

Vanessa Murdock

Yahoo! Research Barcelona
Diagonal 177, Planta 8
08018 Barcelona Spain

phone: (+34) 93 183 8824
E-mail: vanessa.murdock@yahoo.com
Website: <http://www.cs.umass.edu/~vanessa>

RESEARCH INTERESTS

Information retrieval, Sentence retrieval, Internet Advertising, Image and Video Retrieval

EDUCATION

PhD, Computer Science University of Massachusetts, Amherst, Massachusetts

Aspects of Sentence Retrieval September, 2006

Master of Science, Computer Science, University of Massachusetts, Amherst, MA, 2005

Bachelor of Science, Computer Science, Colorado State University, Ft. Collins, CO, 2001

Bachelor of Music, Piano Performance, Texas Christian University, Ft. Worth, TX, 1989

Early Experience Program, University of Denver, Denver, CO, 1984-1987

AWARDS

- Yahoo! Patent Milestone Award, February 20, 2008
- EU Hack Day “Coolest Hack” for FlickrSoup December 2007
 - Released as PhotoSoup on <http://next.yahoo.com> January 2008
 - Released as PhotoSoup on <http://www.facebook.com> May 2008
- AT&T Laboratories Fellowship Program Grant Recipient 2001-2006
- Best Poster Award, Conference on Information and Knowledge Management, 2002.
- University of Massachusetts Graduate School Fellowship 2001

RESEARCH EXPERIENCE

Post-doctoral Researcher

Yahoo! Research Barcelona

Supervisor: Ricardo Baeza-Yates
since September 2006

- **YGeo-Research:** Coordinated the collaboration between Yahoo! Research and Y! Geo team, which included three researchers and three engineers, to evaluate and improve the Yahoo! location identification and geo intent systems. Acted in a consulting role and conducted research in geographical IR.
- **Web Retrieval:** Conducted research in retrieval models for short texts such as user-generated content for images and video, and internet advertising (sponsored search and contextual advertising). Supervised an intern to develop methods to predict the difficulty of a query.

Research Assistant

Center for Intelligent Information Retrieval
Computer Science Department, University of Massachusetts

Advisor: W. Bruce Croft
September 2001 - 2006

- **Sentence Retrieval:** Applied novel translation models to estimate term distributions in sentences in a monolingual context. Investigated smoothing methods for very sparse term distributions. Applied sentence retrieval techniques to rescoring the n-best translations from a machine translation system.
- **Question answering:** Applied Conditional Random Fields to the task of definitional question answering. Constructed a probabilistic framework for QA, investigated modeling the prior probability of an answer.

- **Document retrieval:** Analyzed questions for their task-orientedness, analyzed documents for features relevant to task-oriented questions, and used those features to improve retrieval for task-oriented questions.

Research Internships

Mentor: Srinivas Bangalore

AT&T Labs-Research Florham Park, NJ

- **Summer 2002:** Developed a system to answer definitional questions using multi-document summarization.
- **Summer 2001:** Developed a translation server for Hubbub instant messenger (English and Spanish), a multi-string word-based alignment tool, and several document alignment tools.
- **Summer 2000:** Developed a program to mine parallel bilingual sentences from the Web to train a statistical machine translation model.

SERVICE

- Co-organizer, with Mounia Lalmas, Workshop on Aggregated Search, SIGIR 2008.
- Program Committee Workshop on Information Retrieval for Advertising (IRA), 2008
- Reviewer SIGIR 2008
- Program Committee ECIR 2008 (Best Student Paper Committee)
- Program Committee WSDM 2008
- Reviewer for the Journal of Artificial Intelligence Research 2008
- Reviewer for the Journal of the American Society for Information Science and Technology 2008
- Reviewer for Transactions on Internet Technology 2007
- Reviewer for IEEE Internet Computing 2007
- Reviewer Information Processing & Management 2007
- Paper and Poster Reviewer SIGIR 2007
- Paper Reviewer ECIR 2005
- Poster Reviewer SIGIR 2005
- Paper Reviewer WWW 2004
- Paper Reviewer NAACL/HLT Student Workshop 2003

PRESENTATIONS

“Diversifying Image Search with User Generated Content”, National University of Singapore, Department of Computer Science. Singapore. July 2008.

