#### 

#### **Sid Negi**

6 Heffernan Street, Laverton, Melbourne, Australia 3028

 : 0421557206 (mobile)

: [ssnegi@yahoo.com](mailto:ssnegi@yahoo.com)

Linked In : [Linked In Sid Negi](www.linkedin.com/in/sid-negi-54389251)

Skype ID : Siddartha.Negi

Resume

###### OBJECTIVE

Seeking a position in Data Warehousing in extract, transform & load (ETL) methodology and

BI using IBM Infosphere Datastage, Business Glossary, Information Analyzer, Cognos among others from the stack of ETL tools.

**VISA STATUS Australian CITIZEN Australia Baseline Security Clearance**

## **COMPUTER SKILLS**

**ETL Framework Automation** : Developed IBM Datastage ETL Batch Automation Framework for dataflow automation, efficiency, and financial gain, SCD Type1 & Type2, Fact Load Multiple Instance Jobs, Automating load script

development extracting metadata information from database system tables.

**SQL Tuning** : partition swap for faster updates and inserts, parallel hint in queries, updates and inserts, explain plan for indexing

**ETL tools** : IBM Datastage 11.x, 8.x, 7.5 parallel, Server, IBM Metadata WorkBench, IBM Information Analyzer, IBM Business Glossary, Microsoft SQL Server SSIS, Talend

Reporting Tools : IBM Cognos, Hyperion Performance Suite, Qlikview, Tableau

Scheduler : Autosys, control-M

Operating Systems : Windows, MS Dos , UNIX

### RDBMS : Netezza, Oracle, MS SQL Server, DB2, Teradata

### Programming Languages : SQL, PL/SQL, C, HTML, JAVA 2, Javascript

### Front end IDEs : Developer 2000, Visual Basic 4.0, Oracle JBuilder

CASE Tool : Designer 2000, Erwin, TOAD, IBM Data Architect

**Atlassian tools :Jira, Bitbucket (Git), Confluence**

**AWS Redshift : Knowledge of AWS Redshift SQL (since it is ANSI SQL), scheduling jobs during off times extracting from db etc**

**CODE SAMPLE AVAILABLE :** [**Siddartha Singh Negi Github**](https://github.com/ssnegi1967/Siddartha-Singh-Negi.git)

**[Siddartha Singh Negi drop box](https://www.dropbox.com/sh/7hidjojarstjv1b/AADXKfL4KNev5lrS8-WIBRzha?dl=0)**

[**Siddartha Singh Negi google drive**](https://drive.google.com/drive/folders/1Yu-59CPcj_NuyBOgH1L-qiVhqMoRtF_l?usp=sharing)

**DATA WAREHOUSING IKM TEST : **

**QUALIFICATIONS**

Master of Science in Computer Science ****

Villanova University, Villanova, Pennsylvania, USA.

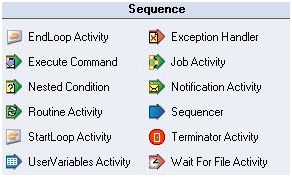
**SOFTWARE LIFE CYCLE** : Well Versed in AGILE methodology for software development using JIRA.

**CAREER SYNOPSYS** :

20 years experience in the field of DWBI. Experience in various capacities including senior software engineer, Design Consultant, Technical Architect. Experience dealing with various stakeholders like business analysts, data analysts, database administrators, unix administrators, testing teams during the software development life cycle. Personal interaction capability with technical and non-technical staff. I have experience in the OLTP snowflake 3NF data model in addition to the star fact and dimensions model for the reporting layer.

**Datastage Experience** : various source systems (RDBMS, Flatfile, XML, CSV). various stages rdbms connectors, transformer. lookup, join, merge, funnel, aggregator, change capture, modify, remove duplicates, sort, switch, pivot, slowly

changing dimension, copy, check sum, dataset, below sequence stages among others.



# **WORK EXPERIENCE**

**AIA Insurance – Senior ETL Developer Jan’20-Apr’20**

**Skills Used : IBM Datastage, Unix, Oracle, IBM Cognos**

The project involves generating csv export files using the reporting data mart consisting of facts and dimensions. Datastage is used to prepare the job to join the tables, parameters like dates, type of information etc.Facts and dimensions are joined to other star schemas using conformed dimensions.

