Abhijit Paul
Senior Business Intelligence Engineer

Cell: +1 910-250-9709

Email: abhijitpaul123@gmail.com

LinkedIn: www.linkedin.com/in/abhijitpaulus

P R OFE S S I O N AL S UM M A R Y

- Over 15 years' experience Designing, Developing and Implementing Data and Analytics Solutions for building Data Lakes, Big Data, Data Warehousing and Risk/Regulatory Reporting.
- Over 15 years' experience in Informatica PowerCenter 7.x - 10.x, involved in building Data Warehouses, Data Marts, ETL Controls and Reports.
- Experience with various SDLC processes including Waterfall, Agile, Kanban methodologies. • Expertise in data warehouse methodologies (Ralph Kimball, Bill Inmon), ODS, EDW.
- Experience in Banking, Insurance and Healthcare domains with ability to quickly adapt to new systems. • Hands-on design/development experience in sourcing/loading data from/to distributed, cloud, legacy systems. • Experienced in integration of various data sources and relational databases Oracle, DB2, Teradata, SQL Server. • Worked on integrating data from flat files including fixed-width and delimited. Good experience working with Excel and COBOL/VSAM files using Informatica and PowerExchange.
- Worked on Teradata utility scripts FastLoad and MultiLoad to load data from source systems to Teradata and creating BTEQ (Basic Teradata Query) scripts to generate Keys.
- Experience creating OLAP Reporting Layer on top of warehouses and creating various reports, covering the full Business Intelligence Cycle.
- Current experience creating On-Prem and Cloud Data Lakes for feeding analytics and Machine Learning/AI systems using AWS Glue, Python, PySpark, Hive and related ecosystems, as well as Informatica PowerCenter, IICS / Informatica Cloud, BDE/BDM and PowerExchange suites and ETL tools like SnapLogic and Pentaho.
- Skilled in Analysis, Requirements, Designing, Developing and Implementing ETL and reporting solutions using Data Warehousing tools Informatica Cloud (IICS), Informatica PowerCenter, Informatica PowerExchange, Informatica Change Data Capture, Big Data using Informatica BDE/BDM, AWS Glue, PySpark, Python, Hive and related ecosystem, Pentaho BI Suite and Oracle Discoverer reports.
- Good knowledge of data modeling concepts, Star Schema/Snowflake modeling, and Facts and Dimension tables for Conceptual, Logical and Physical data modeling.
- Understanding of NoSQL concepts, working knowledge of ML/AI systems using classification and regression modeling, and knowledge of Keras and Tensor Flows.
- Well acquainted with Informatica Designer Components - Source Analyzer, Warehouse Designer, Transformation Developer, Mapplet and Mapping Designer. Worked extensively with complex mappings using multiple transformations like Source Qualifiers, Expressions, Filters, Joiners, Routers, Union, Unconnected / Connected Lookups and Aggregators, Stored Procedures, Rank and Normalizer.
- Strong experience using mapping parameters and variables, session and workflow level parameters and variables, and dynamic generation of parameter files.
- Experience in debugging mappings, identification of bugs in existing mappings by analyzing the dataflow and evaluating transformations.
- Strong understanding of Session logs to identify the performance parameters and errors. • Extensively involved in performance tuning of Informatica jobs and fixing bottlenecks using tuning methods such as session partitioning, reducing lookup sizes, and using persistent caches.
- Good experience documenting ETL process flows for better maintenance of code and analyzing process flows.

• Expertise using SQL and using procedures, functions and packages.
• Experience in Unix shell scripting, CRON, FTP and file management in Unix environments. • Experience in job automation using enterprise schedulers Autosys, CA7, code versioning and migration of code to higher environments using Harvest, ClearCase, Subversion.
• Good experience in using defect tracking tools such as HP Quality Center.
• Expertise in Release Management using version tools and Change Management procedures. • Strong leadership and mentoring skills. Responsible for quality and meeting project objectives. • Proficient working independently and in teams, including onshore-offshore models.

T E C H N IC AL S K IL L S

Operating Systems	Unix, DOS, Windows
Languages	SQL
Tools	Informatica PowerCenter 7.x - 10.4.x
	Informatica IICS
	Informatica Change Data Capture
	Informatica PowerExchange
	Hadoop with Informatica BDE/BDM and PowerCenter
	AWS Glue and PySpark
	Pentaho ETL and Reports (Open Source)
	Epic Clarity ETL
	Oracle Discoverer Reports
	Toad, SQL Explorer
	Autosys and CA7 Schedulers
	Harvest, ClearCase, PVCS Change Management, Jira, Ambari
RDBMS	Oracle 8i/9i/10g, Teradata, SQL Server, DB2, MySQL
Cloud	AWS, Salesforce

P R OFE S S I O N AL E X P E R IE NC E

Cognizant Technology Solutions | *April 2020 - Present*
Client: Truist Bank – Greenville, NC
Senior Informatica / ETL Developer Consultant
• Worked on Informatica Intelligent Cloud (IICS) , Informatica PowerCenter, PowerExchange and Salesforce integration.
• Worked on Merger Project between SunTrust and BB&T Bank to deal with the full Enterprise ETL process merge. • Contributor and Senior Developer. Used many tools such as Relational Junction, SnapLogic, CopyStorm and others, to decide on best enterprise tool for maximum efficiency.

