**SHIBAJI MUKHERJEE**

**Tel No(s):** +91-9845588770

**Mail Address:** [mukher@yahoo.com](mailto:mukher@yahoo.com) , [shibaji@gmail.com](mailto:shibaji@gmail.com)

**Linkedin:** [**http://www.linkedin.com/in/shibajimukherjee**](http://www.linkedin.com/in/shibajimukherjee)

**Professional Summary:**

I am an information technology leader with around 19 years of broad managerial and technical experience in multiple roles, domains and technology.

I have significant research and professional experience on Natural Sciences, Data Analytics, Statistical Modeling, Scientific Computing and Machine Learning.

I have worked with large scale structured and unstructured data-store design, integration, porting and analysis. I have worked with data feeds, data stream, Biological sequence data-stores, retail data, telecom data.

I am a Technology generalist with broad idea of working in multiple domains and technologies, handling complexity, conceptualizing solutions, delivering products and projects.

I have worked in large scale systems in a variety of areas, such as Cloud Infrastructure, Vertical Search Engine, internet/ecommerce, enterprise management systems, insurance, telecom, distributed Transactional systems, messaging services, EAI, B2B and B2C systems.

I have extensive delivery and management experience at middle management level. Management experience includes managing team, accounts, project portfolios, setting up startup delivery units, setting up offshore teams, managing co-located direct reports as well as remote off shore resources.

I have extensive experience of interfacing between CXO/SVP/VP Level and the Implementation Group, as also the R&D and development Group. I have significant Client Management, RFP Management, Estimation and Budgeting experience.

I am a continuous learner and can quickly adapt and upgrade to new technology, requirements and scenario.

**Areas of Expertise/Familiarity:**

* Machine Learning, Big Data, Search Engines:
* Topics: Search Algorithms, Ranking Algorithms, Data Transformation, Natural Language Processing, Markov Models, Recommender Systems, Cloud Infrastructure, Sequence Analysis and Pattern Recognition
* Tools and Libraries: Hadoop/MapReduce Stack, Spark Stack, Streaming Solutions, ML Libraries, Search Libraries, Log Analysis, NoSQL Solutions, Octave/Matlab, Computing Libraries
* Enterprise Solutions: J2EE, CORBA, SOA, Webservices, Exalogic, Oracle RDBMS, Messaging, Caching Solutions
* Cloud, Networking, Virtualization: Oracle Cloud, Nimbula(IaaS), OpenStack, MicroServices, SDN, Infiniband
* Languages: Java, Python, R, Scala, SQL,C
* Portfolio Management: Accounts Management, Delivery Management, Resource Planning and Management, Practice setup, Processes & Metrics, Pre Sales, Client Management, Budgeting and Costing

# Industry Experience:

[Oracle](http://www.oracle.com/), Bangalore, India

Senior Manager, Fusion Middleware, Exalogic, Private Cloud -July 2014 onwards

I manage the Exalogic Private Cloud team, which is the IaaS layer for the Oracle Cloud Solution. This is a team of 9-10 Engineers at various levels. My responsibilities includes delivery, product management, new team setup, team line management, hardware setup and management (in-house commodity cluster setup & rack setup), interfacing with executive management and coordinating with multiple groups in US, Europe and Asia. I successfully built (hired and trained) a team of 9 people in around 6 months time.

Typical projects in the team consists of bug fixing, feature enhancement, exploring new technology, developing POC's and supporting teams consuming the IaaS layer.

[Kinsight Analytics](http://www.kinsightanalytics.com/), Seattle, WA

Consulting Partner, Analytics and Machine Learning- April 2013 to June2014

I worked as a consulting research and technology evangelist in areas of Data Analytics and Machine Learning. My main responsibility was to develop white papers, research reports, industry study of mostly travel and retail domain adaptation and relevance of Analytics and Mathematical Modeling using Machine Learning and Statistical Methods and models. This was an advisory level position.

[Thomson Reuters](http://thomsonreuters.com/), Bangalore, India

Senior Manager, Technology, Search and Index Group- March 2009 to April 2013

I was responsible for managing a team of software developers in planning, directing, and overseeing the entire software development process for the Search and Index team for WestLaw Novus Search-Index offshore development and delivery group. I successfully built (hired and trained) a team of 20 people in around 6 months time.

My primary responsibilities included technical direction on building Java/C++ based enterprise search engine software components, overseeing the incremental/agile software development process, working with various line of business (onshore team, India AVP/VP, Client Consultants, Architect and Technologist group, Program and Portfolio Management, R&D) of the Novus Search-Index group on requirement/design/integration of software components, and addressing customer issues/escalations.

I was responsible for all aspects of resource management, including technical talent acquisition, resource planning, allocation, mid-year and yearly performance evaluation, mentoring/coaching etc.