Some report are prepared using Cognos data sources, packages and report studio.

**Department of Social Services – Senior ETL Developer Jan’18-Nov’19**

**Skills Used – Datastage 11.5, Oracle DB/PLSQL, Unix**

The role involved ETL Design, Development, Testing and monitoring Production Warehouse Loads.

Worked specifically in designing ETL Batch Automation Framework. The ETL Framework has the ability to run multiple applications containing jobs with differing dependencies in parallel. It is also robust and saves development, maintenance and testing time in addition to automating the Data warehouse load leading to cost saving and efficiency to the organization.

I have also designed SCD Type 2 Dimensions, Fact loads, Looping sequences for long running jobs splitting the load into 8 parallel processes based on ROWID. It also involved monitoring, defect fixing of in-production BAU activities.

**FOXTEL-Senior Data Integration Developer Jan’17 – Dec’17**

**Skill Used – IBM Datastage, Oracle DB/PLSQL, MS SQL Server, Unix**

Work involved two phases, one dealing with Oracle PLSQL procedures and another dealing with Datastage. The former was involving the COPS database and the migration of Business and Play customers from the transactional database. The datastage work dealt with a Star schema datamart dealing with Foxtel program scheduling data. It involved a fact table and several dimensions. The dimensions were of type 1. These two pieces of work were accomplished till now.

**MYERS – RXP Services – Datastage Consultant Oct’16 – Dec’16**

**Skills Used – IBM Datastage, Oracle, Unix**

The work involves preparing sales and discounted items quantities and amounts for supplier data. This data is obtained from a star schema fact – dimension data mart. The sales figures are extracted and then aggregated based on different factors. Then the aggregated data is output to a data file containing the data. This data file is separated for each supplier and sent to the suppliers to augment their sales figures.

**NBN Corporation – SMS Management and Technology - Datastage Consultant Feb’16 – Sep’16**

**Skills Used – IBM Datastage, Oracle, Unix, IBM Netezza, MS SSIS,SSRS,SSAS**

The program of work was Change Management Reporting . This included programming oracle source tables into Netezza Staging Layer. Then transforming them to the Common Access Layer where the business logic was applied. Then the CAL tables were brought into the End User Layer for report presentation purpose. History was maintained using Slowly Changing Dimension Type 2 for the dimension. The work also included generic job control sequences to handle the dependencies between the different layers and the dimensions and fact. Prepared a master sequence that managed the job flow in a parallel layout and was robust to send notification in case of failure.

**Qantas Airways - Datastage Consultant Aug’15 – Feb’16**

**Skills Used – IBM Datastage 11.3, Teradata 15.0, Oracle, Unix, MS SSIS,SSRS,SSAS**

The work included builds for access layers for the various business units of Qantas. It was on dual platforms of Oracle and Teradata. It involved building datamarts upon which the reporting layer was placed. Work was completed on 3 access layers which included Checkin, IOC PTS and EMD. These layers had a staging layer which included extract and transform jobs from source to staging based on incremental extract. This was followed by a presentation layer where business logic was applied and data was inserted into fact and dimension tables. This data was modelled on star schema design. The data from presentation layer was utilized for reporting purpose by the business units.

**Sparq Solutions Energex – Consultant Apr’15 – Jul’15**

**Skills Used : IBM Datastage 11.3, IBM Governance Catalogue, Oracle, Unix, IBM Cognos Framework**

**Manager, IBM Cognos Report Studio, MS SSIS,SSRS,SSAS**

This project primarily deals with the energy sector data mart PowerON. It is an existing data mart and my role is to carry out further enhancements and bug fixes. It consists of a star schema with a network fact table dealing with outage information. The enhancement is to develop a joined star schema with some conformed dimensions called outage losses accumulated from the outages. There is also enhancement work with respect to the network outage data mart like fixing decimal metrics in the calculations, adding fact metrics etc. It also included creating Query Subjects in Cognos Framework Manger from the objects developed. These Query Subjects were exported as packages to be used within Congos Report Studio.

