

WITS UNIVERSITY

SCHOOL OF ELECTRICAL AND INFORMATION ENGINEERING

ELEN7046 - SOFTWARE TECHNOLOGIES AND TECHNIQUES

Group Project: Big Data Visualization

Authors:

Gareth STEPHENSON
Matsobane KHWINANA
Sidwell MOKHEMISA
Dave CLOETE
Kyle TREHAEVEN

Student Number:

778919
779053
1229756
1573016
0602877N

SUMMARY

This report presents the work done by Group 2 in response to the project brief for ELEN7046: Software Technologies and Techniques. The report will broadly focus on the following topics:

- The Methodology followed to execute the project;
- The Architecture of the solution developed for the project; and
- The different technologies used to deliver the solution.

June 18, 2016

Contents

1	Abstract	2
2	Literature Review	2
3	Stakeholder Concerns - Big Data	2
3.1	Collection	2
3.2	Processing	2
3.3	Visualization	2
4	Methodology	2
5	Assumptions and Constraints	3
5.1	Tweet Locations	3
5.2	Pros and Cons determinations	3
6	Success Criteria	4
7	Architecture of the Solution	4
7.1	High Level Design: Component Architecture	4
7.2	Detailed Design: Class Diagrams	5
7.2.1	Data Acquisition (Batch)	5
7.2.2	Data Acquisition (Streaming)	5
7.2.3	Data Processing	5
7.2.4	Data Visualization	5
7.3	Operational Model: Infrastructure Design	5
8	Conclusion	6
8.1	Control Situation	6

1 Abstract

TBC

2 Literature Review

3 Stakeholder Concerns - Big Data

3.1 Collection

3.2 Processing

3.3 Visualization

4 Methodology

In order for group two to successfully deliver this project, a development methodology based on IBM Rational Unified Process (RUP) was followed albeit tailored to cater for the specific needs of this project.

The diagram below depicts the IBM RUP model:

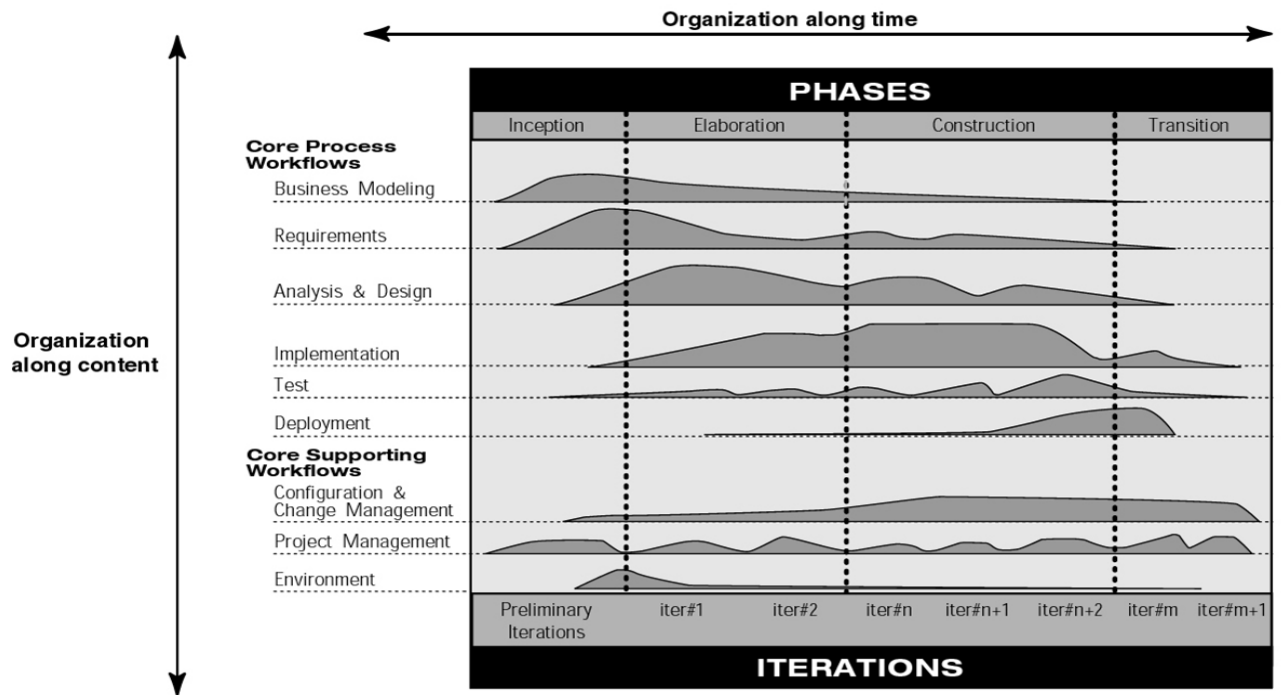


Figure 1: IBM Rational Unified Process (Source: RUP, Best Practices for Software Development Teams)

5 Assumptions and Constraints

5.1 Tweet Locations

The group encountered tweet location issues.

