CLOUDERA SEARCH TRAINING (CLOUD-SEARCH)

Overview of Cloudera Search
Performing Basic Queries

- Executing a Query in the Admin UI
- Basic Syntax
- Techniques for Approximate Matching
- Controlling Output

Writing More Powerful Queries

- Relevancy and Filters
- Query Parsers
- Functions
- Geospatial Search
- Faceting

Preparing to Index Documents

- Overview of the Indexing Process
- Understanding Morphlines
- Generating Configuration Files
- Schema Design
- Collection Management

Batch Indexing HDFS Data with MapReduce

- Overview of the HDFS Batch Indexing Process
- Using the MapReduce Indexing Tool
- Testing and Troubleshooting

Near-Real-Time Indexing with Flume

- Overview of the Near-Real-Time Indexing Process
- Introduction to Apache Flume
- How to Perform Near-Real-Time Indexing with Flume
- Testing and Troubleshooting

Indexing HBase Data with Lily

- What is Apache HBase?
- Batch Indexing for HBase
- Indexing HBase Tables in Near-Real-Time

Indexing Data in Other Languages and Formats

- Field Types and Analyzer Chains
- Word Stemming, Character Mapping, and Language Support
- Schema and Analysis Support in the Admin UI
- Metadata and Content Extraction with Apache Tika
- Indexing Binary File Types with SolrCell

Improving Search Quality and Performance

- Delivering Relevant Results
- Helping Users Find Information
- Query Performance and Troubleshooting

Building User Interfaces for Search

- Search UI Overview
- Building a User Interface with Hue
- Integrating Search into Custom Applications

Considerations for Deployment

- Planning for Deployment
- Determining Hardware Needs
- Security Overview
- Collection Aliasing