

Hrishikesh Rajesh Terdalkar

🏠 hrishikeshrt.github.io | [in](#) [hrishikeshrt](#) | [🔗](#) [hrishikeshrt](#)

CURRENT AFFILIATION

BharatGen, Indian Institute of Technology Kanpur, India

SUMMARY

My research lies at the intersection of *Computational Linguistics*, *Natural Language Processing*, and *Graph Databases*, with a focus on low-resource languages like Sanskrit and other Indian languages. I aim to tackle the unique linguistic and computational challenges they present, advancing language understanding and knowledge extraction from unstructured data. Drawing on my interdisciplinary background, I'm committed to build impactful, real-world solutions. My broader interests include *Databases*, *Software*, *Information Retrieval*, *Artificial Intelligence*, *Human-Computer Interaction*, and *Data Mining*.

EDUCATION

Ph.D. in Computer Science and Engineering

Indian Institute of Technology Kanpur
CGPA: 8.67

2014 – 2023
Kanpur, Uttar Pradesh, India

M.Sc. in Computer Science

Chennai Mathematical Institute
CGPA: 8.88

2011 – 2013
Chennai, Tamil Nadu, India

B.Sc. (Hons.) in Mathematics and Computer Science

Chennai Mathematical Institute
CGPA: 7.69

2007 – 2010
Chennai, Tamil Nadu, India

Higher Secondary School Certificate (Standard XII)

Maharashtra State Board of Secondary and Higher Secondary Education
Percentage: 86.50

2007
Maharashtra, India

Secondary School Certificate (Standard X)

Maharashtra State Board of Secondary and Higher Secondary Education
Percentage: 95.33

2005
Maharashtra, India

THESES

Ph.D. Thesis

Sanskrit Knowledge-based Systems: Annotation and Computational Tools

Supervisor: Arnab Bhattacharya

- Development of *knowledge-based systems for Sanskrit* with the ultimate goal of question-answering.
 - * Framework for Sanskrit QA through the automated construction of Knowledge Graphs
 - * Two web-based annotation tools
 - (1) *Sangrahaka*: for knowledge-graph focused annotation and querying
 - (2) *Antarlekhaka*: for general purpose annotation of multiple NLP tasks in a sequential manner
 - * System to identify Sanskrit meters and utilize Sanskrit prosody for post-OCR correction
 - * Several web-interfaces including three ancient Indian alpha-numeric systems
 - * Tools including a Telegram bot for Sanskrit grammar
 - * Software libraries related to computational aspects of Sanskrit
- All contributions available at <https://sanskrit.iitk.ac.in>

M.Sc. Thesis

On Symmetric Key Cryptography

Supervisor: Rajeeva Laxman Karandikar

- Study and implementation of various Stream Ciphers (RC4, HC128) and Block Ciphers (RC5, IDEA).

BOOKS

- “*Samanvaya: An Interlingua for Unity of Indian Languages*” . Chaitali Dangarikar, Arnab Bhattacharya, Karthika N J, **Hrishikesh Terdalkar**, Pramit Bhattacharyya, Annarao Kulkarni, Chaitanya S Lakkundi, Ganesh Ramakrishnan, Shivani V. ISBN: 978-93-48435-21-7. October 2024. Central Sanskrit University.

