Sr. SCALA Developer Sr. SCALA Developer Sr. SCALA Developer - Instacart, Inc Palo Alto, CA Authorized to work in the US for any employer Work Experience Sr. SCALA Developer Instacart, Inc - San Francisco, CA December 2016 to Present operates as a same-day grocery delivery service company. The Company allows customers to select groceries through an online application platform from various retailers and have them delivered by personal shoppers. Worked on combining a personal touch and cutting-edge technology to offer customers a simple solution to save time and eat fresh food from the most trusted grocery brands.   ? Developed Spark Applications by using Scala, Java and Implemented Apache Spark data processing project to handle data from various RDBMS and Streaming sources. ? Developed Spark Programs using Scala and Java API's and performed transformations and actions on RDD's. ? Involved in converting Hive/SQL queries into Spark transformations using Spark RDD, Scala and Python. ? Develop ETL Process usingSPARK, SCALA, HIVE and HBASE. ? Developed REST APIs using Scala, Play framework and Akka. ? Used ScalaTest for writing test cases and coordinated with QA team on end to end testing. ? Developed REST APIs using Scala and Play framework to retrieve processed data from Cassandra database. ? Developing UDFs in java for hive and pig and worked on reading multiple data formats on HDFS using Scala. ? Involved in converting Hive/SQL queries into Spark transformations using Spark RDDs and Scala. ? Used Scala collection framework to store and process the complex consumer information. ? Used Scala functional programming concepts to develop business logic. ? Developed programs in JAVA, Scala-Spark for data reformation after extraction from HDFS for analysis. ? Developed Spark scripts by using Scala shell commands as per the requirement. ? Processed the schema oriented and non-schema-oriented data using Scala and Spark. ? Developed Scala scripts, UDFFs using both Data frames/SQL/Data sets and RDD/MapReduce in Spark 1.6 for Data Aggregation, queries and writing data back into OLTP system through Sqoop. ? Provided architecture and design as product is migrated to Scala, Play framework and Sencha UI ? Implemented applications with Scala along with Akka and Play framework. ? Expert in implementing advanced procedures like text analytics and processing using the in-memory computing capabilities like Apache Spark written in Scala. ? Auction web app - calculated bids for energy auctions utilizing

Scala, JPA and Oracle. ? Built Kafka-Spark-Cassandra Scala simulator for MetiStream, a big data consultancy; Kafka-Spark-Cassandra prototypes. ? Developed a Restful API using & Scala for tracking open source projects in Github and computing the in-process metrics information for those projects. ? Developed analytical components using Scala, Spark, Apache Mesos and Spark Stream. ? Experience in using the Docker container system with the Kubernetes integration ? Developed a Web Application using Java with the Google Web Toolkit API with PostgreSql ? RedisCreating a dashboard using Flask, Python libraries, and AngularJS to visualize their progress. ? Improve site performance by making better use of caches via MemCached.on Amazon Web Services. ? Used R for prototype on a sample data exploration to identify the best algorithimic approach and then wrote scala scripts using spark machine learning module. ? Developed MapReduce/Spark Python modules for machine learning & predictive analytics in Hadoop on AWS. Implemented a Python-based distributed random forest via Python streaming. ? Utilized Spark, Scala, Hadoop, HBase, Kafka, Spark Streaming, Caffe, TensorFlow, MLLib, Python, a broad variety of machine learning methods including classifications, regressions, dimensionally reduction etc. ? Excellent understanding / knowledge of Hadoop architecture and various components such as HDFS,Hbase, Hive Job Tracker, Task Tracker, Name Node, Data Node and Map Reduce programming paradigm ? Built Kafka-Spark-Cassandra ScalaFX simulator for MetiStream, a big data consultancy; Kafka-Spark-Cassandra prototypes. ? Designed a data analysis pipeline in Python, using Amazon Web Services such as S3, EC2 and Elastic Map Reduce ? Implemented applications with Scala along with Akka and Play framework. ? Expert in implementing advanced procedures like text analytics and processing using the in-memory computing capabilities like Apache Spark written in Scala. ? Designing a highly scalable, highly available, minimum TCO for maximum ROI using big data components like kafka,spark,cassandra,mongoDB, and API. It is python and scala based analytic system with ML Libraries. ? Worked with NoSQL Platforms and Extensive understanding on relational databases versus No-SQL platforms. Background Skills: Scala, Play, Akka, Java, Python, Kafka, Scala, Spark, Hadoop, Hive, Big Data, HBase, HDFS, Sqoop, MapReduce, Pig, Docker, PostgreSQL, JSON, Redis, Memcache, AWS. Sr. SCALA

