Abhinav Java

Research Fellow, **Microsoft Research**

⊙ java-abhinav07.github.io   @ java.abhinav99@gmail.com   🎓 Google Scholar

Education

Delhi Technological University, *India*   08/2018 – 07/2022
B.Tech. in Computer Engineering
> *Research Excellence Award* 2022.

Experience

Microsoft Research, *India*   08/2024 – Present
Research Fellow with Amit Sharma, Nagarajan Natarajan, Srivathsan Koundinyan

University of Virginia, *Remote*   03/2024 – 03/2025
Visiting Researcher with Chirag Agarwal

Adobe Research, *India*   08/2022 – 08/2024
Research Associate I, II with Balaji Krishnamurthy

Adobe Research, *India*   05/2021 – 08/2021
Research Intern with Balaji Krishnamurthy

Massachusetts Institute of Technology, *Remote*   03/2021 – 10/2022
Collaborator with Ayush Chopra, Ramesh Raskar

Publications
                                                                        * denotes joint first authorship

[1] **Characterizing Deep Research: A Benchmark and Formal Definition** 📄○
Abhinav Java*, Ashmit Khandelwal*, Sukruta Midigeshi*, Aaron Halfaker, Amit Deshpande, Navin Goyal, Ankur Gupta, Nagarajan Natarajan, Amit Sharma
NeurIPS 2025 workshop on Scaling Environments for Agents   **[SEA@NeurIPS '25, Under Review ICLR '26]**

[2] **FrugalRAG: Learning to reason and retrieve for multi-hop QA** 📄○
Abhinav Java, Srivathsan Koundinyan, Nagarajan Natarajan, Amit Sharma
NeurIPS 2025 Workshop on Efficient Reasoning   **[ER@NeurIPS '25, Under Review ICLR '26]**

[2] **Understanding Task Transfer in VLMs**
Bhuvan Sachdeva*, Karan Uppal*, Abhinav Java*, Vineeth N Balasubramanian
NeurIPS 2025 Workshop on Unified Representations in Neural Models   **[UniReps@NeurIPS '25, Under Review CVPR '26]**

[3] **Towards Operationalizing Right to Data Protection** 📄○
Abhinav Java*, Simra Shahid*, Chirag Agarwal
North American Chapter of the Association for Computational Linguistics   **[NAACL (Main) '25, MLDPR@ICLR '25]**

[4] **ReEdit: Multimodal Exemplar-Based Image Editing** 📄○
Ashutosh Srivastava*, Tarun Ram*, Abhinav Java*, Avadhoot Jadav, Silky Singh, Surgan Jandial, Balaji Krishnamurthy
Winter Conference on Applications of Computer Vision   **[WACV '25, AI4VA@ECCV '24]**

[5] **All Should Be Equal in the Eyes of LMs: Counterfactually Aware Fair Text Generation** 📄○
Pragyan Banerjee*, Abhinav Java*, Surgan Jandial*, Simra Shahid*, Shaz Furniturewala, Balaji Krishnamurthy, Sumit Bhatia
Association for the Advancement of Artificial Intelligence   **[AAAI '24]**

[6] **Thinking fair and slow: On the efficacy of structured prompts for debiasing language models** 📄
Shaz Furniturewala, Surgan Jandial, Abhinav Java, Pragyan Banerjee, Simra Shahid, Sumit Bhatia, Kokil Jaidka
Empirical Methods in Natural Language Processing   **[EMNLP (Main) '24]**

[7] **One Shot Doc Snippet Detection: Powering Search in Document beyond Text** 📄
Abhinav Java, Shripad Deshmukh, Milan Agarwal, Surgan Jandial, Mausoom Sarkar, Balaji Krishnamurthy
Winter Conference on Applications of Computer Vision   **[WACV '23]**

[8] **Learning to Censor by Noisy Sampling** 📄○
Ayush Chopra, Abhinav Java, Abhishek Singh, Vivek Sharma, Ramesh Raskar
European Conference on Computer Vision   **[ECCV '22]**

Patents

> Personalized form error correction propagation (*US Patent 12394236*) 📄   2025/08/19
> One Shot Doc Snippet Detection (*US Patent 20250005048A1*) 📄   2025/01/02
> Form structure similarity detection (*US Patent 12124497B1*) 📄   2024/10/22
> Systems and methods for data augmentation (*US Patent 20240119122A1*) 📄   2024/04/11

