

# Amrith Krishna

AI Researcher

## Address

Bodhi,  
Symphony Lane,  
Cochin

## Tel

+91 9400354315

## Mail

ak2329@  
cantab.ac.uk  
krishnamrith12@  
gmail.com

## Web

<https://krishnamrith12.github.io>

## Social Presence

LinkedIn [↗](#)  
Twitter [↗](#)  
Stack Overflow [↗](#)  
Github [↗](#)

## Education

**Ph.D**  
2015 - 2019  
IIT Kharagpur  
**M.Tech**  
2013 - 2015  
IIT Kharagpur  
CGPA 9.27  
**B.Tech**  
2008 - 2012  
FISAT  
Aggregate 76.10 %

I was a staff scientist at Uniphore, a conversation automation unicorn. My research focused on Large Language Models (LLMs), Spoken language understanding (SLU), Conversational systems, and broadly on natural language understanding (NLU), processing (NLP) and generation (NLG). During my tenure of 1.5 years, I successfully contributed four of my works to Uniphore's product offerings, filed a provisional patent, and published six research papers in top-tier CORE A/A\* venues.

Previously, I was a postdoctoral researcher at the Department of Computer Science and Technology, University of Cambridge. I was working with Dr. Andreas Vlachos, on a project funded by Facebook research (now meta), with a focus on research in claim verification using knowledge bases. Prior to this, I was a postdoc at ITU Copenhagen for a year. I completed my Ph.D., supervised by Dr. Pawan Goyal, at the Dept. of Computer Science and Engineering, IIT Kharagpur, India. The thesis was titled 'Addressing Language-Specific Characteristics for Data-Driven Modelling of Lexical, Syntactic and Prosodic Tasks in Sanskrit'. My research interests lie in anything that comes under computational linguistics and Natural Language Processing, understanding, and Generation. Specifically, I am interested in information extraction, program synthesis, structured prediction, morphology, and syntax.

I am currently serving as an action editor for the ACL rolling review (ARR). I am serving as a program committee member of several \*CL venues such as ACL, EMNLP, CoNLL, COLING, LREC for the past 5 years. I have also reviewed for Computational Linguistics and LRE journal. I have also served as the organizing committee member of the 2018 and 2019 editions of CODS-COMAD. I am an organizer for IndoML 2022 and 2023.

## Work Experience

### 2022 - 2023 **Staff Scientist, Uniphore**

I worked on problems in spoken language understanding (SLU), Large Language Models, Conversational systems, and Natural Language Understanding, Processing and Generation (NLUP&G). During my tenure of close to 1.5 years, from April 22 to October 23, I have successfully contributed four of my works to Uniphore's product offerings, filed a provisional patent, and published six research papers in top-tier CORE A/A\* venues.

### 2020 - 2022 **Postdoctoral Research Associate, Department of Computer Science and Technology, the University of Cambridge**

Worked on a project funded by Facebook Research on Fact-verification using Knowledge-bases. I was supervised by Dr. Sebastian Riedel (Facebook Research) and Dr. Andreas Vlachos (University of Cambridge). Our work was published in the prestigious transactions of ACL by MIT Press and was presented at EMNLP 2022.

### 2019 - 2020 **Postdoc Researcher, ML and NLP groups, IT University of Copenhagen**

Worked as an Institute postdoctoral researcher, in collaboration with Natalie Schluter (ITU Copenhagen) on data augmentation strategies for text classification and program synthesis for morphological inflection.

## Publications

Singhal, Bhavuk; Gupta, Ashim; Shivasankaran V P; **Krishna, Amrith**. IntenDD: A Unified Contrastive Learning Approach for Intent Detection and Discovery. In Findings of the Association for Computational Linguistics: EMNLP 2023 (Accepted), Singapore. Association for Computational Linguistics (ACL).

Singhal, Bhavuk; Gopalan, Sindhuja; **Krishna, Amrith**; Chetlur, Malolan. Scaling Neural ITN for Numbers and Temporal Expressions in Tamil: Findings for an Agglutinative Low-resource Language. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track: EMNLP 2023 (Accepted), Singapore. Association for Computational Linguistics (ACL).