**Vodafone – Consultant Dec’14 – Mar’15**

**Skills Used : IBM Datastage 11.3, Oracle, Unix, MS SSIS,SSRS,SSAS**

This project involved migration from datastage version 7.5 to 11.3. This was a software upgrade that was a quantum jump in the platform. It involved installation of the new software on a dedicated server. Then the migration of jobs from the old system to the new server. These jobs were compiled and kinks like stages and functions that had become obsolete were replaced with the latest equivalent version. This was followed by unit testing of the modified jobs with sample data. Then system testing end to end was conducted to gauge the smooth process flow. This was conducted in DEV environment followed by SIT for more comprehensive testing. Involved interaction with support personnel, testing team and DBAs for the implementation of this project.

**Oakton Consulting - Sydney Trains Feb’14 –Nov’14**

**Skills Used : IBM Infosphere Datastage, IBM Data Architect, Oracle SQL/PLSQL, Unix Shell Scripting,**

**IBM Cognos Framework Manager, IBM Cognos Report Studio, Microsoft SSIS, Teradata,**

**DB2, MS SQL Server, MS SSIS,SSRS,SSAS**

Designed and developed Batch Control Framework for automating the job run process. This was done in dual technologies of Microsoft SSIS and Datastage. It included a batch information repository that was in the database. Features included in the framework were parallel execution of multiple applications, parallel execution of jobs for performance, re-start ability of application in case of failure, robust emailing system, scheduling of jobs with dependencies. This was developed on multiple database platforms Oracle, Teradata, SQL Server. These versions were deployed at various client sites including Sydney Trains, Canon.

Design and development of Sydney Trains data warehouse. Work involved creating data mart for their incident reporting. It also included performance tuning the existing warehouse system. This was achieved by a variety of methods including sparse lookups, incremental references, Partitioned reads, Partition Scan and Rowid Update DMLs etc.

The data mart was sourced from their data warehouse tables. It was arranged in Star Schema using Kimball methodology. It was integrated with the warehouse load to use the warehouse increment dataset for loading incident data. The functionality was to record the incidents that occurred during a given trip that lead to delays.

The incident factors were calculated based on these incidents by recoding the number of trips affected by the incident and the cumulative delay caused to these trips. This was used in Cognos reporting of incident delay factor affecting delayed trips.

It also included creating Query Subjects in Cognos Framework Manger from the objects developed. These Query Subjects were exported as packages to be used within Congos Report Studio. There was also some reverse engineering of cognos report function logic to be incorporated into the ETL layer.

**Enterprise Business Services - Melbourne**  **Sept’13 – Jan’14**

**Skills Used : IBM Datastage, Oracle, Teradata, Unix**

This was a Design Role for their warehouse. This included ETL design and development of dimension and fact tables of their gas network database. The metadata included suppliers, cleints, measures of their supplies and their billing information. This was designed using Kimball methodology of Star schema with dimensions and facts. The ETL design was in the parallel version. It also included transitioning some jobs from server to parallel. The DW architecture had some jobs on Oracle and some on Teradata platforms.

**Deakin University, Melbourne – Consultant May’13 – Jul’13**

**Skill Used : IBM Infosphere Datastage, IBM Data Architect, Oracle SQL/PLSQL, Unix Shell Scripting**

Developing datamarts for Deakin for their Staff Profile and Payroll system. It involves moving data from their source system to the data warehouse and from the warehouse to the mart.They were based on the Kimball Methodology star schema configuration of dimensions and fact. The dimensions are designed as Type1 and Type2 for maintaining history. The loads are incremental daily and monthly based on a time dimension.

Production rollout included streamlining and optimizing the batch runs of the datamart jobs for fast run time. For this the jobs for the different marts were parallelized at the batch level and the granular level. Incremental jobs were designed to reduce data load processed. Other job specific measures like multi-node partitioning, sparse lookups were adopted for individual jobs.

**Suncorp Bank, Brisbane – Consultant Sep’12 – Apr’13**

**Skill Used : IBM Infosphere Datastage, Metadata Workbench, Cognos, DB2, Netezza, Oracle,**

**Unix Shell Scripting**

Developed data marts Deposit Accounts, Retail Lending, Risk Weight Assets for the banking division of Suncorp. These marts were sourced from the data warehouse and were used for reporting purpose. They were based on the Kimball Methodology star/snowflake schema configuration of dimensions and fact. The dimensions were designed as Type1, Type2 and Type6 for maintaining history. There were some junk dimensions where different combinations of codes were input. The loads were incremental monthly and weekly based on a period dimension.

