

## ADITYA SAKHUJA

655 South Fair Oaks, Apt# C-306  
Sunnyvale, CA – 94086, USA

(240)-751-3140  
aditya.sakhuja[at]gmail.com  
<http://www.sakhuja.us>

### OBJECTIVE

Opportunities in Engineering. Special interests include data processing and analysis, performance optimizations, web search, web technologies, databases and enterprise solutions

### EDUCATION

Masters in Computer Science, Georgia Institute of Technology, Atlanta (Aug 2007 to Dec 2008)  
GPA 3.58  
Bachelors in Computer Engineering, University of Pune, Pune (July 2004 to July 2007)  
First Class with Distinction  
Diploma Computer Engineering, MSBTE, Mumbai (May 2000 to May 2004)  
First Class with Distinction

### TECHNICAL SKILLS

<b>Languages</b>	Python, Java, C, C++
<b>Scripting</b>	Perl, Shell / Awk / Sed, SQL, PLSQL, PIG, SPARQL
<b>Tools</b>	ANT, Svn, CI with Hudson
<b>Web Technologies</b>	PHP, XSLT 2.0, ASP 3.0, JavaScript, VBScript, JSP, XHTML, Semantic Web, XML, mxm12 (adobe flex development), AJAX
<b>Enterprise Systems</b>	Hadoop, Informatica PowerCenter, Apache, Tomcat, Opentaps (ERP + CRM)
<b>RDBMS</b>	Oracle 10g, MySQL, SQL Server
<b>OS</b>	Unix, Mac, Windows
<b>Others</b>	Rational Rose, OmniGraffle, VPSuite (UML), Eclipse, NetBeans, .NET framework, Wireshark

### WORK EXPERIENCE

*Yahoo Search, Yahoo! Inc.* (Jan 2009 – present)

**Search Platforms Group:** Designed and developing a critical component responsible for serving, filtering and transforming data from the hadoop content grid to a variety of consumers.

Responsible for **processing and analysis** of web search data for critical decision making

**Searchmonkey Data Extraction:** Designed and developed a framework for processing a variety of feed sources producing data in our proprietary format, which is pushed into the yahoo search indexing module, to generate more meaningful and intelligent results.

*Yahoo Search Intern, Yahoo! Inc.* (May 2008 – Aug 2008)

**Researched** and **identified** through experimentation the performance bottlenecks in the P2P system used by Yahoo! for data distribution amongst its data centers. Improvements anticipated bringing 4X performance gain in the production setup. This project involved deep understanding of the algorithms underlying the bit torrent protocol.

**Designed** and **implemented** Spatial Alarms using mobile blueprints framework developed at Yahoo!

*Teaching Assistant for Databases Systems, Georgia Institute of Technology*

Responsible for **assisting** the students with their projects, course exams and grading their performance

*Persistent Systems Private Ltd., Pune (Undergrads final year project)* (Aug 2006 – Jun 2007)

Clinical Data Connector for Informatica's ETL Tool – PowerCenter - **Designed** and **developed** a Clinical Data Connector. It performed the extracting of clinical unstructured data from public databases, processing it, cleansing it, transforming it and presenting it to the underlying ETL Tool. The key challenge was the schema generation for the unstructured data sources.

*L&T InfoTech Limited. , Networking & IT Security, Mumbai* (Nov 2003 – May 2004)

Secured file management system

*Larsen & Toubro Limited, HED-IT, Mumbai* (June 2002 – Dec 2002)

Workflow Applications on Lotus Notes

## SELECTED PROJECTS

Georgia Institute of Technology

(July 2007 - Dec 2008)

*Improved Image Search (Web Searching & Text Mining)* - **Designed** and **implemented** personalized relevant image search results through classification and relevance feedback

*Context based Author Ranker & Recommender System (Emerging Database Technologies)* - **Designed** and **implemented** context specific recommendation engine of relevant authors to peer authors in the system.

*DBMS Engine (Databases systems implementation)* - **Implemented** Buffer manager, Indexing and Transaction log management modules of a DBMS engine

*Automated Schema Matching* - **Research work** for automation of the schema matching under Prof. Sham Navathe's guidance

*Co Scheduling algorithm for GTthreads* (<http://www.cc.gatech.edu/~smcmanus/6210/Project1.html>)

**Designed and implemented** a co-scheduling algorithm for GTthreads, a user level library package

*Shared memory based IPC with QoS support* (<http://www.cc.gatech.edu/~smcmanus/6210/Project2.html>)

**Designed and implemented** an IPC system using shared memory and providing QoS support

*MultiRPC support for XML-RPC* (<http://www.cc.gatech.edu/~smcmanus/6210/Project3.html>)

**Designed and implemented** MultiRPC functionality in the open source XML-RPC library. Pthreads library used for threading

*Recoverable Virtual Memory (RVM)* (<http://www.cc.gatech.edu/~smcmanus/6210/Project4.html>)

**Implemented** Recoverable Virtual Memory (RVM) library

## RELEVANT COURSES

High Performance Parallel Computing

Design and Analysis of Algorithms

Database Systems Implementation

Data Communications Lab

Software Architecture

Advanced Operating Systems

Web Search and Text Mining

Emerging Database Technologies

Enterprise Computing

Software Engineering

## AWARDS & ACTIVITIES

USTA member - Played Tennis tournaments for various clubs

Awarded special award by TRDDC (Tata Research Design and Development Center) at Project Exhibition, 2007

Co Curricular Excellence Award – Fourth year undergrads in Computer Engineering, 2007

Academic Excellence Award – Second year undergrads in Computer Engineering, 3rd rank, 2005

Finalist at Bharatiyam '06, an intercollegiate C++ programming contest, 2006

Finalist at Osiris '06, an intercollegiate Java programming contest, 2006

4th Position at Techtalk '05, Robotics National level Competition, 2005

Student In charge for the WiFi network in Undergrads, Computer Engineering, 2007

Won several prizes for swimming at various clubs, India

Member of Under 19 Cricket team – Calcutta club, India

Member of College Cricket team Undergrads, India