# Kevin Christian Wibisono

*Department of Statistics, University of Michigan*
✉ kwib@umich.edu | ⌂ https://k-wib.github.io/ | ⊙ k-wib

## Research Interests

- Language models, transformers and in-context learning.
- Causal inference with text data.
- Regression discontinuity designs.

## Computing Skills

- Proficient in Python (including Tensorflow, PyTorch, Pandas, Numpy and PySpark), R (including tidyverse), and SQL.

## Education

**University of Michigan**                                                             *Ann Arbor, MI, USA*

Ph.D. in Statistics (GPA: 4.00/4.00; Advisor: Dr Yixin Wang)                   *2021 - 2026 (expected)*

- Received the *Rackham International Student Fellowship* for demonstrating exceptional academic and professional promise.
- Received one of the *International Biometric Society Eastern North American Region's Distinguished Student Paper Awards*.

**Columbia University**                                                               *New York, NY, USA*

M.S. in Data Science (GPA: 4.00/4.00)                                                 *2019 - 2020*

- Capstone project: *Improving Automatic Event Understanding Through Sequential and Non-sequential Deep Learning Architectures* (supervised by Yuval Marton, Ph.D. and Asad Basheer Sayeed, Ph.D.).

**National University of Singapore**                                                          *Singapore*

B.Sc. (First Class Honors) in Applied Mathematics and Statistics (GPA: 4.90/5.00)     *2015 - 2019*

- Honors thesis: *Approximate Solutions to the Multivariate Behrens-Fisher Problem* (supervised by Zhang Jin-Ting, Ph.D.).
- Completed the *Special Programme in Mathematics*, a program specially designed for students with a strong passion and aptitude for mathematics.
- Received the *Ho Family Prize* as the best student in Applied Mathematics.

## Publications

### Accepted

Wibisono, K. C. and Wang, Y. (2024). *From Unstructured Data to In-Context Learning: Exploring What Tasks Can Be Learned and When*. Neural Information Processing Systems (acceptance rate: 26%).

Ignaccolo, C., Wibisono, K. C., Plunz, R., and Sutto, M. (2024). *Tweeting During the Pandemic in New York City: Unveiling the Evolving NYC's Sentiment Landscape Through a Spatiotemporal Analysis of Geolocated Tweets*. Journal of Urban Technology.

Wibisono, K. C. and Wang, Y. (2023). *On the Role of Unstructured Training Data in Transformers' In-Context Learning Capabilities*. NeurIPS Workshop on Mathematics of Modern Machine Learning.

Wibisono, K. C. and Wang, Y. (2023). *Bidirectional Attention as a Mixture of Continuous Word Experts*. Uncertainty in Artificial Intelligence (acceptance rate: 31%).

### Submitted or In Preparation

Wibisono, K. C., Mukherjee, D., Banerjee, M., and Ritov, Y. *Estimation and Inference for the Average Treatment Effect in a Score-Explained Heterogeneous Treatment Effect Model*.

Wibisono, K. C. and Wang, Y. *Causal Inference with Text-Based Treatments and Outcomes*.

Wibisono, K. C. and Wang, Y. *Exponential Family Attention*.

# Work Experience

**Uber**                                                                *Sunnyvale, CA, USA*
PHD SOFTWARE ENGINEER INTERN (RIDER STRUCTURAL PRICING)                   *May - Aug 2024*
- Applied causal inference techniques to analyze the paycheck effect and its impact on demand and price elasticity.
- Enhanced Uber's demand model by adding paycheck features, improving predictive performance and business metrics.
- Implemented and refined large language model-based data augmentation methods, achieving notable improvements in test AUC and accuracy based on experiments conducted on Uber's demand data.

**Walmart (Sam's Club)**                                               *Bentonville, AR, USA*
JUNIOR DATA SCIENTIST (FRAUD)                                             *Feb - Jun 2021*
- Improved fraud detection system via model stacking and feature engineering, reducing financial losses by around 30%.

DATA SCIENTIST INTERN (PRICING)                                          *Jun - Aug 2020*
- Developed item-scoring algorithms to inform strategic price investment decisions for each Sam's Club.
- Adapted and implemented natural language processing algorithms to improve item elasticity predictions.

**Portcast**                                                                   *Singapore*
DATA SCIENTIST INTERN                                                     *May - Aug 2018*
- Devised methods to improve existing cargo demand forecasting models of leading shipping companies.
- Enhanced forecasting accuracy via market signal experimentation, reducing mean absolute percentage errors by 5 to 15%.

**Centre for Social Development Asia, National University of Singapore**        *Singapore*
RESEARCH INTERN                                                          *May - Aug 2017*
- Collaborated with a multidisciplinary team of undergraduate students to analyze the state of charity governance in Singapore, resulting in the publication *An Overview of Charity Governance in Singapore*.

**Traveloka**                                                          *Jakarta, Indonesia*
DATA ANALYST INTERN                                                      *May - Jul 2016*
- Developed dashboards to summarize the effectiveness of organic searches in generating hotel sales.

