

Deb Roy

🏢 Big Data Engineer · Internal Revenue Service - Cyber Security

📍 Houston, TX

Professional summary

Big Data Architect/Lead (IRS MBI active clearance)

Over 25 years of software development experience in the corporate world and played various roles.

Experienced with Hadoop based Big Data development , databricks , entity catalog. data lake, AWS cloud computing, full stack web development using Java J2EE/Jakarta,API, Data integration ,Spring Boot/AOP, .NET. Also familiar with Apache Airflow, Python based MVT framework.

Expertised in building large scale 24x7 running ETL data pipeline from scratch.

Expertised with data pipeline migration from on premis to AWS EMR.

Worked for State and Federal Government agencies, National and International Banks, Supply Chain , Truck Manufacturing, Retail , Semiconductor, Tele Communication companies etc.

Expertised in reverse engineering, critical analysis & debugging, cloud migration projects.

📍 Willing to relocate to: Houston, TX

🛂 Authorized to work in the US for any employer

Experience

Big Data Engineer · Internal Revenue Service - Cyber Security

November 2016 – Present · 9 yr 9 mo

Memphis, TN

Leading a MapR based Data lake implementation project using ELK and Spark Streaming.

involved in IRS modernisation efforts to move away from traditional legacy development to AI based development with AWS cloud based computing with data bricks, python notebook based workflow/pipeline implementation.

Syslog data parsing (firewall, web proxy, email etc)(unstructured, semi-structured).

Application server generated event logs parsing (semi structured JSON, unstructured, XML , CSV etc) .

Data correction and validation.

Loading to ElasticSearch and Parquet) for fraud detection and reporting.

Spark jobs orchestration and resource tuning.

The data volume over 3.5 TB per day.

Architected and moved the on premis ETL data pipeline to EMR/ AWS cloud.

DevOps and production support.

Worked on POC for demonstrating a web based dashboard for data arrival activities using Spring Boot, Angular JS framework/ React JS library, data grid .

The environment includes MapR, Elastic Search, MapR Streaming,Kafka, SPARK Streaming, Unix ,SPLUNK, JDK 1.8/17, SCALA, Sbt, IntelliJ, eclipse, git,codecommit, AWS , EMR,API , Data Integration ,s3,Airflow Redhat 7.x, JSON, Parquet ,Tableau, ML lib etc

Sr Software Engineer · Raytheon

November 2014 – November 2016 · 2 yr

Garland, TX

Responsibilities

Responsibilities includes requirement gathering, defining application architectures , producing high level artifacts , develop prototypes with the new software tools meeting the Authority's project requirements.

Working in an Agile development environment and leading a groups of engineers for product deliveries for Authority's JAVA/J2EE based application.

Also time to time work on the software defects on the various systems deployed in the factories,

Involved in technology evaluation and prototype application development with the Open Source JBoss Suite of products , MEAN Stacks and Big Data tool twit Hadoop HDFS, HABSE, PIG, Hive etc.

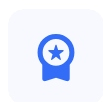
The environment includes Web Sphere Application Server, ND Portal Server, Spring MVC, Oracle, JQuery, Alfresco CMIS, Apache Chemistry, API, Data Integration, Web Services (Rest & SOAP), JBOSS Fuse, Apache Camel, EJB MDB (Messaging), MEAN stack (MongoDB, Express, Angular JS and Node JS) , KML parsing , DWR(Direct Web Remoting), Apache Hadoop, Java API for Map Reduce, HBase, Pig etc

Senior Big Data Engineer

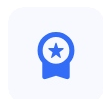
Certifications & licenses

Related licenses and certifications come from the candidate's resume and profile.

Certifications



Confidential Clearance



Driver's License

Education

Related education comes from the candidate's resume and profile.



Masters of Computer Science

Jadavpur University · Jadabpur, WB

1987 – 1989

Skills

Related skills come from the candidate's resume and profile.