Featured Projects and Collaborations

Efficient Multi-Hop Reasoning for Retrieval Augmented Generation   08/2024 – Present
w/ Amit Sharma, Nagarajan Natarajan, Srivathsan Koundinyan   Microsoft Research
> *Led* the development of a project to create efficiently adaptable and low-latency RAG system with SLMs.
> *Proposed* FrugalRAG: by learning efficient use of test-time compute using reinforcement learning
> *Achieved* SoTA recall with 100x ↓ examples, 2x ↓ latency. [**ES-FoMo@ICML'25, ER@NeurIPS'25**]

Rethinking Taskonomy in the era of Large Vision Language Models   04/2025 - Present
w/ Vineeth N Balasubramanian   Microsoft Research
> *Proposed* the first systematic framework to study transferability of perception tasks in Vision-Language Models (VLMs).
> *Introduced* a novel metric to quantify cross-task influence in VLM finetuning. [**UniReps@NeurIPS'25**]
> *Revealed* actionable insights – beneficial task cliques and transfer trends for effective finetuning strategies.

A Model Agnostic Approach to Create Unlearnable Text Datasets   03/2024 – 03/2025
w/ Chirag Agarwal   University of Virginia
> *Developed* the first model-agnostic, efficient unlearnable text data generation framework for user data privacy.
> *Proposed* RegText: a mechanism to perturb sensitive datasets using information-theoretic insights. [**NAACL'25**]
> *Demonstrated* that finetuning large models like GPT-4o on our datasets drops their performance below zero-shot level!

Protecting Private Information in 3D Point Clouds for Perception Tasks   03/2021 – 10/2022
w/ Ayush Chopra, Ramesh Raskar   Camera Culture, MIT Media Lab
> *Designed* an approach to simplify the release of sensitive point cloud data with privacy protection (e.g. medical scans).
> *Implemented* CBNS: an end-to-end differentiable sampler and a novel loss to protect point clouds from adversaries.
> *Demonstrated* improved privacy-utility tradeoffs with CBNS on perception tasks on 3D face datasets. [**ECCV'22**]

Leadership, Invited Talks, and Service

Fireside Chat on Deep Research, Plutos Dev. *Speaker* 📄   Fall 2025

AI Summer School, Delhi Technological University. *Career Talk*   Summer 2025

Oral Talk, AAAI. *On the Efficacy of Prompting for Debiasing LLMs*   Winter 2024

Pre-placement Talk, Delhi Technological University. *Adobe Hiring* 📄   Summer 2024

Team UAS-DTU, Delhi Technological University. *Vice Captain, Software & Airframes Technician*   2019 - 2020
> *Secured* a funding of **$1M** from Adani Group for building Drone Swarm for disaster relief.
> *Won* the best communication architecture prize of **INR 250K** in IAF National Competition. [Media]

Reviewer: ICLR '26, CVPR '25, KDD '24, CVPR'24, U&ME@ECCV'24, WACV'23, ML4PS@NeurIPS'22, ECCV'22

Honors and Awards

NeurIPS Scholar Award, 2025   Awarded funding for travel to NeurIPS 2025 by the conference.

AAAI Oral Talk, 2024   Selected for an invited talk at AAAI 2024, Student Abstract Track.

Research Excellence Award, 2022   Awarded for undergraduate research during undergrad (1 in approx 2000).

Semester Merit Award, 2021   For securing *University Rank 1* in the 5th Semester (1 in approx 2000).

Joint Entrance Exams, 2018   Secured 99.35 percentile in Main and 97.5 percentile in Advanced (Top 1–2% in 1.2 million).

National Talent Search Exam, 2017   Qualified Stage 1 (Top 1-2% in 1.2 million).

Select Coursework

> **Mathematics**: Discrete Math, Linear Algebra, Probability & Statistics, Modeling and Simulation.
> **Machine Learning**: Pattern Recognition, Computer Vision, Deep Learning, Natural Language Processing, CS231n-Stanford (MOOC), Reinforcement Learning, Big Data Analytics
> **Computer Science**: Object Oriented Programming, Data Structures, Algorithms, Software Engineering, Database Management Systems, Operating Systems, Computer Networks