

Taraka Rama Kasicheyanula

CONTACT INFORMATION	Department of Linguistics University of North Texas Denton, TX, USA	<i>Tel:</i> 047-92564918 <i>E-mail:</i> taraka.kasi@gmail.com https://scholar.google.com/citations?user=jhJQuk4AAAAAJ&hl=en
RESEARCH INTERESTS	Computational historical linguistics, Health Informatics, Text Classification, Bayesian data analysis, Deep learning, Treebanking for less resourced languages.	
SERVICE	Reviewer at EACL, ACL, NAACL, COLING, EMNLP, Language Dynamics and Change, PeerJ, SemEval.	
REFERENCES	Gerhard Jäger, Lilja Øvrelid, Richard Johansson, Søren Wichmann, Johann-Mattis List	
RESEARCH EXPERIENCE	<ul style="list-style-type: none">• Post-Doctoral Fellow at BIGMED project¹ from September 2017 to July 2019.• Post-Doctoral Fellow at the <i>Language Evolution: The Empirical Turn (EVOLAEMP)</i>² project (ERC Advanced Grant) from November 2015 to August 2017• Research Assistant in Digital Areal Linguistics Project hosted at Univ. of Uppsala	
TEACHING EXPERIENCE	<ul style="list-style-type: none">• Co-Supervisor of two Masters students at Department of Informatics, Oslo• Co-supervisor of a Bachelors and two PhD students at University of Tübingen• Taught a semester length <i>Hauptseminar</i> on Statistical phylogenetics at University of Tübingen	
COMPUTER SKILLS	<ul style="list-style-type: none">• Languages: C, Perl, Python, Java, and Shell scripting• Tools: R, NLTK, PyTorch, TensorFlow, Keras, Scikit-learn• Operating Systems: Unix/Linux and Windows	
EDUCATION	<p>University of Gothenburg, Gothenburg, Sweden Swedish Graduate School in Language Technology, Sweden</p> <p>Licentiate thesis defended in January, 2014 Doctoral student: September, 2010 – November, 2015</p> <ul style="list-style-type: none">• PhD Topic: “Computational Historical Linguistics”• Supervisors: Lars Borin, Søren Wichmann• Opponent: Gerhard Jäger• Committee: Richard Johansson, Laura Downing, Beáta Megyesi• Pre-defense: Jörg Tiedemann <p>IIIT-Hyderabad, Hyderabad, India</p> <ul style="list-style-type: none">• Masters in Technology, Computational Linguistics, August 2009• Supervisor: B. Lakshmi Bai <p>DA-IICT, Gandhinagar, India</p> <ul style="list-style-type: none">• Bachelors in Technology, Information and Communication Technology, May 2007• Supervisor: Ganesh Devy	

¹<https://bigmed.no/>

²<http://www.evolaemp.uni-tuebingen.de/>

HONORS AND AWARDS

- Ranked **first** in:
 - VarDial 2016, 2017 shared task on dialect classification
 - SemEval-2018 shared task on multilingual emoji classification
 - CLPsych-2018 Shared Task on Predicting Current and Future Psychological Health from Childhood Essays
- Outstanding Reviewer for COLING 2018.
- NVIDIA Academic GPU grant in 2016.
- NSF Travel Grant to attend NAACL-HLT Student Research Workshop 2009 at Colorado, USA.
- Pratibha Scholarship (awarded by Andhra Pradesh State Government) for years 2001 and 2002.
- Gold Medal for standing first in ICSE examinations in the state of Andhra Pradesh for the year 2001.

PROFESSIONAL EXPERIENCE

Digital Areal Linguistics Project, Univ. Uppsala & Univ. of Gothenburg & Max Planck Inst. Leipzig

Research Assistant

October, 2010 - December, 2014

Worked as a research assistant in the *Digital Areal Linguistics*³ project aimed at collecting word lists of 1000 length for ten South Asian languages. Was involved in the digitization of the Linguistic Survey of India volumes for the purpose of investigating the spread of grammatical features in the Himalayas. Also investigated if the word lists provided in the Linguistic Survey of India volumes show traces of genetic relations or borrowing between four different language families in South Asia.

