**Akshay**

**Sr. Data Architect/Data Modeler**

**Email:** [**akshay.md89@gmail.com**](mailto:akshay.md89@gmail.com)

**Cell: (609) 469-4362**

**Summary:**

* Over **9+** years of experience **Data Architect/Data Modeler/Data Analyst**, building and scaling transactional and **analytical database, data warehouse** and **Business Intelligence** solutions.
* Experience in **Big** **Data** **Hadoop** Ecosystem in **ingestion**, **storage**, **querying**, **processing** and **analysis** of **big** **data**.
* Knowledge and working experience on big data tools like **Hadoop, Azure Data lake, AWS** **Redshift**.
* Deep expertise on **Cloud platforms** with preferred knowledge across emerging technologies **(AWS, Azure, Google Cloud, CloudStack/OpenStack, Joyent, RightScale, Scalr)**
* Experience in developing **Map** **Reduce** Programs using **Apache** **Hadoop** for analyzing the **big** **data** as per the requirement.
* Experienced in Technical consulting and end-to-end delivery with **architecture**, **data** **modeling**, **data** **governance** and **design** - development - implementation of solutions.
* Excellent experience in creating **cloud** **based** **solutions** and architecture using **Amazon** **Web** **services** and **Microsoft** **Azure**.
* Experience in Dimensional **Star** and **Snowflake schema, FACT & Dimension tables.**
* Excellent experience on **Teradata** **SQL** **queries**, **Teradata** Indexes, Utilities such as **Mload**, **Tpump**, Fast load and **FastExport**.
* Understanding in development of **Conceptual, Logical and Physical Models** for **Online** **Transaction** **Processing** and **Online** **Analytical** **Processing** (**OLTP** & **OLAP**).
* Expertise in developing solutions around **NOSQL** **databases** like **MongoDB** and **HBase**.
* Experience in working with business intelligence and data warehouse software, including **SSAS**, **Pentaho, Cognos, OBIEE, QlikView, Greenplum Database, Amazon Redshift** and **Azure** **Data** **Warehouse**
* Extensive experienced in **Normalization (1NF, 2NF, 3NF and BCNF)** and **De-normalization** techniques for improved database performance **Data Warehouse/Data** **Mart** environments.
* Good understanding and hands on experience with **AWS** **S3** and **EC2**.
* Good experience on programming languages **Python**, **Scala**.
* Experienced working with **Excel** **Pivot** and **VBA** macros for various business scenarios.
* Experience in automating and scheduling the **Informatica** jobs using **UNIX** shell scripting configuring **Korn**-jobs for **Informatica** sessions.
* Experienced in various databases Design of development and Production environment involving **Oracle, SQL** **server**, **Netezza**, **MY** **SQL**, **DB2**, **MS** **Access**, **Teradata** etc.
* Excellent experience in **Data** **mining** with querying and **mining** **large** **datasets** to discover transition patterns and examine financial data.
* Excellent in creating various artifacts for projects which include specification documents, **data mapping** and **data** **analysis** documents.
* Experienced in Performance **Tuning** on **oracle** databases by leveraging explain plans, and tuning **SQL** queries.
* Excellent experience in writing **SQL** queries to validate data movement between different layers in **data** **warehouse** environment.
* Extensive experience in using **ER** **modeling** **tools** such as **Erwin** and **ER**/**Studio**.
* Excellent knowledge on creating reports on **SAP** **Business** **Objects**, **WEBI** reports for multiple data providers.
* Efficient in analyzing and documenting **business** **requirement** **documents** (**BRD**) and **functional** **requirement** **documents** (**FRD**) along with **Use** **Case** Modeling and Source to Target Mapping document.
* Excellent Team player to work in conjunction with Business analysts, Production Support teams, Subject Matter Experts, Database Administrators and Database developers.

