

# Abdul Wasay

[awasay@seas.harvard.edu](mailto:awasay@seas.harvard.edu) | [scholar.harvard.edu/awasay](https://scholar.harvard.edu/awasay)

**Research Interests.** My research is at the intersection of systems and machine learning. I identify co-design opportunities to develop techniques that enable better data exploration, model design, training, and deployment.

## EDUCATION

---

**Harvard University** Cambridge, MA  
*Doctor of Philosophy in Computer Science* (GPA:3.8/4) August 2014 – June 2021 (expected)  
• Thesis: Computation-Cautious Machine Learning Systems (Advisor: Prof. Stratos Idreos)

**Harvard University** Cambridge, MA  
*Masters of Science in Computer Science* (GPA:3.75/4) August 2014 – May 2016

**Lahore University of Management Sciences** Lahore, Pakistan  
*Bachelors of Science in Computer Science* (GPA:3.9/4) August 2010 – May 2014

## RESEARCH EXPERIENCE

---

**Harvard University: Graduate Researcher** August 2014 – Present  
• Designed and implemented four different systems to accelerate machine learning pipelines advancing the state-of-the-art in data exploration, deep learning model design, training, and deployment (*8 publications, 2 patents*)  
• Collaborated with researchers to automate the design of data systems through reinforcement learning techniques  
• Mentored nine undergraduate researchers from Harvard, CMU, and Peking University (*3 publications, 1 award*)  
• Presented research in various database systems, data science, and machine learning research conferences as well as industry research labs and start-ups

**Hewlett Packard Enterprise: Researcher** Summer 2017  
• Designed a technique to reduce the overhead of storing and querying network traffic data by 10× (*1 patent*)  
• Collaborated with the engineering team to incorporate this research into their product, Arcsight

**École Polytechnique Fédérale de Lausanne (EPFL): Undergraduate Researcher** Summer 2013  
• Developed and evaluated an algorithm to incorporate community structure for social graph de-anonymization

## NOTABLE LEADERSHIP & TEAMWORK EXPERIENCE

---

**Harvard University: Instructor at Harvard Summer School** Summer 2019  
• Designed and taught Harvard's first hybrid course on data systems and machine learning (*course rating: 4.8/5*)  
• Supervised a course staff of two Ph.D. students to run the course

**Ashesi University: Visiting Faculty (Archer-Cornfield Fellowship)** August 2019 – June 2020  
• Designed and taught three undergraduate courses on data systems and machine learning (*100 students total*)  
• Mentored four undergraduate students on their senior-year research projects  
• Lead the curriculum redesign committee to refresh the coursework for the data science and machine learning track

**Harvard College: Resident Tutor at Quincy House** August 2015 – May 2019  
• Advised 7 Harvard College sophomores on academic matters and managed an entryway of 30 students per year  
• Lead the LGBTQ+ Committee (2 tutors, 10 students)  
• Co-founded the InQlusivity initiative to improve inclusion and diversity within the residential house community

## NOTABLE AWARDS

---

|   |             |
|---|-------------|
| • Archer-Cornfield Fellowship                                     | 2019 – 2020 |
| • Second place at ACM SIGMOD Research Competition (student award) | 2020        |
| • Most Reproducible Paper Award at SIGMOD                         | 2017        |
| • Summer@EPFL Research Fellowship                                 | 2013        |
| • Dean's Honor List at LUMS + NOP Scholarship                     | 2010 – 2014 |

## PUBLICATIONS & PATENTS

---

Publications in top-tier deep learning, machine learning, and data systems venues: **ICLR, MLSys, SIGMOD, SIGCOMM, Data Engineering, and BigData.**

Deep Learning: Systems and Responsibility.

**Abdul Wasay**, Subarna Chatterjee, and Stratos Idreos.

ACM International Conference on Management of Data *SIGMOD*, 2021.

More or Less: When and How to Build Convolutional Neural Network Ensembles.

**Abdul Wasay** and Stratos Idreos.

International Conference on Learning Representation *ICLR* 2021.

MotherNets: Rapid Deep Ensemble Learning.

**Abdul Wasay**, Brian Hentschel, Yuze Liao, Sanyuan Chen, and Stratos Idreos.

Conference on Machine Learning and Systems *MLSys* 2020.

Learning Data Structure Alchemy.

Stratos Idreos, Kostas Zoumpatianos, Subarna Chatterjee, Wilson Qin, **Abdul Wasay**, Brian Hentschel, Mike Kester, Niv Dayan, Demi Guo, Minseo Kang, and Yiyu Sun

Bulletin of the IEEE Computer Society Technical Committee on *Data Engineering*, 2019.

The Periodic Table of Data Structures.

Stratos Idreos, Konstantinos Zoumpatianos, Manos Athanassoulis, Niv Dayan, Brian Hentschel, Michael S. Kester, Demi Guo, Lukas Maas, Wilson Qin, **Abdul Wasay**, Yiyu Sun.

Bulletin of the IEEE Computer Society Technical Committee on *Data Engineering*, 2018.

Data Canopy: Accelerating Exploratory Statistical Analysis.

**Abdul Wasay**, Xinding Wei, Niv Dayan, and Stratos Idreos.

ACM International Conference on Management of Data *SIGMOD*, 2017.

Queriosity: Automatic Data Exploration.

**Abdul Wasay**, Manos Athanassoulis, and Stratos Idreos.

IEEE International Congress on *BigData*, 2015.

Rethinking Buffer Management In Data Center Networks.

Aisha Mushtaq, Asad K. Ismail, **Abdul Wasay**, Bilal Mahmood, Ihsan A. Qazi, and Zartash A. Uzmi.

ACM International Conference on Data Communications *SIGCOMM*, 2014.

Network Traffic Data Summarization

**Abdul Wasay**, Alkiviadis Simitsis, Martin Arlitt.

*U.S. Patent* 10,432,539. Issued on October 1, 2019.

## INVITED TALKS

---

|   |      |
|---|------|
| Intel Labs, Morgan Stanley ML Labs, UCSD, Microsoft Research NYC, Boston University | 2021 |
| LUMS, Tableau   | 2020 |
| Google Research Accra, Ashesi University  | 2019 |
| Vertica, NetApp   | 2018 |

## TECHNICAL SKILLS

---

**Selected Languages:** C/C++, Python, Java, SQL.

**Deep Learning Infrastructure:** PyTorch, TensorFlow, Caffe, NumPy, Pandas, Plotly, AWS, GCP, Azure, SLURM.

**Selected Skills:** Computer architecture, machine learning, reinforcement learning, deep learning, data systems implementation, algorithms, data visualization, statistics, empirical research, technical writing, educational psychology, university teaching, public speaking.