- RESTful API
- XML
- APIs
- Maven
- NoSQL
- SQL
- Hadoop
- AWS
- SDLC
- Agile
- Unit testing
- User Interface (UI)
- Java
- Spark
- Data modeling
- JavaScript
- C++
- Git
- Big data
- Scala
- Shell Scripting
- Data visualization
- Computer science
- HTML5
- C#
- Python
- PySpark
- angular JS
- Spring MVC
- MVC
- Oracle
- Project management
- Analysis skills
- Communication skills
- DevOps
- SOAP
- Struts MVC
- Linux
- Django
- Data Integration
- React.js
- Tableau
- C
- Attention to detail
- Managing software development teams
- Large-scale user base software development
- Largest user base software developed for (More than 100,000 users)
- Largest software development team managed (More than 20 team members)
- Software development
- Software development experience (16–20 years)
- Algorithms
- Integration testing
- Pair programming
- TypeScript
- Eclipse
- jQuery
- Product demos
- Negotiation
- Kubernetes
- COBOL
- Enterprise architecture
- Splunk
- ZooKeeper
- Systems engineering
- SVN
- Ant
- Front-end development
- T-SQL
- Microsoft SQL Server
- HBase
- UI development
- Entity Framework
- Gradle
- .NET
- Database administration
- MySQL
- J2EE
- JSP
- SSO
- Web services
- XSLT
- Kafka
- Big data projects
- Programming languages
- HDFS
- Microsoft Excel
- Tomcat
- Solution architecture
- Mobile applications
- JMS
- Apache ActiveMQ
- UX
- Contracts
- Cloud infrastructure
- Encryption
- Team development
- Customer service
- Software implementation
- Data structures
- Presentation skills
- Visio
- ETL pipeline development
- Disaster recovery
- Scalability
- Cloud computing
- ASP.NET
- PL/SQL
- PHP
- Management
- Technical support
- Angular
- DB2
- CSS
- Scripting
- HTML
- IT
- Debugging
- REST
- Rally (agile software)
- GitLab
- Test-driven development
- JPA
- Software architecture
- Hibernate



Languages

English – Expert

Additional information

MAJOR PROJECTS

XO Communication (Plano, Texas) (June 2014 - Nov 2014) (Sr. Java Consultant)

XO Communication (Plano, Texas) (June 2014 - Nov 2014) (Sr. Java Consultant)

XO Traffic and Record Analysis (XTRA) is a Web based Java/J2EE application for analyzing and reporting various anomalies related to customer calls and account related activities.

Involved in enhancement and maintenance of the XTRA application. The environment includes Java J2EE, JPA, EJB 3.0, Struts, Spring, REST Web Services, Eclipse, MySQL, JDK 1.7, GIT, JBoss, JUnit, Ant, Maven, Oracle, AJAX, JQuery, CSS etc.

Barclay Bank, (UK & US) (August 2012- June 2014) (Sr Application Developer) Cloud IT

Worked for the Cloud IT application and time to time played different roles from application architecture to code development in both BCloud Web UI (inside ROLB) and backend BCloud Engine development. Multiple times travelled to Radbroke, UK for on-site development and support during R1 development.

Played leading role in R1 last phase of development for defect tracking, bug fixing and implementation of last minutes changes of the BCloud UI and effectively managed the entire Dallas UI team to deliver the product within the defined time frame. Was also involved in R2 R3 releases for Cloud It back end Java web service API layers to support mobile platforms. The environment includes Java J2EE, Struts, Spring, Java Web Services(SOAP), IBM RAD, Eclipse, IntelliJ 12.x, Net Beans, JDK 1.6, Tomcat, JUnit, Ant, Maven, GIT, Oracle, AJAX, JQuery, CSS etc.

GE Corporate (Nov 2011 July 2012) (Java Application Architect Lead Developer)

Was involved in back-end components development and re-factoring of an existing One GE HelpDesk Chat application. The environment includes Java J2EE, Net Beans, Spring MVC, JSON, Apache JCS, Log4J, Hibernate ORM, JQuery, AJAX, MySQL, Tomcat 6.X, Java Web Services, JUnit, SOAP AXIS API, GlassFish 3.1, SubVersion, NetBeans 7.1, SlickGrid, CSS, HTML etc.

Intel Corp. (July 2007 Sept 2011) (Sr. Software Engineer)

Intel Business Exchange (E-Commerce)

E-commerce based on-line global site for buying, Intel CPU based desktops from multiple store sites across the globe. The environment was Java J2EE, Servlet JSP, Comergent third party Java based E-commerce tools etc..

Intel AppUp Store (E-Commerce)

Multi-tier web based application where the application developers use to submit the application which goes through validation and subsequently land up the App Store. The end users also use various application modules to purchase and download the app from the App Store. I was involved in design, development and maintenance of work areas in App Store for Promos, TC Upload, TC Sync, Dev Catalog, Order Email, IDE Web Services, SPP admin, Consumer Registration, writing test cases etc.

