

## Harshit Tripathi

Ph.: +919035130434

E-mail: [ht1992@gmail.com](mailto:ht1992@gmail.com)

Skills: Python | PySpark | Azure | Databricks | UNIX Shell Scripting | ETL | Oracle | SQLDW

Software Developer with an experience of 5 years in designing end to end data integrations, ETL architecture, business rules & transformations, data/dimensional modelling of Operational Data Store. Extensive knowledge of health insurance, pharmaceutical and financial services industry. Basic knowledge of big data technologies like Hadoop, Spark, Databricks and Hive. Highly Skilled to work in agile environment and adapt the change in requirements quickly.

### JOB SUMMARY

**Designation** Associate Data Engineering L2  
**Organization** Publicis Sapient  
**Duration** December 2019 – Till Date

**Designation** Software Engineer  
**Organization** Tata Consultancy Services  
**Duration** November 2015 – December 2019

### TECHNICAL SKILLS

**Operating Systems:** Windows, UNIX  
**RDBMS:** Oracle 11g/12g, Microsoft SQL Server  
**ETL Tools:** Informatica Power Center 9.1/9.6,10.1,10.2 HF, Informatica Cloud (ICS and IICS), Informatica Developer Big Data Management (BDM) Tool  
**Language:** Python, PySpark, SQL, UNIX Shell Scripting  
**Scheduling Tool:** CA workload automation, Autosys, Oozie, Airflow  
**Ticketing Tools:** Service Now, Jira  
**Other Tools:** Putty, Secure CRT, Toad, Azure  
**Quality Control:** HP ALM 11.0, Bitbucket, XLdeploy, Git

### RESPONSIBILITIES:

- Working as technical back up lead on project involving 6-8 team members
- Engaged in end-to-end project implementation activities which includes requirement gathering, analysis, design, construction, testing and production deployment activities based on Agile Methodology
- Interaction with client/Onsite SMEs to gather the requirements, providing estimate of the effort hours and getting sign off for the deployment request
- Training junior team members on new technologies required as part of project activities

### PROJECTS:

- **Automated SQLDW Data Migration [Tools and s/w: Python, Unix Shell Script]**
  - Created Python and Unix Scripts to migrate SQL Data Warehouse (SQLDW) data from Production to UAT environment.
  - Implemented data migration for both partition and non-partitioned SQLDW tables.
  - Saved manual and monitoring effort of cloud migration team.
- **Automated writing back of data from SQLDW to HIVE [Tools and s/w: Python, PySpark, Unix Shell Script]**
  - Created Python script to retrieve data from SQLDW and store the same on Azure Data Lake Storage (ADLS) layer.
  - Implemented PySpark script to check the partitioning of hive table and backup existing data of the table.
  - Dumped data from ADLS to existing location of hive table.
- **Cloudera to Databricks Migration Project [Tools and s/w: AWS, Azure Cloud, Python, Unix][Ongoing Project- 6 months]**
  - Migrating existing Cloudera resources such as YARN, Oozie, HDFS to Databricks, Airflow and ADLS environments respectively.
  - Integrating existing Informatica BDM mappings with Databricks Clusters.

- **GDPR Automation Project [Tools and s/w: UNIX shell Script]**
  - Automated manual execution of extracting and updating PI data to null values through Linux scripts. Saved manual effort of executing data change requests fortnightly.
- **Medical Inquiry data exchange Process [Tools and s/w: ICRT, UNIX shell Script, Postman]**
  - Designed interface to push and pull data from REST API of third-party system to create a real time Medical Inquiries exchange process using Informatica Cloud Real Time (ICRT).
  - Implemented end to end functionality to populate UTC time zone fields which by default is converted to EST time zones due to informatica server mounting.
  - Recreated interface to invoke API's through Informatica on premise
  - Invoked API's through Unix shell script by capturing session id's that were generated during run time
- **Global Commercial Operations [Activity System]**
  - Core team member engaged in production support and major enhancement activities.
  - Designed ETL maps/logic in Informatica Power Center, Informatica Cloud and Informatica Developer BDM including creation of UNIX shell Scripts, database objects, writing SQL queries.
  - Performing production support activities involving service management using Service Now including incident management, change management, problem management and impact analysis

## EDUCATIONAL QUALIFICATION

<b>Bachelor's Degree</b>	Bachelor's in Electrical and Electronics Engineering [Manipal University, Manipal, Karnataka]
<b>CGPA</b>	8.20
<b>Date</b>	June 2011- June 2015

## ACHIEVEMENTS:

- **Account Recognition Award (March 2020):** Awarded for establishing UAT environment for Cloud Migration Project.
- **On the Spot Award (February 2019):** Awarded for successfully implementing IICs Migration Project.
- **On the Spot Award (January 2019):** Awarded for successfully implementing GDPR Automation Project. Also involved in implementation of Above and Beyond Idea which was appreciated by customer
- **Certificate of Appreciation (February 2016):** Awarded for being part of Empower, TCS initiative program to enhance communication skills of TCS support staff

## PERSONAL VITAE:

Languages Known:	English, Hindi
Interests:	Swimming, Music, Driving