24/7 Innovation Labs,⁴ Bangalore, India

Analytics Consultant

October, 2009 - August, 2010

- Automatic classification of unstructured chat data using repeated-bisection algorithm from unsupervised class of methods. Designed and developed a scalable system for large datasets. The feature extraction and selection was based on NLP techniques such as stemming, lemmatization, POS tagging, dependency parsing, and information gain. The project delivered the possible labels of each category using n-gram models.
- Project for matching the query and responses at a contact center using *Topic Tracking* models. Designed and developed a solution for scoring the responses in a contact center based on customer queries using language models. Applied cross-entropy for estimating the correctness of the agent response to a customer query. This solution was implemented and delivered for an international airlines company and a large UK corporate.
- Was involved right from the beginning to the release of TweetView[®] a service for opinion mining from tweets for major corporations. The service uses a Named Entity recognizer for filtering the non-relevant tweets from relevant tweets.

Language Technologies Research Centre, IIIT-Hyderabad, Hyderabad, India

Research Assistant

July, 2007 - July, 2009

Worked in a wide range of topics in NLP: Cognate identification, phylogenetic inference using Bayesian inference, sentence alignment, tokenization, statistical machine translation, name entity transliteration, letter to phoneme conversion, font converters for Indian languages, and mining parallel corpora for Telugu–Hindi language pair.

SUMMER SCHOOLS, RESEARCH VISITS

- AnthroTree workshop, University of Massachusetts, Amherst, USA. June, 2011.
- Spent the spring of 2012 (February–June) at Department of Linguistics, Max-Planck-Institut für Evolutionäre Anthropologie, Leipzig, Germany.
- European Summer School in Logic, Language, and Information (ESSLLI), Opole, Poland. August 2012.
- Guest researcher at Alpha-Informatica, Groningen, Netherlands. March–April 2013.
- Linguistic Summer Institute, Chicago, USA. July 2015. Courses on Grammaticalization and contact-induced morphological change.

³<https://spraakbanken.gu.se/eng/research/digital-areal-linguistics>

⁴<https://www.247.ai/>

CO-ORGANIZER

- Workshop on using standardized word lists in linguistic data collection, Göteborg, Sweden. October 2010.
- Workshop on comparing approaches to measuring linguistic differences, Göteborg, Sweden. October, 2011.
- International workshop on linguistic microareas in South Asia. Uppsala, Sweden. May, 2014.

CONFERENCES

- TCS NLP Winter School 2008, collocated with IJCNLP 2008, Hyderabad, India.
- Student volunteer at IJCNLP 2008.
- Student Research Workshop at NAACL-HLT 2009, Boulder, Colorado, USA. June, 2009.
- Quantitative Investigations in Theoretical Linguistics (QITL) 4, Berlin, Germany. March, 2011.
- Nordic Conference of Computational Linguistics (NODALIDA), Riga, Latvia. May, 2011.
- International Conference in Historical Linguistics (ICHL-XX), Osaka, Japan. July, 2011.
- Swedish Language Technology Conference (SLTC), University of Lund, Sweden. October, 2012.
- International Conference in Computational Linguistics (COLING), IIT Bombay, Mumbai, India. December 2012.
- Language Diversity Congress, Groningen, Netherlands. July 2013.
- International workshop on linguistic microareas in South Asia, Uppsala, Sweden. May 2014.
- LREC Reykjavik, Iceland. June 2014.
- North American Association for Computational Linguistics, Colorado, USA. June 2015.
- Capturing Phylogenetic Algorithms for Linguistics, Leiden, Netherlands. October 2015.
- International Conference in Computational Linguistics, Osaka, Japan. December 2016.
- European Association for Computational Linguistics, Valencia, Spain. April 2017.
- Treebanks and Linguistic Theories, Prague, Czech Republic. January 2018.
- North American Association for Computational Linguistics (NAACL), New Orleans, USA. June 2018.
- COLING, Santa Fe, USA, August 2018.
- EMNLP, Copenhagen, Denmark, November 2018.
- ACL, Florence, Italy, August 2019.