Developer JobNimbus - American Fork, UT January 2014 to November 2016 The project was to work on JobNimbus CRM (Customer Relationship Management) and Project Management software all wrapped into one. JobNimbus goal was to offer clients the easiest way to manage projects, contacts, and tasks. With detailed reporting and employee management tools.I have worked on customizing workflow and define the sequence of work. With pre-built custom forms, users can create and share invoices. Managed all their contacts at one central place. Contacts can be imported from multiple sales channels including emails and phone calls, that allows project managers to assign tasks to their teams, monitor progress and send notifications to clients and stakeholders.   ? Involved in converting Hive/SQL queries into Spark transformations using Spark RDDs and Scala. ? Developed multiple POCs using Scala and deployed on the Yarn cluster, compared the performance of Spark, with Hive and SQL/Teradata. ? Analysed the SQL scripts and designed the solution to implement using Scala. ? Developed analytical component using Scala, Spark and Spark Stream. ? Involved in performing the Linear Regression using Scala API and Spark. ? Experience in developing and designing POCs using Scala and deployed on the Yarn cluster, compared the performance of Spark, with Hive and SQL/Teradata. ? Performed advanced procedures like text analytics and processing, using the in-memory computing capabilities of Spark using Scala. ? Created various Parser programs to extract data from Business Objects, XML, Informatica, Java, and database views using Scala ? Developed Spark code to using Scala and Spark-SQL for faster processing and testing. ? Designed scalable (scala) Web Architecture hosting reports for the entire application. ? Wrote entities in Scala and Java along with named queries to interact with database. ? Implementing a search microservice (Scala, REST, PlayFramework, ElasticSearch) ? Designed a distributed system using Scala and the AKKA Actor Model that runs on multi-core machines. The server and client(s) are designed to share the work. Server can work independent of a client. ? Wrote a tool for extracting and reconciling trading data from Warehouses. The tool is used for data consistency and data quality monitoring. The tool is written using Java, Scala, Play and Akka Errors flagged by the tool are picked up in Splunk and Sitescope. ? Wrote coding highly flexible, scalable, & distributed applications using Scala. ? Created and consuming

