

Abhijit Joshi

Milpitas, CA | abhijitjoshi@gmail.com | +1-201-356-7910 | www.linkedin.com/in/abhijitjoshi | 15+ years professional experience

Education

Bachelor of Engineering, Shivaji University, Kolhapur, Maharashtra, India. July 2005 to Year of passing – May 2009
Educational IT Projects

1. MATLAB project for audio steganography for final year project. Published paper in esteemed IEEE computer society conference.
2. Java project to implement file splitting application and user friendly UI options.
3. Project for providing integrated hostel management facility to maintain Student data, staff data, stock and medicine details, auto-billing, report tracking functions.

Publications

Please check total 22 publications here
<https://orcid.org/0009-0001-7419-6478>

Affiliations

1. **International journal of computer science and information technology research (IJCSITR)**
Editorial board member
2. **International journal of computer science and engineering research and development (IJCSERD)**
Editorial board member
3. **International journal of research in computer applications and information technology (IJRCAIT)**
Editorial board member
4. **International journal of computer applications (IJCA)**
Editorial board member
5. **International journal of information technology and management information systems (IJITMIS)**
Editorial board member

Certifications

1. NSE's Certification in Financial Markets (NCFM) Beginners Module
2. NSE's Certification in Equity Derivatives Beginners Module
3. Six Sigma Certification - Green Belt
4. IBM Certified Developer – DB2 9.7 SQL Procedure Developer
5. IBM Certified Developer – DB2 9.7 Application Developer
6. IBM Certified Developer – DB2 9.7 Fundamentals
7. Informatica university MDM Hub 10x
8. AWS Certified Solutions Architect
9. GCP Certified Professional Cloud Architect
10. Databricks Data Engineering Associate
11. Databricks Machine Learning associate
12. Databricks Data Engineering professional
13. Databricks Machine Learning professional

Professional Experience Summary

Areas/Applications

- Banking and Financial Services
 - Account and statement linking, Document Management, Party Operational DataBase.
 - Wealth management
 - Lending & fintech
- Media Industry
 - Local Media processes
 - Digital Premium subscriptions

Computers Worked on

Computer System	Programming Languages	Special Software
-----------------	-----------------------	------------------

Windows, Unix, Linux, IBM Mainframe Z/OS, MAC	Python, Terraform, Java, SQL, COBOL	Python, Git, Apache Spark, Apache Airflow, Databricks, DBT, Terraform, GitHub actions. IBM Mainframes, JCL, COBOL, CICS, CICS webservices, Informatica Power Center 9.x/8.x/7.x, Teradata V15, V14, FASTLOAD, MLOAD, BTEQ, TPUMP, Informatica MDM HUB, DB2, Tableau, Autosys, Unix, Data warehouse, Data Mining, AWK, SED
---	-------------------------------------	---

Career Profile

Job 1: Tata Consultancy Services --> September 2009 – September 2017

Project NAO – New Account Opening --> September 2009 – April 07, 2013
Client Name Union Bank, San Francisco
Position Systems Engineer
Location Pune, India

Responsibilities

- Develop of new COBOL-CICS-DB2 modules for Web Options project
- Develop WDSL (web service definition language) for interaction between ASP.net and Mainframe CICS-DB2 code.
- Checking for the generation of reports generated from various systems in project and conveying discrepancies if any to Clients.
- Coordinate with Business Users for UAT. This involves interacting with client for UAT
- Maintenance of existing applications by providing enhancements in functionality.
- Monitoring of running Online and batch applications for smooth functionality.

Project Data Warehousing and Integration --> April 08, 2013 to September 15th, 2017

Client Name Morgan Stanley
Position Architect (Computer Systems Engineers / Architects)
Location New York City, USA

Responsibilities

- Strategically architect, design, develop, and implement efficient information systems and/or operations systems in support of core enterprise functions.
- Understand business requirements, study existing application landscape, and identify redundant/ineffective systems.
- Conceptualize technical solutions to complex problems and maximize benefit of IT systems investments.
- Work closely with clients to gain organizational commitment for all systems and software plans, as well as evaluate and guide the selection of technologies required to complete those plans.
- Coordinate with offshore development team(s) to identify priorities and update scope and delivery schedule.
- Create data migration/system integration strategies.
- Monitor system performance to detect and resolve problems during deployment and support change management.

