

# Adji Bouso Dieng

Department of Computer Science  
Princeton University  
35 Olden Street  
Princeton, NJ 08540

Email: [adji at princeton dot edu](mailto:adji@princeton.edu)  
Homepage: [vertaix.princeton.edu](http://vertaix.princeton.edu)  
Github: [github.com/vertaix](https://github.com/vertaix)  
Personal website: [adjiboussodieng.org](http://adjiboussodieng.org)

## Education

Ph.D in Statistics, Columbia University 2014 – 2020  
Advisors: David Blei, John Paisley  
My PhD thesis won the Savage Award

Master in Applied Statistics, Cornell University Jan 2012 – May 2013  
Advisors: David Lifka, Martin Wells

Diplome d'Ingenieur, Telecom ParisTech Sep 2009 – May 2013  
France's "Grandes Ecoles "

Lycee Henri IV (France's "Classes Preparatoires aux Grandes Ecoles") Sep 2007 – June 2009

Lycee Jacques Decour (France's "Classes Preparatoires aux Grandes Ecoles") Sep 2006 – June 2007

## Employment

Tenure-Track Assistant Professor of Computer Science September 2021 – Present  
Princeton University, Princeton, NJ

Founder and President May 2020 – Present  
The Africa I Know, Inc.  
<https://theafricaiknow.org>

Research Scientist August 2020 – Present  
Google AI, New York, NY  
Team: Google Brain

Research Intern September 2018 – January 2019  
DeepMind, London, UK  
Supervisor: Lei Yu

Research Intern May 2018 – August 2018  
Facebook AI Research, New York, NY  
Supervisor: Yann LeCun

Research Intern June 2017 – Sep 2017  
Microsoft AI & Research, Redmond, WA  
Supervisors: Chong Wang, Jianfeng Gao

Research Intern June 2016 – Sep 2016  
Deep Learning Technology Center, Microsoft Research, Redmond, WA  
Supervisors: Chong Wang, Jianfeng Gao

Junior Professional Associate Market and Counterparty Risk Department The World Bank, Washington, DC	June 2013- July 2014
Intern Doing Business Team International Finance Corporation, Washington, DC	Jan 2013 – May 2013
Research Intern Weill Cornell Medical College, New York, NY Supervisor: Martin T. Wells	May 2012 – Aug 2012

## Awards & Honors

Princeton Precision Health grant award in the amount of \$90K as a Co-PI	2024
Outstanding Recent Alumni Award by Columbia University	2023
AI2050 Early Career Fellow (\$300K)	2022
Annie T. Randall Innovator Award	2022
Princeton Catalysis Initiative grant award in the amount of \$224K as a Co-PI	2022
DataX grant award in the amount of \$125K as a Co-PI	2022
NSF grant award in the amount of \$15M as a Co-PI	2021
Winner of the Savage Award for my doctoral dissertation	2020
Nominated by Columbia University for the ACM SIGAI/AAAI Doctoral Dissertation Award	2020
Nominated by Columbia University for the ACM Doctoral Dissertation Award	2020
Rising Star in Machine Learning (University of Maryland)	2019–
Google PhD Fellowship in Machine Learning (\$35,000 + Tuition and Fees)	2019–
Google sponsored talk award at New York Academy of Sciences ML Symposium	March 2018
Open Philanthropy Project AI Fellowship, Finalist	February 2018
Conference Travel Awards (ICML 2018, NeurIPS 2017, ICLR 2017)	
Microsoft Azure Research Award (\$20K Azure Credit)	November 2016 – November 2017
Columbia University Dean Fellowship (Full Funding)	August 2014 – Present
Cornell Institute for African Development Fellowship (Full Tuition)	August 2013 – May 2013
Pathfinder Foundation Scholarship (€ 60K to study abroad)	2007 – 2011
Senegalese Government Excellence Scholarship	2006 – 2010
Laureate du Concours General (Senegalese Olympiad, Philosophy)	July 2006

