Dan Garrette

dhgarrette@gmail.com http://dhg.ai

EDUCATION

The University of Texas at Austin, Austin, TX

Ph.D., Computer Science Aug 2009 - Apr 2015 M.S., Computer Science Aug 2009 - Dec 2011

Advisors: Jason Baldridge and Raymond Mooney

Illinois Wesleyan University, Bloomington, IL

B.S., Computer Science, with Research Honors

Aug 2003 - Apr 2006

Minor: Cognitive Science (completed in 3 years)

SELECTED RESEARCH EXPERIENCE

Google Research, New York, NY

Research Scientist Oct 2016 - Present

· Machine learning and natural language processing.

University of Washington, Seattle, WA

Post-Doctoral Research Associate. Supervisor: Luke Zettlemoyer May 2015 - Oct 2016

· Semi-supervised/low-resource learning for NLP.

University of Texas at Austin, Austin, TX

Research Assistant. Supervisor: Jason Baldridge

Aug 2011 - May 2015

· Learning NLP models from varieties of weak supervision.

Research Assistant. Supervisors: Katrin Erk and Ray Mooney

Aug 2009 - Aug 2011

· Unifying logical and distributional semantics for natural language representation and inference.

Google, Mountain View, CA

Intern May 2013 - Aug 2013

· Machine learning and natural language processing research related to Google News.

University of Maryland Institute for Advanced Computer Studies, College Park, MD

Research Assistant. Supervisor: Philip Resnik

· Bayesian models of syntactic framing in political writing.

May 2012 - Aug 2012

Publications

[12] <u>Dan Garrette</u> and Hannah Alpert-Abrams. "An Unsupervised Model of Orthographic Variation for Historical Document Transcription". In *Proc. of NAACL*, 2016.

- [11] <u>Dan Garrette</u>, Chris Dyer, Jason Baldridge, and Noah A. Smith. "A Supertag-Context Model for Weakly-Supervised CCG Parser Learning". In *Proc. of CoNLL*, 2015.
- [10] <u>Dan Garrette</u>, Hannah Alpert-Abrams, Taylor Berg-Kirkpatrick, and Dan Klein. "Unsupervised Code-Switching for Multilingual Historical Document Transcription". In *Proc. of NAACL*, 2015.
- [9] <u>Dan Garrette</u>, Chris Dyer, Jason Baldridge, and Noah A. Smith. "Weakly-Supervised Grammar-Informed Bayesian CCG Parser Learning". In *Proc. of AAAI*, 2015.
- [8] <u>Dan Garrette</u>, Chris Dyer, Jason Baldridge, and Noah A. Smith. "Weakly-Supervised Bayesian Learning of a CCG Supertagger". In *Proc. of CoNLL*, 2014.
- [7] <u>Dan Garrette</u>, Jason Mielens, and Jason Baldridge. "Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages". In *Proc. of ACL*, 2013.
- [6] $\underline{\text{Dan Garrette}}$ and Jason Baldridge. "Learning a Part-of-Speech Tagger from Two Hours of Annotation". In Proc. of NAACL, 2013.
 - * Best Talk Award Finalist
- [5] <u>Dan Garrette</u>, Katrin Erk, and Raymond Mooney. "A Formal Approach to Linking Logical Form and Vector-Space Lexical Semantics". Harry Bunt, Johan Bos, and Stephen Pulman (eds) *Computing Meaning*, Vol. 4, 2013.
- [4] Islam Beltagy, Cuong Chau, Gemma Boleda, <u>Dan Garrette</u>, Katrin Erk, and Raymond Mooney. "Montague Meets Markov: Deep Semantics with Probabilistic Logical Form". In *Proc. of *SEM*, 2013.

- [3] <u>Dan Garrette</u> and Jason Baldridge. "Type-Supervised Hidden Markov Models for Part-of-Speech Tagging with Incomplete Tag Dictionaries". In *Proc. of EMNLP*, 2012.
- [2] <u>Dan Garrette</u>, Katrin Erk, and Raymond Mooney. "Integrating Logical Representations with Probabilistic Information using Markov Logic". In *Proc. of the Intl. Conference on Computational Semantics (IWCS)*, 2011.
- [1] <u>Dan Garrette</u> and Ewan Klein. "An Extensible Toolkit for Computational Semantics". In *Proc. of the International Conference on Computational Semantics (IWCS)*, 2009.

