**Nagul Shaik**

**Resume**.[**Nagul@gmail.com**](mailto:Nagul@gmail.com)

**PH: 732-397-9869**

**Professional Career**

I have **12 years** of hands on experience in **Information Technology (IT)** as **Data Warehouse Architect, Data Architect, Sr.Azure Data Modeler** and **Business Intelligence Architect** in heterogeneous business organizations like **County of Santa Clara Health System (CSHS), County of Santa Clara Public Health Department (PHD), The University of Texas System (UT Austin), North Carolina State Department of Transportation (NCDOT), Pharmaceuticals, Banking, Finance, Manufacturing and Telecommunications.**

**Technical Skills in a glance:**

* Proficient in **System** and **Software Development Life Cycle (SDLC)** methodologies **Agile, Scrum, Kanban** and **Iterative.**
* Experienced in **TOGAF** and **Zachman Enterprise Architecture (EA)** frameworks.
* Hands on experience in **relational** and **dimensional data models** design life cycles**.**
* Have good experience and knowledge in **data conversion** life cycle**.**
* Well versed in **normalization (1NF, 2NF** and **3NF)** and **de-normalization** techniques for optimum performance in **relational** and **dimensional** database environments.
* Experienced in designing **Entity Relationship Diagrams (ERD)** and **conceptual, logical** and **physical** data models for **OLTP** and **OLAP** systems using data modeling tools like **Erwin**, **DbSchema,** **ER Studio, SAP** **Power Designer** and **Infosphere Data Architect (IDA)**.
* Experienced in designing OLAP systems using **Kimball, Inmon** and **Data Vault** methodologies for **Business Intelligence (BI)** projects.
* Hands on experience in designing **OLAP** systems using **Star** and **Snowflake** schemas.
* Have hands on experience and solid knowledge in data extraction, data transformation, data profiling, data analysis, data discovery, data masking, data governance, data cleansing, data scrubbing, data validation, data wrangling and Metadata maintenance for Data Warehouse (DWH), Master Data Management (MDM) and Data Conversion (DC) projects.
* Well proficient in **Metadata Management (MM), Reference Data Management (RDM)** and **Test Data Management (TDM).**
* Good experienced in **ETL**, **ELT**, **Data Transformations** and **Source** to **Target** **Data Mapping** concepts.
* Experienced in creating and maintaining **Slowly Changing Dimensions (SCD)** and **Change Data Capture (CDC)**.
* Expertise in **Extract Transform** and **Load** (**ETL**) techniques using **Pentaho Kettle,** **Informatica, SSIS** and **Microsoft** **Data Transformation Service** (**DTS**).
* Have good experience with data profiling tools like Alteryx, Trillium and Informatica Data Quality (IDQ).
* Having very good experience in Health Level Seven (HL7) and FHIR messaging specifications and standards.
* Having hands on experience in EMR/EHR applications like Epic Clarity and Epic Caboodle.
* Solid knowledge with Customer Relationship Management (CRM) and Enterprise

Resource Planning (ERP) applications.

* Experienced in data mining, designing **OLAP, ROLAP** and **MOLAP** databases.
* Hands on experience in **Data Governance** and **Data Stewardship** policies.
* Solid knowledge in **No SQL** data models and databases **Mongo DB, Greenplum, HBase, Couchbase, Cassandra** and **Amazon DynamoDB.**
* Good knowledge and experience in cloud technologies like **Azure Data Lake Store, Azure Data Factory, Azure SQL DB** and **Azure Synapse Analytics (SQL DW), Snowflake cloud data warehouse, AWS Redshift, AWS RDS** and **Google Cloud Platform (GCP).**
* Having exposure to big data technologies **Hadoop** and **Spark** andextensive experience in both batch and stream processing using **Lambda** and **Kappa** architectures.
* Having good knowledge in **IBM Mainframe** systems, **Adabas (Natural)** databases.
* Good in creating **Ad-Hoc** reports, **Indexes**, **Stored Procedures**, **Triggers**, **Functions** and **Packages**.
* Having good experience in writing **SQL** queries to perform end-to-end **ETL** data validations and Solid hands on experience in **Quality Assurance (QA)** techniques, performed **User Acceptance, Unit, System, Manual, Automated** and **ETL Testing’s.**
* Good knowledge and experience in **Microsoft** **Power BI** and **Tableau** tools, created **dashboards, score cards** and **reports**.
* Having good knowledge and experience in **Windows, UNIX** and **Linux** operating systems.
* **Quick learner, self-motivating** and **enthusiastic** to work with new technologies and new people.
* Possessing excellent **verbal** and **written communication skills**.