CONFERENCE PUBLICATIONS

- “*Saṅkhyāpaddhatiḥ: An Interface for Ancient Indian Numeral Systems*” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. **TO APPEAR** In Proceedings of the 1st Annual Academic Conference on Indian Knowledge Systems (IKS), New Delhi, July 2025.
- “*A Case Study of Cross-Lingual Zero-Shot Generalization for Classical Languages in LLMs*” . V S D S Mahesh Akavarapu, **Hrishikesh Terdalkar**, Primit Bhattacharyya, Shubhangi Agarwal, Vishakha Deulgaonkar, Chaitali Dangarikar, Pralay Manna, Arnab Bhattacharya. **TO APPEAR** In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL), Vienna, Austria, July 2025. Association for Computational Linguistics (ACL).
- “*Graph Repairs with Large Language Models: An Empirical Study*” . **Hrishikesh Terdalkar**, Angela Bonifati, Andrea Mauri. **TO APPEAR** In Proceedings of the 8th Joint Workshop on Graph Data Management Experiences & Systems (GRADES) and Network Data Analytics (NDA) @ SIGMOD/PODS 2025, Berlin, Germany, June 2025. Association for Computing Machinery (ACM).
- “*Aganittiyam: Learning Tamil Grammar through Knowledge Graph based Templated Question Answering*” . Mithilesh K, Amarjit Madhumalararungeethayan, Dharanish Rahul S, Abhijith Balan, C Oswald, **Hrishikesh Terdalkar**. In Proceedings of the 38th Pacific Asia Conference on Language, Information and Computation (PACLIC), Tokyo, Japan, December 2024.
- “*Antarlekha: A Comprehensive Tool for Multi-task Natural Language Annotation*” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. In Proceedings of the 3rd Workshop for NLP Open Source Software (NLP-OSS) @ EMNLP, Singapore, December 2023. Association for Computational Linguistics (ACL).
- “*Chandojnanam: A Sanskrit Meter Identification and Utilization System*” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. In Proceedings of the Computational Sanskrit & Digital Humanities: Selected papers presented at the 18th World Sanskrit Conference (WSC), pages 113–127, Canberra, Australia, January 2023. Association for Computational Linguistics (ACL).
- “*Semantic Annotation and Querying Framework based on Semi-structured Ayurvedic Text*” . **Hrishikesh Terdalkar**, Arnab Bhattacharya, Madhulika Dubey, Ramamurthy S, Bhavna Naneria Singh. In Proceedings of the Computational Sanskrit & Digital Humanities: Selected papers presented at the 18th World Sanskrit Conference (WSC), pages 155–173, Canberra, Australia, January 2023. Association for Computational Linguistics (ACL).
- “*A Novel Multi-Task Learning Approach for Context-Sensitive Compound Type Identification in Sanskrit*” . Jivnesh Sandhan, Ashish Gupta, **Hrishikesh Terdalkar**, Tushar Sandhan, Suvendu Samanta, Laxmidhar Behera, Pawan Goyal. In Proceedings of the 29th International Conference on Computational Linguistics (COLING), pages 4071–4083, Gyeongju, Republic of Korea, October 2022. International Committee on Computational Linguistics.
- “*Sangrahaka: A Tool for Annotating and Querying Knowledge Graphs*” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. In Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), pages 1520–1524, Athens, Greece, August 2021. Association for Computing Machinery (ACM).
- “*Framework for Question-Answering in Sanskrit through Automated Construction of Knowledge Graphs*” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. In Proceedings of the 6th International Sanskrit Computational Linguistics Symposium (ISCLS), pages 97–116, IIT Kharagpur, India, October 2019. Association for Computational Linguistics (ACL).

INVITED TALKS

- “*Are LLMs Ready for Sanskrit and Knowledge Graphs?*” . 5 June 2025. Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh, India.
- “*Knowledge Graphs to Large Language Models*” . 28 May 2025. Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan, India.
- “*Hands-on with Language Models*” . Summer School on Generative AI for Image Processing and Healthcare (GAIH), 8 May 2025. Mahindra University, Hyderabad, Telangana, India.
- “*Sanskrit Computational Linguistics: Tools, Annotation and Knowledge Graphs*” . 19 February 2024. Chennai Mathematical Institute, Chennai, Tamil Nadu, India.

POSTERS AND DEMONSTRATIONS

- “Vedapāṭhapaddhati: A Tool to Generate Prakṛtipāṭha and Vikṛtipāṭha forms of Sanskrit Text” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. The 7th International Sanskrit Computational Linguistics Symposium (ISCLS), 2024.
- “Āyurjñānam: Exploring Āyurveda with Knowledge Graphs” . **Hrishikesh Terdalkar**, Vishakha Deulgaonkar, Arnab Bhattacharya. National Youth Conference on Indian Knowledge Systems (NYCIKS), 2023. **Best Poster Award.**
- “Vaiyyākaraṇaḥ: A Sanskrit Grammar Bot for Telegram” . **Hrishikesh Terdalkar**, Mahesh A V S D S, Shubhangi Agarwal, Arnab Bhattacharya. The 18th World Sanskrit Conference (WSC), 2023.
- “PyCDSL: A Programmatic Interface to Cologne Digital Sanskrit Dictionaries” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. The 18th World Sanskrit Conference (WSC), 2023.
- “Jñānasaṅgrahaḥ: A Collection of Computational Applications related to Sanskrit” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. The 18th World Sanskrit Conference (WSC), 2023.
- “KaTaPaYadi System” . **Hrishikesh Terdalkar**, Arnab Bhattacharya. The 6th International Sanskrit Computational Linguistics Symposium (ISCLS), 2019.

EXPERIENCE

Data Engineer

BharatGen, IIT Kanpur

May 2025 – present

- Project: ‘BharatGPT: A Suite of Generative AI Tech for India’.
- Funding Agency: Department of Science and Technology (DST), Government of India.