5.2 Pros and Cons determinations

Rudimentary algorithm for determining Twitter statements (tweets) that are against or for a particular candidate was adopted...

6 Success Criteria

7 Architecture of the Solution

7.1 High Level Design: Component Architecture

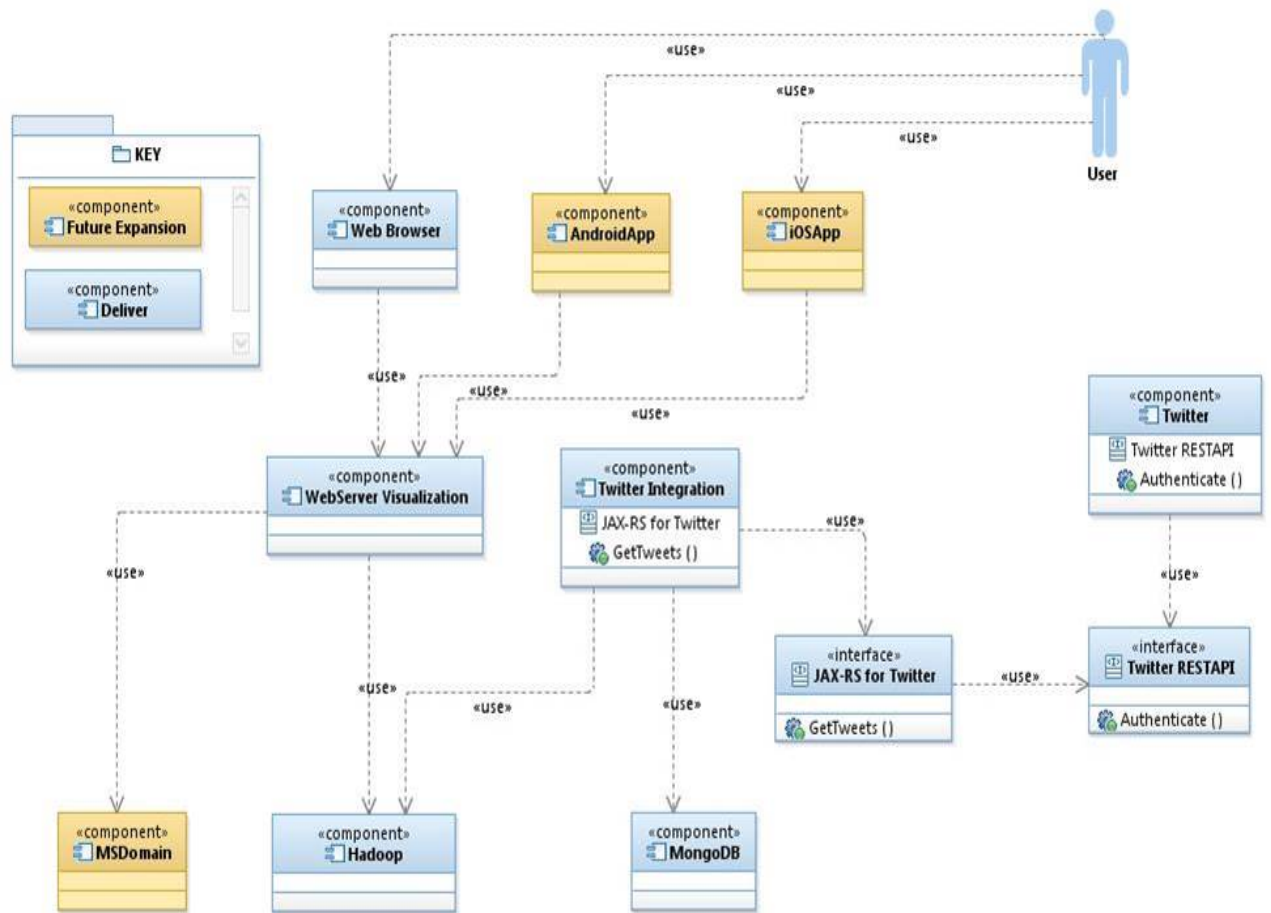


Figure 2: High Level Component Model.

