

Siddhartha Reddy Kothakapu

<http://www.grok.in/>
<http://www.linkedin.com/in/sidsr>

sids@grok.in
+91 998 028 7437

EDUCATION	M.Tech in Information Technology (3.71/4 CGPA) International Institute of Information Technology, Bangalore	Jun 2006
EXPERIENCE	Senior/ Software Engineer , Ziva Software Pvt Ltd, Bangalore <i>Ziva Software</i> is a Venture Capital funded startup based out of Bangalore, India providing products and services in the <i>Mobile Search</i> domain. Established in late 2005, Ziva has provided search solutions to different players in the mobile industry including operators and content owners. In 2007, Ziva launched <i>Zook</i> , it's consumer facing mobile search system. Was the first non-founding employee to join the company. Have been core member of the team, involved in most of the projects and actively participated in defining the development roadmap for them.	Jul 2006 - Dec 2008
PUBLICATION	Measures of "Ignorance" on the Web International Conference on Management of Data, Delhi, India	Dec 2006
SELECTED PROJECTS	Buddhi Online , Ziva Software A system to fetch <i>Yahoo! BOSS</i> Web search results at run time, fetch the result web pages and extract <i>snippets</i> that can possibly provide an answer to the query right away. Developed a prototype of the system as part of a feasibility analysis.	TEAM SIZE: 1 Nov 2007 - Present
	Buddhi , Ziva Software A dream to crawl the Web and automagically extract <i>information content entities</i> off the crawled pages. ◇ Designed and implemented <ul style="list-style-type: none">– a <i>Focused Crawler</i> based on the <i>Nutch</i> open source crawler.– a <i>HTML reducer</i> that transforms complex HTML pages into minimal HTML for easier analysis.– a <i>page abstractor</i> that discerns logical structures (such as tables, lists etc.) from unstructured HTML. ◇ Set up the <i>Hadoop</i> based infrastructure for the crawling and extraction. This was initially set up on a cluster of <i>Amazon EC2</i> machines with <i>Amazon S3</i> as a data store.	TEAM SIZE: 2-3 Apr 2007 - Present
	Zook , Ziva Software Ziva Software's flagship mobile search engine. The main promise is to provide exact answers to questions in the least number of clicks; implying that the search engine provides information right away rather than directing the user to another website that <i>might</i> have the information. ◇ Set up the search system infrastructure. ◇ Designed the <i>User Experience</i> and <i>User Interface</i> of the WAP and SMS interfaces. ◇ Responsible for deploying releases. ◇ Created a HTTP based API for accessing the search system.	TEAM SIZE: 2-3 Jan 2007 - Present

Zivabot, Ziva Software

TEAM SIZE: 1

Aug 2006 - Oct 2006

A scalable, highly-configurable, generic scraping system for crawling specific websites containing useful information and extracting that information.

Was fully responsible for the design and implementation of the system. Configured the system and set it up to work with an initial set of websites which provided base content for the Zook mobile search engine.

Better Web Buys

TEAM SIZE: 2

Feb 2006 - May 2006

A *faceted price comparison shopping search engine* with over 20 million items across more than 100 merchants.

Was completely responsible for the whole project right from downloading feeds, parsing them, storing them in a database, indexing, searching, user experience and setting up the entire system on a cluster of machines. Used *Sphinx* as the backend search system.

yadb – Yet Another DataBase, IIIT, Bangalore

TEAM SIZE: 3

Aug 2005 - Dec 2005

A reasonably *complete RDBMS* with a buffered memory manager, B+ tree indexing and a complete database management module including support for basic SQL, table management, paged storage etc. Part of the *Advanced Databases* course.

Top News, IIIT, Bangalore

TEAM SIZE: 1

Sep 2005 - Oct 2005

A system to bring out the top news of the hour by clustering the news items obtained from the feeds provided by several news sources. Any news being discussed by multiple sources within a given (variable) timeframe is marked as top news. A bi-partite graph based clustering algorithm was used for clustering based on just the title and summary of the news item. Part of the *Web Information Retrieval* course.

SiDuSE, IIIT, Bangalore

TEAM SIZE: 2

Aug 2004 - Sep 2004

A Simple Desktop Search Engine for Linux/UNIX systems. Creates a Trie-based full-text index of the files on the system. Allows advanced searching with *AND*, *OR* & *NOT* operators, including: nested queries, phrase matching and prefix matching.

SOS – S Operating System

TEAM SIZE: 2

Sep 2002 - Oct 2003

A simple monolithic kernel with kernel level threads, kernel level IPC and advanced memory management.

SKILLS

Languages: Perl, Java, C/C++, Ruby, Python

Softwares/Frameworks: Hadoop, Nutch, Lucene, Solr, Sphinx, WEKA, Ruby on Rails, Prototype, Hibernate

INTERESTS

Information Engineering—Retrieval, Extraction & Management, Scalable Architectures, Distributed & Cloud Computing, Usability & User Experience

HONORS

◇ *Valedictorian* for the class for 2006 at IIIT, Bangalore.

Jun 2006

◇ Was awarded the *Honeywell Scholarship* for excellence at IIIT, Bangalore.

Jun 2006