Akshat Saxena

E-Mail : akshat.saxena95@gmail.com
Phone: +91- 7566330280,7987333078
Location Preference: Pune

Targeting Technical role in IT sector in
~ Big Data Engineer ~ System Implementation ~

*Result-oriented professional with more than **4.3 years** of rich experience in **Big Data Engineering** and **ETL Development** using Informatica; scaling new heights of success with key focus on profitability & optimal utilization of resources*

SKILL SET

Big Data Solution Development
New System Implementation
Software Development Lifecycle
Project Execution
Unit Testing
RCA & Troubleshooting
Liaison & Coordination
Strong Interpersonal Skills
Reporting & Documentation

ACADEMIC   DETAILS

☐   BE Electronics and Communication

PROFILE SUMMARY

- Hands on experience Hortonworks/Cloudera Hadoop Components. Working as a Hadoop Developer and Administrator.
- Proficient in implementing complex ETLs and creating web services using Informatica Real Time, Informatica BDM and Nifi Ingestion Pipelines.
- Working knowledge of Microsoft PowerApps, Microsoft Power Automate/Logic Apps and Microsoft Sharepoint.
- Hands on experience with Informatica Data Catalog and Tibco Data Virtualization.
- Visualization using Kibana, Logstash and ElasticSearch
- Creating Proof of Concepts from scratch by illustrating how big data techniques can meet specific business requirements reducing cost and time to market.
- Proficient in interacting with users for requirement gathering and preparing functional specifications and low-level design documents.
- Managed Data platform upgrade initiative to accommodate new features, bug fixes etc. (Cloudera HDP/HDF, Informatica BDM, Informatica PowerCenter, Elastic Search etc.)
- Experience in Software Development Lifecycle beginning from requirement gathering, prototyping, programming, testing to implementation of projects in Agile framework.
- Hands-on different languages used as a part of Integration requirement like Python, Shell Scripts.

IT SKILLS

Hortonworks Components Knowledge: Apache Nifi, HDFS, Hive, Yarn, Ambari, Kafka, Hbase, Pyspark
ETL Tools: Informatica Real Time, Informatica BDE (10.4), Informatica Power Center (9.6.1 & 10.2.0)
Microsoft Tools: Power Apps, Power Automate/Logic Apps and Sharepoint
Data Governance: Informatica Data Catalog, Apache Atlas, Apache Ranger
Data Virtualization: Tibco Data Virtualization
Data Visualization   Kibana, Logstash, Elastic Search
Database: Oracle, SQL Server, PostgreSQL
Languages: Python, PLSQL, Shell
IDE: Eclipse, Putty
Operating Systems: Linux, UNIX, Windows (XP/07/08)

ORGANIZATIONAL EXPERIENCE

AMDOCS: 2016 –Present (MIS DataPlatform Team)

Highlights:

- Onsite experience at Tel Aviv, Israel for requirement analysis and Tool Procurement Finalization.
- Udemy Certified Pyspark and Hadoop Developer.
- Hortonworks certified Hadoop Administrator.
- Awarded EOQ for FY 19-20.
- Promoted to One level of Seniority as "System Analyst" and then to "Senior Software Developer".
- Mentoring the freshers/lateral joining the organization along with their code reviews, setting best practices etc.

ACADEMIC DETAILS

QUALIFICATION	YOP	INSTITUTE	BOARD/UNIVERSITY	PERCENTAGE
B.E.(EC)	2016	LNCT Bhopal	R.G.P.V. Bhopal	89.10%
12th	2012	SJCS,Bhopal	C.B.S.E.	92%
10th	2010	SJCS,Bhopal	C.B.S.E.	85.5%

PERSONAL DETAILS

Date of Birth: 05/01/1995
Address: D104 Daffodils Magarpatta city Pune
Languages Known: English and Hindi

***********End of Resume***********

(Refer annexure for project details)

