

Justin Betteridge

CONTACT INFORMATION

Language Technologies Institute
Carnegie Mellon University
6411 Gates-Hillman Center
5000 Forbes Avenue
Pittsburgh, PA 15213 USA

Voice: (412) 268-6591
Fax: (412) 268-6298
E-mail: jbetter@cs.cmu.edu
WWW: www.cs.cmu.edu/~jbetter

RESEARCH INTERESTS

Semi-supervised information extraction, computational semantics, co-reference resolution, ontology induction

EDUCATION

Carnegie Mellon University, Pittsburgh, Pennsylvania USA

Ph.D., Language Technologies, September 2007 - present

- Advisor: Tom Mitchell
- Thesis topic: Semi-supervised information extraction using the web: recognizing many different coupled entity types and relations in text with a knowledge base bootstrapped from hundreds of millions of web pages.
- GPA: 3.76

M.S., Language Technologies, May 2006

- Advisor: Teruko Mitamura
- Areas of Study: Language & Statistics, Machine Translation, Grammar Formalisms, Machine Learning
- GPA: 3.74

Brigham Young University, Provo, Utah USA

B.S., Computer Science, April 2004

- Minor in Linguistics
- GPA: 3.80

ACADEMIC EXPERIENCE

Carnegie Mellon University Pittsburgh, Pennsylvania USA

Graduate Research Assistant

June 2007 to present

- Helping to develop a Never Ending Language Learner (NELL) as part of Tom Mitchell's Read the Web project
- Refactored prototype system into a maintainable, extensible architecture
- Drastically improved run-time efficiency of various system components through distributed computation on Yahoo!'s Hadoop-based M45 Cluster

Teaching Assistant

September 2008 to December 2008

- Assisted Roni Rosenfeld in 10-601: Machine Learning
- Refined and graded assignments on basic ML theory and implementation.
- Graded midterm and final examinations

Research Associate

May 2006 to May 2007

- Reduced answer type classification error rate in the JAVELIN II question answering system by 16% using a combination of rule-based and machine learning approaches.

Graduate Research Assistant

September 2004 to May 2006

- Extended the KANTOO machine translation system to facilitate knowledge acquisition in the Halo II project and question analysis in the JAVELIN II question answering system.

Brigham Young University Provo, Utah USA

Undergraduate Research Assistant

February 2003 to April 2004

- Helped create a word-dependency transformation of the Penn Treebank.

CONFERENCE
PUBLICATIONS

Andrew Carlson, Justin Betteridge, Richard C. Wang, Estevam R. Hruschka Jr. and Tom M. Mitchell. [Coupled Semi-Supervised Learning for Information Extraction](#). To appear in the Proceedings of the Third ACM International Conference on Web Search and Data Mining (WSDM). 2010.

Tom M. Mitchell, Justin Betteridge, Andrew Carlson, Estevam Hruschka, and Richard Wang. [Populating the Semantic Web by Macro-Reading Internet Text](#). Invited paper, Proceedings of the 8th International Semantic Web Conference (ISWC 2009). 2009.

Andrew Carlson, Justin Betteridge, Estevam R. Hruschka Jr. and Tom M. Mitchell. [Coupling Semi-Supervised Learning of Categories and Relations](#). Proceedings of the NAACL HLT 2009 Workshop on Semi-supervised Learning for Natural Language Processing. 2009.

Justin Betteridge, Andrew Carlson, Sue Ann Hong, Estevam R. Hruschka Jr., Edith L. M. Law, Tom M. Mitchell, Sophie H. Wang. [Toward Never Ending Language Learning](#). AAAI Spring Symposium on Learning by Reading and Learning to Read. 2009.

Nico Schlaefer, Jeongwoo Ko, Justin Betteridge, Guido Sautter, Manas Pathak, Eric Nyberg. [Semantic Extensions of the Ephyra QA System for TREC 2007](#). In Proceedings of the Sixteenth Text REtrieval Conference (TREC). 2007.

Teruko Mitamura, Frank Lin, Hideki Shima, Mengqiu Wang, Jeongwoo Ko, Justin Betteridge, Matthew Bilotti, Andrew Schlaikjer and Eric Nyberg. [JAVELIN III: Cross-Lingual Question Answering from Japanese and Chinese Documents](#). In Proceedings of NTCIR-6 Workshop, Tokyo, Japan. 2007.

Irene Langkilde-Geary and Justin Betteridge. [A Factored Functional Dependency Transformation of the English Penn Treebank for Probabilistic Surface Generation](#). In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC). 2006.

Eric Nyberg, Teruko Mitamura, Justin Betteridge, Simon Fung, and David Svoboda. Poster abstract: [Capturing knowledge from domain text with controlled language](#). In Proceedings of the Third International Conference on Knowledge Capture (KCAP). 2005

AWARDS

Carnegie Mellon University

- Research Assistantship, 2004–2006, 2007–present

Brigham Young University

- University Scholarship (4 year, full–tuition), 1998–1999, 2001–2004

SERVICE Full-time Volunteer Missionary, Church of Jesus Christ of Latter-day Saints, 1999-2001

- Mongolia Ulaanbaatar Mission
- Taught conversational English in schools and universities, 16 hours/week

TECHNICAL SKILLS Programming Languages: Java, Ruby, Lisp, XML, XSLT, SQL, Perl, C/C++, L^AT_EX,
Frameworks: UIMA, Hadoop
Operating Systems: Linux, Apple OS X, Microsoft Windows XP/2000

LANGUAGE SKILLS Advanced competency in spoken/written Mongolian

CITIZENSHIP USA