

Gaurav Kumar

+91 8877386366 gauravkumarsony08@gmail.com

Google Scholar linkedin.com/in/gaurav-kumar17 github.com/Coding-with-Gaurav

Professional Summary

Early-career AI researcher and Junior Research Fellow at Swami Rama Himalayan University, with peer-reviewed publications in natural language processing (NLP), large language models (LLMs), and AI for healthcare, and an international competition placement (3rd/47 teams, SemEval-2026, ACL). Committed to undergraduate and postgraduate teaching in AI, Machine Learning, and Computer Science, with research interests spanning NLP, knowledge graphs, and AI applications in Indian Knowledge Systems.

Education

Master of Computer Applications (Data Science) <i>Dev Sanskriti Vishwavidyalaya, Uttarakhand</i>	2023 – 2025 CGPA: 8.30
Bachelor of Science in Information Technology <i>Patliputra University, Patna</i>	2019 – 2022 Percentage: 70.0%

Research Publications

- **The Spiritual Well-being Index (SWBI): A Multimodal AI Framework for Assessing and Supporting Spiritual Growth** (2026).
Presented at the 3rd MBCC, IIT Mandi, 3–7 June 2026.
- **AI-Monitors at SemEval-2026 Task 4: A Hybrid Embedding and LLM Ensemble for Narrative Similarity** (2026).
Published in the 20th International Workshop on Semantic Evaluation (SemEval-2026), ACL Anthology, San Diego, USA.
- **AI Tongue Diagnosis: Expert System with LLMs for Abdominal Disease Detection Through Mobile App** (2026).
Published in Cognitive Neuroscience and Artificial Intelligence, MBCC 2025, Studies in Rhythm Engineering, Springer, Cham, pp. 14–25.

Work Experience

Junior Research Fellow April 2026 – Present
AI Centre for Research, Swami Rama Himalayan University, Dehradun

- **AI Research — Indian Knowledge Systems (IKS)**
 - Conducting applied research in **Machine Learning**, **NLP**, and interdisciplinary AI domains.
 - Authored and presented a research paper at *3rd MBCC 2026, IIT Mandi* on integrating AI with spirituality at the intersection of IKS and modern technology.
- **Astrowala.world — Learning Management System for Vedic Astrology Mobile Platform**
 - Developing a **React Native Android application** as a learning management system (LMS) for [Astrowala.world](#), a Vedic astrology and spiritual learning platform.
 - Implemented **live class discovery**, **course access management**, and expert-learner interaction features, while applying learning management system (LMS) design principles relevant to e-learning platforms.
- **Mentoring**
 - Mentoring research interns at the lab by supporting their onboarding into applied AI/ML workflows.
- Engaged in **requirement gathering and progress tracking** meetings across active research and development projects.

Junior Research Fellow Jun 2025 – March 2026
Center for Artificial Intelligence & Research (IAI), Dev Sanskriti Vishwavidyalaya, Haridwar

- **HistoAI — Historical Data Extraction and Analysis System**

- Contributed to the **web development, structured data extraction, and system integration** components of a full-stack research platform built with React, Flask, and MongoDB.
- **Ayur Scan — AI-Powered Ayurvedic Tongue Diagnosis**
 - Contributed to an AI-powered Ayurvedic tongue diagnosis data collection application.
- **Research & Conference Contributions**
 - Contributed to interdisciplinary research **presented at the 2nd MBCC 2025, IIT Mandi** on AI applications in healthcare and consciousness studies.
 - Participated in the *International Conference on Faith & Future: Integrating AI with Spirituality*, Dev Sanskriti Vishwavidyalaya, Haridwar — engaging with researchers on AI and Indian Knowledge Systems.
- **Mentoring & Supervision**
 - Mentored a final-year MCA student on an NLP-focused final year project, providing guidance on literature review, experimentation, and report writing.
 - Mentored interns from **Sanskriti Innovation Pvt. Ltd.**, a startup incubated at CAIR Lab, DSVV, supporting their onboarding into applied AI/ML and full-stack project workflows.

Research Intern

Jan 2025 – May 2025

Center for Artificial Intelligence & Research (IAI), [Dev Sanskriti Vishwavidyalaya, Haridwar](#)

- Researched mobile-based tongue image analysis models for Ayurvedic analysis.
- Developed an LLM-powered system for structured data extraction from historical books with intelligent redirection.
- Built full-stack tools using React, Flask, MongoDB, and WebSocket for real-time data processing.
- Collaborated on OCR projects for the digitization of Sanskrit and Vedic texts.

Project Intern

May 2024 – Jul 2024

Center for Artificial Intelligence & Research (IAI), [Dev Sanskriti Vishwavidyalaya, Haridwar](#)

- Explored NLP-based projects, gaining practical exposure to model training, evaluation, and deployment.
- Researched Hindi transcription and language modeling for indigenous language datasets.

Volunteer, IT Cell (Software Development)

Dec 2023 – Feb 2024

All World Gayatri Pariwar, Shantikunj, Haridwar

- Supported internal software development and testing activities to enhance digital infrastructure reliability.

