

## Einat Minkov

---

Language Technologies Institute  
Carnegie Mellon University  
5000 Forbes Ave.  
Pittsburgh, PA 15217

Phone: (617) 417-7878  
einat@cs.cmu.edu  
<http://www.cs.cmu.edu/~einat/>

### Education

**Carnegie Mellon University** – Pittsburgh, PA  
Ph.D., Language and Information Technologies, School of Computer Science, 2008  
Thesis title: “Adaptive Graph Walk Based Similarity Measures in Entity-Relation Graphs”

**Carnegie Mellon University** – Pittsburgh, PA  
M.Sc., Language Technologies, School of Computer Science, 2005  
(GPA: 3.86)

**Tel-Aviv University** – Tel-Aviv, Israel  
M.Sc., Industrial Engineering, *Magna Cum Laude*, 2000  
(GPA: 92/100)

**Tel-Aviv University** – Tel-Aviv, Israel  
B.Sc., Industrial Engineering, *Magna Cum Laude*, 2000  
(GPA: 91/100)

### Research Interests

Information extraction  
Natural language processing  
Link analysis / graph mining  
Information retrieval and question answering  
Machine learning

### Research Experience

**Research Staff** 2008–Present  
Nokia Research, MIT affiliate Cambridge, MA  
Research in the area of natural language interface to databases.

**Research Assistant** 2003–2008  
Carnegie Mellon University Pittsburgh, PA

Advised by Prof. William W. Cohen.  
Research in information extraction, information retrieval, link analysis and machine learning. Early research focused on sequential models (Hidden Markov Models, Conditional Random Fields) and the Named Entity Recognition problem. In thesis research, I proposed a framework for evaluating inter-entity similarity in heterogeneous entity-relation graphs using random graph walks. The derived similarity measure is readily applied in an ad-hoc IR settings. Additionally, learning is applied to adapt the similarity measure for pre-specified tasks. This framework outperformed alternative meth-

ods across multiple domains, including personal information management and information extraction from parsed text.

**Research Intern**  
Microsoft Research

2006  
Redmond, WA

Research in the area of machine translation, applying machine learning techniques to improve on morphology generation, with Dr. Kristina Toutanova and Dr. Hisami Suzuki.

## Teaching Experience

Teaching Assistant  
*Artificial Intelligence* course (undergrad.)  
Instructors: Prof. Martial Hebert and Prof. Michael Lewicki  
School of Computer Science, Carnegie Mellon University, Spring 2007

Teaching Assistant  
*Language and Statistics II* course (graduate)  
Instructor: Prof. Noah Smith  
Language Technologies Institute, Carnegie Mellon University, Fall 2007

## Previous Employment

Amdocs  
Business Intelligence Analyst  
Implemented data mining algorithms for credit risk detection and customer churn prediction at US cellular phone companies.

Israel

2002–2003

3Com  
Business Analyst  
Responsible for sales data analysis and prediction

Israel

2000–2001

## Honors

Military service in the “Talpiot” elite track of the Intelligence, IDF.  
Direct studies towards M.Sc in Industrial Engineering for outstanding students, Tel-Aviv University  
IBM travel award at SIGIR-06 (awarded to three students)

## Professional Service

**Conference program committees:** the 30th annual international ACM SIGIR conference on Research and development in information retrieval *SIGIR-07*, the 31st annual international ACM SIGIR conference on Research and development in information retrieval *SIGIR-08*, the 46th Annual Meeting of the Association of Computational Linguistics *ACL-08*, the 21st International Joint Conference on Artificial Intelligence *IJCAI-09*, Conference on Empirical Methods in Natural Language Processing *EMNLP-09*, the 23rd Annual Conference on Neural Information Processing Systems *NIPS-09*

**Workshop program committees:** *Learning to Rank for Information Retrieval-07*, *Learning to Rank for Information Retrieval-08*, *Learning to Rank for Information Retrieval-09*

**Journal reviewer:** *ACM Transactions on the Web*, *Artificial Intelligence Journal*, *International Journal of AI Tools*, *IEEE Transactions on Knowledge and Data Engineering*

## Publications

### Journal Publications

William W. Cohen, Einat Minkov  
“A Graph-Search Framework for Associating Gene Identifiers with Documents”  
*BMC Bioinformatics* 2006, 7:440

Mark Last, Oded Maimon, Einat Minkov  
“Improving Stability of Decision Trees”  
*International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, Vol.16, No.2, 2002

### Refereed Conference Publications

Einat Minkov, William W. Cohen  
“Learning Graph Walk Based Similarity Measures for Parsed Text”  
*in Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2008. Acceptance rate (oral): 20%.

