

Abhinav Java

Microsoft Research

 Google Scholar  <https://java-abhinav07.github.io/>  java.abhinav99@gmail.com  [abhinav-java-2b57a4187](https://www.linkedin.com/in/abhinav-java-2b57a4187)

Education

Delhi Technological University (DCE), India
Bachelor's of Technology in Computer Engineering
> Research Excellence Award, 2022.








Aug 2018 - Jul 2022
Major GPA: 9.08/10, CGPA: 8.87/10

Experience

Microsoft Research, India Research Fellow with Amit Sharma, Nagarajan Natarajan, Srivathsan Koundinyan	Aug 2024 - Present
University of Virginia Aikyam Lab, Visiting Researcher with Chirag Agarwal	Mar 2024 - Mar 2025
Adobe Inc. Media & Data Science Research (MDSR), India Research Associate II with Mausoom Sarkar, Balaji Krishnamurthy	Aug 2022 - Aug 2024
Adobe Inc. Media & Data Science Research (MDSR), India Research Intern with Mausoom Sarkar, Balaji Krishnamurthy	May 2021 - Aug 2022
Massachusetts Institute of Technology, Remote Collaborator with Ayush Chopra, Ramesh Raskar	Mar 2021 - Oct 2022

Selected Publications

* denotes equal contribution

- FrugalRAG: Learning to reason and retrieve for multi-hop QA** 
Abhinav Java, Srivathsan Koundinyan, Nagarajan Natarajan, Amit Sharma
International Conference on Machine Learning, Workshop on Efficient Systems for Foundation Models [ICML ES-FoMo '25]
- Towards Operationalizing Right to Data Protection** 
Abhinav Java*, Simra Shahid*, Chirag Agarwal
North American Chapter of the Association for Computational Linguistics [NAACL '25, ICLR MLDPR '25]
- ReEdit: Multimodal Exemplar-Based Image Editing** 
Ashutosh S*, Tarun Ram*, Abhinav Java*, Avadhoot J, Silky Singh, Surgan Jandial, Balaji Krishnamurthy
Winter Conference on Applications of Computer Vision [WACV '25]
- All Should Be Equal in the Eyes of LMs: Counterfactually Aware Fair Text Generation** 
Pragyan B*, Abhinav Java*, Surgan Jandial*, Simra Shahid*, Shaz F, Balaji Krishnamurthy, Sumit Bhatia
Association for the Advancement of Artificial Intelligence [AAAI '24]
- Towards Fair Knowledge Distillation with Student Feedback** 
Abhinav Java*, Surgan Jandial*, Chirag Agarwal
International Conference on Machine Learning, Workshop on Efficient Systems for Foundation Models [ICML ES-FoMo '23]
- One-Shot Doc Snippet Detection: Powering Search in Document Beyond Text** 
Abhinav Java*, Shripad Deshmukh*, Milan Aggarwal, Surgan Jandial, Mausoom Sarkar, Balaji Krishnamurthy.
Winter Conference on Applications of Computer Vision [WACV '23]
- Learning to Censor by Noisy Sampling** 
Ayush Chopra, Abhinav Java, Abhishek Singh, Vivek Sharma, Ramesh Raskar
European Conference on Computer Vision [ECCV '22]

Featured Research Projects

Multi Hop Reasoning for Retrieval Augmented Generation Advisors: Amit Sharma, Nagarajan Natarajan, Srivathsan Koundinyan > Highlighted the importance of latency as a metric for multi-hop RAG, and showed that frustratingly simple optimized ReAct prompting matches the performance of newer RAG methods, but at the cost of inference time latency. > Proposed FRUGALRAG, a two-stage RAG framework that achieves state-of-the-art accuracy and efficiency on multi-hop QA benchmarks with minimal search cost [ICML ES-FoMo '25] > Currently, co-leading the development of verifiable reward models for long-form answer generation in language models.	Aug 2024 - Present Microsoft Research
--	--

A Model Agnostic Approach to Create Unlearnable Text Datasets

Mar 2024 - Mar 2025

Advisor: [Chirag Agarwal](#)

University of Virginia

- > Protecting public information on the web from being used to train LMs is critical from a policy perspective.
- > Proposed a simple mechanism to perturb sensitive datasets using information theoretic insights. [NAACL (Main) '25]
- > Demonstrated that finetuning large models like GPT-4o on our datasets drops their performance below zero shot level!