Carried out the following activities

- Designed the ETL processes using Informatica PowerCenter to load data from Oracle, Flat Files, and Mainframe files to target Oracle Data Warehouse project for MSSB (Morgan Stanley Smith Barney) joint venture.
- Extensively used Parameter file to override Mapping parameter, Mapping Variables, Workflow Variables, Session Parameters, FTP Session Parameters and Source-Target Application Connection parameters.
- Wrote Shell scripts and Stored Procedures for regular Maintenance and Production Support to load the warehouse in regular intervals and to perform Pre/Post Session Actions.
- Created various Documents such as Source-To-Target Data Mapping Document, Unit Test Cases and Data Migration Document
- Loading the OLTP data into the Warehouse OLAP and built the Re-Usable Mappings using Informatica PowerCenter..
- Implemented SCD methodology including Type 1, Type 2 to keep track of historical data..
- Coordinate with Business Users for UAT. This involves interacting with client for UAT.
- Support production implementation, deployments and validations

Job 2: CBS Corporate Inc. --> September 2017– July 2021

Project Enterprise Data Management and Data Warehousing --> September 18, 2017 to March 20, 2020
Position Application Developer
Location New York City, USA - September 18, 2017 to March 20, 2020
Responsibilities

- Understand business requirements, study existing application landscape, and identify redundant/ineffective systems.
- Conceptualize technical solutions to complex problems and maximize benefit of IT systems investments.
- Work closely with clients to gain organizational commitment for all systems and software plans, as well as evaluate and guide the selection of technologies required to complete those plans.
- Coordinate with offshore development team(s) to identify priorities and update scope and delivery schedule.
- Create data migration/system integration strategies.
- Monitor system performance to detect and resolve problems during deployment and support change management.

Carried out the following activities

- Developed and deployed ETL job workflow with reliable error/exception handling and rollback framework in information technology domain. Served as the single resource responsible for construction of all objects necessary within Data warehouse solutions and/or analytics platforms. Developed interfaces, data conversion programs, and application extensions.
- Actively involved in business requirements and technical requirements documentation review. Prepared high and low-level design documentation, converted it into solutions by developing, unit testing, provided break/fix support of Informatica ETL and wrote re-usable unit test documents.
- Managed automation of file processing, as well as all ETL processes within a job workflow. Ensured data quality throughout entire ETL process.
- Interacted with other project team members responsible for developing reports, interfaces, data conversion programs, and application extensions.
- Presented technical information in easily understood terms in both verbal and written form. Worked closely with flow IT staff and project management to document standards and procedures for the definition of reference architectures (architecture roadmap).
- He/She utilized skills including:

ETL tools, such as Informatica Power Center, Teradata packages, stored procedures, and triggers; Oracle SQL (DML & DDL), PL/SQL, and packages and full software development lifecycle. Information technology; performance tuning, query plan/explain plan analysis, indexing, and table partitioning; and version control/source code management tools.

- Automate processes to ensure timely delivery of data and take appropriate actions for delays and outages.
- Code bug fixes for existing applications.
- Monitor the production alerts. Fix the production issues.
- Maintenance of existing applications by providing enhancements in functionality.
- Monitoring of running Online and batch applications for smooth functionality.
- Providing timely analysis to the higher management regarding on-going changes in applications, its risk factors and its cost effectiveness.
- Maintaining knowledge flow regarding on-going system components to team members.
- Integrating the code deliveries to Production from various applications.