## Advising & Mentorship

### *Postdoc & Student Advising*

Ines Meraoumia (Postdoc)	March 2024 – Present
Amey Pasarkar (PhD Student)	Spring 2023 – Present
Andre Niyongabo Rubungo (PhD Student)	Fall 2022 – Present
Berlin Chen (PhD Student)	Fall 2023 – Present
Tedi Zadouri (PhD Student)	Fall 2023 – Present
Constance Ferragu (MSE Student)	Fall 2022 – Present
Advika Srivastava (Undergrad)	Fall 2023 – Present
Akshat Agarwal (Undergrad)	Spring 2024 – Present
Aminah Aliu (Undergrad)	Spring 2024 – Present
Arin Gardner (Undergrad)	Fall 2023 – Present
Ava Crnkovic-Rubsamen (Undergrad)	Fall 2023 – Present
Edouard Kwizera (Undergrad)	Spring 2024 – Present
Tinney Mak (Undergrad)	Spring 2024 – Present
Viki Mancoridis (Undergrad)	Fall 2023 – Present

### *Student Mentorship*

Samarth Sinha, University of Toronto	2019–2020
Lucas Gen, University of Chicago	Spring 2022 – Present
Mohammad Reza Rezaei, University of Toronto	Summer 2023 – Present

## Professional Service & Leadership

### *Co-Founder & General Chair*

The Princeton AI Club	November 2021 – November 2022
A community-driven AI network at Princeton University with the goal to create a local community around AI and foster interdisciplinary collaborations.	
<a href="https://aiclub.princeton.edu/">https://aiclub.princeton.edu/</a>	

### *Founder & President*

The Africa I Know <sup>®</sup>	May 2020 – Present
A 501(c)(3) nonprofit organization advocating for STEM education and working to change narratives around African history, knowledge and innovation, and provide opportunities to young Africans.	
<a href="https://theafricaiknow.org">https://theafricaiknow.org</a>	

*Workshop & Conference Chairing*

Tutorials Chair, NeurIPS	2022
Women in Machine Learning Workshop, collocated with NeurIPS	2019
Black in AI Workshop, collocated with NeurIPS	2019

*Journal Reviewing*

Nature Communications	2023
IEEE Transactions on Pattern Analysis and Machine Intelligence	2020
IEEE Transactions on Neural Networks and Learning Systems	2018
Foundations and Trends in Machine Learning	2016

*Conference Reviewing*

Artificial Intelligence and Statistics (AISTATS)	2023, 2018
Conference on Empirical Methods in Natural Language Processing (EMNLP)	2019
Association for Uncertainty in Artificial Intelligence (UAI)	2019
Association for the Advancement of Artificial Intelligence (AAAI)	2019
International Conference on Learning Representations (ICLR)	2018
International Conference on Machine Learning (ICML)	2017, 2018
Neural Information Processing Systems (NeurIPS)	2016, 2017, 2018, 2019, 2020

*Workshop Reviewing*

Workshop on Invertible Neural Networks and Normalizing Flows, ICML	2020
Workshop on Practical Machine Learning for the Developing World, ICLR	2020
Methods for Optimizing and Evaluating Neural Language Generation Workshop	2019
Advances in Approximate Bayesian Inference Workshop	2018
Black in AI Workshop	2018

*Professional Membership*

Association for Computing Machinery  
 Association for Women in Mathematics  
 Bernoulli Society  
 Institute for Mathematical Statistics  
 American Statistical Association

*Workshop & Symposium Organization & Reviewing*

Machine Learning in the Physical Sciences – NeurIPS <a href="https://ml4physicalsciences.github.io/2020/">https://ml4physicalsciences.github.io/2020/</a>	2023, 2022, 2021, 2020
Machine Learning for Materials – ICLR <a href="https://www.ml4materials.com/">https://www.ml4materials.com/</a>	2023
Deep Generative Models for Highly Structured Data – ICLR <a href="https://deep-gen-struct.github.io/">https://deep-gen-struct.github.io/</a>	2022, 2019
Learning Meaningful Representations of Life – NeurIPS <a href="https://www.lmrl.org/2020-organizers">https://www.lmrl.org/2020-organizers</a>	2022, 2021, 2020
Energy-Based Models: Current Perspectives, Challenges, and Opportunities – ICLR <a href="https://sites.google.com/view/ebm-workshop-iclr2021">https://sites.google.com/view/ebm-workshop-iclr2021</a>	2021
Symposium on Advances in Approximate Bayesian Inference <a href="http://approximateinference.org/">http://approximateinference.org/</a>	2021, 2020, 2019