The environment includes .ASP.NET, C#, HTML, Java Script, Web Services (SOAP), WCF, JQuery, Slick Grid, JSON, SOAP, .NET 4.0, AJAX, HTML, CSS, MS SQL Server 2008 2010, AspDotNetStoreFront ecommerce web portal, Code Analysis, MS Test, MS Build Server, Visual Studio 2010, Stored Procedures, external Web Services, XML, TFS, PayPal Cyber Cash payment gateway, Automated Test Build, Cloud Computing(Terramark), Globalization and Localization, Agile development practice.

QPS USA (Oct 2005 July 2007) QPS (System Upgrade) (Architect)

ODS, USA (Oct 2005 July 2007) EOS (System Upgrade) (Architect)

EOS (Employer On-line Services) is a web based application used by the employer administrators as well as ODS internal users for new insurance enrolment, request ID card, termination of benefit, change PCP, subscriber search, group information maintenance and other related purposes. Lead a team of 6 people and mentored to enforce using standard Java J2EE patterns in the Web Based J2EE application development projects. The environment includes Core Java, WSAD 5.1.2, FACETS, Web Services,(SOAP), XML, Web Sphere 5.0, J2EE patterns, JSP, JDBC, Struts MVC,Tiles, XML, DB2 , CSS, etc.

DBIC (Sr. J2EE Developer Architect)

Was responsible for re-engineering a Web based (J2EE) application for Dentist insurance. Application was originally developed in offshore and was very poorly designed with the absence of any J2EE patterns and best practices and became a maintenance nightmare. The environment includes WSAD 5.1.2,IBM DB2 Content management, Slide, Tomcat, WebDav, Web Sphere 5.0, EJB Session beans, J2EE patterns, JSP, JDBC, Struts MVC, Sybase etc.

PacifiCorp, USA (July 2005- Oct 2005) ECS (Architect)

Was responsible for prototype development for P8-Office integration project as well as other part of the FileNet P8 web based content management tool. Also responsible for architecting, prototyping the Penagon project to transfer the contents from Rhythmx to the newly developed FileNet P8 content management system. The environment includes WSAD 5.x, Web Sphere 5.0, Struts 1.x, Validator and FileNet P8 JAVA API etc.

Con-Way Transportation (CNF), USA (Oct 2004 June 2005) Wireless Scale Application (Confirm)
(J2EE, Lead Developer Mentor)

From all the Service Centres PRO weight information wirelessly come to data centre for validation and weight correction. Records are verified against the Shipping records and automatically entered into the correction systems for further processing and a weight certificate is dynamically created (TIFF) and stored into the existing imaging system (using WebDav protocol).

Responsible for object modelling, business process mapping and implementing the J2EE best practices Also mentored other team members on J2EE best practices and one-to-one training on Struts and other JAVA related areas. The environment includes WSAD 5.x, Struts 1.1, JAI, Imagelo, TIFF tags, Apache WebDav Library, JUnit TestCases, Log4J, TIBCO, Oracle , JDBC, CVS, Star Team

US Bank (Sr. J2EE Consultant) (March 2004 Oct 2004) On-line loan application, J2EE Store Forward and CAT Maintenance Tool

Worked as a senior level J2EE developer for an on-line loan application. Also worked, as principal architect for developing a generic stored forward mechanism in Java to be used by loan-lines as well as other J2EE applications. I was also responsible for a database maintenance tool for using on Hibernate and Spring Framework.

Responsible for object modelling, business process mapping and implementing the J2EE best practices. The environment includes UML, RUP, WSAD 5.x, Struts 1.1,EJB, Entity Beans, O R mapping, Hibernate, Middlegen, Spring Framework, JUnit TestCases, Log4J, TIBCO, Rational Rose, ClearCase, JDBC, Stored Procedures, SQL Server.

Con-Way Transportation (CNF), USA (April 2001 March 2004) Base Phase 5 Rating System (J2EE Lead

Con-way Transportation (CNT), USA (April 2001-March 2004) Base Phase Pricing System (J2EE, Lead Developer Mentor)

This was a re-engineering project to convert an existing IBM Main Frame based hybrid application to a J2EE application. Currently, AE NAEs submit Requests for Pricing (RFP) by filling out a Microsoft Word form and submitting it to their component Pricing Department via e-mail. A new web-based application, called Agreement Management developed to allow pricing personnel to view, enter and update Tariffs Pricing Agreements on the computer. Once the Tariffs Pricing Agreements are created updated, the Rule Sets, formerly called Rate Aids, will be created automatically.

