



**MADHUMITHA S R**

Senior Data Engineer

(BigData,Spark,ETL,RDBMS,Azure,SQL)

**(12+ years of experience )**

**Certified SAFe 6 Product Owner/ Product Manager**

Email: [mitha.m2@gmail.com](mailto:mitha.m2@gmail.com)

Address: Chennai, India

Phone: +91 9443307422

DOB: 02/05/1990 | Nationality: Indian

LinkedIn: [Madhumitha S R | LinkedIn](#)

Patent Grant with USPTO: 1

## JOB SKILLS

- Adept professional with extensive experience in **Data warehousing, Python, Spark, ETL & Azure**
- Domain expertise in Banking specialized in **Wealth & Investment Management technology** front for **Consumers & Small Business** optimizing data processes that serve millions of users
- Collaborated with **Agile scrum** cross-functional teams to ensure on-time project deliverables
- Spearheaded **SPARK3 upgrade** program, achieving a **15%** reduction in operational costs annually
- Directed the **Vertica to Hadoop** migration, resulting in a **30%** reduction in overall processing time
- Mentored junior engineers in **Data quality design** approach for implementing **100+** automated checks using **pyspark** framework and ensuring **99% data accuracy** across production datasets.
- Driving Operational Excellence for cross functional team for performance improvement/cost reduction by creating **Autosys JIL validator tool** resulted in save of **1500+** man hours
- Crafted pyspark frameworks enhancing data accessibility saving over **200 man-hours** annually
- Designed **architecture** for the proposal in **Datastage v11.3** with Statement of Work(SOW)
- Facilitated **15+** Design Thinking sessions resulting in the filing of over **20+ patents** with the **USPTO**
- Holds a **Patent Grant** with USPTO and **filed 7 patents** so far which are pending to be granted
- Led technical recruitment across **5+** data engineering domains, hiring **20+** skilled professionals

## TECH. SKILLS

Programming Languages	Python(data processing and ETL workflows), Pyspark, Unix scripting
Big Data Technologies	Apache Spark(batch processing with PySpark, Spark SQL) ,HDFS ,Hive, Impala
Data Warehousing & ETL	Azure Synapse Analytics, Datastage v7.5, v8.1, v9.1, v11.3, v11.5, v11.7
Relational Databases	MySQL, PostgreSQL, Oracle, MS SQL Server, DB2, Teradata
Cloud	Azure
Version Control/ Repo/	GitHub Actions CI/CD, Bit bucket, SVN
Tools & Frameworks	JIRA, PySpark, Data Integration Frameworks
Schedulers	Autosys, Control-M
Software	Microsoft suits
Project Management	SAFe 6 POPM, Technical Documentation, Client Communication

## WORK EXPERIENCE

*July 2019 till date*

### Senior Data Engineer

**Bank Of America Continuum : Chennai, India**

- **Established** and **mentored** a team of **7-members**, overseeing end-to-end project delivery
- Designed & developed rigorous **Data quality checks** assuring **100% unit testing** coverage
- Crafted more than **20+** reusable **spark sql** scripts with streamlined design patterns
- Authored High/Low level Design doc, Data mapping document for DataLake pipelines
- Contributed to **spark3 upgrade** program apart from regular responsibilities & completed a full implementation within a short timeframe of **2 months**
- Formulated performance tuning methodologies for **Datastage** jobs and slashed **50%** of production runtime and significantly **lowering server loads** resulting in improved SLAs
- Assisted **CDP to CDH upgrade** for the application providing support & devised **Sqoop to Hadoop** ingestion framework as part of **Vertica to Hadoop Migration** Effort
- Operational excellence partner for the team enabling performance improvement and cost saving practices for the team across various locations achieved Delivery Head target YoY

### Achievements:

- Recognized as **Delivery Head Innovation champion**; Proposed novel ideas on **Rewards Balance Prediction algorithm** in Global Innovation events gaining support from leadership
- Honored as a **Top Women Inventor in India** for **filing 3 patents** with USPTO on a single idea

Aug 2018 - Jun 2019

**Module Lead (Data Strategy & Insights)**

**L&T InfoTech Ltd: Chennai, India**

- Module Lead for a team of 5, overseeing the migration from **Cognos Data Manager to Datastage v11.5 on Unix OS** achieving **100% on-schedule delivery** and zero critical issues post-migration
- **Architected Solution Proposals** for migration and managed a cross-functional offshore team, improving delivery efficiency by **20%** through Agile Scrum methodologies
- Authored Datastage Design & Mapping Documentation for a high-stakes project with over **10** key stakeholders, meeting strict deadlines and achieving **100%** on-time delivery for all project phases
- Validated data with **100+ complex Oracle & PostgreSQL queries** to maintain **99.9%** data accuracy
- Enhanced **Datastage** job performance, ensuring **100% warning-free** execution & reducing system maintenance time by **15%**

**Achievements:**

- Reduced batch runtime from 3 hours to 20 minutes by implementing a data retention utility

Dec 2015 - Nov 2017

**Design Lead (Data Insights)**

**HCL Technologies: Chennai, India**

- **Design Lead** working within the **Waterfall model** for enhancing **Risk Data Warehouse**
- Led Requirement Gathering from **5+ key stakeholders**, translating needs into actionable plans
- **Architected** Business Solutions in **Datastage v11.3** for project requirements and deliverables
- Proposed **Data Mapping & Calculations** for employee/center-level data using **Oracle**, ensuring accurate financial forecasting and data integrity across multiple business units

**Achievements:**

- Implemented a **Test Validation Framework** using **Oracle queries** to ensure data accuracy
- **Self-identified an audit issue** on current day's batch run, implementing an automation that prevented data loss & saved ~ **1,000 man-hours per year**

Nov 2011 - Nov 2015

**Developer/Module lead**

**Tata Consultancy Services, India**

- Led Migration of Datastage v7.5 to v8.1, automating data validation across multiple platforms using **generic scripts**, reducing runtime by **50%** & improved data processing efficiency
- Developed Automation Scripts to manage the deletion of old datasets based on retention periods, reducing SLA time by 2 hours and enhancing data lifecycle management

**Achievements:**

- Recognized as a **"Problem Solver"** for restoring critical applications during an outage

**EDUCATION**

May 2007- Jul 2011

Bachelor of Engineering- Electrical and Electronics Engineering | Anna University Coimbatore

**LANGUAGES**

English – Fluent ; German – B1 in progress

**CERTIFICATIONS**

- Certified IBM Solution Developer – Infosphere Datastage v8.5
- Certified SAFe® 6 Product Owner/ Product Manager
- Microsoft Azure DP203 – In progress

**PATENTS GRANT**

[US Patent for System and method for processing and intergrating real-time environment instances into virtual reality live streams Patent \(Patent # 11,948,259 issued April 2, 2024\) - Justia Patents Search](#)

**PATENTS FILED**

**7 Patents filed** with USPTO & pending to be granted

**PERSONAL SKILLS**

- Empathic listener skilled in understanding team dynamics, **influencing & lead** by example
- Successfully **led** and managed **cross-functional teams** with diverse skill sets
- Regularly interact with **top management**, providing strategic updates to stakeholders