*Other Volunteering Activities*

Advisor – Science Meets Engineering of Deep Learning Workshop (SEDL), NeurIPS	2019
Deep Learning Indaba Awards Committee	2019
Mentor – Mentorship Roundtable – Women in Machine Learning Workshop (WiML)	2017
Student Volunteer – International Conference on Learning Representations (ICLR)	2017

**Teaching**

Instructor, Princeton University Introduction to Machine Learning	Fall 2024
Instructor, Princeton University Foundations of Probabilistic Modeling	Spring 2024
Instructor, Princeton University Deep Generative Models: Methods, Applications, Societal Considerations	Fall 2023
Instructor, Princeton University Foundations of Probabilistic Modeling	Spring 2023
Instructor, Princeton University Foundations of Probabilistic Modeling	Spring 2022
Teaching Assistant, Columbia University Statistical Machine Learning	Spring 2019
Teaching Assistant, Columbia University Advanced Data Analysis	Fall 2017
Teaching Assistant, Columbia University Statistical Methods for Finance	Spring 2016

Teaching Assistant, Columbia University Probability and Statistics for Data Science	Fall 2015
Teaching Assistant, Columbia University Linear Regression Models	Spring 2015
Teaching Assistant, Columbia University Probability	Fall 2014

## Communication Skills

Human Languages: Wolof(Native), French(Native), English(Fluent).

Machine Languages: Python (Pytorch, Tensorflow, Numpy, Scipy, PySpark), R(Stan), C#, Java, SQL.

Big Data Technologies: AWS, Azure.

Version Control: Git, SVN.

## Press On Research

TechXplore and Princeton Engineering. <i>Researchers leverage LLMs for materials discovery.</i>	March 2024
Schmidt Futures. <i>AI2050 Community Perspective.</i>	Jul 2023
Princeton University. <i>'Learning to see and learning to read': Artificial intelligence enters a new era.</i>	Jan 2023
New Scientist. <i>AI diversity scoring system could help root out algorithmic bias.</i>	Nov 2022

## Talks & Panels

Princeton Materials Institute Annual Symposium	April 2024
NeurIPS Workshop On Synthetic Data	December 2023
Materials Research Society Fall Meeting <i>Symposium On Opportunities In Materials Research</i>	November 2023
Princeton Precision Health Symposium	September 2023
Princeton Catalysis Initiative Annual Symposium	May 2023
Columbia LEAP Lectures in Climate Data Science	May 2023
Applied Machine Learning for Developing Countries ICLR Workshop	May 2023
Princeton Materials Institute Annual Symposium	April 2023
UAI Workshop on Tractable Probabilistic Models	August 2022
AI4ALL Annual Summer Camp, Princeton Chapter	July 2022
Association for Computing Machinery's TechTalk, Online	June 2021
Conference on Empirical Methods in Natural Language Processing	November 2020