Postdoctoral Researcher

LIRIS Research Lab, CNRS, Université Claude Bernard Lyon 1

May 2024 – Apr 2025

- Research: Large Language Models for Knowledge Graph Repair
- Project: ‘Human Factors in Data-Intensive Applications for Health’ (CPJ Data4Health).

Project Scientist

CSE, IIT Kanpur

Jan 2024 – Apr 2024

- Project: ‘An Interpretable Unified Framework Text-to-text Translation among Indian Languages using Sanskrit-based Interlingua Representation’.
- Funding Agency: Ministry of Electronics and Information Technology (MeitY), Government of India.

Senior Student Research Associate

CSE, IIT Kanpur

Aug 2022 – Dec 2023

- Project: ‘An Interpretable Unified Framework Text-to-text Translation among Indian Languages using Sanskrit-based Interlingua Representation’.
- Funding Agency: Ministry of Electronics and Information Technology (MeitY), Government of India.

Project Associate

CSE, IIT Kanpur

Jul 2013 – Apr 2014

- Project: ‘Localization in Wireless Sensor Networks’.

TEACHING

Instructor

École Normale Supérieure de Lyon

Feb 2025 – Apr 2025

- Program: Master 1 Informatique (ICA)
- Course: Databases and Data Mining

Instructor

Université Claude Bernard Lyon 1

Jan 2025 – Mar 2025

- Program: M2: Data and Intelligence for Smart Systems (DISS)
- Course: Natural Language Processing

Senior Tutor CSE, IIT Kanpur <ul style="list-style-type: none"> Managed (Prutor) backend administration for 5 semesters at IIT Kanpur. Duties also involved heading the tutor team, setting question papers and grading. 	Jan 2017 – Nov 2020
Teaching Assistant (MOOC) NPTEL <ul style="list-style-type: none"> NOC17-CS33: Fundamentals of Database Systems 	May 2017 – Sep 2017
Teaching Assistant CSE, IIT Kanpur <ul style="list-style-type: none"> CS315: Principles of Database Systems (Jan-Apr 2019, Jan-Apr 2021) CS685: Data Mining (Aug-Nov 2018) CS774: Optimization Techniques (Aug-Nov 2016) ESC101: Introduction to Programming (Aug-Nov 2014) 	Aug 2014 – Nov 2020
Academic Mentor CSE, IIT Kanpur <ul style="list-style-type: none"> Provided support to students in need of academic assistance 	Jul 2018 – Apr 2020

ADMINISTRATIVE EXPERIENCE

Program Committee Member <ul style="list-style-type: none"> GRADES-NDA Workshop @ SIGMOD/PODS (2025), ICWE (2025), TGD Workshop @ EDBT/ICDT (2025) 	2025 – present
Reviewer <ul style="list-style-type: none"> ARR (2022, 2023, 2024, 2025) 	2022 – present
Web Chair <ul style="list-style-type: none"> Created and managed the website and the proceedings for the conference ISCLS 2024. 	Jan 2023 – Feb 2024
External Reviewer <ul style="list-style-type: none"> CoDS-COMAD (2020, 2021, 2023), CIKM (2021, 2022), AAAI (2022), DASFAA (2022) 	2020-2023
Webmaster <ul style="list-style-type: none"> Managed website of CSE Department, IIT Kanpur. 	Dec 2014 – Jan 2016

PROJECTS

I am associated with or have worked in the following projects:

BharatGPT <ul style="list-style-type: none"> A Suite of Generative AI Tech for India. Department of Science and Technology (DST). GoI. 	ongoing
NLTM Interlingua <ul style="list-style-type: none"> Interlingua-based Translation among Indian Languages. Bhasini Consortium Project. MeitY. GoI. 	ongoing
Ramayana QA <ul style="list-style-type: none"> Automated Question Answering System for Valmiki Ramayana IKS Division. All India Council for Technical Education (AICTE). GoI. 	
Ayurveda KG <ul style="list-style-type: none"> Ontology and Knowledge Graph for an Ayurveda Text. Ashtaadashi, Central Sanskrit University (CSU). 	ongoing

ACHIEVEMENTS

- **Best Software Award** at the 57th Convocation, Indian Institute of Technology Kanpur, 2024.
- **Best Poster Award** at National Youth Conference on Indian Knowledge Systems 2023 **NYCIKS 2023** (Roorkee).
- **Financial support** from Research-I Foundation to attend **ISCLS 2019** (Kharagpur) and **WSC 2023** (Online).
- **Financial support** under ACM SIGSOFT CAPS Virtual Program to attend **ESEC/FSE 2021**.
- **Travel grants** from ACM India to attend **CoDS-COMAD 2017** (Goa), **CoDS-COMAD 2018** (Kolkata) and **CoDS-COMAD 2019** (Hyderabad).
- **TCS Research Scholar Program** Fellowship (Cycle 11), 2015.

