**YOGAMANIKANDAN GURUSAMY Ph:** (M) +61-451985415

**E-mail:** psgyogahcl@gmail.com

**PROFESSIONAL SUMMARY**

A seasoned IT professional with over 14 years of experience in the software development life cycle utilizing Agile methodology. Skilled in designing and implementing data pipeline systems for big data applications using Hadoop, Spark, and AWS. Proficient in Snowflake, Redshift, and real-time data processing pipelines using Apache Kafka and Spark Streaming. Demonstrated expertise in cloud engineering, data services, and CI/CD tools.

**Technical Proficiency**

* **Languages/Technologies:** Python, Scala, Spark, SQL, Unix Shell scripting
* **Big Data Ecosystems:** Hadoop, HDFS, MapReduce, Apache Spark, Hive, Sqoop, Kafka, Oozie
* **Cloud Platforms:** AWS (EC2, S3, RDS, Lambda, CloudFormation, Redshift, ECS, ECR)
* **Data Modeling:** Dimensional data modeling (Kimball methodology), ER modeling, Star schema, Snowflake schema, Data Vault
* **Databases:** Teradata, Oracle, SQL Server, MongoDB, Hive, Snowflake, BigQuery
* **ETL Tools/Frameworks:** Informatica, DataStage, Teradata Control Framework
* **Data Analysis:** Time series analysis, statistical modeling, data visualization
* **Data Visualization Tools:** Tableau, PowerBi
* **Methodologies:** Agile, Scrum, Kanban
* **Version Control:** GIT
* **Orchestration Tools:** Airflow
* **CI/CD Tools:** Buildkite, TeamCity, Jenkins, Bamboo
* **Data Governance Tools:** Alation, Collibra

**WORK EXPERIENCE:**

1. **Data Engineer - Latitude Financial Services, Melbourne (July 2023 – Present)**

* **Project**: Vision Ingestion Transformation
* **Technologies**: PySpark, Snowflake, DBT, Python, AWS, Airflow, BuildKite
* **Responsibilities**:
  + Elicited technical requirements for data pipelines and processing by collaborating with stakeholders.
  + Built and deployed a big data processing solution on AWS using ECS, ECR, S3, and Glue PySpark.
  + Created tables and defined access roles in Snowflake for data storage.
  + Developed and optimized DAGs in Apache Airflow for complex data ingestion workflows.
  + Implemented comprehensive error handling and retry mechanisms.
  + Designed and deployed AWS Lambda functions for automated data transformation.
  + Engineered dynamic scheduling and conditional task executions within Airflow to enhance pipeline reliability and operational efficiency.
  + Designed and deployed containerized applications using Docker for consistent development and scalable operations.

1. **Senior Data Engineer - NAB, Sydney (April 2023 – June 2023)**

* **Project**: ADA – Source Readiness and Ingestion Team
* **Technologies**: Spark, Redshift, S3, Python, AWS, Airflow, Databricks, HVR
* **Responsibilities**:
  + Participated in requirement gathering and analysis working with stakeholders.
  + Designed and implemented a big data processing pipeline using AWS EC2 instances, S3, Hadoop, and Spark.
  + Configured HVR to enable the HVR replication tool to pick files from the source landing location.
  + Created data pipelines in Databricks to load data from HVR staging to Bronze layer tables.
  + Scheduled ingestion data flows in HVR either on event basis or time basis.

1. **Senior Data Engineer - Macquarie Bank, Sydney (July 2021 – April 2023)**

* **Project**: Tax Transformation
* **Technologies**: Spark, Hive, Cloudera, Python, AWS, Airflow, Bamboo, GCP Services, BigQuery
* **Responsibilities**:
  + Participated in requirement gathering and analysis working with stakeholders.
  + Designed and implemented a big data processing pipeline using AWS EC2 instances, S3, Hadoop, and Spark.
  + Configured ELB to distribute incoming traffic across multiple instances in multiple availability zones for improved performance and fault tolerance.
  + Created and managed Auto Scaling groups to automatically scale instances up or down based on demand.
  + Developed Python scripts to connect to source systems using REST APIs.
  + Conducted performance tuning and optimization of SQL queries.

1. **Assistant Consultant / Senior Data Engineer - CBA (TATA Consultancy Services), Bangalore and Sydney (May 2017 – July 2021)**

* **Project**: Data Transformation Program (DTP)
* **Technologies**: Spark, Hive, AWS, Sqoop, Cloudera, Python, Scala, Teradata, TeamCity, Tableau
* **Responsibilities**:
  + Designed and implemented scalable ETL pipelines using Scala scripts and Spark's Data Frames API.
  + Leveraged Kafka for real-time data processing.
  + Managed large datasets using Spark's in-memory capabilities.
  + Designed and implemented cloud-based solutions on AWS.
  + Developed and maintained ETL pipelines using Scala and Spark.
  + Designed and implemented data warehousing solutions using Snowflake.
  + Conducted performance tuning and optimization of SQL queries.
  + Developed Python scripts for data integration and processing.

1. **Senior Associate - Walmart (Cognizant Technology Solutions), Chennai (March 2016 - May 2017)**

* **Project**: DCL RMS (Disney Cruise Line Revenue Management System)
* **Technologies**: Python, Informatica, Teradata, Tableau
* **Responsibilities**:
  + Spearheaded the design and development of ETL pipelines.
  + Created Data Warehouse design and performed data modeling.
  + Optimized data loading and transformation performance.
  + Developed and maintained automation tools using Python.
  + Implemented data quality framework to identify and remove data anomalies.
  + Optimized ETL pipeline performance using techniques such as Hive partitioning and bucketing.
  + Managed a team of 8 developers and provided training on Teradata and HIVE operations.

1. **Senior Software Engineer - CBA (HCL Technologies), Sydney/Chennai (June 2009 - March 2016)**

* **Project**: Core Banking Modernization (CBM) / BAU
* **Technologies**: Teradata, DataStage
* **Responsibilities**:
  + Designed and developed Core Banking Modernization (CBM) ODS.
  + Managed data warehouse design, data modeling, and development.
  + Integrated SAP to different systems.
  + Developed automation tools using Python and Teradata SQL.
  + Successfully modeled solutions for various Home Loan system incidents/issues.
  + Played a Data designer role in enhancing Account Tax data and GCPS data in GDW space.
  + Developed an automation tool using Python and Teradata SQL to identify data anomalies in the Cards System.

**Academic profile:**

* Bachelor of Engineering from **PSG College of Technologies, Anna University** in the year of 2008

**ACHIEVEMENTS AND AWARDS:**

* - Received Quarterly Team Awards - Q1'2019 for the Best Team Collaborator.
* - Awarded “Thanks a Zillion” from TATA Consultancy Services for outstanding project contributions.
* - Innovation Champ Award (Cognizant), Q3 & Q4 (2016-2017).
* - “Leading Lights” award for excellent contribution at Commonwealth Bank of Australia.
* - Appreciated by TATA Consultancy Services for resolving design flaws in different projects.
* - Special mention for managerial capabilities and problem-solving skills during Client Visits.
* - Widely appreciated for adding value to the business unit, saving around 1 million AUD annually.
* - Delivered critical projects by playing roles such as Lead Data Engineer, Solution Designer, and Technical Business Analyst.

# PERSONAL INFORMATION:

Name : Yogamanikandan Gurusamy

Location : Sydney, Australia

Citizenship : India

Language Known : English, Tamil

**Present Address**

Unit 4, 49-53 Essington Street,

Wentworth Ville, Sydney,

NSW – 2145.