**Technical Skills:**

* **Data Modeling Tools:** Erwin r9.6/r9.5, ER Studio 9.7/9.0, Sybase Power Designer.
* **Big Data: Hadoop,** HDFS, Hive, Pig, HBase, Sqoop, Flume, Kafka.
* **Cloud Platform:** AWS, Azure, Google Cloud, CloudStack/OpenStack, Joyent, RightScale, Scalr
* **Database Tools:** Oracle 12c/11g, Teradata 15/14, Netezza, Microsoft SQL Server 2014/2016, and MS Access, PostgreSQL.
* **Quality Assurance Tools:** Win Runner, Load Runner, Test Director, Quick Test Pro, Quality Center, Rational Functional Tester**.**
* **Reporting tools:** SQL Server Reporting Services (SSRS), Tableau, Crystal Reports, Business Objects, MicroStrategy, Business Objects 5.1, Cognos 6.5/7.0
* **ETL Tools:** SSIS, Pentaho, Informatica9.6.
* **Programming Languages:** Java, Base SAS and SAS/SQL, SQL, T-SQL, HTML, Java Script, CSS, UNIX shells scripting, PL/SQL.
* **Operating Systems:** Microsoft Windows 8/7, UNIX , Linux, Redhat
* **Tools & Software:** TOAD, SQL \*PLUS, SQL\*LOADER, MS Office, BTEQ, Teradata SQLAssistant
* **Methodologies:** RAD, JAD, RUP, UML, System Development Life Cycle (SDLC), Waterfall Model.

**State Street, Princeton, NJ Jul 16 – Till date**

**Sr. BI Data Architect/Data Modeler**

**Responsibilities:**

* Working as an Architect and develop scalable, highly available, fault tolerant, secure systems for on-premises, hybrid and cloud-based data systems that meet client business needs.
* As a Architect implement **MDM** **hub** to provide clean, consistent data for a **SOA** implementation.
* Involved in several facets of MDM implementations including Data Profiling, Metadata acquisition and data migration.
* Worked with FACETS Team for HIPAA Claims Validation and Verification Process.
* Extensive experience in all facets of project life-cycle SDLC & Agile Scrum from feasibility analysis and conceptual design through implementation, including documentation, user training and operation support.
* Implemented **Agile** **Methodology** for building Integrated Data Warehouse, involved in multiple sprints for various tracks throughout the project lifecycle.
* Implemented various **Azure** **platforms** such as **Azure SQL Database, Azure SQL Data Warehouse, Azure Analysis Services, HDInsight, Azure Data Lake, Data Factory**
* Involved in developing Database Design Document including Data Model **Conceptual**, **Logical** and **Physical** **Models** using **Erwin** **9**.**64**.
* Responsible for analysis of massive and highly complex data sets, performing **ad**-**hoc** **analysis** and **data** **manipulation** for **data** **integration**.
* Designed and implemented scalable **Cloud** **Data** and **Analytical** **architecture** solutions for various public and private cloud platforms using **Azure**
* Designed and developed data architecture solutions in **big** **data** **architecture** or **data** **analytics**.
* Evaluate **architecture** **patterns**, Define best patterns for data usage, **data** **security**, data **compliance**, Define concept models, logical & physical data model
* Handled importing of data from various data sources, performed transformations using **Hive, MapReduce**, loaded data into **HDFS** and extracted the data from **Oracle** into **HDFS** using **Sqoop**
* Understand transaction data and develop Analytics insights using Statistical models using **Azure** **Machine learning**.
* Applied **Data Governance** rules (primary qualifier, Class words and valid abbreviation in Table name and Column names).
* Designed and documented **logical** and **physical** **database** designs for Enterprise Application **(OLTP), Data Warehouses (OLAP),** **NoSQL** databases.
* Developed **MapReduce** programs to parse the raw data, populate staging tables and store the refined data in partitioned tables in the **EDW**.
* Designed both **3NF** data models for **ODS**, **OLTP** systems and dimensional data models using star and snow flake Schemas.
* Developed and presented **data flow diagrams, conceptual diagrams, UML diagrams, ER flow diagrams**, creating the **ETL Source** to Target mapping specifications and supporting documentation.
* Developed long term **data warehouse** **roadmap** and **architectures**, designs and builds the data warehouse framework per the roadmap.
* Worked on **Metadata Repository (MRM)** for maintaining the definitions and mapping rules up to mark.
* Independently coded new programs and design Tables to load and test the program effectively for the given **POC's** using **Big** **Data**/**Hadoop**.
* Involved in **Normalization**/**De**-**normalization** techniques for optimum performance in relational and dimensional database environments.
* Developed multiple **MapReduce** jobs in **Java** for **Data** **Cleaning** and pre-processing analyzing data in **PIG**.
* Used **windows** **Azure** **SQL** **reporting** services to create reports with **tables**, **charts** and **maps**.
* Performed data modeling to differentiate between **OLTP** and **Data** **Warehouse** data models.
* Developed triggers, stored procedures, functions and packages using cursors and ref cursor concepts associated with the project using **PL/SQL**
* Created **SSIS** Packages for import and export of data between **MS** **SQL** **Server** database and others like **MS** **Excel** and **Flat** **Files**.
* Dimensional modeling of **EDW** following **Kimball** **methodology** with **Erwin** data modeling tool for **Data** **marts** and **data** **warehouses** in **Star** **Schema**, with confirmed dimensions.
* Involved in the hands-on technical delivery of customer projects related to **Azure.**
* Support **Cloud Strategy** team to integrate analytical capabilities into an overall cloud architecture and business case development