“You say ‘Bikini’ I say ‘Biquini’: Textual Search of User-Generated Content for Images”, Future of the Web Workshop. Andorra. April 2008.

“Semantic Associations for Contextual Advertising”, Machine Learning Mini-Summit, Research All-Hands Meeting, Yahoo! Research. Sunnyvale, CA. August 2007.

“Any Input, Any Output”, Future of the Web Workshop. Bertinoro, Italy. June 2007.

“Relevance and Evaluation”, Guest Lecture. Web Mining and Retrieval, Universitat Pompeu Fabra, taught by Ricardo Baeza-Yates. February 2007.

“Aspects of Question Answering”, Seminar at the Universitat Pompeu Fabra, Barcelona. November 2006.

“Question Answering”, Guest Lecture. CS591N Seminar in Computational Linguistics, taught by Andrew McCallum. May 2006.

“Question Answering”, Guest Lecture. CS585 Introduction to Natural Language Processing, taught by Andrew McCallum. December 2004.

“Simple Translation Models for Sentence Retrieval in Factoid Question Answering”, Student Workshop on Information Retrieval, MIT, November 2004.

“Question Driven Information Retrieval”, University of Massachusetts at Lowell, Department of Electrical and Computer Engineering, June, 2004. Host: Charles Thompson.

“Bootstrapping Bilingual Data Using Consensus Translation for a Multilingual Instant Messaging System”, University of Pennsylvania, November, 2002. Host: Aravind Joshi.

PUBLICATIONS

Book

Vanessa Murdock *Exploring Sentence Retrieval* VDM Verlag. ISBN: 978-3-8364-3765-3, 2008.

Refereed Publications

Claudia Hauff, Vanessa Murdock and Ricardo Baeza-Yates. "Improved Query Difficulty Prediction for the Web" *Proceedings of the Conference on Information and Knowledge Management (CIKM'08)*. Napa Valley, October 2008.

Andrei Broder, Massimiliano Ciaramita, Marcus Fontoura, Evgeniy Gabrilovich, Vanja Josifovski, Donald Metzler, Vanessa Murdock, Vassilis Plachouras. "To Swing or not to Swing: Learning when (not) to Advertise" *Proceedings of the Conference on Information and Knowledge Management (CIKM'08)*. Napa Valley, October 2008.

Roelof van Zwol, Vanessa Murdock, Lluís Garcia Pueyo, and Georgina Ramirez. "Diversifying Image Search with User Generated Content" *Proceedings of the Conference on Multimedia Information Retrieval (MIR'08)*. Vancouver, October 2008.

Vanessa Murdock. "Ellen Voorhees and Donna Harman (eds): TREC Experiment and Evaluation in Information Retrieval" *Journal of Information Retrieval*. Volume 11 Number 5. October 2008. (Book review).

Massimiliano Ciaramita, Vanessa Murdock and Vassilis Plachouras. "Online Learning from Click Data for Sponsored Search" *Proceedings of the 17th International World Wide Web Conference (WWW)*. Beijing, April 2008.

Ricardo Baeza-Yates, Aristides Gionis, Flavio Junqueira, Vanessa Murdock, Vassilis Plachouras and Fabrizio Sylvestri. "Design trade-offs for search engine caching". *ACM Transactions on the Web Special Issue on Query Log Analysis: Technology & Ethics*. To Appear 2008.

Massimiliano Ciaramita, Vanessa Murdock, and Vassilis Plachouras. "Semantic Associations for Contextual Advertising". *International Journal of Electronic Commerce Research - Special Issue on Online Advertising and Sponsored Search* February 2008.