Sun Life Financials – Toronto, Ontario CAN | *October 2015 - April 2020*
Senior Informatica Developer / ETL Designer (multiple projects)
Project: Cloud Data Lake

Data from various legacy sources including CIF, OASIS, and Global, were transformed and loaded into an AWS Cloud data lake for applications to access and for reporting.

- Designed file processing and handling of sources, laid down the file creation rules for source systems, designed file validation, rejection and notification process, managed processing of mainframe VSAM files, and interfaced with various data sources in Oracle and DB2.
  - Target database was in Hive and subsequent consumption databases were SQL Server, Oracle, PostgreSQL and others. Designed the process to run as a batch, loaded the data in target in the background, and made data available only on completion of the batch so that the batch run time did not interfere with user work.
- Developed mappings to accomplish above tasks and coordinated delivery by working with offshore teams and other team members.
- Used AWS Glue and PySpark to load data into the cloud.

Project: On-Prem Data Lake
- Individual policies were stored in Sunlife IBM IMS mainframe. Led effort to ingest the data in an on-prem Hadoop Data Lake for analytics.
- Before personal involvement on project, SunLife was using IBM Data Accelerator to extract the data from IMS, which was very costly. Proposed using PowerExchange based on previous experience, as Sunlife already had Informatica as Enterprise ETL.
- In developing a Proof of Concept (POC), was able to extract 3 largest IMS segments with 90% lesser CPU cost, resulting in at minimum 40% savings for the company, including licensing costs.
  - Once data was sent to the lake using the Enterprise ingestion framework, designed process to move data between different Hive layers in PowerCenter BDE/BDM on Hortonworks Hadoop environment.

Project: LICAT
- LICAT was a regulatory project and large initiative. As a regulatory project, it was critical that reports were error free. Designed Excel file processing and error handling activity to ensure there were no errors in the reports. • Involved in designing a process where a batch ran every hour and automatically processed the files that were ready and notified of errors, if any, so that business could correct and send within next hour, and the cycle continued without any logical abends.
- PowerCenter was the ETL tool used, and involved multiple sources and targets.

Project: US Compares
- SunLife had a presence in Canada, USA and Asia. Each month, on a few legacy systems, there were balancing reports that were coded in COBOL. The use of mainframe CPU time was costly for the organization, so the reports were to be converted to PowerCenter.
- As Developer, it was a very challenging task, as COBOL is a procedural language. Incorporating the same logic in PowerCenter is very complex. Hurdles and challenges were overcome and was able to successfully implement the process.

Cognizant Technology Solutions | *February 2012 - October 2015*
Lead Informatica ETL Developer Consultant (multiple projects)

Client: Lincoln Financial Group – Hartford, CT | *November 2012 - October 2015*
Informatica ETL Developer Lead - Unclaimed Property ODS

Lincoln Financial Group is a leading life insurance company in the US. The Unclaimed Property project initiative was to find beneficiaries of unclaimed policies and to escheat the funds to the state, in case beneficiaries were not located. The application interacted with state auditors, exchanged data between 20 legacy systems and third-party business partners, and generated due diligence reports. There were control totals for each data exchange process and balancing was done to automatically generate 'out of balance' notifications.
- Led team of 11 resources, 3 in Hartford, CT and 8 offshore in India.
- Responsible for the ETL design, code development, code promotion and code delivery until code passed UAT. •

Coordinated with BA and Testing teams to resolve defects.
• Worked with DBA and Server Administration teams, Mainframe Support and other Infrastructure teams to request resources and get ETL code promoted and running from end-to-end as a coordinator of the application. • Developed ETL using Informatica 9.1 with SQL Server as the main target database. Read data from multiple sources such as flat files, Oracle, SQL Server databases. The OD was in SQL Server.

- Automated the application as much as possible. Daily jobs were 100 % automated, trying to remove manual dependency on monthly and quarterly jobs.
- Responsibilities included designing ETL solution, estimating target implementation dates, estimating time; ensuring timely delivery of components, and code promotion.
- Managed performance tuning of long running jobs.

Client: Liberty Mutual – Dover, NH | *March 2012 - November 2012*
Informatica ETL Developer / ODS Data Warehouse

- Application was 100% automated loading daily/monthly data, taking care of holidays and dependencies. • Responsible for designing ETL solution, estimating target implementation dates, estimating time, ensuring timely delivery of components, and code promotion.
- Analyzed requirements and prepared low-level ETL design.
- Developed ETL using Informatica 9.1 with Teradata as main target database. Read data from multiple sources. • Conducted performance tuning of long running jobs.