Started working from day one for this full life-cycle project and performed tasks like use case analysis, object modelling, class diagram etc. Involved in prototype and proof-of-concept development in the very early stage of the development cycle and once the project moved to development, was responsible for the entire middle-tier (EJB and Data Layer) development in populating and maintaining the business components to support the front-end Struts based JSP clients. I also mentored a small team of less experienced java developers.

CRM Application (J2EE)

Responsible for the middle tier business object modelling and development for the in-house CRM application. The environment includes EJB, Web Sphere, Core Java, Session Beans, JSP, UML, XML, XSLT, Xpath, FOP, PDF, JDBC, Together J, JDK 1.3, JUnit, Log4J, Oracle 9i Reports, WSAD 5.1, STRUTS 1.x, Validation, Tiles, JSLT, Tag-Libs, Visual Age 3.5, UDB, DB2, IBM Mainframe S390, COOL: Gen proxy.

Intel, USA (July 2000 April 2001) (Sr Developer) Intel Authentication Service (IAS)

IAS is a multi-tier application, concerned with entity authentication. Sensitive data is first encrypted decrypted in the client side before sending receiving over to the network. Using this service the content providers can protect their contents from hackers attack.

Responsible for solving browser related java virtual machine problems as well as code signing using Microsoft's and sun technologies. The environment includes Java Servlet, Core Java, PKI, XML, COM, Applets, Swing, JDBC, JDK 1.3, ASP, VBScript, JavaScript, Oracle8i, IIS, and JBuilder 3.0.

FDBL and New Content Pipeline application

Performance improvement and enhancement of web based Intranet application for Intel Business Link users. FDBL is a document centric application where a document originated from the factory users, moves to Intel outside customers (DELL, IBM etc). Documents once submitted, managed by Documentum. Server side Java Servlets and client side Java applets are used to develop the entire application.

Responsibilities include performance improvement, bug fixing and enhancement of existing application. The environment includes Servlet, Core Java, JDK, JSP, JDBC, Java Beans, Swing, Documentum, ASP, VBScript, JavaScript, and Oracle 8i, IIS, JBuilder 3.0 and Documentum.

Port Of Portland, USA (March 2000 - July 2000) (Lead Developer) Container Status Enquiry

Web based system for on-line enquiry for the latest status of the container imported exported via the port Intranet development for Port employee. The environment includes ASP, VBScript, JavaScript, SQL Server 6.5 and IIS

SERVERS, and MS.

Navistar International, USA (Aug'98 - Feb 2000) (Lead Developer/Mentor) Real Time Shop Floor Data Collection

From the shop floor data collected by PLC are entered into the UDB ODBC database. Using the UDB trigger mechanism an UDF (user defined function), written in Java, is used to post a multicast message. On the client side, signed java applets are used to grab those multicast messages for real time status display inside a web browser. Responsible for developing the multicast enabled application in Java for both client and server side.

Trucks Metrics Web Reporting and Maintenance.

The environment includes Core Java, Crystal Reports, JDK 1.1, Multicast Messaging, DB2, IIS 4.0, CDO mail, NT Service, ASP, and Visual Age for Java. UDB, UDF in Java, UDB Triggers.

US West, USA (Jan'98 - July '98) Intranet development for Call Handling Administration for USWest

Call Handling is a real-time telephone application where the customer calls are routed via the VRU to the appropriate queue for operator assistance. Responsible for the development of the Web based GUI screens for configuring and viewing the Global Conditions, Account Activities, call status and other activities. The environment includes Unix, Perl5, CGI, CGI.pm, DBI.pm, Oracle, JavaScript, Java, HTML.

R R Donnelley, USA(Sept'97 Dec '97) (E-Commerce) Electronic Commerce site development

Customization and implementation of Microsoft Commerce Server for allowing the users to shop on-line Catalogue items by using their credit cards. The environment was ASP, IIS 3.0, Commerce Server 2.0, SQL Server 6.5, Cyber Cash, TaxWare.

Ikonic Interactive, USA(June '96 - Aug'97) - Multiple Projects

IKONIA (Intranet) development for Ikonic Interactive.

IKMENU ActiveX Control .

Online Last Minute Travel web site development for American Express.

Online Travel Contest for American Express

Online Stock for Pacific Stock Exchange.