# Teaching and Mentorship Experience

RESEARCH MENTOR, **University of Michigan**

1. *Undergraduate Research Program in Statistics (URPS)*                 *Winter and Fall 2024*

   - Supervised three undergraduate students in a research project exploring large language models' geographical knowledge. Currently leading a follow-up project on the same topic.

GRADUATE STUDENT INSTRUCTOR, **University of Michigan**

1. *STATS 306: Introduction to Statistical Computing* (upper undergraduate level)          *Fall 2023*

   - Designed and taught weekly lab sections (~40 students), held office hours, prepared and graded homework, and graded exams.

2. *STATS 206: Introduction to Data Science* (lower undergraduate level)     *Fall 2022 and Winter 2023*

   - Taught weekly lab sections (~40 students), held office hours, and graded homework and exams.

3. *STATS 415: Data Mining and Statistical Learning* (upper undergraduate level)        *Winter 2022*

   - Designed and taught weekly lab sections (~40 students), held office hours, and graded homework and exams.

4. *STATS 250: Introduction to Statistics and Data Analysis* (lower undergraduate level)     *Fall 2022*

   - Taught weekly lab sections (~40 students), held office hours, and graded homework and exams.

TEACHING ASSISTANT, **Columbia University**

1. *INAF U6614: Data Analysis for Policy Research Using R* (graduate level)                                    *Fall 2020*
   - Designed and taught weekly lab sections (~20 students), developed course website, held office hours, and graded homework and projects.

2. *CSOR W4231: Analysis of Algorithms* (upper undergraduate/graduate level)                                  *Spring 2020*
   - Held office hours, and graded homework and exams.

TEACHING ASSISTANT, **National University of Singapore**

1. *MA1512: Differential Equations for Engineering* (lower undergraduate level)                      *AY 2018/19 Semester 2*
   - Designed and taught weekly tutorial classes (~60 students), and held office hours.

2. *MA1100: Fundamental Concepts of Mathematics* (lower undergraduate level)                        *AY 2018/19 Semester 1*
   - Designed and taught weekly tutorial classes (~60 students), held office hours, and graded exams.

## Awards

2024 **ENAR Distinguished Student Paper Award**, International Biometric Society

2023 **Rackham International Student Fellowship**, University of Michigan

2020 **Top 1.5%**, Baidu and Xi'an Jiaotong University International Big Data Competition

2019 **Second Runner-Up**, Columbia University Data Science Hackathon
     **Ho Family Prize (Best Graduate in Applied Mathematics)**, National University of Singapore

2017 **Top 250**, William Lowell Putnam Mathematical Competition
     **First Runner-Up**, ALMA College Mathematics Challenge

2015 **Undergraduate Scholarship (Full Funding)**, Singapore Ministry of Foreign Affairs

2014 **Silver Medal**, Asian Pacific Mathematics Olympiad

2013 **Bronze Medal**, International Mathematical Olympiad
     **Silver Medal**, Asian Pacific Mathematics Olympiad

2009 **Satyalancana Wira Karya (Medal for Providing an Example of Meritorious Personality)**, Government of the Republic of Indonesia

## Posters and Presentations

POSTERS

2024 **Neural Information Processing Systems (NeurIPS)**, Vancouver, BC, Canada
     **Midwest Speech and Language Days**, Ann Arbor, MI, USA
     **Michigan Student Symposium for Interdisciplinary Statistical Science**, Ann Arbor, MI, USA
     **NeurIPS Workshop on Mathematics of Modern Machine Learning**, New Orleans, LA, USA
2023 **Michael Woodroofe Memorial Conference**, Ann Arbor, MI, USA
     **Uncertainty in Artificial Intelligence (UAI)**, Pittsburgh, PA, USA
     **ICSA Applied Statistics Symposium**, Ann Arbor, MI, USA
     **Midwest Machine Learning Symposium**, Chicago, IL, USA

PRESENTATIONS

2024 **Joint Statistical Meetings**, Portland, OR, USA

**Michigan Student Symposium for Interdisciplinary Statistical Science**, Ann Arbor, MI, USA

2023 **Joint Statistical Meetings**, Toronto, ON, Canada

# Outreach and Professional Development

SERVICE AND OUTREACH

2023 **Neural Information Processing Systems**, Volunteer

**ICSA Applied Statistics Symposium**, Volunteer

2021 - 2022 **University of Michigan Indonesian Society**, Events Director

2020 **Columbia University Data Science Institute**, Student Ambassador and Mentor

2019 - 2020 **Columbia Indonesian Society**, Assistant Vice President of Training and Development

JOURNALS, CONFERENCES AND WORKSHOPS REFEREED

2024 **Cities**

**Neural Information Processing Systems (NeurIPS)**

**ICML Workshop on In-Context Learning**

**ICML Workshop on Theoretical Foundations of Foundation Models**

2023 - 2024 **Artificial Intelligence and Statistics (AISTATS)**