Software Requirements Specification

for

Chat Analysis System

Version 1.0 approved

Prepared by Group 2

MAKERERE UNIVERSITY, COLLEGE OF COMPUTING AND INFORMATION SCIENCES, SOFTWARE ENGINEERING

28th June 2018

Table of Contents

Table of Contents ii

Revision History ii

1. Introduction 1

1.1 Purpose 1

1.2 Document Conventions 1

1.3 Intended Audience and Reading Suggestions 1

1.4 Product Scope 1

2. Overall Description 2

2.1 Product Perspective 2

2.2 Product Functions 2

2.3 User Classes and Characteristics 2

2.4 Operating Environment 2

2.5 Design and Implementation Constraints 2

2.6 User Documentation 2

2.7 Assumptions and Dependencies 3

3. External Interface Requirements 3

3.1 User Interfaces 3

3.2 Hardware Interfaces 3

3.3 Software Interfaces 3

3.4 Communications Interfaces 3

4. System Features 4

4.1 System Feature 1 4

4.2 System Feature 2 (and so on) 4

5. Other Nonfunctional Requirements 4

5.1 Performance Requirements 4

5.2 Safety Requirements 5

5.3 Security Requirements 5

5.4 Software Quality Attributes 5

5.5 Business Rules 5

6. Other Requirements 5

Appendix A: Glossary 5

Appendix B: Analysis Models 5

Appendix C: To Be Determined List 6

**References**

Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Reason For Changes** | **Version** |
| Chat Analysis System | 28-06-18 | Initial Version | 1.0 |

**1. Introduction**

**1.1 Purpose**

This document defines the requirement specifications of the Chat analysis software system version1.0 for analyzing and visualizing chat data. This document covers the entire system.

**1.2 Document Conventions**

The font used in this document is Times New Roman, of size 12`, line spacing of 1.5. The document is also justified.

Section headings indicated by whole numbers (1, 2, 3, …) and sub-section headings numbered based on the whole numbers of their sections for example, 1.1 belongs to section 1, are written with a bold-faced text for emphasis. *Italic* text is used to label diagrams and interfaces.

**1.3 Intended Audience and Reading Suggestions**

This document is to be read by the project managers, the development team, the documentation writers, marketing staff, testers and the organization stakeholders associated with manufacturing hardware and providing the necessary software. The shareholders will also be able to learn about the project and understand the requirements of the system.

The project managers and the developers need to become well-versed with this requirement document.

The marketing staff should understand the internal and external interface requirements of this document to become used to various product features in order to effectively advertise the product.

The testers need to be well-versed with the system features of the document to develop meaningful test cases and give useful feedback to the developers.

The stakeholders associated with manufacturing hardware needs to review the External or internal Interface Requirements, Nonfunctional and Functional Requirements of this document to know the requirement specifications of the needed hardware.

**1.4 Product Scope**

The Chat analysis software system will analyze and visualize data provided to determine the chat patterns and trends. It will be based on a computer that has RStudio, a free and open source integrated development environment for the R programming language installed on it together with the relevant packages.

The objective of the system is to provide a visual and statistical representation of the analyzed chat data to the National Social Security Fund organization of Uganda. This will assist the organization in strategic decision making.

The benefits of this system include;

* Customer satisfaction since the needs of the customer are taken care of.
* Improved relationship between the customers and the organization since customer complaints are considered and attended to.
* Organization growth since there will be improved service provision thereby attracting more customers.

The goals of the analysis are;

* To examine what the customers think and use this information to improve on the organization’s products and services.
* To determine chat patterns and understand the needs of the customers.

# 2. Overall Description

## Product Perspective

This is a new and self-contained product for Chat Analysis System. It has been developed for the Ugandan National Social Security Fund (NSSF) Organization to use it as a means of analyzing chat data. It has a very active and dependable developer team which will support it and provide immediate and appropriate feedback to the users. It has been developed to run on Windows, Linux, Mac OS X.

