

**COLLEGE OF COMPUTING AND INFORMATION SCIENCES**

**DEPARTMENT OF NETWORKS**

**BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING (YEAR 2) RECESS TERM 2 (BSE 2301)**

**SRS DOCUMENT**

**PROJECT MEMBERS (Group 16)**

|  |  |  |  |
| --- | --- | --- | --- |
| **NAME** | **REG. NO** | **STD. NO** | **SIGN** |
| KAGANZI TIMOTHY | 16/U/5134/EVE | 216007287 |  |
| KASOZI BENJAMIN | 16/U/ 5538/EVE | 216015602 |  |
| OKURUT JULIUS | 16/U/10950/PS | 216014096 |  |
| KUNYA NATHERN | 16/U/6372/EVE | 216006158 |  |

**PROJECT LEADER**

KAGANZI TIMOTHY

**SUPERVISOR**​: ISAAC MBABAZI

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE SOFTWARE ENGINEERING RECESS PROJECT BSE 2301

21​TH​ JUNE, 2018

SOFTWARE REQUIREMENTS SPECIFICATION

FOR

YOUTUBE STATISTICS ANALYSIS

VERSION 1.0 APPROVED

PREPARED BY BSE 2301 GROUP 16

MAKERERE UNIVERSITY

05th JULY 2018

Table of Contents

[1. Introduction 1](#_Toc518038994)

[1.1 Purpose 1](#_Toc518038995)

[1.2 Document Conventions. 1](#_Toc518038996)

[1.3 Intended Audience and Reading Suggestions. 2](#_Toc518038997)

[1.4 Product Scope. 2](#_Toc518038998)

[2. Overall Description. 3](#_Toc518038999)

[2.1 Product Perspective. 3](#_Toc518039000)

[2.2 Product Functions 3](#_Toc518039001)

[2.3 User Classes and Characteristics 4](#_Toc518039002)

[2.4 Operating Environment. 4](#_Toc518039003)

[2.5 Design and Implementation Constraints. 5](#_Toc518039004)

[2.6 User Documentation. 5](#_Toc518039005)

[2.7 Assumptions and Dependencies. 5](#_Toc518039006)

[3. External Interface Requirements. 5](#_Toc518039007)

[3.1 User Interfaces. 5](#_Toc518039008)

[3.2 Hardware Interfaces. 6](#_Toc518039009)

[3.3 Software Interfaces. 7](#_Toc518039010)

[3.4 Communications Interfaces. 7](#_Toc518039011)

[4. System Features. 8](#_Toc518039012)

[4.1 Video categories. 8](#_Toc518039013)

[4.1.1 Description and Priority. 8](#_Toc518039014)

[4.1.2 Stimulus/Response Sequences. 8](#_Toc518039015)

[4.1.3 Functional Requirements. 8](#_Toc518039016)

[4.2 Data Visualization. 8](#_Toc518039017)

[4.2.1 Description and Priority. 8](#_Toc518039018)

[4.2.2 Stimulus/Response Sequences. 9](#_Toc518039019)

[4.2.3 Functional Requirements. 9](#_Toc518039020)

[5. Other Nonfunctional Requirements. 9](#_Toc518039021)

[5.1 Performance Requirements. 9](#_Toc518039022)

[5.2 Safety Requirements 9](#_Toc518039023)

[5.3 Security Requirements 9](#_Toc518039024)

[5.4 Software Quality Attributes 10](#_Toc518039025)

[5.5 Business Rules 10](#_Toc518039026)

[5.6 Special user requirements 10](#_Toc518039027)

[6. Appendix A: Glossary 11](#_Toc518039028)

[6.1 References 12](#_Toc518039029)

# Introduction

## Purpose

This software requirements specification (SRS) specifies the requirements of the YSN (YouTube statistics analysis) which will be used by analysts and other personnel’s who are interested in knowing trending YouTube videos, best performers on YouTube, the user’s views or comments about certain products and therefore be able to make decisions based on the analysis got. This software is also to help individuals understand the user’s reactions and suggestions concerning their products.

## Document Conventions.

This Document was created based on the IEEE template for System Requirement Specification Documents. [1]

## Intended Audience and Reading Suggestions

* YouTube data analysts, who want to use YouTube statistics analysis software for analyzing YouTube videos through the user comments, likes, shares and views.
* Advanced/Professional Users, such as engineers or researchers, who want to use YouTube statistics analysis software for more demanding graph analysis of the user views.
* Programmers who are interested in working on the project by further developing it or fix existing bugs.