RESTful Web services in Scala, Play using AKKA. ? Automated Compute Engine and Docker Image Builds with Jenkins. ? Worked with various schema for Application, Data Processing and Data warehouse that resides in AWS RDS database (PostGreSql), Dynamo DB. ? Interacted with DB sharding, Redis, Jenkins, SOLR, GraphQL, Grafana, Click Tracking for analytics. ? Designed a persistent versus transient architecture - raw Linux server with Spark ML algorithm jobs, test Spark ? Worked in Big Data technologies Hadoop, HDFS, Map Reduce, Kafka, Sqoop, Flume, BigSQL, Hive, Pig, Hbase, and Apache Spark, Spark ML, MLlib as developer and analyst. ? Wrote Spark-ML, Map Reduce programs, HiveQL and PigLatin scripts leading to good understanding in Map Reduce design patterns, data analysis using Hive and Pig. ? Utilized Apache Spark with Python to develop and execute Big Data Analytics and Machine learning applications, executed machine Learning use cases under Spark ML and Mllib. ? Developed different Statistical Machine Learning, Data Mining solutions to various business problems and generating data visualizations using R, SAS and Python and creating dashboards using tools like Tableau. ? Explored MLlib algorithms in Spark to understand the possible Machine Learning functionalities that can be used for our use case ? Utilized Spark, Scala, Hadoop, HBase, Kafka, Spark Streaming, MLLib, Python, a broad variety of machine learning methods including classifications, regressions, dimensionally reduction etc. ? Worked on different data formats such as JSON, XML and performed machine learning algorithms in R and used Spark for test data analytics using MLLib and Analyzed the performance to identify bottlenecks. ? Imported data using Sqoop to load data from MySQL to HDFS on regular basis. ? Developed Scripts and Batch Job to schedule various Hadoop Program. ? Used Spark API over Hortonworks, Hadoop YARN to perform analytics on data in Hive. ? Developed Kafka producer and consumers for message handling. ? Used Amazon CLI for data transfers to and from Amazon S3 buckets. ? Executed Hadoop/Spark jobs on AWS EMR using programs, data stored in S3 Buckets. ? Exploring with Spark improving performance and optimization of the existing algorithms in Hadoop MapReduce using Spark Context, Spark-SQL, Data Frames, Pair RDD's and Spark YARN. ? Deployed MapReduce and Spark jobs on Amazon Elastic MapReduce using datasets stored on S3. ? Used Amazon Cloudwatch to monitor and track

resources on AWS. ? Knowledge of designing and deployment of Hadoop cluster and different Big Data analytic tools including Hive, HBase, Oozie, Sqoop, Flume, Spark, Impala, Cassandra. ? Responsible for developing data pipeline using Flume, Sqoop and Pig to extract the data from weblogs and store in HDFS. ? Involved in converting Hive/SQL queries into Spark transformations using Spark RDDs, Scala. ? Involved in importing the real-time data to Hadoop using Kafka and implemented Oozie jobs for daily imports. ? Automated the process for extraction of data from warehouses and weblogs by developing work-flows and coordinator jobs in Oozie. Background Skills: Scala, Akka, Play, Java, Python, , Spark, Hadoop, Hive, Big Data, Hbase, HDFS, Oozie, Hibernate, Spring, Angular.js, Node.js, bootstrap.js, backbone.js, JSP, Struts, JDBC, HTML, CSS, JQuery, Javascript, XML. Sr. ML Developer Homedepot.com - Atlanta, GA March 2011 to December 2013 Re-platform project is a complete rewrite of the existing site that runs in a Broadvision environment. This site allows customers to browse online and in-store products, add products to a shopping cart, wish list, or gift registry, create a user account, opt for email subscriptions, place orders, compare, and search for products. I have also worked on Tax Exempt application that consists of two main parts; Registration and Maintenance. Tax Exempt Registration is a customer facing application where customers can register for tax-exempt status for their eligible tax-exempt purchases. Maintenance is being used by the Tax Department at the home depot to review, search, and maintain tax exemption reasons and other activities. ? Worked in loading and analyzing large datasets with Hadoop framework (MapReduce, HDFS, PIG, HIVE, Flume, Sqoop, SPARK, Impala, Scala), NoSQL databases like MongoDB, HBase, Cassandra. ? Involved in start to end process of Hadoop jobs that used various technologies such as Sqoop, PIG, Hive, MapReduce, Spark and Shellscripts (for scheduling of few jobs) extracted and loaded data into DataLake environment (AmazonS3) by using Sqoop which was accessed by business users and data scientists. ? Manage and support of enterprise Data Warehouse operation, big data advanced predictive application development using Cloudera &Hortonworks HDP. ? Developed PIG scripts to transform the raw data into intelligent data as specified by business users. ? Utilized Apache Spark with Python to develop and execute Big Data Analytics and Machine learning applications, executed machine

