

Neeraj Somani

Business Data Integration Analyst

Personal Info

Address:

Manhattan, New York

Phone:

+1-2242120651

Email:

Neeraj.somani4@outlook.com

LinkedIn:

www.linkedin.com/in/neerajsomani4

Visa Status:

H1-B

Education and Certifications

Illinois State University:

Aug 2011–Dec 2013

MS Information Technology

Certifications:

Oracle Java SE 8 – OCA certified

Sep – 2018

AWS – Certified Developer Associates

Oct 2018 – Oct 2020

Tech Stack and Skill set

RDBMS: Oracle, MS SQL Server, MySQL

Query Engines: Hive, Impala, DynamoDB

Programming Languages:

Java (certified), PL/SQL (proficient)

Python (learning), Scala (learning)

Big-data Frameworks: Hadoop, Spark

ETL Tools: Informatica, Ab-initio

BI Tools: POWER BI, Tableau (Learning)

Scheduling Tools: Control-M, Tivoli

File Systems: HDFS, AWS - S3

File Formats: JSON, Avro, Parquet, CSV, txt

Strengths

Quick Learning

Precision

Leadership

Communication

Documentation

Personal Learning Experience

- Learned and Certified in AWS cloud serverless architecture framework.
- Practiced Integration of various AWS services like, API Gateway, RedShift, SQS, EC2, CloudFormation etc to build highly scalable, reliable, cost efficient system.

- Installed Cloudera CDH distribution and learned big-data ETL pipeline process using Hadoop/Spark frameworks, HIVE/Impala SQLs, Scala/Python language.
- Processed high volume data in file formats like parquet, Avro, CSV, JSON, and flat file using SparkRDD and DataFrame.
- Understood benefits of advanced Spark APIs like map, flatMap, reduceByKey, groupByKey, AggregateByKey.

Passionate Business Data Integration Analyst with certified knowledge in Java SE 8 (OCA), cloud (AWS-CDA) and big-data technologies having 5+ years of experience in IT industry.

Currently leading a team of 4 Data Analyst under onsite/offshore model while meeting 100% quality deliverable's following AGILE methodologies on strict timelines. Working towards a position where I can assure in building and maintaining highly scalable, robust and cost efficient data driven systems using cutting edge technologies.

Professional Work Experience

Syntel Inc. (Atos | Syntel)

May-2014 - Present

Client III - Moody's Analytics

Mar-2017 - Present

Business Data Integration Analyst

- Building data lake for "Revenue Recognition" reporting project by understanding overall Enterprise Data Warehouse (EDW) architecture and business rules.
- Driving and Designing 10+ ETL data pipelines using Informatica Power Centre data Integration tool to integrate cloud based REST API webservices oriented architecture.
- Collaborating with Business, functional, architect & other technical teams in all phases of SDLC (Software Development life cycle) to build efficient data processing solution.
- Documenting each approach decisions in LLD/HLD (Low/High-level design docs).
- Optimizing the process by distributing the design in small components, like Extract, Parse, Split, Transform, Validate, Load, Archive, Error Handling, and Email Notification.
- Delivering batch process Integrations across sources like (SFDC) salesforce web-services, Leeyo RevPro cloud API, Oracle DB, PeopleSoft, and various file formats.
- Developing JAVA code base using Informatica transformation tools to handle RESTful web API request/response for HTTP methods like GET, PUT, POST, DELETE, etc.
- Working on another project called "Data-Conversion", a complete ETL solution using, MS Excel, PL/SQL, SSIS packages, and SFDC Data Loader utility.
- Writing Stored Procedures to format data as per SFDC objects and generate CSV files to upload to Salesforce using Data Loader utility.
- Writing Linux/Unix shell scripts to perform file transfer and manipulation work.

Client II - QBE Insurance

April-2016 – Feb 2017

Data Quality Analyst | BI Specialist

- Generated Standard Lines monthly financial reports using tools like Microsoft Power BI, Power Pivot, VB Scripting, and Microsoft Office Suite (Excel, PowerPoint, Word).
- Consolidated 2 years' of monthly reports for various Business units and management units to analyse the forecast calculations for current financial year.
- Designed, Developed and Validated MS Excel sheet reports by arranging data in tables, graphs, pivots and calculated using functions like V-Lookups, H-Lookups.
- Researched and Analysed large data sets to extract actionable insights that would help drive business decisions at financial matrix details with SQL scripting.
- Communicated statistical analysis and data values to key stockholders and monthly-presented meaningful insights by using MS PowerPoint.

Client I - Allstate Insurance

May-2014 – Mar 2016

ETL Developer

- Supported EDW with thorough understanding in various domains of insurance industry like, policies, claims, quotes and Agency related data and its complete data flow.
- Analysed business user needs in order to document functional and technical specifications and designed MDM ETL process flow for various insurance domains.
- Developed, tested and reviewed complex Ab-Initio graphs, DML, Pset, XFR, and unix shell scripts for ETL process and act as Team Lead for Onsite/Offshore team model.
- Worked with Data Profiling Team to help in Data Extraction, Cleansing/Validation process. Data sources like oracle databases, flat files, XML, JSON are majorly used.
- Designed ETL process to support 8-way multi-file partition system and parallelism concepts with extensive use of ab-initio Components (Partition by Key, by Expression, by round Robin) to partition the large amount of data file into small multiple data files.
- Involved in creating JSON files from Oracle DB and loading files to HDFS cluster.
- Analysed large data sets by running Hive queries and supported MapReduce ETL jobs running on the Hadoop cluster.
- Defined Test Methodologies, created Test Plans and Test Cases based on negative and positive scenarios for generating sample test reports before releasing to production.

Champro Sports Inc

Oct-2013 – April 2014

Application Developer

- Designed and developed complex Informatica mappings/workflows by using transformations like Lookup, Expression, Update, Sequence generator, Aggregator, and Router to implement complex logics in ETL design.
- Interacted with SMEs to gather all business requirements based on scope of project to co-ordinate the development, test and production activity for E2E project life cycle.