## Product Scope.

YouTube statistics analysis software is one that people can use to analyze YouTube data. It contains shiny app which feature is R based and it will be helpful during the visualization stage using many different techniques. The software also has a data visualization feature that will be of use to the analyst to understand the relationships between the different variables under consideration.

Upon completion of this software, these are some of the benefits that this software will provide to the end users:

* Proper market research.
* Easy decision making.
* Provide a company with an edge over other competitors.
* Provide an insight of the user’s suggestions and reactions about the product or video.

**These are some of the objectives this product is going to implement are:**

* To categorize YouTube videos.
* To analyze what makes certain videos more trending or popular than others and how it affects YouTube.
* To generate statistical information that is important for the various stakeholders.

# Overall Description.

This section contains the complete description of our project. It states the perspective

of the system, how the system is just one part of a complete YouTube statistics analysis software. This section also shows the constraints faced by the system, and the assumptions we made both about the systems it will be interacting with, and the people who will be using the system.

## Product Perspective.

The software is new self-contained product which will be open source under the general Public license. It is a R shiny application that will require the use of various interfaces to visualize data very well. It provides a simple mechanism that enables the users to use it. The various users should also have basic knowledge of how to use the R application.

## Product Functions

The system shall be able to tell whether the video is trending or not depending on the viewer’s reactions.

* Ability to analyze the most trending videos on YouTube.
* Statistical overview of the various YouTube videos.
* Visualize the best performers on YouTube based on likes, comments, dislikes, views.
* Categorize videos based on categories like music, entertainment and trailers.

## User Classes and Characteristics

**YouTube data analysists.**

This product is majorly being developed to help YouTube analysts to easily visualize and generate observations based on comments, user’s views and likes. These are the most highly privileged users of this product.

* The analyst will use the YouTube API to retrieve data about a given channel and save it as a csv file.
* check for errors in the data.
* Use various techniques to analyze the data for certain trends.

**Strategic level managers.**

These will be able to make conclusions and decisions based on what they can visualize from the shiny app and also make future plans from the visualized summary.

* Checks the analyzed data.
* Gets summary of the visualized data from the shiny app and is able to make decisions and future plans for the company.

**Marketing agencies.**

Marketing directors can upload their product on YouTube, get users comments, views, likes and dislikes, extract the data using the YouTube API, from which these mangers can successfully get an insight of what will be the users reaction towards the product they are putting on market.

## Operating Environment.

This product is platform independent and it can run on any machine provided that machine can be able to hold the software called R and Rstudio where this product is going run.

This product is can be used on any operating system of any type be it Linux, windows, mac osx, and many more provided the machine has R and Rstudio already installed.

## Design and Implementation Constraints.

The system is provisioned to be built using R language which is highly flexible which will require the use of various data visualization tools, techniques, and Rstudio. The shiny web application will be used to provide an interface for visualization.

## User Documentation.

The software will be delivered with a user manual and some tutorials that with help to guide the various users. The system is still under the development stage and implementation as soon as it is completed, there will be a project document that will provide an entire description of the system.

For more information, please view the references.

## Assumptions and Dependencies.

Resources: the project will be done by a group of four members each with working hours from 8am to 1pm.

The project will be scheduled to take five to six weeks involving documentation and implementation.

**Dependencies.**

The system will depend on csv file which contain data which is to be analyzed through the use of various data analysis techniques.

The system will also require the use of Rstudio.

# External Interface Requirements.

## User Interfaces.

* Front-end software: shiny app which runs on a localhost machine with no internet connection required.
* Back-end software: shiny server for authentication purposes.

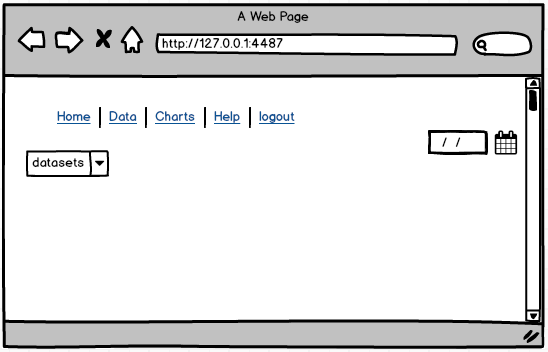


Figure 3.1.1 User Interface.