Einat Minkov, Ramnath Balasubramanyan, William W. Cohen  
“Activity-centric Search in Email”  
*in Proceedings of the 5th Conference on Email and Anti-Spam (CEAS)*, 2008

Einat Minkov, Kristina Toutanova, Hisami Suzuki  
“Generating Complex Morphology for Machine Translation”  
*in Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics (ACL)*, 2007. Acceptance rate: 22%.

Einat Minkov, William W. Cohen  
“An Email and Meeting Assistant using Graph Walks”  
*in Proceedings of the 3rd Conference on Email and Anti-Spam (CEAS)*, 2006. Acceptance rate: 36%.

Einat Minkov, William W. Cohen, Andrew Y. Ng  
“Contextual Search and Name Disambiguation in Email using Graphs”  
*in Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 2006. Acceptance rate: 19%.

Einat Minkov, Richard C. Wang, Anthony Tomasic, William W. Cohen  
“NER Systems that Suit User’s Preferences: Adjusting the Recall-Precision Trade-off for Entity Extraction”  
*in Proceedings of the Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (HLT/NAACL)*, 2006 (short paper). Acceptance rate: 41%.

Einat Minkov, Richard C. Wang, William W. Cohen  
“Extracting Personal Names from Emails: Applying Named Entity Recognition to Informal Text”  
*in Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, 2005. Acceptance rate: 32%.

William W. Cohen, Einat Minkov, Anthony Tomasic  
“Learning to Understand Web Site Update Requests”  
*in Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*, 2005.

Acceptance rate: 18%.

### **Refereed Workshop Publications**

Einat Minkov, Ramnath Balasubramanyan, William W. Cohen  
“Activity-centric Search in Email”  
*in Proceedings of the Enhanced Messaging Workshop at the 23rd Conference of Artificial Intelligence (AAAI), 2008*

Einat Minkov, William W. Cohen  
“Learning to Rank Typed Graph Walks: Local and Global Approaches”  
*in Proceedings of the ACM Workshop on Web Mining and Social Network Analysis, aka the Joint 9th WEBKDD and 1st SNA-KDD Workshop, 2007*

Einat Minkov, William W. Cohen, Andrew Y. Ng  
“Contextual Search and Name Disambiguation in Email using Graphs”  
*in Proceedings of the 1st Workshop on Graph-based Methods for Natural Language Processing (TextGraphs), 2006*

Anthony Tomasic, William W. Cohen, Susan Fussell, John Zimmerman, Marina Kobayashi, Einat Minkov, Nathan Halstead, Ravi Mosur, Jason Hum  
“Learning to Navigate Web Forms”  
*in Proceedings of the 3rd International Workshop on Information Integration on the Web (IIWEB), 2004*

### **Selected Public Talk Venues**

Conference on Empirical Methods in Natural Language Processing, October 2008

IBM Haifa Research Seminar, December 2007

Meeting of the ACM Workshop on Web Mining and Social Network Analysis at KDD, August 2007

Annual meeting of the Association of Computational Linguistics (ACL), June 2007

Annual meeting of the Institute for Operations Research and the Management Sciences (INFORMS), invited talk, November 2006

Conference on Email and Anti-Spam, August 2006

Annual international ACM SIGIR conference on Research and development in information retrieval, August 2006

Meeting of the 1st Workshop on Graph-based Methods for Natural Language Processing (TextGraphs) at HLT/NAACL, June 2006

Information retrieval discussion series, Carnegie Mellon University, April 2006

International Joint Conference on Artificial Intelligence, August 2005