INVITED TALKS

- · Google Research. "Exploiting Universal Grammatical Properties to Induce CCGs". August 2016.
- · Amazon. "Exploiting Universal Grammatical Properties to Induce CCGs". August 2016.
- · Apple. "Exploiting Universal Grammatical Properties to Induce CCGs". August 2016.
- · Lawrence Livermore National Lab. "Exploiting Universal Grammatical Properties to Induce CCGs". August 2016.
- · Allen Institute for AI. "Exploiting Universal Grammatical Properties to Induce CCGs". August 2016.
- · University of Edinburgh. "Learning CCGs from Weak Supervision". June 2016.
- · Workshop on Multilingual and Crosslingual Methods in NLP (at NAACL-2016). "Unsupervised Modeling of Code-Switching and Orthographic Variation". June 2016.
- · Microsoft Research. "Unsupervised Modeling of Code-Switching and Orthographic Variation". May 2016.
- · University of Washington. "Learning CCGs from Weak Supervision". February 2015.
- · Carnegie Mellon University. "Learning CCGs from Weak Supervision". April 2014.

SELECTED OPEN SOURCE PROJECTS

Numerous projects at http://github.com/dhgarrette

Natural Language Toolkit (NLTK) - http://www.nltk.org/

2007 - Present

- · Toolkit for natural language processing in Python.
- · Authored most semantics code: first-order logic, λ -calculus, DRT, inference, etc.

Ocular - https://github.com/tberg12/ocular

2014 - Present

- · State-of-the-art OCR system for transcribing historical texts.
- · Authored extensions/features used in our publications.

Spring Batch – http://projects.spring.io/spring-batch/

2008 - 2009

· Application framework for batch processing in Java.

OTHER SELECTED PROFESSIONAL EXPERIENCE

Accenture, Chicago, IL

Consultant Sep 2008 - Aug 2009

- \cdot Served as one of four committers to the open-source Spring Batch framework.
- Designed and developed new functionality. Identified and fixed bugs.
- Wrote reference documentation and answered questions on the public forum.
- · Assisted Accenture projects as a Subject Matter Expert for Spring Batch.
- · Designed Spring Batch training curriculum and led on-site training in the US and India.
- · Wrote documentation on Accenture's approach to Java development for use company-wide.
- · Certified by Accenture as a Technology Architect.

Analyst May 2007 - Sep 2008

- · Developed large-scale applications in Java.
- · After just four months, was made a sub-team leader, formally leading a group of developers.
- · After 15 months, promoted to Consultant. (Normal requirement is 2 to 3 years).

TransUnion, Chicago, IL

Intern Summers 2001-2005

· Added features to the application that interacts with the credit database. Coded in C.

Programming Languages

Skilled: Scala, Python, Java (ordered by frequency of use)

Previous Experience: C, C++, Haskell, MATLAB, Prolog, R, SQL (ordered alphabetically)

ACADEMIC SERVICE

Conference reviewing: EMNLP (2013, 2014, 2015, 2016), ACL (2012, 2016), NAACL (2016), EACL (2014, 2016),

COLING (2014, 2016), WNUT (2016), IWCS (2013, 2015), Texas Linguistic Society (2012)

Journal reviewing: Computational Linguistics (4), Language Resources and Evaluation (1)

Conference website developer/maintainer: IWCS-2013

UTCS Admissions Committee – Master's Degree Program, Spring 2015

Honors and Awards

Post-Graduate

Best Reviewers List: NAACL-2016

Graduate

Best Talk Award Finalist, NAACL-2013

National Defense Science and Engineering Graduate Fellowship (NDSEG) - 2010-2013

Student Travel Grants: EMNLP-2012, CoNLL-2014

Undergraduate

Alumni Academic Scholarship

Jennings Music Scholarship

Upsilon Phi Epsilon - Computer Science Honor Society

Dean's List - 6 of 6 semesters

SELECTED TEACHING EXPERIENCE

The University of Texas at Austin

Instructor, Natural Language Processing (CS 378 / LIN 353N)

Fall 2013

· Upper-division undergraduate Computer Science and Linguistics course

Guest Lecturer, Graduate Computational Linguistics (LIN 386M)

Fall 2011

Undergraduate Research Advising

Kelsey Taylor Ball

Brianna Connelly

Matthew Ebeweber

2014 - 2015

2014

The United States Peace Corps

Volunteer, Ghana, West Africa

2006 - 2007

- · Taught math and English in an underperforming rural junior secondary school.
- · Planned and taught HIV/AIDS presentations in rural communities and schools.

Illinois Wesleyan University

Teaching Assistant

2004 - 2006

- · Courses on algorithms (in C and Haskell), data structures (in C), and computer architecture.
- · Ran study sessions for classes and tutored students.

ACTIVITIES

Post-Graduate

Seattle ScienceSlam - Presenter: Learning to read 16th century books (Voted Best Talk of the event) Jan 2016

Graduate

Organizer of the Natural Language Learning reading group

2011 - 2015

Undergraduate

John Wesley Powell Student Research Conference - Presenter

Symphonic Winds - Member, Percussion Section Leader for 5 semesters

Civic Orchestra, Opera Orchestra, Titan Band, Percussion Ensemble - Member