**Education Qualifications:**

**Bachelor of Technology in Electronics and Communications Engineering.**

**Core Technologies and Tools:**

|  |  |
| --- | --- |
| **Data Modeling Tools** | **Erwin 9.6,** **DbSchema 8.2,** **SAP Power Designer 16.5, ER/Studio 9, Infosphere Data Architect (IDA),** **Oracle SQL Data Modeler, UML and Microsoft Visio 2016.** |
| **ETL Tools** | **Pentaho Data Integration (PDI), Kettle, Informatica 9.5, SQL Server Integration Services (SSIS), Ab Initio, OBIEE, IBM Data Stage.** |
| **Data Quality Tools** | **Query Surge 4, HPQC-ALM, Informatica Data Quality (IDQ), Data Quality Services (DQS), Data flux, Trillium, IBM InfoSphere Quality Stage and IBM Info Sphere Information Analyzer.** |
| **Database Tools** | **Microsoft SQL Server 2012/2016, Azure SQL DB, Azure Synapse Analytics (SQL DW), AWS Redshift, Postgresql, SQL Server Management Studio (SSMS), Toad for DB2,Oracle SQL Developer 4.1, Hadoop, Spark**, **SAP HANA, Db Visualizer Pro 9.2, Oracle 11g/10g, IBM DB2, Teradata, MS Access, IBM Mainframe and Adabas (Natural).** |
| **Reporting Tools** | **Power BI, Tableau, Looker, SQL Server Analysis Services (SSAS), TIBCO Spotfire, Cogno’s Power Play Studio, Qliksense, Qlikview and Crystal Reports.** |
| **Programming Languages** | **ASCII-SQL, PL/SQL, SnowSQL(CLI), Python, XML, XSD, WSDL, NIEM schema HL7,FHIR, SAS, Spark SQL, R Studio, HTML.** |
| **Project Tracking**  **Tools** | **Microsoft Office Suite 2016, Microsoft Project 2016, JIRA, Azure DevOps, SVN Version Control, Team Foundation Server (TFS), IBM Rational Clear Case and Clear Quest.** |
| **Operating Systems** | **Microsoft Windows (10, 8), NT/2000/Vista, Unix, Linux.** |

Professional Experience:

County of Santa Clara Health System and County of Santa Clara Public Health Department (PHD), San Jose, CA Oct 18 - Present

Role: Data Warehouse Architect

The County of Santa Clara Health System is local healthcare safety net and provider of comprehensive care, service and program to the residents of Santa Clara County. The Health System includes the Public Health Department, Behavioral Health Services Department, Santa Clara Valley Medical Center hospital and clinics, O'Connor Hospital, St. Louise Regional Hospital, Emergency Medical Services Agency, Custody Health Services Department and Valley Health Plan. Santa Clara County health officials dedicated to the health and well-being of all communities of County.