Vanessa Murdock, Massimiliano Ciaramita and Vassilis Plachouras. "A Noisy-Channel Approach to Contextual Advertising." *Proceedings of the Workshop on Data Mining and Audience Intelligence for Advertising (ADKDD)*. August 2007.

Ricardo Baeza-Yates, Aristides Gionis, Flavio Junqueira, Vanessa Murdock, Vassilis Plachouras and Fabrizio Sylvestri. "The Impact of Caching on Search Engines". *Proceedings of the Thirtieth International ACM SIGIR Conference*. Amsterdam, July 2007.

Carlos Castillo, Debora Donato, Aristides Gionis, Vanessa Murdock, Fabrizio Sylvestri. "Know Your Neighbors: Web Spam Detection using the Web Topology". *Proceedings of the Thirtieth International ACM SIGIR Conference*. Amsterdam, July 2007.

Vanessa Murdock, Diane Kelly, W. Bruce Croft, Nicholas J. Belkin, Xiaojun Yuan. "Identifying and Improving Retrieval for Procedural Questions." *Information Processing and Management*, Volume 43 Number 1, pages 181-203. January 2007.

Georges Dupret, Vanessa Murdock and Benjamin Piwowarski. "Web Search Engine Evaluation using Clickthrough Data and a User Model". *Proceedings of the Workshop on Query Log Analysis, WWW*. May 2007.

Vanessa Murdock and W. Bruce Croft. "A Translation Model for Sentence Retrieval." *Proceedings of Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*. Vancouver, B. C., Canada. October 2005.

Vanessa Murdock and W. Bruce Croft. "Simple Translation Models for Sentence Retrieval in Factoid Question Answering." *Proceedings of the Workshop on Information Retrieval for Question Answering, SIGIR*. Sheffield, England. July, 2004.

Diane Kelly, Vanessa Murdock, Xiaojun Yuan, W. Bruce Croft, Nicholas Belkin. "Features of Documents Relevant to Task- and Fact- Oriented Questions." *Proceedings of the Conference on Information and Knowledge Management (CIKM'02)*. McLean, VA, November 4-9, 2002. **(Best Poster Award)**.

Srinivas Bangalore, Vanessa Murdock and Giuseppe Riccardi. "Bootstrapping Bilingual Data Using Consensus Translation for a Multilingual Instant Messaging System." *Proceedings of the International Conference on Computational Linguistics (COLING 2002)*. Taipei, Taiwan, August 2002.

Vanessa Murdock and W. Bruce Croft. "Task-Orientation in Question Answering." *Proceedings of the Twenty Fifth International ACM SIGIR Conference*. Tampere, Finland, August 2002. (Poster)

James Bieman and Vanessa Murdock. "Finding Code on the World Wide Web: A Preliminary Investigation." *Proceedings of the International Workshop on Source Code Analysis and Manipulation (SCAM)*. 2001.

Unrefereed Publications

Diane Kelly, Vanessa Murdock, and Xin Fu. "Using Interactions to Improve Translation Dictionaries: UNC, Yahoo! and ciQA." E. Voorhees and L. P. Buckland (Eds.) TREC 2007, Proceedings of the Sixteenth Text Retrieval Conference. Washington, D.C.: GPO. 2008.

Andres Corrada-Emmanuel, W. Bruce Croft, and Vanessa Murdock. "Answer Passage Retrieval for Question Answering." CIIR Technical Report IR-283. 2003.

Vanessa Murdock and Srinivas Bangalore. "Some Experiments in Bootstrapping Machine Translation Models." AT&T Technical Report TD-556QLY. AT&T Labs-Research, December 2001.

Miscellaneous Publications

Vanessa Murdock. "Aspects of Sentence Retrieval" Dissertation Abstract. SIGIR Forum 41(2): 127 (2007).