Gupta, Ashim; **Krishna, Amrith**. Adversarial Clean Label Backdoor Attacks and Defenses on Text Classification Systems. Proceedings of the 8th Workshop on Representation Learning for NLP (RepL4NLP 2023). pages 1–12, Toronto, Canada. Association for Computational Linguistics (ACL).

**Krishna, Amrith**; Gupta, Ashim; Garasangi, Deepak; Sandhan, Jeevnesh; Satuluri, Pavankumar; Goyal, Pawan. Neural Approaches for Data Driven Dependency Parsing in Sanskrit. In Proceedings of the Computational Sanskrit Digital Humanities: Selected papers presented at the 18th World Sanskrit Conference 2023, pages 1–20, Canberra, Australia (Online mode). Association for Computational Linguistics (ACL).

**Krishna, Amrith** ; Riedel, Sebastain ; Vlachos, Andreas. ProofVer: Natural Logic Theorem Proving for Fact Verification. Transactions of the Association for Computational Linguistics 2022; 10 1013–1030. doi: [https://doi.org/10.1162/tacl\\_a\\_00503](https://doi.org/10.1162/tacl_a_00503) [Link](#) (Presented at EMNLP 2022)

Maheshwari, Ayush; Singh, Nikhil; **Krishna, Amrith**; Ramakrishnan, Ganesh. 2022. A Benchmark and Dataset for Post-OCR text correction in Sanskrit. In Findings of the Association for Computational Linguistics: EMNLP 2022, pages 6258–6265, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Roy, Aniruddha; Thakur, Rupak Kumar; Sharma, Isha; Gupta, Ashim; **Krishna, Amrith**; Sarkar, Sudeshna; and Goyal, Pawan. 2022. Does Meta-learning Help mBERT for Few-shot Question Generation in a Cross-lingual Transfer Setting for Indic Languages?. In Proceedings of the 29th International Conference on Computational Linguistics, pages 4251–4257, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Kumar, Rishabh ; Adiga, Devraj ; Ranjan Rishav; **Krishna, Amrith** ; Ramakrishnan, Ganesh ; Goyal, Pawan; Jyothi, Preethi. Linguistically Informed Post-processing for ASR Error correction in Sanskrit. Proceedings of the Interspeech 2022, Incheon, South Korea, 2022 (accepted) [Link](#).

Adiga, Devraj ; Kumar, Rishabh ; **Krishna, Amrith** ; Jyothi, Preethi ; Ramakrishnan, Ganesh ; Goyal, Pawan. Automatic Speech Recognition in Sanskrit: A New Speech Corpus and Modelling Insights. Proceedings of The Findings of ACL, 2021.

Sandhan, Jivnesh ; **Krishna, Amrith** ; Gupta, Ashim ; Behera, Laxmidhar ; Goyal, Pawan. A Little Pretraining Goes a Long Way: A Case Study on Dependency Parsing Task for Low-resource Morphologically Rich Languages. Proceedings of the The 2021 EACL Student Research Workshop.

**Krishna, Amrith** ; Gupta, Ashim ; Garasangi, Deepak ; Satuluri, Pavankumar ; Goyal, Pawan. Keep it Surprisingly Simple: A Simple First Order Graph Based Parsing Model for Joint Morphosyntactic Parsing in Sanskrit. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)

**Krishna, Amrith** ; Santra, Bishal; Gupta, Ashim; Satuluri, Pavankumar; Goyal, Pawan. A Structured Prediction Framework Using Energy Based Models for Sanskrit. Computational Linguistics, 2021; 46 (4): 785–845. MIT Press.

Gupta, Ashim; **Krishna, Amrith** ; Goyal, Pawan; Hellwig, Oliver. Evaluating Neural Morphological Taggers for Sanskrit. Proceedings of the 17th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology (SIGMORPHON 2020). 58th Annual Meeting of the Association for Computational Linguistics (ACL).

