Big Data Engineer Big Data Engineer Big Data Engineer Canton, OH   Around 8 years of professional IT experience in the fields of Big Data, Business Intelligenceand Java in Financial, Insurance and Digital Services Industries   Hands-on experience with major components in Hadoop Ecosystem like Map Reduce, HDFS, YARN, Hive, Pig, HBase, Sqoop, Oozie, Cassandra, Impala and Flume   Experience with new Hadoop 2.0 architecture YARN and developing YARN Applications on it   Experience with Apache Spark's Core, Spark SQL, Streaming and MlLib components   Experience with distributed systems, large-scale non-relational data stores and multi-terabyte data warehouses   Firm grip on data modeling, database performance tuning and NoSQL map-reduce systems   Experience in processing semi-structured and unstructured datasets   Responsible for setting up processes for Hadoop based application design and implementation Experience in managing HBase database and using it to update/modify the data   Experience in running MapReduce and Spark jobs over YARN   Experience with Cloudera, Hortonworks and MapR distributions   Handling data in various file formats such as Sequential, AVRO, RC, Parquet and ORC   Strong knowledge on the scalability and applications of Spark and its components - Core, SQL and Dataframes   Hands-on experience in complete project life cycle (design, development, testing and implementation) of Client Server and Web applications   Involved in developing complex ETL transformation & performance tuning   Extensively worked with Teradata utilities like BTEQ, Fast Export, Fast Load, Multi Load to export and load data to/from different source systems including flat files   Hands on experience using query tools like Teradata SQL Assistant, TOAD, PLSQL developer and Query man   Experience in Object Oriented Analysis and Design (OOAD) and development of software using UML Methodology, good knowledge of J2EE design patterns and Core Java design patterns   Experience using middleware architecture using Sun Java technologies like J2EE, JSP, Servlets, and application servers like Web Sphere and Web logic Authorized to work in the US for any employer Work Experience Big Data Engineer SunGard Financial Systems - New York, NY August 2014 to December 2015 Description: Kiodex Risk work bench is a SaaS solution that helps trading firms, banks/FCM's, and corporations manage commodities exposures through front, middle and back office functions including quantifying portfolio

risks via different exposure reports. Responsibilities: Extracted data from relational databases such as SQL Server and MySql by developing Scala and SQL code Uploaded it to Hive and combined new tables with existing databases Developed code to pre-process large sets of various types of file formats such as Text, Avro, Sequence files, XML, JSON and Parquet Configured big data workflows to run on the top of Hadoop which comprises of heterogeneous jobs like Pig, Hive, Sqoop and MapReduce Loaded various formats of structured and unstructured data from Linux file system to HDFS Used Combiners and Partitioners in MapReduce programming Written Pig Scripts to ETL the data into NOSQL database for faster analysis Read from Flume and involved in pushing batches of data to HDFS and HBase for real time processing of the files Parsing XML data into structured format and loading into HDFS Scheduled various ETL process and Hive scripts by developing Oozie workflow Utilized Tableau to visualize the analyzed data and performed report design and delivery Created POC for Flume implementation Involved in reviewing both functional and non-functional aspects of the business model Championed to communicate and present the models to business customers and executives, using the same Environment: Hadoop, HDFS, Map Reduce, Sqoop, HBase, Shell Scripting, PIG, HIVE, Oozie, Core Java, Hortonworks Distribution, LINUX Business Intelligence/ETL Developer Synacor - Buffalo, NY January 2013 to July 2014 Description: Synacor is a provider of private label digital Internet services to North America's cable and telecom companies. It's a global technology company providing Internet solutions to enable ISPs, media companies and advertisers to build close relationships with consumers on the Internet. Theintranet application represents a near term opportunity in which thecompound would move quickly into Development. The purpose of this application is to create a simplified process where ideas are generated, tracked and reviewed using existing governance. Responsibilities: Involved in design & development of operational data source and data marts in Oracle Reviewed source data and recommend data acquisition and transformation strategy Involved in conceptual, logical and physical data modeling and used star schema in designing the data warehouse Designed ETL process using Teradata to load the data from various source databases and flat files to target data warehouse in Oracle Used Power mart Workflow Manager to

design sessions, event wait/raise, and assignment, e-mail, and command to execute mappings

Created parameter based mappings, Router and lookup transformations   Involved in migration projects to migrate data from data warehouses on Oracle/DB2 and migrated those to Teradata Optimized mappings using transformation features like Aggregator, Filter, Joiner, Expression and Lookups   Created daily and weekly workflows and scheduled to run based on business needs Environment: Data modeling, SQL Server SSIS, SSRS, Oracle 10g, Teradata 6, XML, TOAD, SQL, PL/SQL, IBM AIX, UNIX Shell Scripts, Web Intelligence, DSBASIC, Cognos, Erwin, STAR team, Remedy, Maestro job scheduler, Mercury Quality Center, Control-M Java Developer Emorphosys August 2011 to December 2012 Description: Emorphosys Software Solutions is a global technology based product development company which is a pioneer in creating simple, scalable products that focus on business transformation through innovation. It provides technology based solutions in the E-Learning, Digital Media and Enterprise Content Management domains.   Responsibilities: Involved in the core product development using J2EE, JSF and Hibernate   Actively involved in the full life cycle Object Oriented application development - ObjectModeling, Database Mapping, GUI Design   Used JavaScript to perform client side validations and Struts-Validator framework for server-side validation   Worked on requirement gathering, high level design and Waterfall model to get best result   Created data access using SQL and PL/SQL stored procedures   Used Hibernate annotations with Java for various stages in the application   Built web services upon SOAP to export and import attachments from file to associated applications   Developed DAO (data access objects) using Spring Framework   Deployed the components in to WebSphere Application server Used HTML/CSS and JavaScript for UI development   Written sql queries including Joins, Triggers, Stored procedures, Views using MySql   Implemented the JSPs and EJBs in the JSF Framework to handle the workflow of the application   Developed Unit Test Cases, used JUnit for unit testing of the application Environment: Java, J2EE, Struts, SQL, JAX RPC, XML, RAD, Websphere, MQ, Agile, JSPS, SOAP Education Bachelors in Computer Science VIGNAN University - El Segundo, CA September 2017 to Present Skills DATABASES (4 years), JAVA (4 years), PL/SQL (4 years), SQL (4 years), XML (4 years) Additional Information TECHNICAL SKILLSET: Big Data HDFS,

MapReduce, Hive, Pig, ZooKeeper, Apache Spark, Core, MlLib, Spark SQL and Dataframes Utilities Sqoop, Flume, Kafka, Oozie and AutoSys No SQL Databases Hbase, Cassandra Languages C, C++, Java, J2EE, PL/SQL, MR, Pig Latin, HiveQL, Unix shell scripting and Scala Operating Systems Sun Solaris, RedHat Linux, Ubuntu Linux and Windows XP/Vista/7/8 Web Technologies HTML, DHTML, XML, AJAX, WSDL, SOAP Databases and Datawarehousing Teradata, DB2, Oracle 9i/10g/11g, SQL Server, MySQL Tools and IDE Maven, Toad, Eclipse, NetBeans, ANT, Hudson, Sonar, JDeveloper, Assent PMD, DB Visualizer

Name: Tina Jensen

Email: sandra60@example.org

Phone: +1-316-717-1523x407