**Responsibilities**

* Worked with end users **(Epidemiologists)** team to get business requirements, use cases, prioritize project goals and maintain **project tasks traceability matrix**.
* Thoroughly understood **PHI** and **HIPAA** rules and regulations to meet **organization compliance** and **patient data privacy requirements**.
* Interacted with multiple **SME’s** to understand source system business process.
* Designed data warehouse project Data Flow Diagram (DFW) using Microsoft Visio tool.
* Identified EMR, Epic Clarity, Epic Caboodle source systems and performed data discovery, data analysis and data profiling.
* Defined and created data modeling design specifications and standards document for data vault data models.
* Designed Raw Vault (RV) and Business Vault (BV) data models using DbSchema tool for patients, encounters, providers, procedures, medications, labs, vitals, problem list, diagnosis, claims, coverage, enrollment, member, pharmacy subject areas.
* Generated DDL script and forward engineered Raw Vault (RV), Business Vault (BV) and Star schema data models into target databases.
* Designed Star schema data model using Erwin tool for billing, payment and insurance (VHP, SCFHP, Medicaid/Medi-Cal and Medicare) subject areas.
* Defined and maintained Raw Vault (RV), Business Vault (BV) and Star schema data models metadata, business and technical glossary, data dictionary and User Defined Properties (UDP).
* Reverse Engineered (RE) Epic Caboodle data model using ER/Studio tool.
* Created Source to Target Mapping (STTM) document with data quality and data transformation rules.
* Heavily worked with Health Level Seven (HL7) and FHIR messaging specifications and standards documents to identify mapping for EMR/EHR data elements like patients, encounters, vitals, labs, diagnosis, problem list and claims.
* Created ETL transformations to move data from HL7/FHIR applications to SQL server database.
* Developed ETL process and loaded data from source and stage systems into target systems using Lambda architecture.
* Designed and developed Azure Data Lake Store (ADLS), Azure Data Factory (ADF), Azure SQL DB and Azure Synapse Analytics services.
* Reviewed Epic Clarity compass comparison document and updated Raw Vault (RV), Business Vault (BV) and Star schema data models.
* Designed and deployed Data Vault (DV) security model.
* Created and maintained **reference data tables** on target database.
* Prepared, created and executed **Unit** and **System** test cases for **HUB, LINK** and **SAT** (Satellite) tables.
* Executed **Data Validation** techniques on target schema data to measure **data** **accuracy** and **data** **quality**.
* Created **Business Intelligence (BI)** reports and performed **multidimensional data analysis** using **Microsoft Power BI** and **Looker BI** tools.
* Worked with DBA team to deploy **Raw Vault (RV), Business Vault (BV)** and **Star schema data models into PRD databases.**

**Environment:** DbSchema 8.2, Erwin 9.6, ER/Studio, Microsoft Visual Studio (SSIS), Microsoft Azure DevOps, Power BI 2.81, Tableau 2018.1, Looker, IBM Mainframe, Adabas (Natural),Redgate tool suite, Health Level Seven (HL7),FHIR, Epic Chronicle, Epic Clarity, Epic Clarity Data Dictionary, Epic Caboodle, QNXT,HIPAA, Datavault Builder, Snowflake Cloud Data Warehouse, Lambda, Team Foundation Server(TFS), SQL,PL/SQL,Unix, Linux, UML, Agile, Microsoft Visio 2016, Microsoft Project 2016, Microsoft Office Suite 2016.

The University of Texas System (UT Austin), Austin, TX Apr 17 - Sep 18

Role: Data Warehouse Architect

The University of Texas at Austin is a public research university, it is a flagship institution of the [University of Texas System](https://en.wikipedia.org/wiki/University_of_Texas_System) (UT System). UT Austin is one of the most selective universities in the region and consistently ranked among the country’s top public universities. It is a major academic research institution in Texas State. It also offers innovative programs for promoting academic excellence and leadership development.

**Responsibilities**

* Coordinated with business architect team and **IRRIS (Institutional Reporting, Research and Information Systems)** team to get high level business requirements.
* Designed **Business Process Diagrams (BPD)** and circle diagrams using **Microsoft Visio** tool.
* Defined **Grain**, identified **Dimensions** and **Measures** for each subject area in an **Enterprise Data Warehouse (EDW)** project.
* Performed **Data Analysis** and **Data Profile** on source tables to identify **Critical Data Elements (CDE)**.
* Worked with **IBM legacy systems** in **application and data modernization projects.**
* Validated source data to identify **data anomalies, data redundancy** and understand **data quality**.
* Worked with **Collibra Reference Data Accelerator** to create reference data.
* Designed **conceptual, logical** and **physical data models** using **Erwin v9.6** tool.
* Maintained different versions of data models using **Erwin Mart** feature.
* Created reference tables with **reference data**, **conformed dimensions**, **Slowly Changing Dimensions (SCD)** anddifferenttype’s **fact tables.**
* Designed data models for **Master Data Management (MDM**) project.
* Worked with big data technologies **Hadoop, Spark** implementation team.
* Designed **Data Lake centralized repository** system to store structured and unstructured data.
* Designed **Data Lake system** using **Data and Analytics as a Service (DAaaS) model**.
* Designed dashboards, scorecards using Power BI tool.
* Created data base objects like **Indexes, Views, Stored Procedures** and **Tablespaces**.
* Reviewed **business user reports** and **cubes** using **Cogno’s Power Play studio tool.**

