

# Abhijit Das

D103, Rohan Avriti, 4th Cross,  
ITPL Main Road, Krishnarajapuram  
Bangalore, Karnataka

## Career Summary

Results-Focused professional, offering an experience of over 14+ years (**6+ years into HADOOP & 2+ year into AWS**) in the **IT Sector**. Possessing excellent analytical skills, summed up with proficiency in **Data Lake, Data Engineering and Data Analytics using AWS, Java, Python, Hadoop, Hive, and Spark**. Testing Implementation techniques and host level applications teamed with quality management control.

**Practical exposure of working in Agile methodology** and various phases of SDLC. Hands-on expertise of working in team building, **Enterprise Platform Applications and Integration Tools**.

Innovative and exceptionally creative, to foster ideas that impel the organization towards a result-oriented direction. Impeccable Client Management and leadership skills with impressive professional-speaking abilities. Driven by new challenges and adept at adapting to any cultural and business environments.

### Career Objective

Dedicated and disciplined Software professional, by contributing to the company's growth and in turn ensuring personal growth within the organization. I believe that my technical, functional and communication skills will enable me in facing the challenging career ahead

### Core Competencies

- **AWS, EC2, EMR, IAM, Glue, DMS, Lambda, Java, Python, Hive, RBAC, Spark, Snowflake and Redshift**
- **Leading a team of 5 - 10 developers**
- **Troubleshooting and Bug Fixing**
- **Client Relationship Management**

## Professional Experience

*McKinsey & Company (Serving Notice Period)*

**September' 2021 - Till Date**

*AspireNXT (P) Ltd.*

**April' 2020 – September' 2021**

*Zaloni Technologies India (P) Ltd.*

**December' 2010 – March' 2020**

*System Consultant Information India (P). Ltd*

**August' 2008 – December' 2010**

## Key Responsibilities & Achievements

- My role is of **Data Architect** and to design/review cloud based big data solutions using AWS
- Understand and convert business requirements into architecture and design.
- Create documentation and maintain it throughout the release life cycle.
- **Maintain good quality code and code review using GitHub, Crucible and Sonar**
- Helping product owner / Business Analyst with the requirements, by preparing **USE CASE diagram, TDD, LLD and FLOW DIAGRAM**

- Helping project manager with the estimates using Project management tool like Open Project and Jira
- Maintaining requirement, architecture and flow document
- Demo the feature to the stakeholders as per **Agile Methodology**

### Achievement

- Achieved the award "*Start of the Month*" from Zaloni Technologies, for the month of February' 2013. This award is given to the best performer of the organization every month.
- Achieved the award "*Spot Award*" from Zaloni Technologies, for the month of May' 2015 and April' 2017. This award is given to the best performer across specific team of the organization every month

### Technical Skills

| Category              | Software / Tool / Framework                          |
|-----------------------|--|
| Programming Languages | Java, Scala, Python                                  |
| Framework             | MapReduce, Sqoop, Flume, Spark                       |
| Databases             | Oracle, MySQL, PostgreSQL, H2, Hive                  |
| Cloud Services        | IAM, Glue, Lambda, DMS, EMR, Redshift, S3, Snowflake |

### Projects Completed

|   |  |
|---|--|
| <b>Project Title</b>                      | ETL Tool and Data Migration  |
| <b>Duration of Work</b>                   | January-2022 to till date  |
| <b>Role</b>                               | Lead Data Engineer   |
| <b>Software / Tools / Technology Used</b> | Airflow, Snaplogic, Glue, Oracle, Snowflake, DMS, S3, IAM Roles, Lamda, PySpark. |

**Project Description:** This ETL migration is designed to migrate Snaplogic and Oracle data warehouse into AWS environment using AWS Services like IAM, DMS, Glue and Lamda. Generic re-usable Patterns are being developed to migrate 1000+ snaplogic pipelines across the enterprise to AWS, either using Glue or DMS. Data from Oracle is getting migrated into Snowflake. Data transformation is done using PySpark code in Glue Jobs and quality of the data is checked via the lambda code.

|                                       |  |
|---------------------------------------|--|
| <b>Project Title</b>                  | Oracle Data Migration  |
| <b>Duration of Work</b>               | July-2020 to September-2021  |
| <b>Role</b>                           | Data Architect   |
| <b>Software/Tools/Technology Used</b> | AWS DMS, S3, Glue, Lambda, SNS Notification, AWS IAM, Redshift (Spectrum), Spark |

**Project Description:** This data migration pipeline is designed to migrate the client data from legacy data source to Redshift using DMS, Glue and Labmda. Data analytics is done using QuickSight on top of the flatten data in Redshift.