[Accenture](http://www.accenture.com/in-en/Pages/index.aspx), Bangalore, India

Manager/ Senior Architect, Technology Architecture Capability Group- Nov 2007 to March 2009

I worked as a senior architect in a horizontal capacity overseeing enterprise, technical and solutions aspect of application projects primarily in the telecom domain. My main architect responsibility was to design, review code and implementation details, help in critical trouble shooting, interface with various functional groups, external vendors, third party tools providers and across project groups.

I was responsible for interacting with the senior delivery and business management group, represent the company in technology forums and deliver tech talks.

I also managed an architect team and was responsible for hiring, mentoring, review, coaching and certifying architect skill level.

[SUN Microsystems & SeeBeyond](http://www.oracle.com/us/sun/index.html), Bangalore, Hyderabad, India

Engineering Manager, EAI & B2B Group-May 2004 to Sept 2007

I was responsible for managing a team of software developers in planning, directing, and overseeing the software development process for the EAI B2B team for JCAPS offshore development and delivery group. I successfully transitioned the team from Hyderabad to Bangalore and re staffed the gap positions.

My primary responsibilities included giving technical direction on building Java/SOA based integration software components, overseeing the incremental/agile software development process, working with various line of business (onshore team, Client Consultants, Program and Portfolio Management) of the B2B group on requirement/design/integration of software components, and addressing customer issues/escalations.

I was responsible for all aspects of resource management, including technical talent acquisition, resource planning, allocation, mid-year and yearly performance evaluation, mentoring/coaching etc.

I worked as the group manager for the SeeBeyond B2B group and setup the team and operations in Hyderabad for this group. I was responsible for quickly staffing the team, training the team, coordinate with the onshore product team and be a single point of contact for the B2B offshore team for all aspects of delivery, development and operations. I was also involved significantly in the Hyderabad development centre setup.

Note: SeeBeyond was acquired by SUN Microsystems, which was further acquired by ORACLE.

[Data-Core Systems](http://datacoresystems.com/), Calcutta, India/White Plains, NY

Senior Project Manager-March 1999 to May 2004

I worked for multiple accounts, technology and consulting group at Data-Core. I was responsible for setting up the New Technology Practice (Java/J2EE/CORBA) group in the company, I hired and trained 50+ resources for this group and set the direction, focus and vision for this group. I worked as Lead Architect for multiple Client Server/Legacy Integration/CORBA distributed solution projects for the company both in India and US in domains like Telecom/Travel/Mortgage/Insurance/Financial Applications /Port Management.

I also worked as Accounts Manager, Solutions Architect, worked on multiple RFP's and represented the company in various proposal discussion and bidding. I worked both as an Independent contributor and also managed multiple teams of sizes 5 to 50 spanning multiple groups and worked closely with the top management including the CEO office.

[I-Kinetics](http://www.insideview.com/directory/i-kinetics-inc), Calcutta, India/Burlington,MA

Software Engineer-July 1998 to March 1999

I worked as a Software Engineer in the database driver development group. My responsibilities included Java/C/PL/SQL coding, testing, database schema design, and support. I was also involved significantly in the Calcutta development centre setup.

Note: I-Kinetics was acquired by [eForce Global](http://www.eforceglobal.com/).

[Dias Infotech](http://www.interrait.com/), Calcutta, India

Software Engineer- September 1997 to July 1998

I worked as a Software Engineer in the web application development group. My responsibilities included Java/C/PL/SQL coding, testing, database schema design, and support.

Note: Dias InfoTech is a sister concern of [Interra](http://www.interrait.com/)

[British Institute](http://www.thebritishinstitutes.com/index.html), Calcutta, India

Senior Programmer R&D-July 1996 to May 1997

I worked mainly in their Software development group managing the UNIX environment and Database setup. I also developed POC's and designed and developed content for various technology training programs.

**Research Experience**: **(**2006 onwards research experience, PhD--Jadavpur University, Computer Science Department registered**)**

**Research Guide:** Prof. Sushmita Mitra, Machine Intelligence Unit, ISI, Calcutta

**Topic*:*** Biological Sequences and Structures: Some Studies with Machine Learning

**Publications: (**<http://www.researchgate.net/profile/Shibaji_Mukherjee>)

* Discovering Patterns & Structure in Travel Data--Kinsight Analytics, 2013
* Product development outsourcing and ODC setup, a personal perspective -- Seventh Annual International Smart-Sourcing Conference,2008, Hyderabad,India
* Fuzzy Measures in Hidden Markov Model - Neural Network World, Vol 10,2005
* Hidden Markov Model, Grammars and Biology: A tutorial –Journal of Bioinformatics and Computational Biology, Vol 3, No 2, 2005
* Tutorial on Distributed Computing, Sixth International Conference on Information Technology (CIT 2003) , Bhubaneswar, December, 2003
* Algorithms for Resource Discovery in Distributed Networks – Electronic Notes in Discrete Math’s, Vol 15, 2003
* O(d, d) Invariant solutions of Space Time Dependent String Vacua –FIZIKA B 2(1993)