**Environment:** Erwin 9.6, ER/Studio, Oracle SQL developer data modeler 4.0, Oracle11g, Oracle Exadata X6,Azure SQL DB, Azure Synapse Analytics (SQL DW)**,** Hadoop, Spark, Legacy IBM Mainframe System, Adabas (Natural), FASET System, SIS System, IBM Data Stage, IBM InfoSphere Quality Stage, Oracle SQL, Collibra, Lambda, Power BI,Tableau and Cogno’s Power Play Studio, Looker, JIRA Software, Agile, Unix, Linux, SQL,PL/SQL,UML and Microsoft Visio 2010, Microsoft Project 2010, Microsoft Office Suite.

Pharmaceutical Product Development (PPD Inc.), Morrisville, NC May 15 - Mar 17

**Role: Sr.Data Architect**

Pharmaceutical Product Development (PPD) is a global Contract Research Organization (CRO) providing discovery, development and post-approval services for pharmaceutical, biotechnology, medical device, academic and government organizations. PPD is mainly known for conducting drug development for its clients, it also invests in developing new products through its compound partnering program.

Responsibilities

* Accumulated data model business requirements, documented functional and technical specifications.
* Designed **Conceptual Data Model (CDM), Logical Data Model (LDM) Entity Relationship Diagrams (ERD)** using **Erwin tool**.
* Designed **Star schema data marts** and **designed dimensional logical** and **physical models** using **Kimball** and **Data Vault methodology.**
* Conducted data model review sessions with **Business architect team**.
* Worked on **Metadata Management** across heterogeneous models.
* **Transformed logical data model into physical data model** using **Infosphere Data Architect (IDA)** tool**.**
* **Compared data model** to **database** physical schema using **compare feature** in **Infosphere Data Architect (IDA)** tool**.**
* Generated physical data model from database by using **reverse engineering** feature in

**Infosphere Data Architect (IDA) tool.**

* Documented **design specifications of data marts, data warehouse logical** and **physical data models**.
* Created and maintained **Reference Data tables** in clinical trials data environment**.**
* Enforced enterprise **naming standards and data type standards** in logical and physical data models using **Infosphere Data Architect (IDA) tool.**
* Hands on experience in **maintaining and working** on **different versions of data models** in a team environment**.**
* Designed **Slowly Changing Dimension (SCD) tables**.
* Created **Source to Target Mapping (STTM) document for ETL development.**
* Deployed Physical data models in target database and also deployed in virtualization environment using **Db Visualizer Pro tool**.
* Designed and created reports using **Power BI** tool.
* Designed and updated metadata for **Data Lake model**.
* Executed data definition and data manipulation operations for data analysis.
* Created query pairs in **Query surge tool** to test data in virtualization layer.
* Used **Informatica Data Quality (IDQ 9.5.x)** tool for **data verification** and **data cleansing.**
* Developed mappings using **Informatica Data Quality (IDQ 9.5.x)** transformations.
* Performed data quality comparison between **development** and **virtualization environment** using **Query surge tool.**
* Executing **unit** and **system testing for dimensional** and **fact tables** using **HPQC- ALM tool**.
* Developed SQL scripts to **validate data** in target database**.**
* Worked with **IBM legacy systems** in **data platform modernization project.**
* Executed **Inserting, Updating** and **Deleting (I, U, D) test case against target database.**
* Exported and prepared detailed unit and system test execution, bugs and defect reports using **HPQC-ALM** tool.

**Environment:** Erwin 9.6, Infosphere Data Architect (IDA), Oracle SQL developer data modeler 4.0, HPQC 12, Query Surge 4.0, Oracle11g, Oracle Exadata X6, Teradata12, Hadoop, Spark, IBM IMS database, Postgre SQL, MS Azure SQL server, SAP HANA, Informatica Powercenter 9.5, IBM Data Stage, Legacy IBM Mainframe System, Ab-Initio, OBIEE, Oracle SQL developer 4.1, Db Visualizer Pro 9.2, Power BI,TIBCO Spotfire, Collibra, CTMS system, Spark SQL, Oracle Clinical system, Health Level Seven (HL7),FHIR, SQL,PL/SQL,Python, R Studio, Scrum, Unix, Shell script, Cisco Tidal scheduler, Microsoft office suite 2013.