**Environment:** ERWIN r9.6, Netezza,Azure,Amazon Redshift, Oracle12c, OLAP, OLTP, T-SQL, SQL Server 2016, SSIS, SSRS, Linux, MDM, Hadoop, MapReduce, Pig, HBase, Redshift, Java, PL/SQL.

**Akorn Pharmaceuticals, Warren, NJ Aug 15 – Jun 16**

**Sr. BI Data Architect/Data Modeler**

**Responsibilities:**

* Act as the technical lead during the architect phase, working in conjunction with a Project Manager to create the development plan.
* Designed the **Logical & Physical Data Model** using **ERWIN** **9**.**5** with the entities and attributes for each subject areas.
* Involved in several facets of **MDM** implementations including **Data Profiling, Metadata** **acquisition** and **data** **migration**.
* Responsible for technical **Data governance**, **enterprise** wide **Data** **modeling** and **Database** **design**.
* Included migration of existing applications and development of new applications using **AWS cloud** **services**.
* Defined best practices for **data** **modeling** and extraction and ensure **architectural** **alignment** of the **designs** and development.
* Evaluate and developed mastery of emerging technologies in the cloud space, especially **AWS**
* Architected, created and moved data to new front end using **SQL** **Azure** as backend.
* Involved in integration of various relational and non-relational sources such as **Oracle**, **XML** and **Flat** **Files**.
* Connected to **Amazon Redshift** through Tableau to extract live data for real time analysis.
* Involved in **OLAP** model based on Dimension and **FACTS** for efficient loads of data based on **Star** **Schema** structure on levels of reports using multi-dimensional models such as **Star** **Schemas** and **Snowflake** **Schema**.
* Handled importing of data from various data sources, performed transformations using **Hive**, **MapReduce**, loaded data into **HDFS** and Extracted the data from **Oracle** into **HDFS** using **Sqoop**.
* Documented logical data integration into **AWS Data Lake** from **Data** **warehouses**
* Responsible for delivered the **logical** and **physical** **data** **models** for **AWS** **Data** **Lake** for sales components.
* Developed **Star** and **Snowflake** **schemas** based dimensional model to develop the Data warehouse.
* Wrote and executed various **MySQL** database queries from **python** using **Python**-**MySQL** connector and **MySQL** package.
* Involved in **Normalization (1NF/2NF/3NF), De-normalization** techniques in **relational/ dimensional database** environments.
* Worked on analyzing source systems and their **connectivity, discovery, data profiling** and **data** **mapping**.
* Driven the technical design of **AWS** **solutions** by working with customers to understand their needs
* Generated **ad**-**hoc** **SQL** **queries** using joins, database connections and transformation rules to fetch data from **Teradata** database.
* Designed and implemented basic **PL/SQL** queries for testing and sales report/data validation.
* Used **Microsoft** **azure** using **PowerShell** to generate query results from **SQL** **server**.
* Extensively used **MS** **Visio** for representing existing and proposed **data** **flow** **Diagrams**.
* Implementation of **Business** **Rules** in the Database using **Constraints** & **Triggers**.
* Designed and architecting **AWS Cloud** solutions for data and analytical workloads such as warehouses, **Big Data, data lakes**, real-time streams and **advanced** **analytics**
* Interacted with End-users for gathering **Business** **Requirements** and Strategizing the **Data** **Warehouse** processes
* Write complex **Netezza** views to improve performance and push down the load to database rather than doing it in the ETL tool.
* Worked with **Teradata** **RDBMS** using **Fast load, Fast Export, Multi load, Tpump**, and **Teradata** **SQL** **Assistance** and **BTEQ** **Teradata** **utilities**
* Defined different data integration and validation frameworks including **data** **validation**, **Error** **checking** **process**, **lineage**, **recovery** and **reconciliation**.
* Extensively used **MS** **Access** to pull the data from various data bases and integrate the data.