**Education:**

**M.S (Physics),** Northeastern University, Boston, USA, 1995

**M.Sc (Physics),** Calcutta University, 1992

**Certificate Courses:**

Finance for Executive Decision Making Skills, IIM Calcutta, 2012

Essential Project Management, QAI, India, 2003

Various MOOC Courses from Coursera/edX/Stanford Online/Udacity/NPTEL

**Professional Society Membership:**

## ACM, Professional Member

AMS, Member

**Workplace Recognitions:**

1. Financial reward for successful project account management at DataCore, 2002-3

2. Financial reward from SUN, Hyderabad, key employee retention program, 2005-6

3. Selected in SUN Labs SEED mentoring program, (A program for growing top talents in SUN globally) 2007

4. Accenture certified Senior Technical Architect, an Accenture global recognition, 2008

5. Financial reward from Thomson Reuters, Project Bonus program, 2010-12

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## 

## Project Details

**Products**:

Oracle Compute/Nimbula: Exalogic is a hardware and software solution for extreme performance and for running customers own private and hybrid cloud. Exalogic is hardware and software engineered together to provide extreme performance, reliability and scalability for Oracle, Java and other applications, while delivering lower TCO, reduced risk, higher user productivity and one-stop support. Exalogic Elastic Cloud is based on Nimbula (Oracle Compute), which is an extremely scalable, high perforrmant IaaS solution.Oracle Compute uses ZFS Storage and Infiniband Network to deliver high performance solutions. Nimbula controls the complete lifecycle for the Hardware infrastructure and provides suitable interfaces for lifecycle management and the PaaS layer to integrate vertically with it to provide a complete cloud solution along with the SaaS Layer.

Novus (Westlaw) is a general Purpose Information Retrieval System which supports multiple content types including Legal, Tax and Accounting, Business Information News, Financial, Public Records, Trademark and Patent records

The platform is serving 9 countries and 17 business segments with a user base of around 200K+ and revenue of around 3 billion USD.

The platform is highly scalable, inverted index based search engine runs on 25K+ servers searching TBs of data, developed over 15+ years with 2M+ lines of code mainly in Java, C/C++.

The system uses various Information Retrieval and Machine learning techniques for Parsing, indexing, ranking, NLP, word sense disambiguation, word decompounding, learning to rank, query intent and collaborative recommendation. Other major areas of focus are Index compression, Real Time Load, Semantic Search and Componentization.

JCAPS/ICAN is a SOA enabled Enterprise Application Integration (EAI) solution product suite. It’s a comprehensive suite in the sense that it provides the design, development, runtime, and deployment and management facilities in one single product suite through a series of constituent products. The suite is SOA based and provides web services endpoints to all the services developed as components. It has got a UDDI registry to do services management. The tool comes with the standard WSDL builders and the XML schema builders/editors and web services publishing framework through an integrated development environment.

The B2B & Layered product group team is engaged in development, designing and

sustenance of the B2B stack and the ETL stack. The work involves developing

various component for the stack and protocol specific messaging components to cater

to vertical industry segment(AS2/CIDX/RNIF/ebXML).

OpenJDBC is a CORBA compliant JDBC Driver developed for Data access in a heterogeneous environment and from Legacy Applications. The product implements the JDBC Interface and Uses JDBC API Calls and Orbixweb ORB to access data from remote data servers through dynamically loaded Database specific Plug-in. The driver is designed to support a three tier Enterprise System

**Applications**:

Industry Solutions:

Exalogic Product Release: New features are added to each release of Exalogic to map with the customer requirement, product strategy and technology adoption. The work is divided into multiple short sprint projects, with each focusing on specific areas. Some of the specific areas the Nimbula team focused on were Orchestration, Monit integration, Infiniband Replacement with Flat Networks, public Cloud services integration and public/private cloud interoperability and switching.

Regression Framework/Log Analytics: A regression test framework for Nimbula was developed using Python Nose framework. The scope of the framework was to provide easy API's for writing test cases for specific scenarios. A log analysis framework for troubleshooting and analysis across nodes and various logs was developed using Logstash/ElasticSearch/Kibana. The test framework was also integrated with the Jenkins build.

Coherence Caching Solution: The Novus Search Service relies on an RDBMS (Oracle) to complete its searches. Each instance of the Search Controller caches its own data. With the number of Search Controllers per blade increasing from 2 to 24, the data cache was newly designed to be shared across multiple instances of the Search Controllers within the same blade.