TECHNICAL SKILLS

Languages	: Python, C, C++, Shell Scripting, Haskell, R, PHP, JavaScript, HTML, CSS
Databases	: MySQL, Neo4j, MongoDB
Libraries	: Flask, SQLAlchemy, Pandas, Scikit-learn, Gensim, NLTK, spaCy, PyTorch, NumPy
Dev Tools	: git, Python packaging, Make, docker
Typesetting	: LaTeX, Markdown, reStructuredText

RESEARCH SOFTWARE AND TOOLS

GitHub Profile Statistics	★ 160+ 🍴 45+ 🔄 3000+	🔗 hrishikeshrt
PyPI Profile		🔗 hrishikeshrt
Antarlekhaka		🔗 Antarlekhaka/code
<ul style="list-style-type: none">• Annotation of a comprehensive set of computational linguistic tasks.• General purpose, multi-task, sequential annotation.		
Sangrahaka		🔗 hrishikeshrt/sangrahaka
<ul style="list-style-type: none">• Annotation towards creation of knowledge graphs.• Querying of knowledge graphs using natural language query templates.		
Jnanasangraha	🌐 https://sanskrit.iitk.ac.in/jnanasangraha	
<ul style="list-style-type: none">• Collection of several web-based Sanskrit computational tools.• Includes three of ancient Indian alpha-syllabic numeral systems, a Sankrit prosodical meter identification and utilization system and various utility functions related to Sanskrit alphabet.		
Vaiyyakarana		🔗 vyakarana_bot
<ul style="list-style-type: none">• Telegram bot for the learners of Sanskrit grammar.		
Sanskrit Audio Alignment	🌐 https://hrishikeshrt.github.io/audio_alignment	
<ul style="list-style-type: none">• Interface for Sanskrit corpus recitation using forced-alignment to align text and audio.		

EXTRACURRICULAR

- I am a frequent blood donor with 45 donations so far and actively promote blood donation. I have also served as the Overall Co-ordinator and PG Wing Head of *Raktarpan*, IIT Kanpur and helped organize more than 20 blood donation camps in and around campus.
- I am interested in Sanskrit language and grammar. I hold a Diploma in Sanskrit Language under the Non-Formal Sanskrit Education programme offered by Central Sanskrit University. I have also completed Kovida, an advanced-level correspondence course offered by Samskrita Bharati.

REFERENCES

Arnab Bhattacharya

Professor, Indian Institute of Technology Kanpur, India

🏠 <https://cse.iitk.ac.in/users/arnabb/>

📄 RM409, Dept. of Computer Science and Engineering,
Indian Institute of Technology Kanpur,
Kanpur, Uttar Pradesh, India – 208016.

✉ arnabb@iitk.ac.in, arnabb@cse.iitk.ac.in

☎ +91-512-259-7650

Angela Bonifati

Distinguished Professor, Claude Bernard University Lyon 1, France

🏠 <https://perso.liris.cnrs.fr/angela.bonifati/>

📄 Université Claude Bernard Lyon 1
LIRIS - Département Informatique (Campus La Doua, Bât. Nautibus)
23-25 avenue Pierre de Coubertin
69622 Villeurbanne Cedex (France).

✉ angela.bonifati@liris.cnrs.fr

☎ +33-472-44-5824

Andrea Mauri

Junior Professor, Claude Bernard University Lyon 1, France

🏠 <https://www.andreamauri.me/>

📄 Université Claude Bernard Lyon 1
LIRIS - Département Informatique (Campus La Doua, Bât. Nautibus)
23-25 avenue Pierre de Coubertin
69622 Villeurbanne Cedex (France).

✉ andrea.mauri@liris.cnrs.fr

☎ +33-472-43-1636

Pawan Goyal

Associate Professor, Indian Institute of Technology Kharagpur, India

🏠 <https://cse.iitkgp.ac.in/~pawang/>

📄 Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur,
Kharagpur, West Bengal, India – 721302.

✉ pawang@cse.iitkgp.ac.in

☎ +91-3222-282370

Amba Kulkarni

Professor, University of Hyderabad, India

🏠 <https://sanskrit.uohyd.ac.in/faculty/amba/>

📄 Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad, Telangana, India – 500032.

✉ apksh@uohyd.ernet.in, ambapradeep@gmail.com

☎ +91-40-23133800