

# SHRUTI RIJHWANI

srijhwan@cs.cmu.edu | [shrutirijhwani.com](http://shrutirijhwani.com)

## EDUCATION

---

- Aug '16–Aug '18 *Master of Science in Language Technologies*  
**Carnegie Mellon University**  
GPA: 4.12 (4.0+ grading system)
- Aug '11–Aug '15 *Bachelor of Engineering in Computer Science*  
**Birla Institute of Technology and Science, Pilani**

## RESEARCH EXPERIENCE

---

- Aug '16–Present **Graduate Research Assistant, Language Technologies Institute, Carnegie Mellon University**  
*Advisors: Jaime Carbonell and Graham Neubig*  
I work on unsupervised and language-independent cross-lingual Entity Detection and Linking. The effort is part of the DARPA Low Resource Languages for Emergent Incidents (LORELEI) project, which focuses on rapid development of NLP tools for languages with little or no training data.
- Jan '17–Present **Independent Study, Language Technologies Institute, Carnegie Mellon University**  
*Advisor: Taylor Berg-Kirkpatrick*  
**Historical Collation:** Unsupervised large-scale historical collation of *Don Quixote* editions printed between the 15th and 19th centuries using OCR and a beam-search approximation for the string search problem.  
**Printer Attribution:** OCR combined with statistical image analysis in order to cluster unique character typesets, to identify the printers of 130,000 covertly published documents from the early modern era.
- Jul '15–Jul '16 **Research Fellow, Microsoft Research, Bangalore, India**  
*Advisor: Monojit Choudhury*  
**Code-switching language identification:** Developed an unsupervised and language-independent technique for word-level language identification of text data with no prior linguistic information.  
**Machine translation** of code-switched data.  
**Language Preference** of Hindi-English bilinguals for sentiment and opinion expression.
- Jan '15–May '15 **Research Intern, Microsoft Research, Bangalore, India**  
*Team: Advanced Development Group*  
**People Search:** Developed retrieval and ranking techniques based on fuzzy matching of misspelled people names in e-mail search queries for Email Insights, a publicly released Microsoft Garage product.  
**Neighborhood Deduplication:** Worked on detecting duplicates among unstructured and transliterated neighborhood address strings using fuzzy search, synonym learning and subset grouping. Bing teams are using the technique on datasets containing millions of Indian, US and German addresses.
- May '14–Jul '14 **Bing Ads Intern, Microsoft India Development Center, Bangalore, India**  
*Advisor: Mehul Parsana, Principal Research Manager at Microsoft*  
**POS-tagging search queries** using weakly-labeled search logs and Conditional Random Fields.

## PUBLICATIONS

---

Patrick Littell, Tom McCoy, Na-Rae Han, Shruti Rijhwani, Zaid Sheikh, David R. Mortensen, Teruko Mitamura and Lori Levin. “Parser combinators for Tigrinya and Oromo morphology”. *Language Resources and Evaluation Conference (LREC)*. 2018. (Accepted)

Shruti Rijhwani, Royal Sequiera, Monojit Choudhury, Kalika Bali and Chandra Sekhar Maddila. “Estimating Code-Switching on Twitter with a Novel Generalized Word-Level Language Detection Technique”. *The 55th Annual Meeting of the Association for Computational Linguistics (ACL)*. 2017. [\[pdf\]](#)

Koustav Rudra, Shruti Rijhwani, Rafiya Begum, Kalika Bali, Monojit Choudhury and Niloy Ganguly. “Understanding Language Preference for Expression of Opinion and Sentiment: What do Hindi-English Speakers do on Twitter?”. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2016. [\[pdf\]](#)

Michael Yoder, Shruti Rijhwani, Carolyn Rosé and Lori Levin.

“Code-Switching as a Social Act: The Case of Arabic Wikipedia Talk Pages”.

*The Second Workshop on Natural Language Processing and Computational Social Science*. 2017. [\[pdf\]](#)

Shruti Rijhwani, Royal Sequiera, Monojit Choudhury and Kalika Bali.

“Generalized and Unsupervised Word-Level Language Detection”.

*The First Women and Underrepresented Minorities in NLP Workshop*. 2017. [\[pdf\]](#)

Paul Michel, Abhilasha Ravichander and Shruti Rijhwani.

“Does the Geometry of Word Embeddings Help Document Classification? A Case Study on Persistent Homology-Based Representations”.

*The Second Workshop on Representation Learning for NLP*. 2017. [\[pdf\]](#)

Sunayana Sitaram, Sai Krishna Rallabandi, Shruti Rijhwani and Alan W Black.

“Experiments with Cross-lingual Systems for Synthesis of Code-Mixed Text”.

*9th ISCA Speech Synthesis Workshop*. 2016. [\[pdf\]](#)

Shruti Rijhwani, Royal Sequiera, Monojit Choudhury and Kalika Bali.

“Translating Code-Mixed Tweets: A Language Detection Based System”. (Demonstration Track)

*Workshop on Indian Language Data: Resources and Evaluation at LREC ‘16*. 2016. [\[pdf\]](#)

Abhilasha Ravichander, Shruti Rijhwani, Rajat Kulshreshtha, Chirag Nagpal, Tadas Baltrusaitis, Louis-Philippe Morency.

“Preserving Intermediate Objectives: One Simple Trick to Improve Learning for Hierarchical Models”.

*arXiv preprint*. [\[pdf\]](#)

## ENGINEERING EXPERIENCE

---

May ‘15–Aug ‘15	<b>Google Summer of Code, MIT Media Lab</b> Built SVG-capture and Undo features for the Blocks Editor, a visual programming editor.
Nov ‘13–Feb ‘14	<b>Lead Developer, Lost Child Art Studios, Pilani, India</b> <i>Web design and development startup</i> . I built four webapps end-to-end, using various web technologies.
May ‘13–Jul ‘13	<b>Web Design Intern, Tune Patrol Media Labs Pvt. Ltd., Bangalore, India</b> <i>Tune Patrol is an online marketplace for independent Indian musicians to expand their audience</i> . Developed the user interface for music concerts and song playlist features for the product website.
May ‘12–Jul ‘12	<b>Software Engineering Intern, Tata Consultancy Services, Bangalore, India</b> <i>Multinational information technology consulting and solutions firm</i> . I built a database management system.
May ‘11–Jul ‘11	<b>Web Development Intern, InfoToros, Bangalore, India</b> <i>Startup that provides complete web solutions</i> . I designed and developed a client’s business website.

## TEACHING AND LEADERSHIP ROLES

---

Dec ‘14–May ‘15	<b>Co-founder, getPlaced()</b> getPlaced() was a group that conducted programming bootcamps for computer science careers, attended by over 1000 engineering undergraduates in colleges across India.
Aug ‘14–Dec ‘14	<b>Co-founder, Computer Programming Special Interest Group, BITS Pilani, India</b> Co-founded and taught a lecture series on algorithms with an audience of over 300 students.
Aug ‘13–May ‘14	<b>Captain, BITS Pilani Swimming Team</b> Captained the swim team to a championship win at BOSM 2013, an international collegiate sports meet.
Jan ‘13–May ‘13	<b>Editor, English Press, BITS Pilani, India</b> Coordinated the publication of the daily newspaper and numerous blog articles, and conducted four literary events during APOGEE 2013, an all India inter-collegiate technical symposium.

## OTHER ACTIVITIES

---

**Linguistic Olympiads:** *Volunteer*, International Linguistic Olympiad 2016, Panini Linguistic Olympiads 2016, 2015

**Extra-Curricular:** Swimming, SCUBA diving (PADI Rescue Diver), Piano Performance (Grade 4 Trinity)