Teaching Experience

• Guest Lecturer & Team Lead, Social Internship (MCA Curriculum)

One Month

Surat, Gujarat — [Dev Sanskriti Vishwavidyalaya](#)

- Delivered 50+ interactive lecture sessions across schools, colleges, and universities in Surat, including SVNIT, Veer Narmad South Gujarat University (VNSGU), the Centre for Hindu Studies, and an IAS/IPS training centre.
- Taught mixed-discipline student groups (BCA, MCA, B.Com, M.Com, BBA, MBA, Biotechnology, B.Sc/M.Sc Physics & Chemistry, Master of Social Work), adapting delivery to each audience.
- Designed and delivered sessions on time management, personality development, de-addiction awareness, environmental responsibility, scientific spirituality, and Indian culture.
- Led a 3-member team to plan session flow, allocate topics, and coordinate scheduling across institutions.

• Computer Science Teacher, St. Agnesh English School

2021 – 2023

- Taught computer science to students in grades 6–10, covering programming fundamentals and problem-solving.
- Designed lesson plans and classroom activities to build foundational programming and digital literacy skills.

- Instructed students in computer fundamentals, MS Office applications, and introductory programming.
- Conducted interactive sessions to strengthen student understanding, participation, and communication skills.

Projects

AI-Monitors: Hybrid Embedding & LLM Ensemble for Narrative Similarity

SemEval-2026 Task 4, ACL

Prompt Engineering · LLM Evaluation · Ensemble Methods · NLP

- Designed and evaluated structured prompting strategies for **Large Language Models** to solve narrative similarity reasoning tasks.
- Built a hybrid **ensemble of sentence embedding and LLM-based models**, selecting components by error diversity rather than accuracy alone, achieving **75% test accuracy** near the human ceiling of 78%.
- Contributed to a team system ranked **3rd out of 47 international teams** at SemEval-2026 Task 4, ACL 2026, San Diego.

Technologies: LLMs, Prompt Engineering, Sentence Transformers, Ensemble Methods, Python.

HistoAI: Historical Data Extraction and Analysis System

Team Project

Full-Stack Development · LLM · RAG · Knowledge Graphs · OCR

- Built a **role-based access control (RBAC)** web platform unifying four independently developed AI modules (OCR, RAG Chatbot, Knowledge Graph, Data Extraction) under a single authentication system with **JWT-secured routes** and **React conditional rendering** for role-specific UI.
- Developed a **metadata-driven duplicate prevention system** using structured upload forms (title, author) for content-level deduplication; managed a **centralised book repository** with **Celery + Redis** async task queue, **WebSocket** real-time progress, and automated email notifications.
- Built a **structured data extraction pipeline** using OCR text chunking and a **locally hosted LLM**; iteratively benchmarked multiple open-source LLMs to eliminate hallucination.
- Implemented **source-linked output verification** across all AI tools where every result (extracted data, knowledge graph node, chatbot response) is traceable back to the **originating PDF page**, enabling users to validate outputs against the source.
- Led codebase as **developer lead** - managed **Git branch strategy** (feature → dev → main), reviewed pull requests, and coordinated **agile sprint-wise delivery** across the cross-functional team.

Technologies: React, Flask, MongoDB, OCR, RAG, Redis, Celery, JWT, WebSocket, JSON, GitHub.

Contactless Fingerprint Matching & Liveness Detection - UIDAI Open Proposal

React Native · Computer Vision · Biometric Authentication · Flask

- Built a **contactless fingerprint capture** app that matches finger images against a database and returns **matched identity with confidence score** for the **UIDAI Open Innovation Proposal**.
- Developed a **real-time liveness detection** app that classifies face input as **real or spoofed** using an on-screen face-centering guide - preventing fraudulent biometric access.
- Owned complete **mobile-side development** for both apps, integrating **Flask API endpoints** for end-to-end biometric operations.

Technologies: React Native, Face Recognition, Flask, Python, REST API, Computer Vision.

Ayur Scan — AI-Powered Tongue Data Collection for Ayurvedic Diagnosis

Mobile App Development · Computer Vision · TFLite · Ayurvedic AI

- Built a **Flutter-based Android application** enabling Ayurvedic therapists to collect structured patient records (name, age, gender, symptoms, clinical observations) alongside tongue imagery to build a dataset for an **AI-based abdominal disease diagnosis expert system**.

- Integrated an on-device **TFLite binary classification model** to validate tongue images at capture time — ensuring submissions are clear, well-framed, and tongue-only — preventing low-quality data from entering the dataset.
- Designed the app in direct support of the **published** research paper *AI Tongue Diagnosis: Expert System with LLMs for Abdominal Disease Detection through Mobile App*, serving as the **primary data collection tool** for the study.

Technologies: Flutter, Firebase, Cloudinary, MongoDB, TFLite, Python.

Knowledge Graph Generation from Text Documents using LLMs

LLMs · Knowledge Graphs · NLP · Information Retrieval

- Built a **knowledge graph pipeline** from unstructured text documents using **locally hosted LLMs** to perform **entity recognition and relationship extraction**, enabling faster information retrieval and structured exploration of complex document collections.
- Developed an interactive **graph visualisation interface** with intuitive node-edge exploration; linked every graph node and edge directly to the **originating PDF page or section**, allowing users to navigate from graph to source evidence in single click.