These were data marts based on business requirements for financial data in the warehouse. These requirements were customized in the mart.

My role included development, unit testing, bug fixing and enhancements for these marts. I was also guiding onshore and offshore resources. I had to interact with the data analysts in the design of the data model and the eventual delivery of the mart.

**General Motors**  **Feb’11 – Aug’12**

**Skills Used : Oracle, Unix, Datastage, Autosys**

Datawarehouse project for GM inventory & payroll system. Mainframe feed into warehouse in the form of flat files. Data also being received from the web systems.

GMAC project involving design, development of retail and wholesale warehouse.

Origination warehouse project for GM.

Inteis project for ALLY bank for Annualized Metrics.

**Assurant Health Corporation Feb’08 – Feb’11**

**Skills Used : oracle, sqlserver, datastage, unix, autosys**

Datawarehouse project pulling data from mainframe flat files, sqlserver, oracle source systems and moving them to oracle datawarehouse. Smaller datamarts based on different subject areas getting data from the warehouse. Design, development, migration and production support work done.

**Vodafone**  **Jan’07 – Jan’08**

**Skills Used : Oracle, Datastage**

Data warehouse populated by moving data from the client source Oracle database to the data

warehouse using Datastage, SQL and PLSQL scripts. Scheduling of table data loads and

transformation of data for reporting purpose done.

**Information Analyst ,** **Electronic Data Systems Corporation (EDS) Apr’05 –Aug’06**

**On site for Towers Perrin, Philadelphia, USA**

**Oracle Data warehouse development**

**Skills Used : Oracle, Datastage, Unix**

Loading of data warehouse using PLSQL and SQL scripts. Client and Line of business was used to

define schema integrity to store data. This was an ongoing development effort for Towers Perrin

clients. Unix shell scripts were used to call database stored procedures and packages to move data from the online system to the warehouse using the loading system described below. Datastage was used to extract , transform and load data from the online database into the warehouse.

**Enterprise Data Loader Program**

**Skills Used : Oracle, Unix**

Development and administration of unix and Oracle based job scheduling system to process data warehouse load jobs. The job process steps to be executed are stored in a database repository schema. A unix based recursive script runs the jobs at the time defined in the database. There was a provision to run the job processes in parallel when that is possible saving time. The system is also robust with email notification available at the operating system and database level.

**Distributed environment CPU calculation**

**Skills Used : Oracle, Unix**

Application code was moved onto a new unix server. The database is sitting on another dedicated server. So command line jobs need to be submitted by the application server to the database server. Hence there was a need to calculate the CPU available on the database server before submitting jobs from the application server. This was done by having a recursive CPU calculating script execute on the database server at 10 second intervals and insert current CPU available value into a table on the database. The applications scripts File Loader and EDL were modified to read the value in the table to get the most recent available CPU before submitting jobs for execution on the database server.

**Hyperion 8.5 Install and Administration**

**Skills Used : Hyperion Performance Suite**

This included installing the foundation on a Windows 2003 machine. Also installed was an additional

job service on an unix server as part of a distributed environment. We used a tomcat server as an application server to connect to the foundation. Administration involves adding data sources, starting services and solving any technical problems experienced by users.

**Image transfer** **program**

**Skills Used : Oracle, Unix**

Developed to move data image graph files from on site production server to new servers located at Tulsa, Oklahoma as these servers are behind a local firewall and existing server cannot write output file to them. The program was written in MKS Korn shell and uses the windows scheduling system to check for newly written files. Then these files are written to a holding directory and moved to remote servers by opening ftp sessions. Html files used by the Hyperion reporting system look for these files and display them to users. The application was administered to change password, create new client accounts etc.

**Disaster Recovery Test**

**Skills Used : Oracle, Unix**

Completed testing of the disaster recovery systems located at Allen, Texas USA. The DR process is backing up files from existing production drives onto a different set of servers located at Allen. We changed the settings on the DR servers to be able to start the required services that comprise the Brio Portal. After starting the portal we tested reports and brio query documents belonging to different clients in co-ordination with the client delivery teams.