This shared cache across multiple Search Controllers within the same blade was implemented using Oracle Coherence. One of the key focuses when designing the new caching mechanism was code reusability, maintainability and ability to revert to using the old code to help mitigate risks. The solution was implemented in three modes to accommodate the requirements of moving to a new caching solution, ability to use the old solution and scope for redundancy if any if the solution don't work in production.

Enterprise Telecom Solution- The project is divided into a set of functional domains like billing, provisioning, customer care, long distance, ESD. I have worked for ESD and Digital Subscription domain.

The Digital Subscription domain deals with provisioning, customer support, billing, reconciliation, customer relationship management for the Internet digital subscriber domain.

The ESD Architecture and Technical group is responsible for overall technical architecture and support of the ESD platform. The platform is based on SOA, Web Services, Java, J2EE, Oracle, Sybase, C++, Mainframe, Jlogrule, Orbix, Websphere MQ.

Areas covered by this support include vendor provided and in house group developed framework and tools used for coding, configuration, deployment, and management of applications. Lastly the group assists application teams by providing design and API reviews, code reviews, and debugging support.

JTB (Brunei Telecom) has a COBOL based Telecom billing system which was replaced with a modern Telecom billing solution.. This project aims at mapping the new System with the existing COBOL based system and migrate base data from old system to new system both in bulk and in incremental mode. Various components were integrated to provide a smooth transition and synchronized operation. This was a collaborative project between two companies and one consulting firm.

GTFS is an Insurance Domain project and the aim of the project is to build up a J2EE Framework based application using Oracle 9iAS on Linux to handle all insurance related process flows. The project involves large scale structured and unstructured data migration to Oracle 9i, system integration, hardware, software solutions specifications, production rollout and support.

CPT project aims to develop a software solution for process management and data management for Calcutta Port Trust so as to automate the working of various departments of CPT. The project involves large scale structured and unstructured data migration to Oracle 8i, system migration, system integration, hardware, software solutions specifications, production rollout and support.

EuroNet is a travel and hospitality segment solution for a leading Travel and booking agency for European railways.   
The project involved designing and developing a 3 Tier CORBA based OLTP system to replace the existing Informix based system. The project also involves large scale data migration from Informix to Oracle, business process mapping and production rollout.

Exploratory/R&D Projects

Fuzzy Ranking/Ranking DSL: The purpose of this feature is to provide a framework that allows the partner to create custom rank expressions that can be executed by Novus Search at run time. The framework included a set of pre-defined ranking related basic functions and supported mathematical operations. Novus Search enabled the partner application to specify the usage of custom rank rules through enhanced query syntax of the search request(A suitable DSL). During Search Execution time, Search Engine performed custom ranking on the search results by evaluating the partner specified algorithmic ranking expressions.

AlgoRank parser parsed the algorithm given by the user for their custom sorting logic. AlgoRank parser in search engine parsed the given algorithm and provided Search Query Tree which was converted into a form of tree for psort key computation.

The implementation of this parser consisted of 3 parts, namely Tokenizer, Lexer and Parser, which was implemented using JavaCC.

Travel Recommender System: Product recommendation is a major feature of any travel portal. We conducted a study to evaluate the search/recommendation features of the major travel portals and the technology behind it. We further explored the scope of sentiment analysis, twitter data feed analysis and application of some Machine Learning techniques for various sources of information analysis. The outcome was a concept paper.

Placement Algorithm: Placement of Virtual machines is an important aspect in any cloud Infrastructure. There is a huge amount of exploration that is going on in this field both in industry and in academics . A huge attention is given to this field as a small improvement in the placement algorithm can cause a huge amount of savings to the cloud provider and thus is a prime area to work on. In this project, firstly the existing Virtual machine placement algorithm of the Oracle compute cloud was understood from the already existing code. The code relating to the Placement algorithm was separated and made standalone. Unit testing of the current algorithm was done to test the existing algorithm for various important and worthy test cases.

Various research articles were surveyed to get an idea as to what other features the

current algorithm may not be considering and to identify the potential improvements that could be brought on to the existing algorithm. Some of the main sources that provided valuable information were openstack, cloudstack and many academic research papers.

**Research:**

Thesis Synopsis:

The analysis of biological sequences is a topic that involves both some biological knowledge and some advanced computational techniques. This discipline has emerged as a strategic frontier between biology and computer science.

Machine learning approaches like Neural Networks, Hidden Markov models, and Stochastic Grammars are ideally suited for areas where there is a lot of data but little theory which, is the exact situation in molecular biology.

Our research area is mainly concerned with

i) Analysis and extension of Algorithms applied to Biological sequences

ii) Application of Higher Order Grammar for biological structure prediction

iii) Pattern Recognition