

Raja Kumar

✉ raja.kumar@cnrs.fr | 🏠 | 🌐

Education

University of Toulouse

Oct 2025 - Sep 2028

DOCTORAL SCHOOL: MATHEMATICS, COMPUTER SCIENCE AND TELECOMMUNICATIONS

- NeuroAI PhD researcher at CNRS-ANITI (CerCo, Toulouse), exploring brain-inspired multimodal intelligence

Indian Institute of Technology Bombay

Aug 2019 - Jun 2024

Bachelor's: Materials Science Master's: Healthcare Informatics

CPI: 8.65/10

- Secured **AA (highest)** grade in both stages of the **Master's Thesis** on Natural Language Processing for Mental Health

Publications

Raja Kumar, Julie Hunter, Rufin VanRullen; **GloW-ify: Multimodal Instruction Tuning with Text and Vision Recoverable from a Shared Space**, *NAACL'27* [[Under Review](#)]

Raja Kumar, Julie Hunter, Rufin VanRullen; **How to GloW-ify your MLLMs?**, Unifying Representations in Neural Models (UniReps) Workshop, *NeurIPS'26* [[Under Review](#)]

Raja Kumar, Julie Hunter, Rufin VanRullen; **ReCap: Strengthening Grounding Through Re-Captioning in Instruction-Tuned Vision-Language Models**, *TMLR'26* [[Under Review](#)]

Raja Kumar, Raghav Singhal, Pranamy Kulkarni, Deval Mehta, Kshitij Jadhav; **M3CoL: Harnessing Shared Relations via Multimodal Mixup Contrastive Learning for Multimodal Classification**, *TMLR'25*; version accepted at UniReps Workshop, *NeurIPS'24* [[Paper](#)] [[Code](#)] [[Project Page](#)]

Raja Kumar, Kishan Maharaj, Ashita Saxena, Pushpak Bhattacharyya; **Mental Disorder Classification via Temporal Representation of Text**; *In Findings of Association for Computational Linguistics EMNLP'24* [[Paper](#)] [[Code](#)]

Kishan Maharaj, Ashita Saxena, **Raja Kumar**, Abhijit Mishra, Pushpak Bhattacharyya; **Eyes Show the Way: Modelling Gaze Behaviour for Hallucination Detection**; *In Findings of Association for Computational Linguistics EMNLP'23* [[Paper](#)] [[Code](#)]

Research Experience

Brain-Inspired Multimodal Representation Learning | CNRS-ANITI

Toulouse, France

Guide: Prof. Rufin VanRullen & Dr. Julie Hunter

2025 - Present

- Top-down: adapting existing multimodal models (MLLMs, VLMs, VLAs) with Global Workspace Theory principles for more efficient cross-modal representation learning.
- Bottom-up: building architectures around a shared Global Latent Workspace, where vision, language, and other specialized modules communicate through a common representational space.

Multimodal Mixup Contrastive Learning | Monash University & IIT Bombay

Mumbai, India

Guide: Prof. Kshitij Jadhav & Dr. Deval Mehta

2024

- Developed a novel multimodal contrastive loss incorporating mixup training to improve representation learning for complex real-world multimodal data relations, beating SOTA on four diverse public multimodal classification benchmarks
- Designed and implemented a multimodal learning framework incorporating unimodal prediction modules, a fusion module, and a new Mixup-based contrastive loss for continuous representation updating

Mental Disorder Identification through Linguistic Markers | IIT Bombay

Mumbai, India

Guide: Prof. Pushpak Bhattacharyya

2023 - 2024

- Proposed an idea of including temporal information for detecting mental disorders using time-series-based approaches
- Developed a framework for mental disorder identification via foundational deep learning models which surpasses the performance of BERT-based approaches by 5% in the F1 score on mental conditions: Depression, Self-harm, and Anorexia
- Conducted in-depth analysis on the errors and investigated the potential for cross-domain mental health data usage

Cognitive-Driven Hallucination Detection in LLMs | UT Austin & IIT Bombay

Mumbai, India

Guide: Prof. Abhijit Mishra & Prof. Pushpak Bhattacharyya

2023 - 2024

- Identified global and local attention biases in hallucination detection, inspiring a gaze-based detection framework
- Curated eye-tracking data with 500 instances for hallucination detection task and developed a BERT-based framework
- Proposed a novel attention bias framework inspired by human behavior and obtained a balanced accuracy of 87.1%

Professional Experience

CerCo-CNRS | Research Engineer

2024-2025

- Conducted a literature survey of recent and existing MLLMs to characterize their learning paradigms and identify which Global Workspace Theory principles they most critically lack.
- In parallel, developed and studied the Shimmer project, focusing on understanding the framework and building its transformer-based version as a bottom-up Global Latent Workspace.

Assert AI | Computer Vision & AI Intern

Summer 2022

- Developed an approach for food grain quality assessment: contour detection for separation, SVM for 89% accuracy
- Deployed customized YOLOv4 models for object detection tasks in surveillance leveraging the Nvidia Jetson series GPUs
- Generated tailored datasets and trained YOLOv4 models for diverse object detection and classification scenarios

Academic Services

Introduction to Artificial Intelligence | Lecturer, M2 (Neuro, Bio, Medical)

2026

- Taught M2 Medical students (Spring 2026) and M2 Neuroscience students (Fall 2026)
- Covered NLP fundamentals, recurrent neural networks, Transformers, and large language models (LLMs)

DH 302: Introduction to Public Health Informatics | Head Teaching Assistant

Autumn 2023

- Responsible for managing logistics and assisting the professor in ensuring the smooth functioning of the course
- Assisting in evaluation of answer scripts, designing projects, and conducting tutorials for a batch of 210+ students

Key Courses Undertaken

Computer Science	Deep Learning for NLP, Speech & NLP, Machine Learning, Image Processing, Speech Processing
Maths	Probability and Statistics, Optimization in ML, Applied Linear Algebra, Data Structures
MOOCs	Quantum Computing: Qubit, Specialization in DL: deeplearning.ai, Python Specialization

Research Interests

AI/ML NeuroAI, Multimodal Representation Learning, Language Processing, Deep Learning

Technical Skills

Softwares	Python, MATLAB, C++, Spice, LaTeX
Libraries	Pytorch, TensorFlow, Scikit-Learn, NumPy, Pandas, Matplotlib