

Danny Shacham

Profile

- Skilled and experienced Data wrangling person. My main areas of expertise are **Machine Learning, Data Mining, Information Extraction and Natural Language Processing**.
- In my heart I'm a software engineer. I write production code. All the time.
- I do science, and this requires both empirical creativity as well as empirical inquiry.
- I distill social meaning from the data that is available, and know how to tell a statistically grounded, eschewing complexity and spurious correlations story through data.
- I find data relationships using statistics, visualization tools and machine learning techniques. I create great recommendation algorithms, often relying on intuition mixed with clever feature selection.
- I design product changes and verify impact using measurements. I'm responsible for experimental design and final interpretation of data, combined with a deep understanding of large-scale highly available systems reality.
- I apply supervised and unsupervised techniques in context of large data sets and latent variables.
- The abilities I bring with me, which are crucial to the success of work I do, are high level self-learning, working alone and in a team, high skills and initiative.

Experience

5/2011 - newBrandAnalytics

NLP and ML Lead

- I lead the data extraction team, working on the most challenging NLP projects exist :).
- Architected and built the team and core NLP product from the prototype stage to a fully functional production module.
- The team is building a text analysis engine, that involves many top notch NLP algorithms, such as classification, opinion extraction, aspect-base analysis, sentiment analysis, CRF, and many more.
- We also deal with large ontologies, semantic analysis and data mining.
- I wrote production code more than 50% of my time.

3/2008- 5/2011 Aol Answers (formerly Yedda.com)

Senior Software Engineer

- I work to solve complex algorithmic and machine learning problems and play a critical role in the group as a creative research engineer.
- Architected and built highly-scalable, high-performance machine learning modules, using, among others, memcached, Solr-Lucene and more.
- Responsibility on software and algorithms modules.
- My main interest is the relevancy of the Aol Answers system, search optimizations and personalization.

7/2006- 3/2008 IBM Inc.

Staff Member

- Senior Software Engineer at the Search Technologies Group.
- Part of the team that develops IBM's search engine (Juru ©) written in Java.
- Designed and implemented major performance improvements in the internal Search engine of Portal.
- Part of the Portal Search engine group that integrates the core Search engine to IBM's WebSphere portal, using J2EE technologies.

7/2005- 7/2006 TimeBridge Inc., A collaboration startup.

Software Engineer

- Senior developer in the development team of TimeBridge.
- In charge of the web application side of the product, developed in PHP, using AJAX and Web 2.0 techniques.
- Part of the team developing TimeBridge client, written in C#.
- Researcher of possible architectures and solutions for several aspects of the product.

2002- 2005 Infosoc, The Center for the Study of the Information Society

Software Engineer

- Programmer in the development team of the QSIA System (<http://qsia.org>), and the SHIL website (<http://shil.shil.info>). Both Systems are client-server web applications, built over J2EE platform, using Tomcat servers.
- My work also touches other aspects of the development cycle, starting at characterizing customer requirements and designing the architecture.

2001 – 2002 The Interdisciplinary Center (IDC) at Herzliya, Israel

Java course tutor.

Education

2002–2006 Haifa University

- **M.Sc** in computer science, at the field of Natural Language Processing.
- Thesis: Morphological Disambiguation of Hebrew.

2000–2002 The Interdisciplinary Center (IDC) at Herzliya, Israel

- **B.A.**, in Computer Science. (Graduated cum laude)

Publications

- Gennadi Lembersky, Danny Shacham and Shuly Wintner. **Morphological Disambiguation of Hebrew: A Case Study in Classifier Combination.** *Natural Language Engineering, accepted for publication.*
- Danny Shacham and Shuly Wintner. **Morphological Disambiguation of Hebrew: A Case Study in Classifier Combination.** In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 439-447, Prague, June 2007.
- Idan Szpektor, Ido Dagan, Alon Lavie, Danny Shacham and Shuly Wintner. **Cross Lingual and Semantic Retrieval for Cultural Heritage Appreciation.** In *Proceedings of the ACL-2007 Workshop on Language*

Technology for Cultural Heritage Data (LaTeCH 2007), pages 65-72, Prague, June 2007.

- Danny Shacham. Morphological Disambiguation of Hebrew. University of Haifa MSc. Thesis, 2007.