+302-207-0005
info@uiskills.com

HADOOP DEVELOPER:

Hadoop Developer is a programmer who is actively involved in the development of Big Data applications. Developer has a vast knowledge of the various components of Hadoop framework. Hadoop developer job responsibilities which include design and develop Hadoop system with strong documentation skills. The job of a Hadoop developer is almost similar to the software developer but in the Big Data domain.

Hadoop Developer is responsible for:

- Hadoop development and implementation.
- Loading from disparate data sets.
- Pre-processing using Hive and Pig.
- Designing, building, installing, configuring and supporting Hadoop.
- Translate complex functional and technical requirements into detailed design.
- Perform analysis of vast data stores and uncover insights.
- Maintain security and data privacy.
- Create scalable and high-performance web services for data tracking.
- High-speed querying.
- Managing and deploying HBase.
- Being a part of a POC effort to help build new Hadoop clusters.

+302-207-0005
info@uiskills.com

- Test prototypes and oversee handover to operational teams.
- Propose best practices/standards.

Course Content:

Hadoop Architecture:

Introduction to

- Parallel Computer vs. Distributed Computing
- How to install Hadoop on your system
- How to install Hadoop cluster on multiple
- Hadoop Daemons introduction: NameNode, DataNode, JobTracker, TaskTracker
- Exploring HDFS (Hadoop Distributed File System) Exploring the HDFS Apache Web UI
- NameNode architecture (EditLog, FsImage, location of replicas) Secondary NameNode architecture
- DataNode architecture

MapReduce Architecture:

- Exploring JobTracker/TaskTracker

+302-207-0005
info@uiskills.com

- How a client submits a Map-Reduce job
- Exploring Mapper/Reducer/Combiner
- Shuffle: Sort & Partition
- Input/output formats
- Job Scheduling (FIFO, Fair Scheduler, Capacity Scheduler) Exploring the Apache MapReduce Web UI

Hadoop Developer Tasks:

- Writing a map-reduce programme
- Reading and writing data using
- Java Hadoop Eclipse integration
- Mapper in details
- Reducer in details
- Using Combiners
- Reducing Intermediate Data with Combiners
- Writing Partitioners for Better Load
- Balancing Sorting in HDFS
- Searching in HDFS
- Indexing in HDFS

+302-207-0005
info@uiskills.com

- Hands-On Exercise

Hadoop Administrative Tasks:

- Routine Administrative Procedures
- Understanding dfsadmin and mradmin Block Scanner, Balancer
- Health Check & Safe mode
- DataNode commissioning/decommissioning
- Monitoring and Debugging on a production
- cluster NameNode Backup and Recovery
- ACL (Access control list) Upgrading Hadoop

HBase Architecture:

- Introduction to HBase
- HBase vs. RDBMS
- Exploring HBase Master & region server
- Column Families and Regions
- Basic HBase shell commands.

+302-207-0005
info@uiskills.com

Hive Architecture:

- Introduction to Hive
- HBase vs Hive
- Installation of Hive
- HQL (Hive query language)
- Basic Hive commands

Pig Architecture:

- Introduction to Pig
- Installation of Pig on your system
- Basic Pig commands
- Hands-On Exercise

Sqoop Architecture:

- Introduction to Sqoop
- Installation of Sqoop on your system
- Import/Export data from RDBMS to HDFS
- Import/Export data from RDBMS to HBase

+302-207-0005
info@uiskills.com

- Import/Export data from RDBMS to Hive
- Hands-On Exercise

Mini Project / POC (Proof of Concept):

- Facebook-Hive POC
- Usages of Hadoop/Hive @ Facebook
- Static & dynamic partitioning
- UDF (User defined functions)

Our learning methods include:

- Comprehensive course selection of Instructor-Led Training
- Logistical convenience and interactive classroom experience of Online Training
- Flexible pacing and instructor-guided support of Mentored Learning
- Self-paced convenience of Online ANYTIME

In addition:

- Interview preparation with mock interview drills
- Effective resume building

- Process of applying jobs at the right places

Reach us:

Call : +1 302-207-0005

Email ID: info@uiskills.com

Website: www.uiskills.com