

Big Data Integration and Analysis

Documentation Strategy

Lab Mentor: Mohamed Nadjib Mami

Team Members:

Gaurav Kumar

Héctor Ugarte

Miguel Mármol

Tina Boroukhian

Methodology

Agile software development: Based on some



Extreme Programming practices:

Pair programming: 2 programmers work as a pair together on one workstation. The driver, writes code while the observer reviews each line of code as it is typed in. The two programmers switch roles frequently.

Continuous integration: The development team should always be working on the latest version of the software.

Coding standard: The standard specifies a consistent style and format for source code.

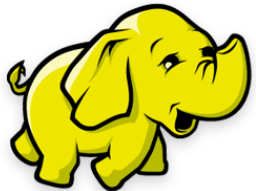
Testing: All code must have unit tests

Technical System Documentation Plan

...describes handling, functionality and architecture of the software product...

- Framework architecture: Hadoop MapReduce and Spark architecture.
- Coding: Code comments, naming conventions, simplicity, readability.
- Testing: Testing strategy, results.

```
/**  
 * Code Readability  
 */  
if (readable()) {  
    be_happy();  
} else {  
    refactor();  
}
```



MapReduce

Spark



End-user System Documentation(User Manual) Plan

...give assistance to people using the system...

- For users.
- Running the application.
- HowTo do the tasks.
- Examples.

...Introduction.
Configuration.
Doing data integration.
Doing data analysis...

