

Laurel Hart

Computational Linguist

✉ laurel.elise.hart@gmail.com

🌐 konahart.com

in laurelehart

🔗 konahart

Education

- 2013—2014 **Professional Master of Science in Computational Linguistics**,
University of Washington, Seattle, WA.
- Completed intensive, graduate-level program in 1 year
 - Focused on Natural Language Processing (NLP), Machine Learning (ML), and language modeling
 - Specific topics included Machine Translation, Information Retrieval, and Question Answering
- 2009—2013 **Bachelor of Arts in Applied Linguistics**, *Portland State University*, Portland, OR.
- Minor in Computer Science
 - Honors College, undergraduate thesis: “The Linguistics of Sentiment Analysis”

Experience

- 2015 **Artificial Intelligence Engineer**, *ToyTalk, Inc.*, San Francisco, CA.
- Used NLP techniques to improve interaction with users
 - Added features and fixed bugs in in-house content creation tools
- 2014 **Research Assistant**, *University of Washington - CSE*, Seattle, WA.
- Designed and implemented Java training data pipeline for multi-linguistic relation extraction system, MultiR
 - Worked with large corpora:
 - Gigaword English, Spanish, and Chinese corpora: over 9 million documents
 - Freebase: currently over 2 billion facts
 - Utilized Stanford’s CoreNLP package for tokenization, Named Entity Recognition, parsing, etc.
- Winter and Summer **NLP Intern**, *BCL Technologies*, San Jose, CA.
- 2012
- Conducted research on Semantic Role Labeling for Modern Standard Arabic
 - Authored demonstration paper published in *Proceedings of COLING 2012* (see Publications)
 - Created annotated linguistic data
 - Wrote grant proposals, company awarded \$80,000 SBIR toward Multilingual Author Identification
 - Other research areas: text summarization, spoken language identification, automated financial data extraction

Technical skills & selected projects

- ML** kNN, Decision Trees, Naïve Bayes, MaxEnt, and SVM (Python)
- NLP** HMM-based Part-of-Speech tagger, Word Sense Disambiguation system, CKY parser, tokenization, Context Free Grammars incorporating features and semantic representations
- Current** Promptbot: an IRC bot for writers; used upwards of 40 users across 2 networks (Python, ongoing)
- Recent** Training pipeline for a semi-supervised multilingual relation extraction system (Java); product review sentiment analysis research (Python)
- Languages** Python, Java, C++
- Tools** NLTK, Mallet, Stanford CoreNLP, Maven, Git, Bash, L^AT_EX

Publications

- 2012 Hart, Laurel, Hassan Alam and Aman Kumar, “**Revisiting Arabic Semantic Role Labeling using SVM Kernel Methods**” *Proceedings of COLING 2012*. December, 2012: 207—14.
<http://aclweb.org/anthology//C/C12/C12-3026.pdf>