

# Siddhartha Reddy Kothakapu

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

sids@sids.in  
+91 998 028 7437

EXPERIENCE	<b>Technical Guru</b> , Ziva Software, Bangalore <i>Jul 2006 - Present</i> <i>Ziva Software</i> is a startup based out of Bangalore, providing products and services in the <i>Mobile Search</i> domain. Ziva provides search solutions to several players in the mobile industry including operators, manufacturers and content owners. In 2007, Ziva launched <i>Zook</i> , its consumer facing mobile search engine. Was the first non-founding employee. Have been a core member of the team and was involved in most of the projects. Helped define both development and strategic roadmap for the company.
EDUCATION	<b>M.Tech in Information Technology</b> (3.71/4.00 CGPA) <i>Jun 2006</i> International Institute of Information Technology, Bangalore
SKILLS	<b>Languages:</b> Perl, Clojure, Java, PHP, Ruby, Javascript, C/C++ <b>Frameworks/Libraries/Tools:</b> Hadoop, Nutch, CouchDB, MySQL, Lucene, Solr, Sphinx, WEKA, Ruby on Rails, Sinatra, jQuery, RabbitMQ <b>Others:</b> HTML, CSS, JSP, Bash Shell Scripting
INTERESTS	Engineering Information Systems—Retrieval, Extraction & Management, Machine Learning, Scalable Architectures, Distributed & Cloud Computing, Usability & User Experience.
SELECTED PROJECTS	<b>Buddhi Online</b> , Ziva Software      TEAM SIZE: 1-2 <i>Nov 2008 - Present</i> Real-time extraction of actionable <i>information snippets</i> from web pages. The extracted snippets are provided as exact answers to search queries. Conceptualised, designed and developed the complete system — first as a prototype and then as a production system that is currently powering Zook. <b>Buddhi</b> , Ziva Software      TEAM SIZE: 3-4 <i>Apr 2007 - Present</i> A dream to crawl the Web and automagically extract <i>information-content entities</i> off the crawled pages. That is, Buddhi extracts (semi-)structured data out of <i>any</i> web page without any a priori information about it. Have been part of the project right from its inception and have been managing it for the past few months. <ul style="list-style-type: none"><li>◊ Developed the initial prototype for a feasibility study.</li><li>◊ Designed the <i>Map-Reduce</i> based architecture for the system.</li><li>◊ Implemented a <i>focused crawler</i> based on the <i>Nutch</i> open source crawler. The crawler has been deployed and regularly crawls large portions of the Web.</li><li>◊ Set up the <i>Hadoop</i> based infrastructure for the crawling and extraction.</li></ul> <b>Zook</b> , Ziva Software      TEAM SIZE: 2-3 <i>Jan 2007 - Present</i> Ziva Software's flagship mobile search engine. Zook's 1-Click Search provides actionable information nuggets instead of a set of links to websites that <i>might</i> have that information. Have been involved in all aspects of Zook's development and deployment. <ul style="list-style-type: none"><li>◊ Set up the search system infrastructure.</li><li>◊ Designed the <i>UI</i> and <i>UX</i> of the WAP and SMS interfaces.</li><li>◊ Responsible for deploying releases.</li><li>◊ Created a HTTP API for accessing the search system.</li></ul>

**Zivabot**, Ziva Software

TEAM SIZE: 1

Aug 2006 - Oct 2006

A scalable, highly-configurable and generic *scraping system* for crawling specific websites containing desired information and then extracting that information into a pre-defined schema.

Was responsible for the design and implementation of the system. Configured the system and set it up to work with an initial set of websites; this provided the base content for the launch of Zook.

**Better Web Buys**

TEAM SIZE: 2

Feb 2006 - Jun 2006

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

Was completely responsible for the implementation right from fetching the product feeds, parsing and storing them in a database, indexing, searching, user experience and setting up the system on a cluster of machines. Used *Sphinx* as the backend full-text indexing 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 moment by *clustering the news items* obtained from the feeds provided by several news sources. Any news item being discussed by multiple sources within a given (variable) timeframe is marked as top news. Final project for 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. Apart from allowing the basic *AND*, *OR* & *NOT* operators, the system supports complex querying through nested queries, phrase matching and prefix queries. Part of the *Mathematical Foundations of Computing* course.

**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.

PUBLICATION **Measures of “Ignorance” on the Web**

Dec 2006

International Conference on Management of Data, Delhi, India

## HONORS

- ◇ *Valedictorian* for the class for 2006 at IIIT-Bangalore.
- ◇ Awarded the *Honeywell Scholarship* for academic excellence at IIIT-Bangalore.
- ◇ Chosen as the *employee of the year* for 2007-08 at Ziva Software.
- ◇ Adjudged the *best “Really Needs an Interface”* hack at the Yahoo! Open Hackday India 2007.
- ◇ Elected the president of the IEEE student chapter at the J.B. Institute of Engineering & Technology. Was responsible for organising two of the institute’s technical festivals in this position.

