|  |
| --- |
| Individual Report |
| Big Data Visualization using Commodity Hardware and Open-Source Software – Data Visualization Component |
| Course: ELEN7046 – Software Technologies and Techniques |

|  |
| --- |
| Dave Cloete (Student Number 1573016)  6-30-2016 |

Contents

[Abstract 1](#_Toc455072023)

[1. Introduction 2](#_Toc455072024)

[2. Background 2](#_Toc455072025)

[3. Project Scope 2](#_Toc455072026)

[3.1. Functional 2](#_Toc455072027)

[3.2. Non-functional 2](#_Toc455072028)

[4. Approach 2](#_Toc455072029)

[5. Challenges 3](#_Toc455072030)

[6. Design Overview 3](#_Toc455072031)

[6.1. Conceptual 3](#_Toc455072032)

[6.2. Technologies Used 3](#_Toc455072033)

[6.3. The Visualisation Framework 3](#_Toc455072034)

[6.4. Data Models 3](#_Toc455072035)

[6.5. Controllers 3](#_Toc455072036)

[6.6. Class Diagram 3](#_Toc455072037)

[6.7. Graphing Components 3](#_Toc455072038)

[7. Recommendations 3](#_Toc455072039)

[8. Conclusion 3](#_Toc455072040)

[References 4](#_Toc455072041)

# Abstract

One of the Twit-Con-Pro project’s primary focuses where to prove that big data processing and visualisation is possible using free open-source software combined with inexpensive hardware and not reserved for enterprise domain.

Twit-Con-Pro sources data from Twitter and attempts to visualise sentiment around a topic, the specific topics chosen for this project was US and SA Elections. Categories representing the political parties and candidates were defined in order to visualise the positivity and/or negativity around each category as well as the combined sentiment across the categories.

This report focus on the design and development of the Data Visualisation component of the Twit-Con-Pro project. The Data Visualisation component is a web based component made possible with technologies such as HTML, JavaScript, D3 (Data Driven Documents), Bootstrap, Node.js and the Node.js Express Framework.

# Introduction

Data Visualisation is key to understanding large sets of data. Graphically representing data allows the human mind to better comprehend abstracted views and makes it easier to identify trends, patterns and anomalies not easily identifiable when looking at data in more traditional forms such as spreadsheets.

The Twit-Con-Pro Data Visualisation Component does just that. It essentially translate data from the Data Processing Component into vector graphics through various graphs. The types of charts chosen are:

* Word Cloud
* Bar Charts
* Streamgraphs
* Heat Maps

Furthermore, this is all done using open-source software such as Node.js, D3, and commodity hardware, a Raspberry-PI 3.

# Background

Initially the project team decided to do small POCs (Proof of Concepts) to understand which technologies can be used to develop a data visualisation solution on the Raspberry-PI hardware. Investigation shown various options existed on the latest Raspberry-PI 3 using the Raspbain OS which is a flavour of Debian part of the Linux Operating systems family.

# Requirements

The following is a list of the key requirements.

## Functional

* Represent the total amount of tweets mentioning a category in comparison
* Represent data over time in daily and hourly increments
* Represent the sentiment for a category in a positive and negative light (Con-Pro)

## Non-functional

* The concept of representing sentiment should be generic, in other words, the solution should be able to represent topics other than the US and SA Elections
* The charts should be scalable / responsive
* The charts should be represented side by side so that comparisons can be drawn across charts
* The solution had to run on commodity hardware such as the Raspberry-PI
* Data was to be present in such a way that comparisons can be drawn

# Approach

A web based component was chosen as client-browsers devices would then share the processing required to render charts. Research also shown availability of various open-source web based technologies existed.

A responsive dashboard containing more than one chart was decided on as it would allow cross comparison between charts.

# Challenges

# Design Overview

todo

## Conceptual

todo

## Technologies Used

Todo

such as D3, Chart.js, FusionCharts and InfoVis.

## The Framework

todo

## Data Models

todo

## Controllers

## Class Diagram

todo

## Graphing Components

# Recommendations

Various graphs

The Summary bar char information could better be represented in a pi chart.

# Conclusion

todo

# References

**There are no sources in the current document.**