

Summary

Senior technical manager and engineer with over 10 years of experience building web-scale systems for search and information retrieval. Extensive background as a developer and technical lead in multiple languages, with experience designing, building, and deploying software used by hundreds of engineers and millions of users from the ground up. Seasoned manager with experience building and leading large teams across multiple sites on multiple projects at once.

Experience

Google, Inc.

New York, NY • June 2004 - Present

Senior Staff Engineer and Engineering Manager

Built structured data engineering group from 5 to 35 engineers. Architected and built Google's core representation for structured annotations on documents, now in use across Google in products ranging from Web, Maps, and Product Search to Flu Trends. Responsible for both long-term strategic planning as well as direct team leadership, development, and recruiting.

Key accomplishments:

- **Annotations Framework:** Designed and implemented a system for annotating structured information on documents. Built team of 12+ engineers from scratch to develop and implement new extraction techniques and drive product integrations across Google. Extractors range from simple (dates, numbers, measurements) to complex (geographic information, named entities, sentiment analysis). This system now drives features based on structured data across all of Google, including:
 - Places Mode, Rich Snippets, Recipe Search, and many others on Web Search.
 - "Web Pages" and "Reviews" sections on Google Maps and Place Pages.
 - Dozens of core ranking and snippet improvements across google.com search, ads, and other products.
- **Structured Data and NLP Strategy:** Responsible for Google's overarching strategy for structured data, information extraction, and document understanding. Developed and executed both short- and long-term plans for project selection, hiring, and integration across all of Google's products.

Engineering lead for M&A in the space of semantics and structured data. Co-sponsored acquisition of Metaweb Technologies in July 2010. Responsible for all areas of due diligence on the deal, post-deal strategic planning, and on-boarding the team of 25 employees.

Public face for our semantic search efforts, including a recent talk at the Google NY Meetup entitled The Structured Search Engine.

- **Question Answering and NLP:** Led the development of a world-class query and document understanding framework that has answered more than a billion questions on google.com. Built industry-leading systems for extracting facts from documents, indexing structured data, and parsing queries. Grew and led 10-person team in rapid, iterative development cycle with nightly builds and evaluations and weekly pushes.
- **Google Squared:** Lead engineer and manager for team of 10 engineers building a next-generation search product. Responsible for initial strategy and presentation to executives. Built team, developed "startup-like" atmosphere within Google, and drove launch in June 2009. Post-launch, continued to drive team on rapid, iterative development cycle, launching

new features on a bi-weekly basis. Drove the integration of Squared technologies into google.com, leading to the re-launch re-launch of question answering and better organization of the search page.

- **Google Goggles:** Helped instantiate the New York Google Goggles team based on a demo by the lead engineer. Built team from a single engineer to 9 in New York, working closely with a team distributed across four other offices.

MIT Computer Science and Artificial Intelligence Laboratory

Cambridge, MA • July 2003 - May 2004

Research Assistant

Worked on the Haystack project, a Java- and XML-based information management client. Research included designing and implementing data model, algorithm, and user interface for learning patterns for semantic information extraction on the World Wide Web. System allowed users to highlight examples of data they wished to extract, forming a pattern which could be re-used to extract new information at a later date. Semantic labels were applied to these patterns, allowing them to be integrated into the larger framework of the Semantic Web.

Publication: Thresher: Automating the Unwrapping of Semantic Content from the World Wide Web, WWW 2005, May 10-14, 2005, Chiba, Japan.

Advent, Inc.

New York, NY • August 2001 - May 2004

Senior Developer / Team Lead

Managed design, development, and implementation of "DAX" (Data Acquisition and Transformation), a system for consolidation and aggregation of disparate financial data using J2EE components. Design included distributed, multi-threaded, agent-based architecture with centralized administration and auditing features. Standardized internal business object representations to allow handling of a wide range of inputs and outputs with little or no new code. Included work with Enterprise Java Beans (EJBs), JSP, JMS, XML, BEA WebLogic, ATG Dynamo, and Oracle.

Storefront Media, Inc.

Cambridge, MA • Sept. 1999 - January 2001

Co-founder and CFO

Co-founded e-commerce solutions startup. Lead development of "Natural Fit" software to match a given shopper's measurements with the correct clothing size. Also aided in development of collaborative filtering recommendation application. Development work included C++, JavaScript, TCL, and PL/SQL with Oracle. Managed relations with main client, including technical and product marketing meetings. Presented technology to various potential investors. Co-wrote business plan (technical and financial sections) as well as developing sections of main technical documents.

Education

Massachusetts Institute of Technology

Cambridge, MA • July 2003 - June 2004

Master of Engineering Degree in Computer Science, June 2004. Cumulative GPA: 4.8/5.0
Thesis: *Tree Pattern Inference and Matching for Wrapper Induction on the World Wide Web*

Advisor: David Karger

Massachusetts Institute of Technology

Cambridge, MA • August 1997 - June 2001

Bachelor of Science degree in Computer Science, June 2001. Cumulative GPA: 4.5 / 5.0

Proficiencies

- Wide array of programming languages, libraries, and markup, including C++, STL, Python, Javascript, Java/J2EE/GWT, Perl, SQL, XML, CSS, and many more.
- Expert skills across a wide array of technologies, including highly-scalable, distributed systems, multi-tier indexing and serving frameworks, web applications, and databases, as well as specific systems such as MapReduce/Hadoop, BigTable, and Sawzall.
- Expert knowledge of search ranking, information retrieval, machine learning, information extraction, and natural language processing.
- Management and team building in a high-impact, fast-moving, distributed environment. Recruiting, career mentoring and growth, communication across sites and organizations, managing both senior and junior engineers. Proven track record of delivering large, complex projects on schedule.
- Skilled presenter and speaker, both internally and externally. Strategy presentations to executives and external talks on a wide array of topics, from core engineering to Google.org initiatives.

Miscellaneous

- William A. Martin Memorial Thesis Award for Outstanding Computer Science Thesis, 2004.
- Inventor or co-inventor on more than 25 patent applications, including 8 issued.
- Fully built, installed, and maintained own Linux distribution and supporting software from source code.
- Winner of 2003 MIT "Real Complex Planes" competition (category: longest duration).
- ACM Member since 2005.
- Eagle Scout.