Interactive television software development for Classified Content Management (CCM) for Orlando Sentinel

The environment was MS SQL Server 6.5, Active Server Pages (ASP), IIS 3.0, Java Script, Visual C C++ 4.0, VB Script, Perl, Java, and HTML

Lotus (IBM), USA (August '95 -May'96) Porting and development of ccMail Router to NT

Responsible for porting the MS DOS based ccMail router to Windows NT 3.5 environment. The environment is Windows NT 3.5 Visual C C++ 2.0, MASM.

Vantive Corporation, USA (August'95 -Oct'95)

Porting and testing the Vantive products under multiple platforms. The environment was RS 6000, HP

...,writing the entire program using multiple programming languages like C, C++,
9000,SUN workstation, Unix C,ORACLE 7.x,SYBASE 10.

HCL-HP (Oct '94 - June '95) X Protocol Display Server Development in Windows NT

A client server based application development for a X protocol display server in Microsoft Windows NT 3.1 from the MIT sample server code. The NT based server can be used like a pure X server running under any Unix workstation. The NT based X server can handle any unix client as long as they talk in X protocols. Responsible for implementing images with different image formats, like XYPixmap, XYBitmap and ZPixmap and also responsible for implementing GC equivalent attributes in DC(device context). The environment was Windows NT 3.1 SDK, Win Socket, C ,Unix ,XLib,MIT source code for sample server.

Bay Networks, USA (March '94 - July ' 94) Platform Interface of Optivity 5.0 for Novell NMS, Installation Testing

Integrating Optivity NMS 5.0 product with Novell Netware Management Systems by using the Application Integrator for OLF creation and introduction. Code support for a generic launcher for all the Smart Launched Applications. Programming the DLLs for determining the device types and spawning appropriate views. Discovering 28K devices and support for 5K 28K devices in find node utility. Installation and Setup for entire Optivity NMS 5.0 product using Microsoft Setup Toolkit. The environment was Windows 3.1 SDK, Novell NMS SDK, Microsoft Setup Toolkit, C 7.0,AT 486.

AT T , USA (Oct' 93 - Feb '94) TCI Field Access System

This was a Pen Computer based project for a client server-based application. TCI technicians remotely access the customer report data by a Pen Computer and goes to the customer location. They send back the final report to the centralized server. Responsible for developing a windows DLL for the terminal security and also for the data conversion in G2 Format. The environment was Visual Basic 3.0,Unix C, Visual C ++,Pen Computer ,AT 486.

FCI , USA (Jan '93 - Sept '93) Development of a INFOConnect MS Windows application

A client server based application for designing a front end under MS Windows for STAR 1100 which is a Tape Management System running under Unisys 2200 400. Responsible for Development of the Communication between PC and STAR 1100 using IDK, writing the installation program using MS Windows Setup Toolkit, generating SQL statement from various user inputs.

The environment was MS Windows 3.1 SDK,Visual Basic,Window Maker,Code Base 5.0,Boroland C 3.0,INFOConnect Development Kit,AT 486, Unisys 2200 400.

ATC, USA (Oct '91 - Dec'92) Development of Multi Media Kids Activity SW under MS Windows

A window based painting and drawing application specially designed for small kids. There were various different tools for image drawing with rectangles, circles etc and as well supported a number of advanced features like image cutting and trimming. It also includes sound with various application commands. Responsible for development of functional specifications, system design, coding and debugging. The

environment was MS Windows SDK 3.x, Multimedia SoundBlaster, MS C 6.0,DLLs .

ORG SYSTEMS, India (March '91 - Sep '91) Hotel Management Systems

Responsible for the full life-cycle development for a PC Based Hotel Management System to cover all the Front Office activities. The environment was MS C 5.0, Btrieve, ORACLE,PC AT 486.

Indian Customs, India (May '90 - Feb '91) Question Banking System Customer Support to Indian Customs
On line system for customs appraisers were developed for maintaining a database of question and answers. Group of Sperry Systems were networked to provide all India basis information of Bill of Entries. The work involved maintaining the activities on the application, designing new Mapper runs as per the requirements and training the new users. The environment was Mapper, on Sperry 11,CMS 1100, OS 1100.

BIS , India (March '89- April' 90) Publication of Indian Standards Handbook

Indian Standards Data Base Cross Reference ISAM to UNIFY Database Porting Publication of Indian Standards Handbook . Extracted ISI standards information from the mainframe database into PC Ventura based desktop publishing system. A new system was developed to keep track of Cross Reference of the Indian standards. Modification of an IS Standard notifies the bodies who are related to this standard. Involved in system study, requirement specs, coding and user training.