Project: Data Engineering and Solution Architecture with GCP and Apache Airflow --> April 1, 2020 to July 2021

Position: Senior Data Engineer

Location: New York City, USA - April 1, 2020 to July 2021

Responsibilities:

- Translate business requirements into data solutions, optimizing or decommissioning inefficient systems.
- Design data migration and system integration strategies.
- Coordinate with offshore teams to manage project scope and delivery.
- Develop ETL workflows using Apache Airflow, ensuring data quality.
- Utilize Databricks and DBT for data modeling and transformation.
- Present technical information in an accessible manner to various stakeholders.
- Automate processes for timely data delivery and troubleshooting.

Carried out the following activities

- Developed and deployed robust ETL workflows, including interfaces and data conversion programs.
- Drafted and implemented detailed design documentation and unit test documents.
- Collaborated with project team members on reports, interfaces, and application extensions.
- Used skills in GCP technologies, Databricks, DBT, SQL, and version control.
- Monitored production alerts, troubleshoot issues, and enhanced application functionality.
- Provided analysis to upper management on application changes, risks, and cost-effectiveness.
- Maintained knowledge flow about ongoing system components within the team.

Job 3: Oportun Inc. --> August 2021 – Till date

Project	Minerva2.0 Data Platform --> August 2021 – Till date
Position	Staff Data Engineer
Location	Remote, Milpitas California - August 2021 – Till date

- Managed and architected solutions on Databricks, Airflow, DBT, and AWS, optimizing platforms and code to achieve significant cost savings: \$100K monthly on Databricks and \$30K monthly on AWS costs related to Databricks services, resulted in 50% savings.
- Employed Terraform for efficient infrastructure management and automation, contributing to substantial cost efficiencies.
- Led 20 person cross-functional team of Data engineers, DataOps and Database admins in data and infrastructure initiatives, focusing on cost-effective solutions and performance optimization.
- Implemented MLOps and CI/CD methodologies to streamline project delivery, further reducing operational costs.
- Created, deployed, and decommissioned secure, scalable cloud instances using IaaS concepts in a DevOps environment. Practiced test-driven development with CI/CD processes, resulting in robust, scalable systems and rapid cloud infrastructure provisioning with notable cost reductions.

Trainings Conducted

Year	Title	Presented to	Location	Organised by
2011	Session on using EASYTRIVE tool for file sorting.	NAO team members	Pune, India	Internal
2012	Session on SOAP UI for CICS Web Services	NAO team members	New York City, NY-USA	Internal
2015	Session on implementation of Data Warehouses	PADT Area	TCS	PADT
2016	Session on Document Management data warehouse	PADT Area	New York City, NY-USA	Internal
2018	Airflow data pipelines using Python	CBS IT Team	New York City, NY-USA	CBS IT Team
2020	Google Cloud platform management	CBS IT Team	New York City, NY-USA	CBS IT Team
2022	AWS Cloud platform management	Oportun DataOps Team	San Carlos, CA	DataOps
2023	Databricks infrastructure management through Terraform	Oportun DataOps Team	San Carlos, CA	DataOps

Trainings Attended

Year	Title	Location	Organised by
2010	System Z: VSAM-E1-Training	Pune	TCS
2011	Session on IT security	Pune	TCS
2012	Session on Time Management, Communication Skills, Team Building	Pune	TCS
2013	1. Session DB2 2. Informatica PowerCenter 3. Teradata	Pune	TCS
2014	Session on IPMS, IQMS, ICALMS	Pune	TCS
2015	Training on fundamentals of Capital markets – Options and Equity trading	Pune	TCS
2016	Infosec Training: Social Engineering and How to Avoid Becoming a Victim	New York, USA	Morgan Stanley
2016	Banking and Capital Markets – Quiz	New York, USA	TCS
2016	Informatica MDM Hub 10 certification	New York, USA	Morgan Stanley

2018	Google Cloud Platform Essentials for machine learning	New York, USA	Google
2019	Google Cloud Platform BigQuery Fundamentals	New York, USA	Google
2021	Databricks Fundamentals	San Fransisco, CA	Databricks
2022	Databricks Unity Catalog	San Fransisco, CA	Databricks
2023	AWS solution architecture	Remote	AWS