Platform/Framework/Skills for Climate Analysis Implementation

* **Hadoop Distribution**: **Hortonworks 2.4.3**
* **Programming**: **Scala/Java/Java Script/HTML/CSS**
* **ETL:** **Spark Core/Spark SQL/Spark Streaming**
* **Data Ingestion**: Apache Nifi/linux script/wget
* **Data Storage:** **HDFS/MySQL/Cassandra/Redis/HBase**
* **Machine learning**: **Spark ML**
* **Data Format:** **Parquet/Avro**
* **Scheduler and Workflow:** **Oozie/Falcon**
* **Ad-Hoc Query:** **Spark SQL/Hive/Presto**
* **Data Access:** **JDBC/Spark Thrift Server/Hive Thrift Server/REST API/GoogleAPIS/Phoenix**
* **Notebook:** **Zeppelin**
* **Streaming Process:** **Spark Streaming/Kafka**
* **Visualization/Charting:** **DataMap/Kibana**
* **Search Engine:** **ElasticSearch**
* **Web Framework:** **Spring Boot/Spring MVC/Tomcat/Spring REST/Spring** Data
* **IDE:** **Intellij/Eclipse**