



---

# BHASHA 2025

1st Workshop on  
Benchmarks, Harmonization, Annotation, and  
Standardization for Human-Centric AI in Indian Languages

---

23rd December, 2025 (IJCNLP-AACL)

# Motivation

- India has a population of 140+ crores
- 22 official languages
- Digital resources range from fairly high to very low
- Dedicated workshop for development
  - Datasets
  - Benchmarks
  - Evaluation metrics
  - Tools
- Exclusively for Indian languages

# Paper Statistics

- Through OpenReview
- 26 submissions
  - 1 desk rejected
- All papers received at least 3 reviews
- 19 reviewers
- 11 accepted (42%)
  - 8 from India
  - 1 from Japan
  - 1 from Canada
  - 1 from USA

[illegible]

# Best Paper

- Multi-Feature Graph Convolution Network for Hindi OCR Verification
  - Shikhar Dubey, Krish Mittal, Sourava Kumar Behera, Manikandan Ravikiran, Nitin Kumar, Saurabh Shigwan and Rohit Saluja
  - IIT Mandi

# Program

- Website: <https://bhasha-workshop.github.io/>
- Two Keynotes
  - Monojit Choudhury: 09:15-10:15
  - Kalika Bali: 16:00-17:00
- Shared Task Overview: 10:15-10:30
- Research Paper Sessions
  - 11:00-11:50
  - 14:00-14:30
- Posters and Demos: 14:30-15:30
- BharatGen Sponsor Session: 11:50-12:30

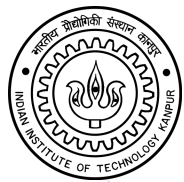
# Shared Tasks

- 2 shared tasks
- Grammar Error Correction
  - 5 Indian languages – Hindi, Bangla, Telugu, Tamil, Malayalam
- Word Grouping
  - 1 Indian language – Hindi
- Participating teams
  - 14 in Task 1
  - 2 in Task 2
- 6 shared task papers accepted

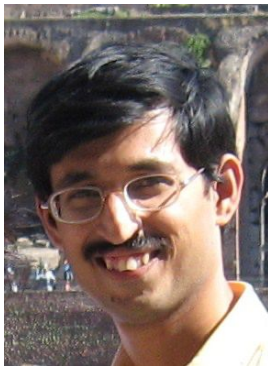
# Acknowledgments

- IJCNLP-AAACL Workshop Chairs
  - Sowmya Vajjala
  - Lizhen Qu
- IJCNLP-AAACL policy for supporting free registration of workshop organizers
- BharatGen for support
  - Certificates
  - Mementos





# BHASHA Workshop Organizers



**Arnab  
Bhattacharya**



**Kripabandhu  
Ghosh**



**Pawan  
Goyal**



**Saptarshi  
Ghosh**



**Ganesh  
Ramakrishnan**

**Hrishikesh  
Terdalkar**



**Shubham  
Nigam**



**Pramit  
Bhattacharyya**



**Karthika  
N J**



**Manoj  
Balaji**



# Thank you!

