

## Education

Jul 2023	<b>Ahmedabad University</b>	Ahmedabad, India
Aug 2019	B.Tech in Computer Science & Engineering (CSE) Thesis: Modelling the spread of Hate Speech on Twitter   Advisors: Prof. Amit Nanavati, Dr. Seema Nagar	

## Experience

Present	<b>Tattle Civic Tech</b> [🌐]	Remote
Aug 2023	Software Development Engineer - I and Data Engineer   Advisors: Tarunima Prabhakar, Denny George → I help build citizen centric open-source tools and datasets to understand and respond to <b>inaccurate</b> and <b>harmful</b> content. → I help build a browser extension <b>Uli</b> , that redact's slurs and abusive content, archive's problematic content, and collectively pushes back against <b>online gender-based violence</b> . Uli de-normalizes everyday violence that people of marginalized genders experience online in India. [🌐] [Media Coverage: UNFPA Report] → I lead the development of <b>Feluda</b> , a configurable engine for analyzing multilingual and multimodal content. Feluda helps fact-checkers and researchers in combating misinformation. [🌐] [📄] [Media Coverage: UNDP] → Feluda is being used as the back-end analysis engine on a national Deepfake detection <b>WhatsApp helpline</b> ( <b>Deep-fakes Analysis Unit</b> ) by the <b>Misinformation Combat Alliance</b> for <b>2024 Indian Lok Sabha elections</b> . This helpline is funded by Meta. [Media Coverage: Nieman Lab   The Hindu   World Economic Forum   Meta] → I build data pipelines to cluster large amount's of Audio and Video data into social media thematic labels like politics, humor, memes, devotional content etc. This work helps fact-checker's better understand, vizualize and analyze the content received on their helpline. → Worked with <b>MLCommons</b> to develop an <b>AI safety benchmark dataset</b> for hate and sex related crimes. As a part of this project, I also conducted a <b>survey</b> on evaluating Indic LLM's for natural language tasks and online harms. [📄 Dataset] [📄 Report] [📄]	
May 2022	<b>Ahmedabad University</b> [🌐]	Ahmedabad, India
May 2021	Undergraduate Research Fellow   Advisors: Prof. Shilpa Pandit, Prof. Keyur Joshi Projects: Analysing people's psychological and physiological effects and experiences in the second wave of COVID-19, Perceivable Image Blur Quality Index	

## Publications

S=In Submission, C=Conference, W=Workshop, P=Preprints, J=Journal

- [S.2] **Improving Content Moderation for Indian Languages using Related Tasks**  
Aatman Vaidya, Dipankar Srirag, Aditya Joshi  
[In preparation]
- [S.1] **Strategies to Mitigate Spread of Toxicity on a Social Network**  
Aatman Vaidya, Harsh Bhagat, Seema Nagar, Amit Nanavati  
[In preparation]
- [P.1] **Analysis of Indic Language Capabilities in LLMs** [📄]  
Aatman Vaidya, Tarunima Prabhakar, Denny George, Swair Shah  
Report Submitted to **MLCommons** [📄]
- [W.1] **The Uli Dataset: An Exercise in Experience Led Annotation of oGBV** [📄]  
Arnav Arora, Maha Jinadoss, Cheshta Arora, Denny George, Brindaalakshmi...Aatman Vaidya, Tarunima Prabhakar  
Workshop on Online Abuse and Harms at North American Chapter of the Association for Computational Linguistics  
★ **Outstanding Paper Award** [NAACL'24]
- [C.3] **Analysing the Spread of Toxicity on Twitter** [📄]  
Aatman Vaidya, Seema Nagar, Amit Nanavati  
ACM International Conference on Data Science and Management of Data [CODS-COMAD'24]
- [C.2] **Forecasting the Spread of Toxicity on Twitter** [📄]  
Aatman Vaidya, Seema Nagar, Amit Nanavati  
IEEE International Conference on Cognitive Machine Intelligence [IEEE CogMI'23]
- [C.1] **Overview of the 2023 ICON Shared Task on Gendered Abuse Detection in Indic Languages** [📄]  
Aatman Vaidya, Arnav Arora, Aditya Joshi, Tarunima Prabhakar  
International Conference on Natural Language Processing [ICON'23]

## Talks

### “Evaluating Indic Language Performance in LLMs”

› AI for Global Development by [Agency Fund](#) *Invited Speaker* [[📄 Slides](#)]

Mar’25 (Bengaluru, India)

### “Modelling Hate Speech on a Social Network”

› Social Network Analysis Class (CSE533) *Invited Speaker* [[📄 Slides](#)]

Apr’24 (Ahmedabad University, India)

### “Building NLP classifiers to detect Hate Speech”

› ACM Winter School on Network Science *Invited Speaker*

Dec’23 (Ahmedabad University, India)

### “Tools to respond to Online Gender Based Violence”

› [Digital Citizen Summit](#) *Invited Speaker*

Nov’23 (Remote)

## Achievements

**Bosch Future Mobility Challenge | Semi-Finalist** [[📄](#)] Programmed and Engineered an **autonomous driving car** to navigate a miniature city. One of the 24 teams out of 118, and **only team** to represent **India** for the finals in Cluj, Romania. [[📄](#)] [[📺](#)]

**Outstanding Paper Award** Workshop on Online Abuse and Harms at NAACL’24

**Ingenious Hackathon 3.0 | Winner** Open Source Track

## Academic Service

**Shared Task Organiser** ICON’23 (Gendered Abuse Detection in Indic Languages) [[📄](#)]

**Reviewer** ACL Rolling Review’24

## Teaching and Leadership Roles

**Discrete Mathematics (MAT 101), Ahmedabad University** *Teaching Assistant*

Nov’22 - Mar’23

› Responsibilities included evaluating assignments, and helping students with the coursework.

**Center for Learning and Empowerment (NGO)** *Volunteer Teacher* [[📄](#)]

Jul’23 - Apr’23

› Taught 20+ secondary school kids mathematics for an entire academic year, at a tribal village in Jharkhand, India.

**Reading Group, Ahmedabad University** *Organiser* [[📄](#)]

Dec’21 - May’22

› Organized a weekly Reading Group focused on reading recent and classical papers in Computer Science.

**Programming Club, Ahmedabad University** *Content Head and Event Organiser* [[📄](#)]

Jul’20 - May’22

› Co-organized and curated events around Web Development, Machine Learning and Competitive Programming.

**Google Android Study Jams Facilitator** *Instructor* [[📄](#)] [[📄](#)]

Dec’21 - Jan’22

› Conducted 3 sessions teaching Android Development in Kotlin to 40+ students.

**Usable Pasts, Sustainable Futures** *Student Volunteer* [[📄](#)] [[📄](#)] [[📄](#)]

Jan’23

› Worked under Dr. Tejaswini Niranjana to understand how historically effective practices around Food, Clothing and Housing (*roti-kapda-makan*) can help mitigate climate change and public health issues.

## Skills

<b>Research Areas</b>	Online Safety, Evaluation, Language Models & their Societal Impact, Content Moderation
<b>Languages</b>	Python, JavaScript, Elixir C/C++, Kotlin, GraphQL
<b>Frameworks/ Libraries</b>	PyTorch, Tensorflow, Nodejs, OpenCV, networkx, CameraX, Elasticsearch, Sequelize
<b>Databases and Developer Tools</b>	MySQL, PostgreSQL, Docker, AWS (S3, EKS, EC2), aws-cli, Kubernetes (kubectl), CI/CD, Android Studio, Firebase