

# PERAMPALLI SHRAVAN NAYAK

## PERSONAL DATA

EMAIL: [shravannayak.p@gmail.com](mailto:shravannayak.p@gmail.com)  
WEBSITE: [bajuka.github.io](http://bajuka.github.io)  
LINKEDIN: [www.linkedin.com/in/shravannayak99](https://www.linkedin.com/in/shravannayak99)  
GOOGLE SCHOLAR: <https://scholar.google.com/citations?user=vT4f678AAAAJ>

## EDUCATION

SEP 2023 | Master's in COMPUTER SCIENCE (THESIS)  
- CURRENT | University of Montreal, Montreal, GPA: 4.3/4.3  
Advisor: Prof. Aishwarya Agrawal

JULY 2017 | Bachelor's in ELECTRONICS ENGINEERING  
- MAY 2021 | IIT (BHU) Varanasi, Varanasi, GPA: 9.41/10  
Head of Programmers Club, Department Football team

## WORK EXPERIENCE

MAR 2024 | Diffusion Model Alignment  
- CURRENT | Applied Science Intern: AMLRT Mila, Montreal  
- Creating a preference dataset to align diffusion models for generating images of public spaces that are inclusive, diverse and accessible  
- Investigating fine-tuning techniques (DPO, D3PO, DRaFT) on diffusion models to enhance diversity and inclusivity based on multi-criteria human preferences

SEP 2023 | Benchmarking Cultural Understanding in VLMs  
- CURRENT | Research Assistant: Mila - Quebec AI Institute, Montreal  
ADVISOR: [Prof. Aishwarya Agrawal](#)  
- Developed CULTURALVQA, a visual question-answering benchmark to assess VLM's geo-diverse cultural understanding across 11 countries  
- Benchmarked VLMs like GPT-4V and Gemini to show disparities in cultural understanding across regions and highlighted areas of improvement

SEP 2021 | Machine Translation for Low Resource Languages [\[Github\]](#)  
- MAR 2023 | Research Assistant: University of Toronto  
ADVISORS: [Prof. En Shiun Annie Lee](#), [Prof. Surangika Ranathunga](#)  
- Conducted an extensive analysis of mBART's and mT5's translation performance on eight diverse low-resource languages  
- Executed large scale empirical tests using LLMs to identify the best fine-tuning and pre-training methods for low-resource language translation

JUL 2021 | Software Engineer II at MICROSOFT, Hyderabad  
- AUG 2023 | Part of the PowerPoint Mobile team  
- Played a key role in crafting user experiences, developing features like background customization, text highlighting, and online video insertion  
- Reduced overall memory consumption of the app, thereby reducing ungraceful app exits due to hangs by 7.1% and improved automation reliability by 16%

MAY 2020 | Machine Dubbing [\[Github\]](#)  
- DEC 2022 | Research Intern: University of Hamburg, Germany  
ADVISORS: [Prof. Timo Baumann](#), [Prof. Chris Beermann](#),  
- Curated single-speaker German audio-video-text corpus from Angela Merkel's podcasts  
- Developed video aware text-to-speech systems for lip-synchronous speech generation  
- Investigated machine translation systems in the context of Dubbing

## SELECTED PUBLICATIONS

---

Refereed	<b>BENCHMARKING VISION LANGUAGE MODELS FOR CULTURAL UNDERSTANDING</b>
Papers	<b>Shravan Nayak</b> , Kanishk Jain, Rabiul Awal, Siva Reddy, Sjoerd van Steenkiste, Lisa Anne Hendricks, Karolina Stanczak, Aishwarya Agrawal <i>Under Review</i>
	<b>IMPROVING ADVERSARIAL ROBUSTNESS IN VISION-LANGUAGE MODELS WITH ARCHITECTURE AND PROMPT DESIGN</b> Rishika Bhagwatkar, <b>Shravan Nayak</b> , Pouya Bashivan, Irina Rish <i>Under Review</i>
	<b>LEVERAGING AUXILIARY DOMAIN PARALLEL DATA IN INTERMEDIATE TASK FINE-TUNING FOR LOW-RESOURCE TRANSLATION</b> <b>Shravan Nayak</b> , Surangika Ranathunga, Sarubi Thillainathan, Rikki Hung, Anthony Rinaldi, Yining Wang, Jonah Mackey, Andrew Ho, En-Shiun Annie Lee. <i>Practical Machine Learning for Developing Countries (PML4DC) workshop, ICLR 2023.</i>
	<b>A DEEP DIVE INTO NEURAL SYNCHRONY EVALUATION FOR AUDIO-VISUAL TRANSLATION</b> <b>Shravan Nayak</b> , Christian Schuler, Debjoy Saha, Timo Baumann. <i>Proceedings of the 2022 International Conference on Multimodal Interaction, ICMI 2022.</i>
	<b>PRE-TRAINED MULTILINGUAL SEQUENCE-TO-SEQUENCE MODELS: A HOPE FOR LOW-RESOURCE LANGUAGE TRANSLATION?</b> En-Shiun Annie Lee, Sarubi Thillainathan, <b>Shravan Nayak</b> , Surangika Ranathunga, David Ifeoluwa Adelani, Ruisi Su, Arya D. McCarthy. <i>Findings of the Association for Computational Linguistics, ACL 2022.</i>
	<b>MERKEL PODCAST CORPUS: A MULTIMODAL DATASET COMPILED FROM 16 YEARS OF ANGELA MERKEL'S WEEKLY VIDEO PODCASTS</b> Debjoy Saha, <b>Shravan Nayak</b> , Timo Baumann. <i>Language Resources and Evaluation Conference, LREC 2022.</i>
	<b>THE TWO SHADES OF DUBBING IN NEURAL MACHINE TRANSLATION</b> Alina Karakanta, Supratik Bhattacharya, <b>Shravan Nayak</b> , Timo Baumann, Matteo Negri, Marco Turchi. <i>Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020.</i>
	<b>SEE ME SPEAKING? DIFFERENTIATING ON WHETHER WORDS ARE SPOKEN ON SCREEN OR OFF TO OPTIMIZE MACHINE DUBBING</b> <b>Shravan Nayak</b> , Timo Baumann, Supratik Bhattacharya, Alina Karakanta, Matteo Negri, Marco Turchi. <i>Deep Video Understanding Workshop, ICMI 2020.</i>

## AWARDS

---

2024	UdeM Excellence Scholarship
2023	Mila Scholarship
2020	DAAD WISE Scholarship
2019	Indian Academy of Sciences, Summer Research Fellowship

## INVITED TALKS

---

MAY 2023	Demystifying LLMs, IIT (BHU) VARANASI
NOV 2022	<b>Pre-Trained Multilingual Sequence-to-Sequence Models for NMT: Tips, Tricks and Challenges</b> , TORONTO MACHINE LEARNING SUMMIT
AUG 2021	<b>Research Scholarships for Undergrads</b> , COMPUTER VISION TALKS