ANNEXURE ORGANIZATIONAL PROJECTS

Project:	**Datalake Project**
Sub Project:	Generic Data Ingestion from the sources like SQL server, Oracle, Files, API, Hbase etc.
Exposure:	Nifi, Atlas, Python.
Description:	1) Created a generic flow with Nifi to ingest data from various sources into datalake propogating from different zones in the lake architecture. Also, incorporated Ranger policies later. Incorporated various CR to handle file ingestion process. 2) Implemented ORC design for hive tables via Nifi. Create a Rest API which is used as a backend for Watson Analytics to ingest the data to datalake.
Sub Project: **Exposure:**	**S3 Bucket Integration with Lake using Nifi** AWS S3, Nifi, Python, Samba
Description:	Using python and boto3, generated temp credentials for AWS and integrate it with Nifi Using python and samba client, pushed the data to file servers
Project:	**Data Virtualization using Tibco**
Sub Project:	Internal Mobility Project and One Source of Truth
	1) Create a virtualized REST API to provide integration from Success Factor to Service Now. 2) Building a one source of Truth/Data to be consumed cross Amdocs.
Project:	**Data Cataloging using Informatica DataCatalog**
	1) Onboarded different resources like Hive, Tableau, SAP, Informatica to Catalog. **2)** Create a synchronization process from DataCatalog to Atlas and vice a versa. **3)** Enable SSO with Azure AD for Catalog.
Project:	**DataIngestion Reconciliation Process**
Exposure:	Microsoft PowerApps, PowerFlow , Sharepoint, Graph, Nifi
Description	We are creating an automated platform using PowerApps whose backend will be Sharepoint List. It is an approval flow for any new data source to be a part of DataLake. It will be primarily used by Business Analyst team for Requirement creation, then by Information Security to validate Risk and Data Analyst team for Ingestion. Sharepoint is integrated with Lake using Nifi and Microsoft Graph API. Python is used to create schemas for the tables.
Project:	**DataIngestion Reconciliation Process**
Exposure:	Microsoft PowerApps, PowerFlow
Description	Create an App which will be used by higher management for sending NDA agreements while acquiring companies. It provides the complete details about the companies acquired, participants details, Sending Notification and terminations.
Project:	**DataPaltform Tool Analysis**
Sub Project:	Participated in various POC before procuring any new tool and technology.
Exposure:	Informatica Data Catalog, Tibco Data Virtualization, Cloudera Data Analytics Studio, Rundeck
Sub Project: **Exposure:** **Description:**	**Hadoop Administration** All Hadoop components. **1)** Working as an Administrator to manage the clusters (POC, DEV and PRODUCTION). **2)** Setup complete GR environment with HDP 2.6.5 version. **3)** Upgraded HDP environment from HDP 2.6.5. to HDP 3.0.1 and then to HDP 3.1.4 **4)** Setup HDF 3.2 env and upgraded till 3.5.1
SubProject: **Exposure:** **Description:**	**Attendance Data** Informatica BDM,Python Ingesting data for employees in out for around 120 sites (both database and file systems) into datalake with Informatica BDM so as to prepare dashboard.

SubProject:	**Datalake Phases**
Exposure:	Informatica BDM, Sqoop, Oozie, Nifi, Yarn, Ambari, Hive, HDFS
Description:	1) Written sqoop scripts for different type of datasources (SQL server, oracle etc.).
	2) Scheduling of sqoop workflows using Oozie.
	3) Ingested data using Nifi from sources like (SAP HANA, oracle, Files, sql server) and provided Nifi services for ingesting data.
	4) Created SQL based authentication on Hive tables.
	5) Knowledge of Ambari Administration.
	6) Nifi Flow Designing
	7) Rest API development using Nifi
	8) Migration of environments from POC to dev to Production.
SubProject:	**Engage**
Exposure:	Informatica Powercenter and Informatica Realtime
Description:	1) Created MDM for employee data from different sources of data and handled different CR's on top of it
	2) Created data sources for creating Rest services on top of the data
	3) Created SOAP web services on top of MDM data for the consumers.
	4) Created ETLs for different consumer in order to provide Data Integration.
SubProject:	**xSite, Joopy, Compass, IS-QC Data Integration**
Exposure:	Informatica ETLs
Description:	Data Integration for different sources involving different databases like MS-SQL, SQL server, Oracle with transformation logics in between.