Tata Consulting Services | *March 2004 - February 2012*
Senior Informatica Consultant (multiple projects)

Client: PNC Bank – Cleveland, OH | *August 2010 - February 2012*
Informatica Technical Project Lead

- Worked as Project Leader for Risk projects in Informatica Center of Excellence (COE).
- Managed team of 14 people, 3 onsite and 11 offshore associates.
- Responsible for designing ETL solution, estimating target implementation dates, estimating time; assigning tasks to the team, ensuring timely delivery of components, and code promotion.
- Improved 70% load time by modifying design, using persistent cache for large lookups, removing non-indexed queries, and using Multi-Threading, designing partition and enable parallel loads.
- Automated daily and monthly jobs, involving 10 sources (VSAM, COBOL, Flat Files and Relational), designed stage and fact load jobs, dimension refresh and aggregates.
- Worked with Cognos reporting team to derive ETL solution based on reporting requirements, specially indexes and aggregates.

- Developed synergy between onsite/offshore teams by distributing tasks in a way that ensured continuous development activities, thereby reducing delivery times.
  - Automated application to manage holidays, dependencies and reduce manual intervention as much as possible.

Client: Experian USA | *July 2009 - August 2010*
ETL Designer Consultant (remote India)

- Designed system end-to-end, from tool selection to report delivery.
- Led a team of 5 associates and implemented the project in 1 year.
- Mapped AS-IS and TO-BE systems, analyzed tools required to deliver the requirements. • Suggested out-of-the-box free BI tool versus existing tools and saved license costs for the client. • Prepared architecture where 57000 users uploaded mappings and data files that were loaded using ETL. ETL gathered logs and sent PDF reports to stakeholders using Pentaho classes.
- Gathered and understood requirements from BRD and prepared Technical Design Documents. •

Performed data modelling for reports.

• Designed ETL to bring data from flat file to relational database.

• Designed ETL to bring data from relational database to dimensional and summary tables. • Designed reports and developed report queries.

• Integrated reports with the WebLogic application server.

• Performed debugging and worked with QA Analysts to debug ETL and reports.

Client: Kaiser Permanente - Pleasanton, CA | *January 2006 - June 2009*
ETL Administrator / Module Lead Consultant

• Gathered requirements from clients on tables required for reporting, and loaded on demand. • Installed, maintained, upgraded EPIC Clarity ETL environment, created/modified scheduling jobs. • Monitored/investigated errors that caused Extract or Load failures. Researched and resolved issues with ETL processes or the resulting data.

• Analyzed, maintained and resolved data integrity issues.

• Acted as primary liaison between reporting team, A&I, DBA team, Business team and EPIC software vendor, in addressing specific problems.

• Developed/maintained strong professional and technical relationships with assigned clients. Became the internal expert for the client's data requirements.

• Provided project management, particularly focused on meeting critical timelines, project communication, and hands-on implementation for assigned projects..

• Worked with ETL Developers and QA Analysts to create reliable and dependable ETL processes. • Researched/resolved issues with ETL processes or resulting data. Managed new console setup. • Automated workflows to minimize babysitting activities for job monitoring.

Client: Sunlife Financials - Toronto, ON, Canada | *October 2004 - December 2005*
ETL Developer Consultant (remote India)

• Responsible for ETL development and reports development on client project.

• Analysis of client business process, designed reports, warehouses meeting specifications. • Managed design/development of ETL to load data from Feeders to Oracle GL, AP and PO. • Managed design/development of ETL to load data from Oracle GL, AP, PO to database.

• Managed design and development of ETL to load data from Oracle General Ledger, AP, PO and Dimensional Database to mainframe.

• Designed Discoverer reports from the dimensional database.

• Designed Discoverer Reports from GL, AP, PO to reconcile/verify ETL data loads in dimension. • Created end-user layer (Universe) for reporting.

• Performed unit and integration testing, bug fixing and performance tuning.

Client: McGraw Hill Education – New York, NY | *March 2004 - October 2004*
ETL Developer

• Performed analysis of existing business process of client and prototyping of the new system. • Evaluation of Informatica PowerCenter 7.1 to be used as an ETL tool.

• Oracle Discoverer Reports Developer for the project.

E D UC A TI O N

• **Master of Business Administration, Information Technology** | Indian Institute of Information Technology and Management, Gwalior-Deemed University 2003

• **Bachelor of Engineering, Electronics and Telecommunication** | Bhilai Institute of Technology, Pt Ravi Shankar Shukla University Raipur 2001

• **Information Technology, Data Science Certification** | University of Toronto