**Krishna, Amrith**; Vidhyut, Shiv; Chawla, Dilpreet; Sambhavi, Sruti; Goyal Pawan. SHR++:

## Current Research Areas

- Program Synthesis
  - Fact verification
- Imitation Learning
- Natural Language Understanding
- Textual Entailment
  - Sanskrit Computational Linguistics

## Relevant Courses

- Natural Language Processing
- Language Processing in E-Learning
- Information Retrieval
  - Machine Learning
  - Complex Networks

## Teaching Assistantship

- Deep Learning
- Natural Language Processing ([Jupyter Notebooks](#))
- Information Retrieval ([Jupyter Notebooks](#))
- Computing Lab- I and II

An Interface for Morpho-syntactic annotation of Sanskrit Corpora. In International Conference on Language Resources and Evaluation 2020, Marseille (accepted). [Online tool](#)

Sandhan, Jivnesh; **Krishna, Amrith**; Goyal, Pawan; Behera, Laxmidhar. Revisiting the Role of Feature Engineering for Compound Type Identification in Sanskrit. October 2019. Proceedings of the Sixth International Sanskrit Computational Linguistics Symposium. IIT Kharapur, India. [paper](#)

**Krishna, Amrith** ; Sharma, Vishnu Dutt ; Santra, Bishal ; Chakraborty, Aishik; Satuluri, Pavankumar ; Goyal, Pawan. Poetry to Prose Conversion in Sanskrit as a Linearisation Task: A case for Low- Resource Languages. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers, ACL 2019). Florence, Italy. [paper](#)

**Krishna, Amrith** ; Santra, Bishal ; Bandaru, Sasi Prasanth ; Sahu, Gaurav ; Sharma, Vishnu Dutt ; Satuluri, Pavankumar ; Goyal, Pawan. Free as in Free Word Order: An Energy Based Model for Word Segmentation and Morphological Tagging in Sanskrit. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018. Brussels, Belgium. [Code](#) | [paper](#)

**Krishna, Amrith** ; Majumder, Bodhisattwa P. ; Bhat, Rajesh ; Goyal, Pawan. Upcycle Your OCR: Reusing OCRs for Post-OCR Text Correction in Romanised Sanskrit. Proceedings of the 22nd Conference on Computational Natural Language Learning, CoNLL 2018.Brussels, Belgium. [Code](#) | [paper](#)

Reddy, Vikas; **Krishna, Amrith**; Sharma, Vishnu; Gupta, Prateek; M R Vineeth; Goyal Pawan. Building a Word Segmenter for Sanskrit Overnight. In International Conference on Language Resources and Evaluation 2018, Miyazaki, Japan. [Code](#) | [paper](#)

**Krishna, Amrith**; Majumdar, Bodhisattwa; Goyal Pawan. An 'Ekalavya' Approach to Learning Context Free Grammar Rules for Sanskrit Using Adaptor Grammar. In 17th World Sanskrit Conference – July 9-13, 2018 Vancouver, BC.

**Krishna, Amrith**; Satuluri Pavankumar; Goyal, Pawan (2017). "A Dataset for Sanskrit Word Segmentation" in Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, 55th Annual Meeting of the Association for Computational Linguistics (ACL), Vancouver,Canada 2017. [Dataset](#).

**Krishna, Amrith**; Satuluri, Pavankumar; Ponnada, Harshavardhan; Ahmed, Muneeb; Arora, Gulab; Hiware, Kaustubh; Goyal, Pawan (2017); A Graph Based Semi-Supervised Approach for Analysis of Derivational Nouns in Sanskrit. in Proceedings of TextGraphs 11 Workshop, ACL 2017 ,Canada 2017.[link](#).

**Krishna, Amrith**; Satuluri, Pavankumar; Sharma, Shubham; Kumar, Apurv and Goyal, Pawan (2016). Compound Type Identification in Sanskrit: What Roles do the Corpus and Grammar Play? International Conference on Computational Linguistics WSSANLP, Workshop at COLING 2016, Osaka, Japan, Dec. 11-16. [link](#).