Learning use cases under Spark ML and Mllib. ? Involved in designing and deployment of Hadoop cluster and different Big Data analytic tools including Pig, Hive, HBase, Oozie, ZooKeeper, SQOOP, flume, Spark, Impala, and Cassandra with Horton work Distribution. ? Installed Hadoop, Map Reduce, HDFS, and AWS and developed multiple MapReduce jobs in PIG and Hive for data cleaning and pre-processing. ? Developed Spark/Scala, Python for regular expression (regex) project in the Hadoop/Hive environment with Linux/Windows for big data resources. ? Assisted in upgrading, configuration and maintenance of various Hadoop infrastructures like Pig, Hive, and Hbase. ? Used SparkAPI over Hortonworks Hadoop YARN to perform analytics on data in Hive. ? Improved the performance and optimization of the existing algorithms in Hadoop using Spark Context, Spark-SQL, Data Frame, Pair RDD's, Spark YARN. ? Deployed this application which uses J2EE architecture model and Struts Framework first on Weblogic and helped in migrating to JBoss Application server. ? Worked in Java, J2EE, XSL, XML, Oracle, DB2, Struts, spring, Hibernate, REST Web services, Model driven architecture and software configuration management tools. ? Developed Application based on J2EE using Hibernate, Spring, JSF frameworks and SOAP/REST web services, Web Sphere Integration Developer (WID) Tool to develop WPS components. ? Used Spring Framework for Dependency injection and integrated with the EJB using annotations. ? Responsible for analysis, design, development and integration of UI components with backend using J2EE technologies such as Servlets, Java Beans and JSP. ? Design and development of the exception management workflow using Oracle BPM. ? Deployed the applications in Linux servers using deployment scripts. ? Designed and developed programs in C++ to integrate as per the users requirements. ? Created/translated PL/I programming into SAS, which were used as part of the process used to standardize military personnel records ? Created PL/SQL stored procedures for new Oracle Forms and Reports development   Background Skills: Machine Learning, MapReduce, HDFS, PIG, SPARK, Impala, Scala, Java, J2EE, Spring, Struts, JSF, JSP, EJB, DOJO, JQuery, Sencha ExtJS, JavaScript. Java Developer WellsFargo - San Francisco, CA July 2009 to February 2011 Wells Fargo & Company I have worked on its application that primarily helps the Operational/Risk/Compliance team to report & analyze operational loss events and their

impacts across the bank operations globally. Loss events are either grouped as linked losses or non-linked losses according to the nature of the events. Also worked on its features to record Root causes which resulted in the loss event and come up with Action plans to address such loss events. The loss events are then associated with multiple Processes, Risk & Control measures based on the loss and insurance recovery amounts.   ? Integrated Hibernate ORM with Spring-Hibernate framework to facilitate DML and DQL queries and represent object-database mapping. ? Involved in transforming the Use Cases into Class Diagrams, Sequence Diagrams and State diagrams. ? Involved in development of Web Services, creation of WSDL and schemas. ? Extensively participated on working with Spring framework. Involved in writing JSP and Servlets. ? Involved in development of Web Services, were developed to receive client requests. ? Implemented Spring JDBC template, Spring Exception Strategy and AOP. ? Involved in setting up WebSphere Application server and using Ant tool to build the application and deploy the application in WebSphere Application server. ? Worked with the creation of Store Procedures. Involved in writing SQL queries, Stored Procedures to accomplish complex functionalities. ? Part of team creating quality working J2EE code to design, schedule, and cost to implement use cases. ? Developed Reusable classes in the middleware using Hibernate. ? Involved in writing lots of JSP for maintains and enhancements of the application. Worked on Front End using Servlets and also backend using EJB and Hibernate. ? Worked on Presentation Layer using Struts Tiles, JSPs and Servlets. ? Created quality working J2EE code to design, schedule, and cost to implement use cases. ? Setting of DB2 build settings in RAD application development server. ? Involved in writing the database integration code using Hibernate. ? Creation of managed server and JDBC connections. ? Worked on the application using Rational Application Developer. Designed & Developed Application flow UML diagrams of the application using Rational Rose   Background Skills: J2ee, Java, Jsp, Servlet, Jdbc, Struts, Junit, Log4j, Javascript, Websphere Application Server, Axis, Wsad, Xml, Xslt, Ant, Sql, Sql Query Analyzer, Jprobe, Cvs, Opprox Reports, Windows'xp, Unix-Ibm Aix. Java Developer OxfordHealthcare - Broken Arrow, OK February 2007 to June 2009 OxfordHealthcare The Project was to enhance and integrate Commercial off the shelf (COTS) web based Electronic Health Record