Technologies: Ollama, LangChain, Pyvis, NetworkX, Streamlit, Python.

Skills

Programming & Core CS: Python, JavaScript, Dart, Data Structures & Algorithms

AI, ML & NLP: LLMs, Prompt Engineering, RAG, Sentence Transformers, Ensemble Methods, Knowledge Graphs, Entity Recognition, Ollama, LangChain, TFLite

Computer Vision: OpenCV, Tesseract OCR, Image Processing

Web & Mobile Development: React.js, Node.js, Flask, Flutter, React Native, REST API, WebSocket, Tailwind CSS

Databases & Backend Infrastructure: MongoDB, Redis, MySQL, Celery, JWT, Firebase

Deployment & DevOps: GitHub Pages, Vercel, Render, Docker, GitHub Actions, Cloudinary, MongoDB Atlas

Research, Data & Visualization: NetworkX, Pyvis, Streamlit, Jupyter Notebook, Latex, JSON, CSV

Version Control & Collaboration: Git, GitHub (branch strategy, pull requests, code review), Postman

Certifications

- [Introduction to Data Science](#) – Coursera
- [Crash Course on Git, GitHub, and Markdown](#) – Udemy
- [Introduction to Generative AI](#) – Google Cloud Skills Boost
- [Big Data Certification](#) – Udemy
- [Data Analytics with Tableau](#) – Jobaaj Learnings

Academic Activities & Professional Development

- **3rd International Mind, Brain & Consciousness Conference (MBCC 2026)** [MBCC 2026 [🔗](#)]
IIT Mandi — 3–7 June 2026
 - **Presented original research** on the Spiritual Well-being Index (SWBI), a multimodal AI framework integrating NLP, affective computing, and Indian Knowledge Systems, representing AI Centre for Research, Swami Rama Himalayan University.
 - Engaged in interdisciplinary sessions on **Neuroscience, Consciousness Studies, and AI**, networking with researchers from IITs and other national and international institutions.
 - Proceedings forthcoming — see Research Publications.

- **Workshop on Art of Effective Manuscript Writing** [\[Conducted by Dr. Khanna\]](#)
Swami Rama Himalayan University, Dehradun — May 2026
 Conducted by **Dr. Vinay Kumar Khanna** (FNASc, FIANS), President, Indian Academy of Neurosciences.
 - Gained practical training in **scientific manuscript writing, critical appraisal of literature**, and **journal selection** strategies for publication.
 - Developed understanding of the **editorial and peer-review process** from an expert with 120+ research publications.
- **SemEval-2026 Task 4 — Narrative Story Similarity** [\[SemEval 2026\]](#)
ACL 2026, San Diego, USA
 - **Ranked 3rd out of 47 international teams** in a competitive NLP shared task on narrative reasoning, approaching the human annotator ceiling of 78%.
 - Gained hands-on experience in **prompt engineering, LLM evaluation, and ensemble design** through systematic hypothesis-driven experimentation.
 - Served as a **peer reviewer for 2 workshop papers**, contributing to the academic review process in narrative understanding and NLP research.
- **2nd International Mind, Brain & Consciousness Conference (MBCC 2025)** [\[MBCC 2025\]](#)
IIT Mandi — 2025
 - **Presented peer-reviewed research** on AI-based tongue diagnosis for abdominal disease detection, representing CAIR Lab, Dev Sanskriti Vishwavidyalaya.
 - Engaged in interdisciplinary sessions on **Neuroscience, Ayurveda, and AI for healthcare**, networking with researchers from IITs, national and International institutions.
 - Research subsequently published as a book chapter in *Cognitive Neuroscience and Artificial Intelligence* (Springer, 2026) – see Research Publications.
- **Winner — 24-Hour Coding Hackathon** *CAIR-LAB, Dev Sanskriti Vishwavidyalaya*
 - Developed a **web-based learning platform** under time constraints, demonstrating rapid prototyping and full-stack development skills.
 - Collaborated in a team to ideate, build, and present a working solution within 24 hours.
- **Greenovation Ideathon** [\[Greenovation Hackathon\]](#)
ICFAI University, Dehradun
 - Proposed an **AI-driven solution** to support the circular economy and sustainable resource management initiatives.
- **International Conference on Scientific Spirituality – 2024** [\[ICSS\]](#)
Dev Sanskriti Vishwavidyalaya, Haridwar
 - Volunteered in **conference logistics, speaker coordination, and registration management** alongside master's studies, supporting smooth execution of an international interdisciplinary event.
 - Facilitated **session handling** and assisted in moderating interdisciplinary discussions spanning **AI, consciousness studies, Indian Knowledge Systems (IKS), and scientific spirituality**.
- **6th International Conference on Yoga, Yagya & Ayurveda – 2023** [\[ICYA 2023\]](#)
Shriram Yoga Society, AIIMS Rishikesh
 - Provided **technical support and session coordination** for an international academic conference at a premier medical institution.