**Krishna, Amrith**; Santra, Bishal; Satuluri Pavankumar; Bandaru, Sasi Prasanth; Faldu , Bhumi;Singh, Yajuvendra; Goyal, Pawan;"Word Segmentation in Sanskrit Using Path Constrained Random Walks" in COLING 2016. [link](#).

**Krishna, Amrith**; Mallick, Madhumita; Mitra, Bivas; "SleepSensei - An automated sleep quality monitor and sleep duration estimator." in In IoT of Health Workshop, Mobisys 2016 Chakraborty, Tanmoy; **Krishna, Amrith**; Singh, Mayank; Ganguly, Niloy; Goyal, Pawan and Mukherjee, Animesh. "FeRoSA: A Faceted Recommendation System for Scientific Articles." in PAKDD 2016. [link](#). | [www.ferosa.org](http://www.ferosa.org).

**Krishna, Amrith** and Goyal, Pawan. Towards automating the generation of derivative nouns in Sanskrit by simulating Panini. 16th World Sanskrit conference, 2015. [link](#).

**Krishna, Amrith**; Bhowmick, Plaban; Sahu, Archana; Ghosh, Krishnendu; Roy, Subhayan. "Automatic Generation and Insertion of Assessment Items in Online Video Courses." In

Proceedings of the 20th International Conference on Intelligent User Interfaces Companion, pp. 1-4. ACM, 2015. [link](#).

## Programming

Python ★★★★★

C ★★★★★

PHP/HTML5 ★★★★★

CSS/JS ★★★★★

## Libraries

PyTorch

Fairseq

HF transformers

Scipy/Numpy

Scikit-Learn

Pandas

## Awards and Recognition

- Our proposal, extending on my PhD thesis, 'An Interpretable Unified Framework for Translation among Indian Languages using Sanskrit-based Interlingua Representation', was awarded a grant worth Rs. 3.5 Crore by MeitY.
- Our work on Computational Sanskrit, published in the computational linguistics journal, received [media coverage](#) in prominent national media including India Today, Indian Express, Times Group, Outlook, etc.
- Recipient of numerous travel grants from various entities including Microsoft, EMNLP, ACM-IARCS, CNeRG, and the Institute Best Conference travel grant.
- Secured a grant of INR 5,00,000 for the project 'IndicView', from National Level Google IIT Pilot program. September 2014 - 2015.
- Stood among the top 0.080 % students in the Graduate Aptitude Test in Engineering, GATE 2013 CS with All India Rank 180 of 2,24,160 candidates
- Received best demo award in IBM Day, IIT Kharagpur on 29th August 2015
- Recipient of 'Special Recognition award' for excellence in B.Tech from FISAT in 2012.

## Academic Services and Volunteering

- **Consultant** - Actively Consulting for the project titled 'An Interpretable Unified Framework for Translation among Indian Languages using Sanskrit-based Interlingua Representation' with IIT Kharagpur as the lead institute.
- **Organizer, IndoML** - Served as Tutorial chair for IndoML 2022, IIT Gandhinagar Currently serving as a Datathon Chair for IndoML 2023, IIT Bombay.
- **PhD Supervisor, IIT Bombay** - I am officially a co-supervisor to a PhD student at IIT Bombay, in addition to mentoring numerous undergraduate and master students at various IITs and other institutes of national eminence.
- **Action Editor** - Currently serving as an action editor for ACL Rolling Review for an year.
- **Reviewer** - Computational Linguistics Journal, MIT Press and LRE Journal, Springer.
- **Program Committee Member** - Program committee member of AAI, ACL, EMNLP, CoNLL and several other \*CL venues.
- **Organising Committee Member (Web-Chair) - CODS-COMAD 2018, 2019: [Link](#).**
- **Web Chair - 6th ISCLS 2019 ([Link](#)), Data Science in India, KDD 2017:**

## Education

2015 - 2019 **PhD Student, Dept. of CSE, IIT Kharagpur**

Broad Area of Research: Sanskrit Computational Linguistics.

Thesis Advisor: Prof. Pawan Goyal. Department of CSE, IIT Kharagpur.