AI4ALL Annual Summer Camp, Princeton Chapter	July 2020
ICML Workshop on Invertible Neural Networks and Normalizing Flows	July 2020
Johns Hopkins University CS Seminar, Baltimore, MD	March 2020
University of Pennsylvania CIS Seminar, Philadelphia, PA	March 2020
Princeton University CS Colloquium, Princeton, NJ	March 2020
Brown University CS Colloquium, Providence, RI	March 2020
Carnegie Mellon University CS Colloquium, Pittsburgh, PA	March 2020
Yale University Seminar, New Haven, CT	March 2020
New York University Courant Institute of Mathematical Sciences, New York, NY	February 2020
California Institute of Technology CMS Seminar, Pasadena, CA	February 2020
University of Chicago CS Seminar, Chicago, IL	February 2020
University of Illinois at Urbana Champaign CS Seminar, Urbana & Champaign, IL	February 2020
Twitter Cortex, New York, NY	December 2019
Google Brain, Mountain View, CA	December 2019
Microsoft Research, Redmond, WA	December 2019
Carnegie Mellon University Machine Learning Seminar, Pittsburgh, PA	November 2019
University of California Berkeley EECS Seminar, Berkeley, CA	October 2019
IPAM Workshop on Interpretable Learning in Physical Systems, Los Angeles, CA	October 2019
Deep Learning Indaba, Two-Hour Lecture on Deep Generative Models, Nairobi, Kenya	August 2019
Deep Learning Indaba, Women in Machine Learning and AI, Nairobi, Kenya (panel)	August 2019
Yahoo Research Seminar Series, New York, NY	July 2019
New York Machine Learning and Artificial Intelligence Meetup, New York, NY	June 2019
Latinx in AI and Black in AI joint Workshop, ICLR 2019, New Orleans, LA (keynote)	May 2019
NYU Text as Data Seminar Series, New York, NY	March 2019
NYAS Machine Learning Symposium, New York, NY (short spotlight talk)	March 2019
South England Natural Language Processing Meetup, UK	January 2019
Microsoft Research Cambridge, Cambridge, UK	January 2019
International Conference on Machine Learning, Sweden	July 2018
Columbia University Data Science Institute Student Seminar, New York, NY	April 2018
Columbia University Minghui Memorial Conference, New York, NY (short talk)	April 2018
Tufts University CS Colloquium, Medford, MA	April 2018

Harvard University NLP Group Meeting, Cambridge, MA	April 2018
Stanford University NLP Seminar, Stanford, CA	April 2018
NYAS Machine Learning Symposium, New York, NY (short spotlight talk)	March 2018
Women Techmakers 2018 Summit, Google, New York, NY (Panel)	Mach 2018
Machine Learning and Friends Seminar, UMass, Amherst, MA	February 2018
Black in AI Workshop, Long Beach, CA (short spotlight talk)	December 2017
MSR AI, Microsoft Research, Redmond, WA	August 2017
SSLI Lab, University of Washington, Seattle, WA	August 2017
DeepLoria, Loria Laboratory, Nancy, France	April 2017
AI With The Best, Online	April 2017
Columbia University Minghui Memorial Conference, New York, NY (short talk)	April 2017
OpenAI, San Francisco, CA	January 2017
IBM TJ Watson Research, Yorktown Heights, NY	December 2016

## Publications & Preprints

A. Li, Z. Ding, **A. B. Dieng**, R. Beeson. *Efficient and Guaranteed-Safe Non-Convex Trajectory Optimization with Constrained Diffusion Model*. Under Submission.

DOI: <https://arxiv.org/abs/2403.05571>

S. Jiang, **A. B. Dieng**, M. Webb. *Property-Guided Generation of Complex Polymer Topologies Using Variational Autoencoders*. Under Submission to npj Computational Materials.

DOI: [10.26434/chemrxiv-2024-h2xgs](https://doi.org/10.26434/chemrxiv-2024-h2xgs)

A. Pasarkar, **A. B. Dieng**. *Cousins Of The Vendi Score: A Family Of Diversity Metrics For Science And Machine Learning*. Under Submission To AISTATS 2023.

Link: <https://arxiv.org/abs/2310.12952>

A. N. Rubungo, C. Arnold, B. Rand, **A. B. Dieng**. *LLM-Prop: Predicting Physical And Electronic Properties Of Crystalline Solids Using Their Text Descriptions*. Submitted To ICLR 2023.

Link: <https://arxiv.org/abs/2310.14029>

A. Pasarkar, G. Bencomo, S. Olsson, and **A. B. Dieng**. *Vendi Sampling For Molecular Simulations: Diversity As A Force For Faster Convergence And Better Exploration*. Journal Of Chemical Physics, 2023.

<https://doi.org/10.1063/5.0166172>

D. Friedman and **A. B. Dieng**. *The Vendi Score: A Diversity Evaluation Metric for Machine Learning*. Transactions On Machine Learning Research, 2023.