North Carolina State Department of Transportation (NC DOT) Raleigh, NC Nov 14 - May 15

Role: Data Architect

The North Carolina Department of Transportation (NCDOT) is responsible for issuing driver licenses, building, operating, repairing highways and maintaining different modes of transportations across North Carolina State.

**Responsibilities:**

* Interacted with the project architect team on regular basis to gather requirements and discuss project milestones.
* Created **Entity Relationship Diagrams (ERD), Functional Diagrams** and **Data Flow Diagrams (DFD’s).**
* Created **Logical** and **Physical** data models using **Erwin** and **Power designer** tools**.**
* Document **Mater Data Management (MDM) Master Data Services (MDS) model standards** and **naming conventions.**
* Designing **Enterprise Information Management (EIM) Logical data model** using **Erwin** tool and deploying Physical data model on target database.
* Worked in customer and product Master Data Management (MDM) projects.
* Deployed Master Data Services (MDS) model on target database.
* Creating **XML, XSD, WSDL** and **NIEM schemas for Service Oriented Architecture (SOA)** services.
* Heavily worked on **data validation** and **Data quality methods,** performed **data profiling** using **SQL** and other methods.
* Worked with **Data Quality Services (DQS)** tool to create data quality rules.
* Performed data stewardship roles like Implement data standards, create metadata and manage databases.
* Maintained security and data integrity of the database.
* Executed **complex queries** on source and target data base for data analysis and data validation.
* Involved in the maintenance of the database.
* Document the conversion processes, ensuring **data accuracy** and **Integrity.**
* Heavily used **data conversion process techniques** and **methods** to move data from **legacy systems** to **SQL server database** environment.
* Worked with the business in helping define & develop business rules, value maps, source to target data mapping documents and assisting **ETL** development team in development activities.

**Environment:** Erwin 9.5, ER/Studio, Oracle SQL developer data modeler, SQL server 2012, Teradata, Postgresql, Oracle10g, SAP HANA, MS Azure SQL server, SQL BI, Data Quality Services (DQS), IBM Data Stage, SQL server management studio 2012, Spark SQL, R Studio, MS Visio, SQL,PL/SQL,Python, XML, XSD, WSDL and NIEM schema, Agile, Unix, Linux, Shell scripting, Microsoft Excel, Access 2010.

John Deere Moline, IL Jun 14 - Nov 14

Role: Sr.Data Modeler/Sr.Data Analyst

John Deere & Company is an American corporation, it manufactures agricultural, construction and forestry machines, diesel engines, transmissions used in heavy equipment’s and lawn care equipment’s. John Deere also provides financial services and other business activities for customers.

**Responsibilities:**

* Interacted with the business users on regular basis to consolidate and analyze project requirements.
* Created **conceptual, logical** and **physical data models** using **Infosphere Data Architect (IDA) tool.**
* Used **Normalization methods** up to **3NF** and **De-Normalization** techniques for effective performance in **OLTP systems.**
* Enforced **Referential Integrity (R.I)** for consistent relationship between parent and child tables.
* Generated Data Definition Language (DDL) script from physical data model using Infosphere Data Architect (IDA) tool.
* Extensively used Infosphere Data Architect (IDA) tool to create web reports from models.
* Reverse Engineered (R.E) target databases, identified new data elements in the source systems then added new data elements to the existing data models
* Compared data with original source documents and validated for Data Accuracy.
* Extracted large volumes of data feed from different data sources, performed transformations and loaded the data into various targets.
* Worked with **Data Stewardship** team to maintain data standards, performing **data profiling** operations against databases.
* Developed **Data Migration** and **Data** **Cleansing** rules for the integration architecture **(OLAP, ODS, DW).**
* Experienced in development of **ETL mapping** and **scripts**.
* Heavily used **Pentaho Data Integrator (PDI)** and **Informatica Powercenter** for migrating data from legacy database to new database.
* Involved in **Star** and **Snow flake** schema for dimensional data models.
* Worked on multiple **data marts in Enterprise Data Warehouse project (EDW).**
* Involved in designing **OLAP** data models extensively used **Slowly Changing Dimensions (SCD).**

