

# ARCHANA WAGHWASE

Data Engineer

Mumbai, Maharashtra | [linkedin.com/in/archana-waghware-b43aab160](https://www.linkedin.com/in/archana-waghware-b43aab160)

| waghware1996@gmail.com | 9321011504

## PROFILE

Results-oriented Data Engineer specializing in addressing complex challenges within big data frameworks. Proficient in building big data solutions using various frameworks and cloud computing technologies.

## WORK EXPERIENCE

### RELIANCE JIO

Data Engineer

Sep 2021 - Current

- **Ingestion and KPI's:** Ingested diverse data types from multiple sources into Jio's big data lake using big data frameworks, and developed KPIs to enhance customer experience and help the business make data-driven decisions.
- **Project Parivartan:** Developed Spark-Scala code to capture Jio Setup Box and Fiber related events, storing the data into Hive tables. Established an Oracle view on top of these Hive tables, serving as the backend for the MyJio app. This integration empowers customers to self-assess Internet connectivity, Wi-Fi coverage, STB device health, and remote connectivity, enabling real-time issue resolution through DIY assistance. Furthermore, customers can schedule engineer visits and track them via Google Maps, offering convenient engagement anytime, anywhere through the MyJio app.
- **RFH Ingestion:** Developed Ingestion Pipelines to Anonymize PI information present in the DICOM Images. Developed ICU LOS prediction, ICU data ingestion, non-VIP patient data ingestion pipelines, Reports Anonymization, and HL7 message parsing pipeline using bigdata technologies like Kafka, NiFi, PySpark, Hive, and HDFS.
- **Research and POC:** Assessed Hive's ability to generate external tables across different data sources. Successfully established external Hive tables connecting to various platforms, including relational databases like PostgreSQL, MYSQL, ORACLE and streaming services such as Kafka.
- **JFS:** Working on the JFS project for Jio Finance Services, utilizing Azure HDInsight to manage and streamline data ingestion across various verticals. This involves implementing efficient big data solutions to enhance data processing and analytics capabilities.
- **CDP Migration:** Migrated Ingestion and KPI flows from HDP to CDP Environment.
- **Team Collaboration, Recruitment and Mentoring:** Conducted Fresher's interviews and mentored them to be part of JIO Data Platforms Group.

## LTI

Data Engineer

Sep 2019 - Sep 2021

- **AI-EXIT:** Converted ETL flows and components of Ab Initio into Spark Scala for various products in the financial area as part of the AI-EXIT framework.
- **Data Reconciliation using PySpark:** Developed code for data reconciliation to verify data during data migration using PySpark to generate reports on the basis of verification.
- **Celonis ETL POC:** Designed an ETL pipeline to read JSON files and generate equivalent jobs in Spark using Java. These jobs dynamically create SQL queries comprising ETL components such as joins and filters. Results are then stored in a table or saved as a CSV file, intended to be mapped as a data model in Celonis.

## EDUCATION

GOVERNMENT COLLEGE OF ENGINEERING, JALGAON

B. TECH(E&TC) (2014-2018)

CGPA – 8.3

## SKILLS

- Apache Spark
- Apache Hive
- Hdfs
- Apache Kafka
- Spark-Scala
- Pyspark
- Apache NiFi
- Elastic Search Index
- GCP (Google Cloud Storage, Dataproc, Pub-Sub, BigQuery)
- Azure (HDInsight, Azure Synapse, Azure Data Factory)

## AWARDS

- Secured 3rd position in Hackathon (Hack the future competition held by LTI in partnership with IBM).
- Spot Award by Jio for contributing towards the company's growth.

## LANGUAGES

- English
- Hindi
- Marathi