

# Harshit Joshi

harshitjoshi.com | LinkedIn | Github  
josharshit@gmail.com | harshitj@stanford.edu

## EDUCATION

**STANFORD UNIVERSITY**  
PHD IN COMPUTER SCIENCE  
2023-29 | Stanford CA  
Cum. GPA: 4.1/4

**UNIVERSITY OF DELHI**  
BTECH IN INFORMATION  
TECHNOLOGY  
2017-21 | New Delhi  
Cluster Innovation Center  
Department Rank: 2  
Cum. Percentage: 87.78%

## LINKS

Github:// [duskybomb](#)  
LinkedIn:// [harshitjos](#)  
Twitter:// [duskybomb](#)

## SKILLS

### PROGRAMMING

Over 5000 lines:

Python • R •  $\LaTeX$

• PyTorch

Over 1000 lines:

Java • C++ • C#

• MySQL • Kafka • MongoDB

Familiar:

Javascript • F#

## RESEARCH FOCUS

LLM-driven workflows, LLM Agents,  
NLP, PL4ML, ML4Code

## SERVICES

### Volunteer

EMNLP 2020, EACL 2021, ICWSM  
2021, AAAI 2024

### Program Committee

AAAI 2024, ICML CA2MH 2021

## INVITED TALKS

AAAI 2024, OOPSLA 2022, WSDM  
2021, EMNLP 2020, PyData Delhi  
2019, University of Delhi (2018-2023)

## EXPERIENCE

### MICROSOFT RESEARCH | RESEARCH FELLOW (PROSE)

November 2021 - Present | Remote | Redmond, US

- Designed and trained domain-specific LLMs. Covered by [media outlets](#) (AAAI '24).
- Built SOTA production-ready neurosymbolic program repair framework (OOPSLA '22)
- Created multilingual repair framework, RING, built on top of LLMs (AAAI '23).

### SUPEDIO GMBH | RESEARCH SOFTWARE ENGINEER

January 2021 - October 2021 | Remote | HQ: Dresden, Germany

- Built automation tools, reducing person-hours employed per day by 400%
- Designed symbolic framework for the pdf data extraction for production.

### CRONCYCLE LTD. | SOFTWARE ENGINEERING INTERN

Jan 2019 – Oct 2019 | New Delhi, India | HQ: UK

- Ported batch jobs to live production, latency by 5 minutes.
- Built a new data pipeline to increase RSS data dump by 10%.

## RESEARCH

### IIITD - MIDAS LAB | RESEARCHER

May 2020 – Present | New Delhi, India

Working with **Ramit Sawhney** and **Prof Rajiv Ratn Shah** on Suicide Risk Assessment on social media using NLP and data mining techniques (EMNLP, WSDM, EACL, ICWSM, NAACL).

### DEFENCE R&D ORGANISATION (DRDO) | RESEARCHER

May 2019 – October 2019 | New Delhi, India

Worked on navigation of bot in Gazebo environment. Used CityScape Dataset for Image Segmentation and used the features in Gazebo environment, building algorithms for depth perception and creating disparity maps.

## AWARDS

2020	Honorable Mention	COMAP Mathematical Contest in Modeling
2019	Invited Talk	PyData Delhi 19 Conference
2019	National	Mathematical Finance Scholar Summer Fellowship 2019
2018	International	Google Summer of Code 2018
2017	Honourable Mention	ACM-ICPC 2018 (Regionals)

## PUBLICATIONS

1. **H. Joshi**, A. Ebenezer, J. Cambroner, S. Gulwani, A Kanade, V. Le, I. Radicek and G. Verbruggen. *FLAME: A small language model for spreadsheet formulas* Under Submission, 2023 <https://arxiv.org/abs/2301.13779>.
2. **H. Joshi**, J. Cambroner, S. Gulwani, V. Le, I. Radicek and G. Verbruggen. *Repair Is Nearly Generation: Multilingual Program Repair with LLMs*. AAAI 2023 <https://arxiv.org/abs/2208.11640>.
3. R. Bhavishi\*, **H. Joshi\***, J. Cambroner, A. Fariha, S. Gulwani, V. Le, I. Radicek and Ashish Tiwari. *Neurosymbolic Repair for Low-Code Formula Languages*. In In Proceedings of the ACM on Programming Languages (OOPSLA) 2022.