7.2 Detailed Design: Class Diagrams

7.2.1 Data Acquisition (Batch)

7.2.2 Data Acquisition (Streaming)

7.2.3 Data Processing

7.2.4 Data Visualization

7.3 Operational Model: Infrastructure Design

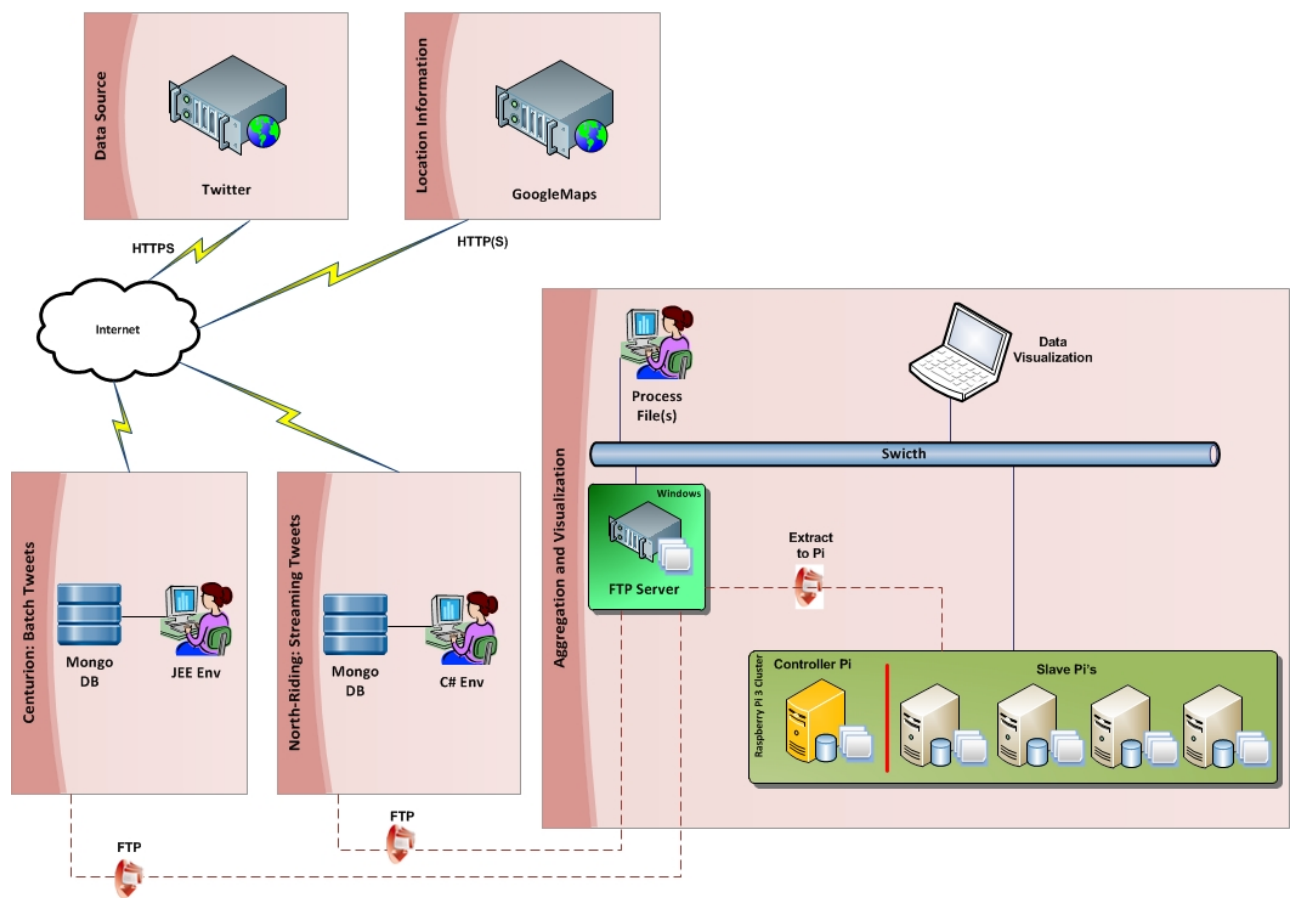


Figure 3: Operational Model: Physical

- **Personnel Shortfall:** Inexperience in the management team is a potential risk, due to the possible oversight and inaccuracies . Leach does not have

sufficient IS Management experience(*pg 502 paragraph 4*), the project may suffer if leach continues at his current position

8 Conclusion

8.1 Control Situation

Before we come to a conclusion and define the proposed actions we must first define the archetypical control situation. Based on..

References

[Van Vliet, 2008] Van Vliet (2008). Software Engineering Principles and Practice

[Frederick P. Brooks, 1975] Frederick P. Brooks (1975). The Mythical Man-Month