Link: <https://arxiv.org/abs/2210.02410>

K. Kim, J. Oh, J. R. Gardner, **A. B. Dieng**, H. Kim. *Markov Chain Score Ascent: A Unifying Framework of Variational Inference with Markovian Gradients*. Neural Information Processing Systems (NeurIPS), 2022.

Link: <https://arxiv.org/abs/2206.06295>



F. L. Ruta, A. J. Sternbach, **A. B. Dieng**, A. S. McLeod, and D. N. Basov. *Quantitative nanoinfrared spectroscopy of anisotropic van der Waals materials*. Nano letters, 2021.

Link: <https://pubs.acs.org/doi/abs/10.1021/acs.nanolett.0c02671>

S. Sinha and **A. B. Dieng**. *Consistency Regularization for Variational Auto-Encoders*. Neural Information Processing Systems (NIPS), 2021.

Link: <https://arxiv.org/abs/2105.14859>

**A. B. Dieng**, F. J. R. Ruiz, D. M. Blei, and M. Titsias. *Prescribed Generative Adversarial Networks*.

Link: <https://arxiv.org/abs/1910.04302>

**A. B. Dieng** and J. Paisley. *Reweighted Expectation Maximization*.

Link: <https://arxiv.org/abs/1906.05850>

**A. B. Dieng**, F. J. R. Ruiz, and D. M. Blei. *Topic Modeling in Embedding Spaces*. Transactions of the Association for Computational Linguistics (TACL), 2020.

Link: <https://arxiv.org/abs/1907.04907>

**A. B. Dieng**, F. J. R. Ruiz, and D. M. Blei. *The Dynamic Embedded Topic Model*.

Link: <https://arxiv.org/abs/1907.05545>

**A. B. Dieng**, K. Cho, D. M. Blei, and Y. LeCun. *Learning with Reflective Likelihoods*.

Link <https://openreview.net/pdf?id=SJlh2jR9FX>.

D. Tran, A. Kucukelbir, **A. B. Dieng**, M. Rudolph, D. Liang, and D.M. Blei. *Edward: A Python library for probabilistic modeling, inference, and criticism*.

Link: <https://arxiv.org/abs/1610.09787>

**A. B. Dieng**, Y. Kim, A. M. Rush, and D. M. Blei. *Avoiding Latent Variable Collapse With Generative Skip Models*. Artificial Intelligence and Statistics (AISTATS), 2019.

Link: <https://arxiv.org/abs/1807.04863>

**A. B. Dieng**, Y. Kim, A. M. Rush, and D. M. Blei. *Avoiding Latent Variable Collapse With Generative Skip Models*. Workshop on Theoretical Foundations and Applications of Deep Generative Models, ICML, 2018.

Link: <https://arxiv.org/abs/1807.04863>

**A. B. Dieng**, R. Ranganath, J. Alotaib, and D. M. Blei. *Noisin: Unbiased Regularization for Recurrent Neural Networks*. International Conference on Machine Learning (ICML), 2018.

Link: <https://arxiv.org/abs/1805.01500>

F. R. Ruiz, M. Titsias, **A. B. Dieng**, and D. M. Blei. *Augment and Reduce: Stochastic Inference for Large Categorical Distributions*. International Conference on Machine Learning (ICML), 2018.

Link: <https://arxiv.org/abs/1802.04220>

D. Ciao, T. Ma, **A. B. Dieng**, D. M. Blei, and F. Wang. *Readmission Prediction via Deep Contextual Embedding of Clinical Concepts*. Plos One, 2018. Named as one of the **best articles published in 2018 on clinical information systems** by the International Medical Interpreters Association (IMIA).

**A. B. Dieng**, C. Wang, J. Gao, and J. W. Paisley. *TopicRNN: A Recurrent Neural Network with Long Range Semantic Dependency*. International Conference on learning Representation (ICLR), 2017.

Link: <https://arxiv.org/abs/1611.01702>

**A. B. Dieng**, D. Tran, R. Ranganath, J. W. Paisley and D. M. Blei. *Variational Inference via  $\chi$  Upper Bound Minimization*. Neural Information Processing Systems (NIPS), 2017.

Link: <https://arxiv.org/abs/1611.00328>

Last updated: April 5, 2024