

Anish Nair

A9.com
130 Lytton Ave
Palo Alto, CA

email: an@anishnair.com
phone: (323) 620-0501

Areas of specialization

Search Relevance, Machine Learning, Natural Language Processing, Computational Advertising, Big Data Infrastructure (Hadoop, Pig, S4), Distributed Systems

Education

- | | | |
|------|---|---|
| 2004 | MS in Computer Science
University of Southern California | Research Area: <i>Computational Linguistics</i> |
| 2002 | BEng in Computer Engineering
Mumbai University | |

Appointments held

- | | |
|--------------|---|
| 2011-present | Software Development Engineer, A9.com, Palo Alto, California <ul style="list-style-type: none">› Machine Learned Ranking for Product Search on Amazon.com |
| 2007-2011 | Scientist, Yahoo! Labs, Santa Clara, California <ul style="list-style-type: none">› Part of applied research team running Yahoo!'s search ads marketplace - a \$2 billion business› Responsible for tuning and optimization of Yahoo!'s search ad auctions› Developed and deployed search ad algorithms for 20 international markets - \$80MM revenue gain› Part of team that developed S4: the open-source distributed stream processing platform (http://s4.io)› Various problems in machine learning, stream processing, language processing and econometrics› Part of team awarded Yahoo! Superstar (highest form of recognition in the company) twice |
| 2004-2006 | Research Engineer, Ordinate Corporation, Menlo Park, California <ul style="list-style-type: none">› Developed novel algorithms for automatic grading of spoken utterances› Developed and maintained production code for live (over-the-phone) grading |
| 2002-2004 | Research Assistant, University of Southern California, Los Angeles, California <ul style="list-style-type: none">› Research problems in computational linguistics, unsupervised learning, folk psychology |

Honors & awards

- | | |
|------|-------------------------|
| 2010 | Yahoo! Superstar (Team) |
| 2008 | Yahoo! Superstar (Team) |

Selected projects

- 2008-2011 **Search Advertising Auctions** *Yahoo! Labs, Santa Clara*
- › Worked on optimizing ranking, pricing, placement and personalization; within the context of search advertising auctions
 - › Responsible for the problem of deciding where to place ads on the search results page
 - › Attempted to model user and advertiser behavior to better understand long term effect of ads
 - › Worked closely with Yahoo!'s economists to understand the 'health' of the search advertising marketplace
 - › Monitored and tuned ad footprint on Yahoo! search pages
 - › Worked on personalizing the search advertising experience from the users perspective
 - › Ran numerous experiments on live traffic, many of which are now deployed on all of Yahoo! search traffic
- 2008-2011 **S4: Distributed Stream Computing Platform** (<http://s4.io>) *Yahoo! Labs, Santa Clara*
- › Part of team that designed and implemented S4: a scalable, distributed, pluggable platform for real-time processing of stream data
 - › Similar in spirit to real-time Map Reduce; but without the disadvantages of batch processing
 - › Participated in initial design and built search advertising applications to demonstrate the platform. Applications built and deployed include:
 - › Online parameter optimization: A control system that adapts model parameters based on changes in the system environment
 - › Real-time search traffic shaping for optimizing revenue and user engagement
 - › User session based ad retrieval model
 - › Real-time click feedback system
 - › Real-time metrics dashboard
- 2007 **Click prediction in Search Advertising** *Yahoo! Labs, Santa Clara*
- Part of a team of four responsible for launching click prediction models for Yahoo!'s search advertising product (Panama) in 20 non-US markets. We designed a scalable and completely adaptive statistical model, implemented online and offline components, tested models on real user traffic and deployed into production; all within a year
- 2004-2006 **Automatic grading of spoken utterances** *Ordinate Corporation, Menlo Park*
- › Designed novel algorithms for automatic grading of open ended spoken utterances
 - › Incorporated techniques from machine learning, speech recognition and psychometrics
 - › Built acoustic and linguistic models in multiple languages, designed experiments to validate models, deployed models to production
 - › Developed and maintained production grading code
- 2003 **Unsupervised Decipherment** *Information Sciences Institute, USC, Marina Del Rey*
- Applied unsupervised learning methods to the problem of converting random Hindi encodings to a standard encoding, with a view to create crawler-friendly Hindi text
- 2002-2003 **The Theory of Mind in Language Project** *Institute for Creative Technologies, USC, Marina Del Rey*
- Assisted in the construction of a lexical-semantic resource for identifying references to commonsense psychology concepts in English text and used the resource to derive proof for psychological postulates from text corpora

Publications & talks

- 2010 *S4: Distributed Stream Computing Platform*
Leo Neumeyer, Bruce Robbins, Anish Nair, Anand Kesari
2010 IEEE International Conference on Data Mining Workshops (ICDMW 2010), Sydney, Australia
- 2010 *Transitive history-based query disambiguation for query reformulation*
Karim Filali, Anish Nair, Chris Leggetter
SIGIR 2010, Geneva, Switzerland
- 2010 *Generalized Utility in Web Search Advertising*
Stefan Schroedl, Anand Kesari, Anish Nair, Leo Neumeyer, Sharath Rao
Workshop on Data Mining and Audience Intelligence for Online Advertising, Washington DC, USA
- 2006 *Unsupervised Analysis for Decipherment Problems*
Kevin Knight, Anish Nair, Nishit Rathod and Kenji Yamada
Proceedings of Coling/ACL2006, Sydney, Australia
- 2005 *Item difficulty estimation using Linguistic and Acoustic Predictors*
Anish Nair
Proceedings of the IEICE Thought and Language Workshop, Tokyo, Japan
- 2005 *Automatic Grading of Spanish Story Retellings*
Anish Nair, Jennifer Balogh, Jared Bernstein, Jian Cheng, Matthew Lennig
Presented at the Language Testing Research Colloquium, Ottawa, Canada
- 2004 *Expressions related to knowledge and belief in children's speech*
Andrew Gordon and Anish Nair
Proceedings of the 26th Cognitive Science Conference, Chicago, Illinois, USA
- 2003 *Literary evidence for the cultural development of a theory of mind*
Andrew Gordon and Anish Nair
Proceedings of the 25th Cognitive Science Conference, Boston, Massachusetts, USA
- 2003 *Recognizing Expressions of Commonsense Psychology in English Text*
Andrew Gordon, Abe Kazemzadeh, Anish Nair and Milena Petrova
Proceedings of the 41st ACL Conference, Sapporo, Japan

Patents

- 2010 *Method for reducing North Ad Impact in Search Advertising*
Leo Neumeyer, Anish Nair, Stefan Schroedl, Anandsudhakar Kesari. Patent Pending
- 2010 *Multiple Cascading Auctions in Search Advertising*
Leo Neumeyer, Anish Nair, Stefan Schroedl, Anandsudhakar Kesari. Patent Pending
- 2010 *Method and System for providing Contents based on Past Queries*
Karim Filali, Anish Nair, Chris Leggetter. Patent Pending
- 2004 *Spoken Language Proficiency Assessment by Computer*
Anish Nair, Brent Townshend, Matthew Lennig. Patent Pending

Skills

Programming

Languages

Java
C++
Perl
Python
some Scala
some Ruby

Data Processing

Hadoop
Pig
Matlab

Web

PHP
Ruby on Rails
XHTML
CSS

Platforms

Linux
Mac OS X

Specialized

- › Large-scale user testing for the web
- › Behavioral modeling of web users
- › Data pipelines on Hadoop
- › Stream processing using S4

Spoken Languages

English, Hindi, Malayalam

Status

Nationality: **Indian**
US Immigration: **H1-B**

References

Available on request