Data transformation is done using PySpark code in Glue Jobs and quality of the data is checked via the lambda code.

|                                       |   |
|---------------------------------------|---|
| <b>Project Title/Name</b>             | <b>RG Data Hub Hydration</b>  |
| <b>Duration of Work</b>               | August-2019 to March-2020   |
| <b>Role</b>                           | Module Lead   |
| <b>Software/Tools/Technology Used</b> | Core Java, Hive, EMR, EC2, AWS IAM, S3 and project management tool (Smartsheet) |

**Project Description:** Data Hub Hydration is a tool to ingest conventional data to data lake and do the necessary transformation on those data for analysis. Secure information is tokenized and finally the data is stored in trusted zone.

|                                       |   |
|---------------------------------------|---|
| <b>Project Title/Name</b>             | <b>TMX Enterprise Data Platform Management</b>                                  |
| <b>Duration of Work</b>               | January-2018 to July-2019   |
| <b>Role</b>                           | Module Lead   |
| <b>Software/Tools/Technology Used</b> | Core Java, Hive, EMR, EC2, AWS IAM, S3 and project management tool (Smartsheet) |

**Project Description:** EDP (Enterprise Data Platform) is a tool to ingest conventional data stored in tape(s) and disk to data lake and do the necessary transformation on those data for analysis.

|                                       |   |
|---------------------------------------|---|
| <b>Project Title/Name</b>             | <b>MICA - Data Analytics Tool (Product)</b>   |
| <b>Duration of Work</b>               | October-2015 to December-2017   |
| <b>Role</b>                           | Module Lead   |
| <b>Software/Tools/Technology Used</b> | Core Java, Liquibase, <u>MySQL / H2 / PostGreSql</u> , <b>Hive, MapR, Spark, Sqoop</b> and Project Management Tool (Open Project) |

**Project Description:** Mica aims to be an application that leverages Data Management Tool to Browse, Annotate, Transform and Visualize data in the big-data lake. Its core users are data scientists, analysts, architects and data governance teams. Mica will leverage and extend the core modules of Data Management Tool but will be a separate application.

|                                       |  |
|---------------------------------------|--|
| <b>Project Title/Name</b>             | <b>PCCM</b>                                      |
| <b>Duration of Work</b>               | Jan-2015 to September-2015                       |
| <b>Role</b>                           | Development Lead                                 |
| <b>Software/Tools/Technology Used</b> | Core Java, <b>Hive, MapR</b> and Eclipse IDE 4.2 |

|                           |                      |
|---------------------------|----------------------|
| <b>Project Title/Name</b> | <b>UDW Offload</b>   |
| <b>Duration of Work</b>   | Jan-2014 to Nov-2014 |

◆ +91-9707409925 ◆ [abhijitcaps@gmail.com](mailto:abhijitcaps@gmail.com) ◆

**Role** Java Developer using Hive and Hadoop(Pivotal)  
**Software/Tools/Technology Used** Core Java, Hive, Pivotal and Eclipse IDE 4.2  
**Project Title/ Name** **Zaloni Data Platform (ZDP -> Arena)**  
**Duration of Work** August-2012 to December-2013  
**Role** Java Developer using Spring and Hibernate  
**Software/ Tools/ Technology Used** Core Java, Spring, Hibernate, HTML, MySQL 5.x and Eclipse IDE 4.2

**Project Title/ Name** **ECOnirmaan Prescriptive Tool**  
**Duration of Work** August-2011 to July-2012  
**Role** Java Developer using Hibernate  
**Software/ Tools/ Technology Used** Core Java, Flex 4, Hibernate, MySQL, Eclipse IDE 3.5 and IReport

### Education

---

**Bachelor of Engineering - Computer Science** **2008**  
*Sri Siddhartha Institute of Technology, Tumkur (VTU)*

**Post Graduate Diploma in Business Administration - Operations** **2013**  
*Symbiosis Centre for Distance Learning, Pune.*

### Personal Details

---

**Date of Birth** 04/06/1986  
**Languages Known** English, Hindi, Assamese and Bengali  
**Nationality** Indian  
**Passport Details** Passport Number - R2184618, Valid till year 2027  
**Personal Interests** Sports, Music and Reading Assamese Novel