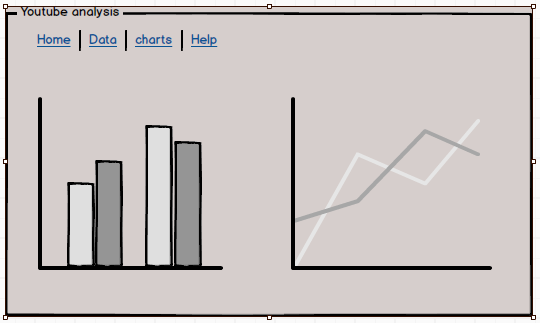


Figure 3.1.2 Interface showing graphs that visualize the data.

## Hardware Interfaces.

* Windows.
* A browser which supports shiny.

## Software Interfaces.

Following are the software used for the YouTube statistics analysis software.

|  |  |
| --- | --- |
| **Software used** | **Description** |
| Operating system | We have chosen Windows operating system for its best support and user-friendliness.  But someone can easily choose any operating system that suitable for him/her. |
| Database | To save the authentication details for our users, we are going to use shiny server. |
| Shiny | To implement the project, we have chosen shiny that comes embedded in Rstudio for its more interactive support. And this will run in any browser on any machine. |

## Communications Interfaces.

This project supports all types of web browsers, using the HTTP protocol for communication over the localhost since we don’t require any internet connection to run the app.

# System Features.

## Video categories.

### Description and Priority.

The system will be able to categorize videos based on the likes, shares, number of comments that videos has from that channel. By categorizing them helps the analyst to know which kind of videos are highly trending and which are not, the best performers on YouTube.

### Stimulus/Response Sequences.

Stimulus: select the csv file to be analyzed.

Response: the data has been well selected.

Stimulus: select which visualization technique to be used.

Response: videos sorted according to the likes, number of comments and views.

### Functional Requirements.

* Input: csv file containing the data to be analyzed.
* Output: plots containing the relationships, correlations between different variables of concern.

## Data Visualization.

### Description and Priority.

This is the use of various visualization techniques that will be used to analyze the various trending YouTube videos. They include bar graphs, line graphs, box plots and scatterplots.

### Stimulus/Response Sequences.

Stimulus: load data to be analyzed.

Response: plots of the variables of interest.

### Functional Requirements

* Input: csv file containing the data to be analyzed.
* Output: plots containing the relationships, correlations between different variables of concern.

# Other Nonfunctional Requirements

## Performance Requirements.

The system should be interactive and the delays involved should be minimal or less. So, in every action- response to the system, there are no immediate delays. [[2]](#_References) For example, when the user is loading the dataset for visualization in shiny app, there shouldn’t be any delay in producing the expected output.

## Safety Requirements

During user authentication, the information should be securely transmitted to the server (shiny server) without any changes in information.

The dataset extracted should be saved as a csv file for easy importing into R and for easy manipulation using the shiny app.

## Security Requirements

The system will require only the analyst to load the various datasets which will limit access to the system. Only the analyst will be able to load the datasets into the system.

## Software Quality Attributes

* **AVAILABILITY:** The software should be readily available to the user on his/her local machine.
* **CORRECTNESS:** the datasets should have correct information concerning a certain video from a given YouTube channel.
* **MAINTAINABILITY:** the user should be able to load a given dataset of his choice and can be able to do any visualization and make conclusions based on the visualized data.
* **USABILITY:** the shiny app should be easily accessible to authorized users and it should be highly responsive to the changes made by the user.

## Business Rules

Illegal duplication of the reports should be strictly dealt with. The analyst should have full access to the dataset which is saved as a csv file.

## Special user requirements

#### 5.6.1 User training.

The users should be trained on how to use shiny app correctly and effectively.

And they should be also taught how to extract their own data from different YouTube channels using the YouTube API and how to visualize it in the shiny app.

#### 5.6.2 Installation.

A conducive environment for the efficient running of shiny app is always recommended. The server should be handled with great care and all unauthorized access should be restricted.

# Appendix A: Glossary

csv: comma separated values.

Shiny: a feature in Rstudio that is used in making web-based applications in R.

YouTube analysists: a person who has general knowledge about YouTube statistics and play with the data on YouTube to acquire the required output.

HTTP: Hyper Text Transfer Protocol.

API: Application Programming Interface.

# References

|  |  |
| --- | --- |
| [1] | 2 may 2015. [Online]. url available: [http://www.vceit.com/p/SRS-sample.htm.](%20http:/www.vceit.com/p/SRS-sample.htm.) |
| [2] | [Online]. url available: [https://en.wikipedia.org/wiki/Software\_requirements\_specification.](%20https:/en.wikipedia.org/wiki/Software_requirements_specification.) |