

Loganathan Ramasamy | résumé

✉ loganathanspr@gmail.com • 🌐 loganathanspr.github.io • in loganathanspr

Education

Charles University in Prague

PhD, Mathematical Linguistics

Prague, Czech Republic

2010–2014

- Thesis topic: “Cross-lingual transfer strategies for parsing under-resourced languages”
- Advisor: *assoc. prof. (doc.) Ing. Zdeněk Žabokrtský*

Amrita Vishwa Vidyapeetham

MS by Research, Computational Informatics

Coimbatore, India

2006–2008

Industry experience

Merck Sharp & Dohme

Machine Learning Engineer

Prague, Czech Republic

Jan 2019–present

- Developer/Lead for a solution that assists system admins in identifying non-compliant (by various internal and external policy regulations) content in unstructured content platforms.
 - Developed an end-to-end machine learning (ML) solution using TensorFlow to recognize document labels from images within documents.
 - Organizer for stand-up meetings & I was an internal lead to productize the solution.
 - Major developer by code contributions.
- Developer for a specialized search solution that uses natural language processing (NLP) to improve the search results.
 - Responsible for adding new search features, data sources and managing data ingestion pipeline.
 - Responsible for managing Elasticsearch indices and creating Kibana dashboards.
- I was a representative/lead for various internal and vendor technology assessments for NLP/ML solutions.

Merck Sharp & Dohme

Scientific Software Engineer

Prague, Czech Republic

Jan 2016–Dec 2018

- Developer/Business Analyst for an internal search engine (approx. 20TB of data).
 - Responsible for creating NLP strategies for improving the search.
 - Developed named entity recognition (NER) proof of concepts for the search index data.
 - Gained experience using big data technologies such as *Hive*.
- Worked as a developer for various text mining related projects and gained experience using domain specific NLP/text mining platforms such as *I2E*.
- Experience working with small/large teams in a highly interdisciplinary environment.

Consulting/Freelancing

Amazon: Worked as a linguistic expert for a Tamil speech project.

Jan 2020 - Mar 2020

STTS, Sweden: Worked as a computational linguist for Tamil speech projects.

Mar 2019 - May 2019

Research experience

Charles University in Prague

Research Assistant

Prague, Czech Republic

Aug 2010–Dec 2015

- Extensive experience working with parallel corpus for machine translation and multilingual parsing.
- Created a parallel corpus (about 170K sentence pairs) involving an under-resourced language by crawling the web.
- Created Prague style dependency treebank for Tamil - license acquired by Google.
- Created a spelling error corpus by parsing Wikipedia data sets.

University of L'Aquila

Young Indian Researcher

L'Aquila, Italy

Oct 2009–Jun 2010

- Explored parallels between statistical machine translation and translation between architecture description languages.

Amrita Vishwa Vidyapeetham

Research Associate

Coimbatore, India

Jun 2004–Jul 2010

- Teaching assistant for the course on natural language processing and worked on the following funded projects:
 - “English to Indian Languages: Machine Translation System (EILMT)”
 - “Study for Development of Methodology Detection Tools for Digital Contents Plagiarism and Piracy Software”

Selected publications

L. Ramasamy, D. Mareček, and Z. Žabokrtský, “Multilingual dependency parsing: Using machine translated texts instead of parallel corpora,” *The Prague Bulletin of Mathematical Linguistics*, vol. 102, pp. 93–104, 2014.

D. Zeman, O. Dušek, D. Mareček, M. Popel, L. Ramasamy, J. Štěpánek, Z. Žabokrtský, and J. Hajič, “Hamledt: Harmonized multi-language dependency treebank,” *Language Resources and Evaluation*, vol. 48, no. 4, pp. 601–637, 2014.

L. Ramasamy, O. Bojar, and Z. Žabokrtský, “Morphological processing for english-tamil statistical machine translation,” in *Proceedings of the Workshop on Machine Translation and Parsing in Indian Languages (MTPIL-2012)*, pp. 113–122, 2012.

Computer skills

Programming: Python, Perl, Java

Web: HTML, JavaScript

NLP tools: NLTK/spaCy, Stanford tools, Moses

Machine learning: scikit-learn, TensorFlow

Others: Docker, Elasticsearch, Hive, Jenkins CI/CD

OS: Mac OSX, Linux, Windows

Languages

English: full professional proficiency

Tamil: native proficiency

Czech: A2

German: A1

Awards

European Commission

- Marie Curie Fellowship: Early Stage Researcher (ESR)

2010-2013

Ministry of Education, Universities and Research (MIUR), Italy

- Young Indian Researcher

2009-2010