(EHR) NextGen Software system and provide a better approach to documentation than conventional, paper-based system. The project involved building a web-based application, and was created to provide online access to the components of disease management programs and staging process. Provides direction necessary to clinical and support staff to improve financial, revenue performance, and the overall patient experience.   ? Worked with Spring Batch Used Spring ORM module to integrate with Hibernate. ? Developed the web pages using JSP, CSS and HTML ? Developed the RTM interface module to map the requirements to the test-case and Test design modules ? Used several J2EE Design Patterns (Session Fa ade, Aggregate Entity) for the Middle Tier development. ? Developed EJBS (Session and Message-Driven Beans) in (RAD) for handling business processing, database access and asynchronous messaging. ? Made extensive use of Java Naming and Directory Interface (JNDI) for looking up enterprise beans. ? Developed Message-Driven beans in collaboration with Java Messaging Service (JMS). ? Involved in writing JSP/HTML/JavaScript and Servlets to generate dynamic web pages and web content. ? Wrote various stored procedures in PL/SQL and JDBC routines to update tables. ? Wrote various SQL queries for data retrieval using JDBC. ? Involved in building and parsing XML documents using SAX parser. ? Exposed business logic as a web service and developed WSDL files for describing these web services. ? Extensively used SOAP formatted messages for communication between web services. ? Developed the application on IBM WebSphere Application Server. ? Developed the plug-in interfaces for the TMS features (TEE, Requirements, Version Control) ? Developed Form Beans, which are used to store data when the user submits the HTML form ? Coded various Java beans to implement the business logic ? Development of GUI using AWT ? Involved in creating the tables using SQL and connectivity is done by JDBC ? Involved in generating the reports regarding the marks they secured in the online test once they press the submit button in the test using HTML and JSP. ? Apache Tomcat is used as an Application Server.   Background Skills: J2ee, Java, Jsp, Servlet, Jdbc, Struts, Junit, Log4j, Javascript, Dhtml, Websphere Application Server, Axis, Xml, Xslt, Ant, Sql, Opprox Reports, Windows'xp. Education Bachelor's Skills Apache hadoop hdfs (8 years), Apache hbase (8 years), Aws. (8 years), Hadoop (8 years), Java (10+

years), Scala (6 years), Play (4 years), Python (6 years), Kafka (3 years), Spark (4 years), Hive (4 years), BigData (5 years), Hbase (4 years), HDFS (3 years), Sqoop (4 years), MapReduce (4 years), Pig (4 years), Docker (3 years), PostgreSQL (4 years), JSON (3 years), Redis (5 years), AWS (6 years), MapR (3 years), MongoDB (4 years), Spark (3 years), Scala R (3 years), Redis (4 years), HDFS (3 years), Apache Mesos (3 years), PMML (4 years), Spring Security (2 years), Spring Batch (4 years), Solaris (4 years), AWS EMR (4 years), Azure (3 years), AWS Aurora (3 years), Apache mesos (4 years), Spring batch (4 years), Spring Security (3 years), Spring.Net (3 years), S3 (2 years), Microservices (4 years), Scala (3 years), Akka (3 years), Puppet (3 years), Chef (3 years), Ansible (2 years), Asp.Net (3 years), Mongodb (3 years), Akka-Http (3 years), AWS EMR (3 years), Kafka (2 years), Python (3 years), Asp.Net (3 years) Additional Information Areas of Expertise include: ? Scala ? Akka ? Hive/Big Data/Hbase ? Java ? Play ? AWS ? Spark ? Hadoop ? HDFS

Name: Thomas Holmes

Email: perryjessica@example.net

Phone: 357.321.8360x3008