**Environment:** Infosphere Data Architect (IDA), Power designer 16.5, IBM DB2, Teradata, Oracle10g, PostgreSQL, SQL server 2012, MS Azure SQL server, SAP HANA, Informatica power center 8.6, Pentaho Data Integrator (PDI), IBM Data Stage, OBIEE, Visual studio, Crystal reports, SQL server management studio 2012, Informatica Data Quality (IDQ), Trillium, Spark SQL, R Studio, SQL,PL/SQL,Unix, Linux, Toad for DB2, Scrum, Microsoft Excel, Access 2010.

JP Morgan Chase Columbus, OH May 11 - May 14

Role: Data Modeler/Data Analyst

JP Morgan Chase is an American multinational banking and financial services holding company. It is the largest bank in United States. It is a major provider of financial services and it also providing services in investment management, wealth management, asset management and treasury and securities services.

**Responsibilities:**

* Demonstrated strong analytical skills in identifying and resolving data exchange issues.
* Used **E.F Codd’s Normalization methods1NF**, **2NF**, **3NF** and **De-Normalization** techniques for effective performance in **OLTP** and **OLAP** systems.
* Designed data models in **Dimensional Modeling (D.M)** environment.
* **Executed SQL queries** to retrieve data from databases for analysis purpose**.**
* Created **data model reports** and **Data Dictionary (D.D)** byusing **Erwin.**
* Have deep understanding of **Know your Customer (KYC) procedures** and **standards.**
* Executed user acceptance testing for **Know your Customer (KYC) applications** and **KYC project’s.**
* Worked with **Anti-Money Laundering (AML) partners** and **AML advisory team.**
* Designed **CDM, LDM, PDM** data models for **AML applications.**
* Reviewed **KYC documentation,** worked as liaison betweenbusiness and compliance **AML teams.**
* Created relational data modelsand **populated reference data.**
* Developing normalized Logical and Physical database models to design OLTP system for enterprise applications.
* Performing Forward Engineering (FE) operations to create a Physical data model with DDL that best suits requirements from the Logical data model.
* Developed data mapping, data governance and transformation and cleansing rules for the Master Data Management (MDM) Process.
* Executed enterprise data governance strategies in line with the organization’s business strategy and objectives.
* Conducted meetings with **DRB** team for **metadata** approval**.**
* Extensively used data quality tools to maintain consistent data.
* Convert data from multiple sources and run validation test on converted data.
* Thorough understanding and experience with the entire **data migration** process from analysis of existing data, cleansing, validation, translation tables, conversions and subsequent upload into new platform.

**Environment:** CA Erwin 9.5,Power Designer 16.5, Oracle10g, SQL server 2012, PostgreSQL, IBM IMS database, IBM DB2, Teradata, Informatica power center 8.6, OBIEE, Visual studio, Crystal report’s, SQL Server Management Studio (SSMS) 2012, Trillium, Informatica Data Quality (IDQ),IBM discovery, R Studio ,Toad for DB2, SQL,PL/SQL,Unix, Agile, Microsoft Excel, Access 2010.

**N.C.E.T, India June 07 - Aug 09**

**Role: Database Analyst**

**Responsibilities:**

* Gathered high level requirements and converted into technical and functional requirements.
* Involved in all phases of **SDLC** phases including **gathering,** **analysis, designing, developing, testing, implementation** and **maintenance.**
* Created business process flow diagrams using **MS Visio** tool.
* Created **conceptual,** **logical** and **physical** data models using **Erwin** tool.
* Extensively used **normalization** techniques to design **OLTP data models.**
* Performed **forward engineering (F.E)** to generate **DDL script** and deploy into target database.
* Performed data element mapping between source and targets databases.
* Worked with **data extraction** and **conversion** techniques for data migration project.
* Created and maintained data model **data dictionaries.**

**Environment:** Erwin 6.0,UML 6.0,SQL Server Management Studio (SSMS) 2005, Microsoft Visio, Microsoft office 2003, MS Access 2003,Toad for DB2, SQL,PL/SQL,UNIX, Windows NT, Scrum.