**Environment:** Erwin9.5, AWS, Amazon Redshift, Teradata15, OLAP, OLTP, Hive, HDFS,Netezza, Hadoop, Spark, ETL, PL/SQL, MDM, MS Visio, OLTP, BTEQ.

**BMO Harris Bank, Chicago, IL May 14 – Jul 15**

**Sr. Data Architect /Data Modeler**

**Responsibilities:**

* Understand the high level design choices and the defined technical standards for software coding, tools and platforms and ensure adherence to the same .
* Used **Agile** Methodology of **Data** **Warehouse** development using **Kanbanize**.
* Analyze business requirements and build **logical** **data** **models** that describe all the data and relationships between the data
* Designed both **3NF** data models for **ODS**, **OLTP** systems and dimensional data models using **Star** and **Snow** **Flake** **Schemas**
* Provided suggestion to implement multitasking for existing **Hive** **Architecture** in Hadoop also suggested **UI** customization in **Hadoop**
* Architect and lead significant data initiatives in various data dimensions **Master Data, Meta Data, Big Data** & **Analytics**
* Involved in Planning, Defining and Designing database using **ER** **Studio** on business requirement and provided documentation.
* Translate business and data requirements into **logical** **data** **models** in support of **Enterprise** **Data** **Models**, **Operational** **Data** **Structures** and **Analytical** **systems**.
* Partner with **DBAs** to transform **logical** **data** **models** into **physical** **database** designs while optimizing the performance and maintainability of the physical database
* Work with **Data Management** to establish **governance** **processes** around **metadata** to ensure an integrated definition of data for enterprise information, and to ensure the **accuracy**, **validity**, and **reusability** of **metadata**.
* Migrated **SQL** **Server** **Database** to **Microsoft** **Azure** **SQL** **Database**
* Developed Full life cycle of **Data Lake, Data Warehouse** with **Big** **data** technologies like **Spark** and **Hadoop**.
* Applied all phases of the **Software Development Life Cycle**, which include requirements definition, analysis, review of design and development, and integration and test of solution into the operational environment
* Responsible for full data loads from production to **AWS** **Redshift** staging environment.
* Worked on **Azure** **Power** **BI** Embedded to integrate the reports to application.
* Developed **Map Reduce** programs to cleanse the data in **HDFS** obtained from heterogeneous data sources to make it
* Lead database level tuning and optimization in support of application development teams on an **ad-hoc basis**.
* Created **data** **schema** and **architecture** of **data** **warehouse** for standardized data storage and access
* Used **data profiling** automation to uncover the characteristics of the data and the relationships between data sources before any data-driven.
* Used **Azure** **reporting** **services** to upload and download reports
* Develop **test scripts** for testing sourced data and their **validation** and **transformation** when persisting in data stores that are **physical representations** of the **data** **models**
* Designed and documented **Use Cases, Activity Diagrams**, **Sequence** **Diagrams**, **OOD** (Object **Oriented Design) using UML** and **Visio**.
* Completed enhancement for **MDM (Master data management)** and suggested the implementation for **hybrid MDM (Master Data Management)**
* Designed processes and jobs to source data from Mainframe sources to **HDFS staging** zone
* Integrated data from multiples sources including **HDFS** to **Hive** **Data** **warehouse**.
* Worked very close with **Data** **Architectures** and **DBA** **team** to implement data model changes in database in all environments.
* Generate **DDL scripts** for database modification, **Teradata, Macros, Views** and set tables.

**Environment:** ER Studio 9.0, Hive, Hadoop, MDM, AWS, Redshift, HDFS, Teradata 14, PL/SQL, Informatica 9.0, Oracle 10g, UNIX

