

Daniel Paes

Profile:

Performing at the Information Technology market for 10+ years with large consulting experience Daniel acquired remarkable skills at both technical and leadership aspects. Big Data professional with solid analytical skills and strong technical background on data pipelines within batch, microbatch (near real-time) and streaming data, with devops experience. Constantly chasing advancement, Daniel's objective is to bring solutions to a continuous evolution of the business data architecture and environment.

Work Experience

National Bank of Canada: March 2018 - Present

Data Engineer

- Building ETL solutions using spark on scala Hadoop environment (Hortonworks HDP);
- Generating ETL using spark and Spark Streaming to ingesting semi structured and structured data;
- PoC on behavioral analysis using sentiment analysis;
- PoC of data pipelines with the usage of microservices.

Technologies used: MariaDB, Scala, Spark, Spark-Streaming, HDP 2.5, Python 2.7, textblob, Jenkins, Jira, Hive, Hadoop, Docker, Nexus, Kafka, Agile (Scrum and Kanban), Shell-Script, Unix environment.

Videotron Canada, Montreal: May 2017 – March 2018

Senior Developer Technical Lead

- Optimization and creation of PL/SQL and Datastage jobs for data integration;
- Model all DWH stages (Staging and ODS areas, EDV on 3NF, and OLAP and physical models);
- Estimating data storage and efforts to define big data implementation environment;
- Implemented PoCs of data consumption flows using scala and spark;
- Analysed non structured data on sentiment analysis using nlkt.

Technologies used: Netezza, Scala, Spark, PLSQL, HDP 2.5, SAS Studio, Business Objects, Python 2.7, NLTK, Oracle 10g, Data warehouse modeling Kimball approach, Agile (Scrum), Shell-Script, Unix environment.

Reitmans Canada, Montreal: May 2016 – May 2017

Data Engineer

- Created and maintained ETL processes within Informatica Powercenter to extract data using java UDFs to create connectivity drivers from MongoDB, Google Analytics and Big Query to integrate with existing data on Oracle databases;
- Analysed data using pyspark and pandas on datalab and loaded them into Google Cloud SQL and Google Big Query using python bigquery and mysql packages;
- Managed preemptive and compute engines on Google Cloud Platform.

Technologies used: Informatica 9.5.1, Java, Scala, Spark, MongoDB, Tableau Desktop, Python, scikit, NLTK, Oracle 10g, Data warehouse modeling Kimball approach, Agile (Scrum), Shell-Script, Unix environment. Microstrategy 9, Google Cloud Platform, Google Analytics

EXIA, Saint-Hyacinthe: November 2014 - May 2016

On-site Consultant – ETL Developer

- Developed Informatica components using Velocity framework, Java UDF transformations and PL/SQL packages under Agile Management concepts;
- Designed data warehouse model based on data vault and Inmon modeling methodologies.

Technologies used: Tableau Desktop, Control-M Scheduler, Informatica 9.5.1, Oracle 10g, flat-files, Shell-Script, Unix server, Java, Data warehouse modeling Data vault approach, Agile (Scrum)

Informatica, Rio de Janeiro: Sept 2011 - Nov 2014

On-site Consultant – Data Integration Specialist

- Delivered Quality Assurance on the data, creating templates for data validation, and on the code creating user guides documentations and development templates;
- Analyzed, defined and delivered performance tuning time solutions, which were able to optimized existing Informatica mappings, increasing performance and reducing processing time up to 50%;
- Provided internal training to technical users on Informatica development and coordinated a small development team.

Technologies used: Informatica 9.5.1, Control-M Scheduler, Informatica Data Quality, Oracle 10g, flat-files, Shell-Script, Unix server, Data warehouse modeling Kimball approach, business object

Accenture, Rio de Janeiro: Jan 2010 - Sept 2011

On-site Consultant – Senior ETL Developer

- Responsible for optimization tuning and restructuration of existing Informatica processes;
- Planned, designed and created functional documentation to integrate web services data on Oracle database using Informatica 8.5.1 following Velocity methodology;
- Unix, informatica's metadata and Control-M administration;
- Provided internal training on ETL data warehousing model, and hands-on Informatica training.

Technologies used: Informatica 9.5.1, Control-M Scheduler, Teradata 10 BTEQ, T-Pump, Fast-Load, Multi-Load, Fast-

Value Team, Rio de Janeiro: Jan 2006 - Dec 2009

On-site Consultant – BI Developer

- Analyzed existing oracle production code to generate new Informatica ETL processes;
- Changed Business Objects and Microstrategy data models to reflect the creation of new attributes, metrics and hierarchies;
- Migrated On line Transactional Process reports into On line Analytical Process dashboards in Microstragy Architect;
- Created technical mapping specification and process possible troubleshooting actions.

Technologies used: Informatica 7.5.1, Oracle 9g, Business Objects 6.5, flat-files, Data warehouse modeling Inmon approach, Shell-Script, Unix server, microstrategy 7

PROFESSIONAL DEVELOPMENT

- Data Engineering on Google Cloud Platform specialization
- Agile Project Management (April 2015 – 4 hours): Concepts, roles and responsibilities, key issues, vocabulary, basic techniques.
- Datawarehouse modeling (November 2008 – 8 hours): Basic concepts of datawarehouse modeling showing differences between Inmon and Kimball approaches. ETL life cycle, ODS/Staging concepts, types of dimensions and facts, difference between surrogate keys, business keys and primary keys, modelization of data Marts and Datawarehouse.
- ITIL V2 Foundation (2008): Certified on ITIL disciplines Service Support, Service Deliver, ICT infrastructure Management, Security management, Application management and Software Asset Management and planning to implement service management.

EDUCATION

- Francisation (2016 – 2017): French communication, Listening and Pronunciation, grammar, vocabulary and Conversation
- EC English School (2014 – 2014): IELTS Preparation, English for business
- Alliance Française (2010 – 2012): French communication, Listening and Pronunciation, grammar, vocabulary and Conversation
- Estácio de Sá University (2004 – 2008): Bachelor Degree, Information Technology