**Dynamic Dashboard**

**Skills Used : Hyperion Performance Suite**

This was designed to generate a jpeg image of the messages that the different users of a client can see, for updates on the status of data processing. So if a file was submitted and is being processed the users should be able to know the status of the file. This was done by creating an SQR report to input the message that needs to get sent out. This input was inserted into a table in the client schema via the DO-SQL procedure in SQR. After inserting the data into the table a BQY job that created the image on the last 20 rows in the table can be called via a link in the report. When this BQY is run a jpeg file is created and deposited on the external file server. This image is visible by uploading an html file from the personal page content folder under broadcast messages to the personal pages of the users in the client group. This page is written so that it refreshes every 10 minutes to get the latest message.

**Software Developer,** **Towers Perrin**, **Philadelphia, USA** **Aug’99 – Mar’05**

**Hyperion Performance Suite 8.2.1 Installation and Administration**

**Skills Used : Hyperion Performance Suite**

Involved with the installation, migration of the client delivery reporting system (outlined below) to the new Hyperion Performance Suite. The new system was installed on a fresh set of servers. This involved moving the old client content to the new system. At present we have both the environments operational and are gradually phasing out the old system usually client at a time. Involved in resolving technical issues with the new system.

**Disaster Recovery System**

**Skills Used : Oracle, Unix**

Involved with the development and testing of disaster recovery servers in line with production servers. This includes content on windows and unix servers and software upgrades to the production servers.

Testing of servers is from time to time to see that the recovery servers are available when needed.

**File Loader Program**

**Skills Used : Oracle, Unix**

Development and administration of an automated system to receive data files from clients on a Windows machine. A MKS unix based shell process is run by the windows scheduler every 5 minutes to check for data files. If more than one file is received multiple ftp sessions are opened to move the file to a unix server where recursive script keeps checking for files. The data in this file is processed by appropriate shell programs and then moved to the data warehouse. The files received are zipped using the winzip command line interface onto a storage drive on the windows machine. This program automated the file movement process and made it easy for us to be able to keep track of the different files that have been loaded by different users at different times.

**Client Delivery Reporting (CDR) using Brio.**

**Skills Used : Brio**

Design of a complicated system using Brio’s limited front-end capabilities for all our

clients. Web access was the key in this implementation.

Consequently, security was a big issue too. Giving clients the ability to see their reports and to be able

to schedule them over the web in a secure environment was successfully carried out. Clients were

given access to dynamic reports (YTD, MTD & Adhoc) and Prior Period report instances. The data

for these reports was extracted from the **Integrated database management System**.

Customizing the look and feel of the portal led us to introduce a lot of changes to the JavaScript &

html files used for the purpose. Coding had to be carried out in the native SQR of the Brio Reporting

Engine to be able to enhance reporting capabilities.

**Data warehousing for the Integrated Database Management System (IDMS)**

**Skills Used : Datastage, Oracle Designer**

System was developed to include data from various databases like Health & Welfare, Deferred Benefits etc into a composite data model for the entire system using DataStage. After loading all available history records a process was developed to load data on a daily basis. A module was also developed to check data consistency and correct incongruities across different databases. This system was developed by a four member team including myself.

**Data modeling for the Generic Pension Fund Scheme.**

**Skills Used : Oracle Designer**

Planning of schemas, objects therein, object outlay.

Entity relationship modeling, Business & Functional Hierarchy modeling using Oracle Designer.

Study of samples of data of the various data files for compatibility levels with the generic model being

developed and solutions desired to meet current standards.

Include Object Oriented and other features available with current Technology in the data model to

improve robustness, integrity, security and other features of data.

Implementing the generic data model in the Oracle 8i database.

PROJECTS at Villanova University:

Designed and developed a program in C, for editing text on the system output.

The program provided the user with all the editing options such as cut, copy, paste and inserting text into the file etc. This was a project for Data Structures and Algorithm course.

Developed a program in C which analyses the different Sorting and Searching algorithms by calculating their best, worst and average case efficiencies in C Language with respect to time and space, and found their optimal efficiencies under various conditions.

Developed a program in C, which basically simulates the super market check-in system using the queue data structure.

Developed a web server prototype using C for the Operating Systems course.

Developed a Photo-Album implementation for the Object Oriented Programming Course in JAVA.