**E-trade Financial, Alpharetta, GA Jan 12 – Apr 14**

**Sr. Data Analyst /Data Modeler**

**Responsibilities**

* Performed in team responsible for the analysis of business requirements and design implementation of the business solution.
* Developed **logical** and **physical** data models for central model consolidation.
* Worked with **DBAs** to create a best fit physical data model from the logical data model.
* Conducted data modeling **JAD** sessions and communicated data-related standards.
* Used **Erwin r8** for effective model management of sharing, dividing and reusing model information and design for productivity improvement.
* Used **Star/Snowflake** schemas in the **data** **warehouse** architecture.
* Redefined many attributes and relationships in the **reverse** **engineered** model and **cleansed** **unwanted** **tables**/**columns** as part of **data** **analysis** responsibilities
* Developed process methodology for the **Reverse** **Engineering** phase of the project.
* Used **reverse** **engineering** to connect to existing database and create graphical **representation (E-R diagram)**
* Utilized **Erwin's reverse engineering** and target database **schema** **conversion** process.
* Involved in **logical** and **physical** **designs** and transforms logical models into physical implementations.
* Created **3NF** business area data modeling with de-normalized physical implementation data and information requirements analysis using **ERWIN** **tool**.
* Involved in extensive **data** **analysis** on **Teradata**, and **Oracle** **Systems** **Querying** and Writing in **SQL** and **Toad**.
* Involved using **ETL** tool **Informatica** to populate the database, data transformation from the old database to the new database using **Oracle** and **SQL** **Server**.
* Creation of database objects like tables, views, Materialized views, procedures, packages using **Oracle** tools like **PL/SQL, SQL\* Plus, SQL\*Loader** and Handled Exceptions.
* Used **Informatica Designer**, **Workflow** **Manager** and **Repository** **Manager** to create source and target definition, design mappings, create repositories and establish users, groups and their privileges
* Involved in **Data** **profiling** in order to detect and correct inaccurate data and maintain the **data** **quality**.
* Developed **Data Migration** and **Cleansing** rules for the Integration Architecture **(OLTP, ODS, DW).**
* Involved in the creation, maintenance of **Data** **Warehouse** and repositories containing **Metadata**.
* Developed Star and Snowflake schemas based dimensional model to develop the data warehouse.
* Involved in the study of the **business** **logic** and understanding the **physical** **system** and the terms and condition for database.
* Worked closely with the **ETL SQL Server Integration Services (SSIS)** Developers to explain the **Data Transformation**.
* Creating reports using **SQL Reporting Services (SSRS)** for customized and **ad-hoc Queries**.
* Created documentation and test cases, worked with users for new module enhancements and testing.
* Created simple and **complex** **mapping** using **Datastage** to load Dimensions and **Fact** **tables** as per **Star** **schema** techniques.
* Designed and Developed **Oracle database Tables**, **Views**, **Indexes** with proper privileges and Maintained and updated the database by deleting and removing old data.
* Generated **ad-hoc** reports using **Crystal** **Reports**.

**Environment:** Erwin r8, Informatica 9.1, Windows XP, Oracle10g, SQL Server 2012, MS Excel, MS Visio, Oracle10g, Microsoft Transaction Server, Crystal Reports, SQL\*Loader

**Wells Fargo, Pune, IN Aug 08 – Nov 11**

**Data Analyst/Data Modeler**

**Responsibilities**

* Analyzed **data** **sources** and requirements and business rules to perform **logical** and **physical** data modeling.
* Analyzed and designed best fit **logical** and **physical** **data** **models** and relational database definitions using **DB2**.
* Conducted source **data analysis** of various data sources and develop source-to-target mappings with business rules.
* Maintained existing **ETL procedures**, fixed bugs and restored software to production environment.
* Involved in different stages of **SDLC such** as translating business requirements to high level and low-level design, Coding, **Unit testing,** deployment and post-deployment support activities
* Worked with **Data Warehouse** Extract and load developers to design mappings for **Data Capture, Staging, Cleansing, Loading**, and **Auditing**.
* Developed **enterprise data model** management process to manage multiple data models developed by different groups
* Designed and created **Data Marts** as part of a data warehouse.
* Transformed project data requirements into project data models for **OLAP** and **OLTP** systems using **Erwin**.
* Effectively used **triggers** and stored **procedures** necessary to meet specific application's requirements.
* Created **SQL scripts** for database modification and performed multiple data modeling tasks at the same time under tight schedules.
* Reviewed new data development and ensured that it is consistent and well integrated with existing structures.
* Wrote complex **SQL queries** for validating the data against different kinds of reports generated by **Business Objects XIR2**.
* Worked on **PL/SQL** collections, index by table, arrays, bulk collect, FOR ALL, etc.
* Involved in reviewing business requirements and analyzing data sources form **Excel/Oracle SQL Server** for design, development, testing, and production rollover of reporting and analysis projects.
* Document and publish test results, troubleshoot and escalate issues
* Worked on **SAS** and **IDQ** for **Data** Analysis.
* Using **Erwin** modeling tool, publishing **of** a **data dictionary**, **review** of the model and dictionary with subject matter experts and generation of data definition language.

**Environment:** Erwin 7.0, Oracle 9i, SQL Server 2005, XML, MS Excel, MS Access, MS Visio, PL/SQL, SSIS, Metadata.