian, C. Ré, <i>Shoring Up the Foundations: Fusing Model Embeddings and Weak Supervision</i> , in Uncertainty in Artificial Intelligence (UAI), 2022. Oral presentation (best student paper runner-up) . [Paper] [Code] N. Roberts*, X. Li*, B. Huang, D. Adila , S. Schoenberg, C. Liu, L. Pick, H. Ma, A. Albarghouthi, F. Sala, <i>AutoWS-Bench-101: Benchmarking Automated Weak Supervision with 100 Labels</i> , in Neural Information Processing Systems (NeurIPS), 2022. [Paper]	
WORKSHOP PUBLICATIONS	Dyah Adila , C. Shin, Y. Zhang, F. Sala, <i>Can Language Models Safeguard Themselves, Instantly and For Free?</i> , in ICML 2024 Next Generation of AI Safety Workshop. [Paper] Dyah Adila , C. Shin, L. Cai, F. Sala, <i>Zero-Shot Robustification of Zero-Shot Models</i> , in NeurIPS 2023 R0-FoMo Workshop. Oral presentation (best paper honorable mention) . (Workshop version of ICLR 2024 paper). Dyah Adila , and Dongyeop Kang, <i>Understanding Out-of-distribution: A Perspective of Data Dynamics</i> , in NeurIPS 2021 Workshop: I Can't Believe It's Not Better!, PMLR 2022. [Paper]	
AWARDS	Qualcomm Innovation Fellowship Finalist Best paper award honorable mention at NeurIPS R0-FoMo ICCV DataComp winning team in small scale filtering track UAI Best Student Paper Runner-up	2024 2023 2023 2022
INDUSTRY EXPERIENCE	Google Research, USA <i>Student Researcher</i> • Eliminated catastrophic forgetting in LLM fine-tuning via function-preserving expansion.	Summer 2025

Amazon Web Services AI, USA

Fall 2023, Summer 2024

Applied Scientist Intern

- Unsupervised denoising of retrieval augmented generation (RAG) using weak supervision.
- Attention steering for bias mitigation in foundation models. **Results published in ICML 2024.**

Traveloka, Indonesia

Jul 2017 - Jul 2019

Software Engineer

- Led the development of company-wide React Native user interface library (runs on Android and iOS), which speed up development time by 2x.
- Built Traveloka's customer-facing and business-facing mobile applications.

JPMorgan Chase, Singapore

May 2016 - Jul 2016

Software Engineer Intern

- Built a real-time log monitoring tool to keep track of daily transactions using Java.
- Built an API for multiple currency transfer application.

Seagate Technology, Singapore

Jan 2016 - May 2016

Research Engineer Intern

- Built a continuous integration framework to automate Seagate's build and test pipelines.

TECHNICAL SKILLS

- **Competent:** Python, PyTorch, TensorFlow, Java, Unix, Google Cloud Platform, JavaScript
- **Familiar:** SQL, AWS, Apache Spark