Multimodal Foundation Model and AEM Forms

May 2021 - Aug 2024

Advisors: [Mausoom Sarkar](#), [Balaji Krishnamurthy](#)

Adobe MDSR

- > Led the development of large document snippet search datasets to extend traditional document matching to multimodal setting (3 US Patents).
- > Designed an architecture fusing vision, text, spatial cues to improve performance on PubLayNet (+12.83 mAP) [WACV'23]
- > Drove the Multimodal Foundation Model initiatives, created task-enriched document instruction tuning datasets and optimized LLaVA attention for high-resolution image understanding in [Automated Forms Conversion Service](#).

Protecting Private Information in 3D Point Clouds for Perception Tasks

Mar 2021 - Oct 2022

Advisors: [Ayush Chopra](#), [Ramesh Raskar](#)

Camera Culture, MIT Media Lab

- > Perception tasks on 3D point clouds (e.g. medical scans) rely on a small, task-relevant subset of points.
- > These critical points may reveal sensitive attributes (e.g. race), hindering data sharing b/w untrusted sources.
- > Introduced an end-to-end task *invariant differentiable point cloud sampler* to enhance privacy-utility tradeoffs. [ECCV '22]

Patents

- > Mask-CLIPstyler: Localized text-based style transfer in images
- > One Shot Document Snippet Search (P11686-US · Filed Jun 30, 2023)
- > Personalized Form Error Correction Propagation (P12004-US · Filed Apr 27, 2023)
- > Form structure similarity detection (P11882-US · Filed Mar 27, 2023) 🔗
- > Systems and methods for data augmentation (P11364-US · Filed Oct 11, 2022) 🔗

Honors and Awards

Research Excellence Award, 2022 Awarded for undergraduate research during B.Tech (1 in 2000 students)

Semester Merit Award, 2021 For securing University Rank 1 in the 5th Semester.

University Football Team, 2018-2019 Won gold medal at IIT Kanpur Sports Fest.

National Talent Search Exam, 2017 Qualified Stage 1 (5000 out of 1.2 Million across India)

JENESYS Program, 2016 Selected by Central Board of Education for youth football-cultural exchange to Japan.

Teaching and Leadership

Invited Talk at AI Summer School DTU, India

Jul 2025

- > Will deliver a talk on multimodal learning to undergraduate students at Delhi Technological University.

Pre Placement Talk at DTU Adobe, MDSR

Jul 2023

- > Delivered a talk on introduction to research as an undergraduate student. [Slides]
- > Facilitated an interactive panel with professors and students, fostering dialogue on research opportunities.

UAS-DTU (University Tech Team) [🔗, 🔗] Vice Captain, Software, Airframes Technician

2019 - 2020

- > Secured funding from Adani Group by writing grants proposals and delivered presentations for lab (\$1M).
- > Participated in the IAF Mehar Baba Prize, and collectively developed a swarm solution with a 50km range.
- > Won the best communication architecture prize (INR 250K) for its use in disaster relief situation. [Media]
- > Taught basic programming and introduced deep learning to newcomers at Team UAS-DTU.

Academic Service and Select Coursework

Conference Reviewer KDD '24, CVPR'24, U&ME@ECCV'24, WACV'23, ML4PS@NeurIPS'22, ECCV'22

Invited Talk Oral talk at AAAI '24 - Prompting for Debiasing LMs (Student Abstract) 🔗

Select Coursework **Machine Learning:** Pattern Recognition, Computer Vision, Deep Learning, Natural Language Processing, CS231n-Stanford (MOOC), Deep Learning by Andrew Ng (Coursera). **Mathematics:** Discrete Math, Linear Algebra (NPTEL), Probability and Statistics (NPTEL). **Computer Science:** Object Oriented Prog., Data Structures, Algorithms, Software Eng.