2013 - 2015 **M.Tech in CSE**

**IIT Kharagpur**

[CGPA 9.27](#)

## Talks

"An Introduction to NLP", Talk at ATMA, NIT Calicut (July 2023).

- "An Introduction to Data programming", Talk at ASAS, Cochin, India (Feb 2023).

- "A Graph-Based Framework for Structured Prediction Tasks in Sanskrit", NLIP Seminar Series, University of Cambridge.

- "Introduction to Machine Learning and Deep Learning". Talk at ASAS, Cochin, India.

- Lectures on Deep learning with Python at Machine Learning for Cyber Physical Systems, Executive course for TATA Steel, Centre for AI, IIT Kharagpur.
- Introduction to Jupyter, Foundations of Artificial Intelligence and Machine Learning, Centre for AI, IIT Kharagpur.
- Synthesising Grammars & Programs for Natural Language, Paradigm Shift in Indian Linguistics and its Implications for Applied Disciplines, IAS Shimla, 30 Oct - 1 Nov 2017.
- "Hangman with Language Models", Hands-on Lab session at the ACM Summer School on Natural Language Processing and Machine Learning, IIT Kharagpur, 9th July 2017
- "A Dataset for Word Segmentation in Sanskrit", Workshop for Bridging the gap between Sanskrit Computational Linguistics tools and management of Sanskrit Digital Libraries, ICON, Dec. 18th 2016, IIT-BHU.
- "Automated Sanskrit Text Segmentation Aided by Statistical Analysis", Talk at ASTRA International Conference 2016, Deccan College, Pune.10th January 2016.
- "Named Entity Recognition in Bhagavatham with Rich Linguistic Features.", Talk at ASTRA International Conference 2016, Deccan College, Pune.10th January 2016.

## Technical Skills

**Programming Languages** - Python, R, PHP, HTML, CSS/JS

**Operating System** - Linux ,Windows, MacOS

**Libraries** - Langchain, Haystack, PyTorch, Fairseq, Huggingface Transformers, Pandas, Scikit-Learn, Scipy, Numpy, DGL

**Natural Languages** - Malayalam (Native), English, Hindi, Sanskrit

## Teaching and Leadership Experience

Nov'21-Mar'22 **Supervisor, Undergraduate Part II Project at the University of Cambridge**  
Co-supervisor for one undergraduate part II project

Jan'20-Jun'20 **Supervisor, Research Project, Master thesis, and Bachelor thesis at ITU**  
Supervisor for a bachelor thesis project, a master thesis project, and a research project. The projects are "Contextual Bandits - Solving text-based tasks" and "Learning Rational Transductions" and "Program Synthesis as Learning Rational Transductions" respectively.

Aug'19-Nov'19 **Lecturer, Introduction to Data Science and Programming**  
Course taught along with Natalie Schluter and Michael Szell at ITU. Link for Live coding sessions and exercises available at - [Github Link](#).

Jul'15-May'18 **Webmaster - CNeRG**

Jan'17-May'17 **Teaching Assistant, NPTEL Online Course on Natural Language Processing**  
Teaching Assistant for the online course hosted at NPTEL. [Course Link](#).

Jul'15-Nov'18 **Mentor - B.Tech and M.Tech Projects**  
Mentored projects for 7 B.Tech Projects and 6 M.Tech Projects

July'14-Nov'18 **Teaching Assistant, During M.Tech and PhD at IIT Kharagpur**  
*NLP*: Lecture Session on Lexical Semantics, Information Extraction & Entity Linking. [Kaggle Link](#). | *Deep Learning*: Tutorial Session on Neural Sequence Generation with CNTK.

## References

Dr. Pawan Goyal  
Associate Professor  
Dept. of Computer Science & Engg.  
Indian Institute of Technology, Kharagpur.  
Email: pawang@cse.iitkgp.ac.in

Dr. Andreas Vlachos  
Associate Professor,  
Dept. of Computer Science & Technology  
University of Cambridge.  
Email: av308@cam.ac.uk