## 

## Product Functions

System Context Diagram

## E:\Recess2\Real\Screenshot (2).png

*Figure1: System context diagram*

## 2.3 User Classes and Characteristics

**System Administrator** The administrator is expected to have full knowledge of how the system works.

He /she maintains the system and has full access to the data store.

He/she has authority to make necessary updates to the system.

**Data Analyst (DA)**

The Data Analyst has access to the data store and the visualized data for example pie charts, bar charts, and is able to make necessary changes or updates in that data. Access to the data store by the DA is granted by the System Administrator.

**Company Executives**

These only have access to visualized data and important summarized information from the analyzed data.

**Marketing Staff**

These should have access to visualized data.

## 2.4 Operating Environment

The system will operate effectively on a computer whose RAM is at least 2GB with an average processing speed of at least 2.0GHz. The system should have Linux operating system version 12.04 and above or Windows 7 and above on which R programming language and R Studio have been installed.

The system interface will require an up-to-date browser such as Chrome version 59.0.3071 and above, Firefox version 52.0 and above.

## Design and Implementation Constraints

The system is limited by a number of factors as highlighted below;

**Hardware Constraints:**

The system will be able to support less than 20 users at a time with repository of 10-60GB,

having a minimum configuration of 1GB – 2GB RAM, 1-2 Cores (1.86 GHz), 30-180 GB SATA (hard disk) and configuration for better performance 2GB – 4GB RAM.

**Time Constraint:**

The system will be delivered by 16th August 2018.

**Technology Constraints:**

The system will require use of web server technologies such as R Shiny Server, Statistical Analysis, Visualization tools and must have a database.

**Cost Constraint:**

The design phase may require maintenance of the hardware needed for implementation of the system. This may be costly.

**Programming Standards:**

The system will be maintained by the development team and incase of any query about the system, it should be submitted to the team.

**Communication Protocols:**

The system will use Two-tier architecture where the client will directly communicate with the server.

**Language Requirement:**

The system will strictly use English as the standard language.

## User Documentation

* System User Manual (PDF Document)
* System Tutorials (PDF Document)

The above documents will be delivered on delivery of the system.

## 2.7 Assumptions and Dependencies

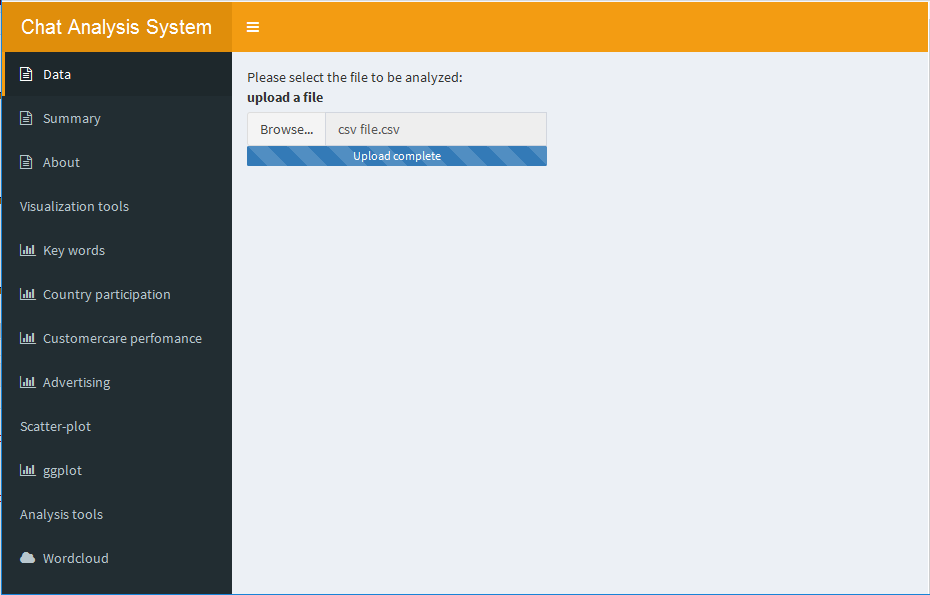
The system is developed in R using R Studio therefore the user should have R installed on their computer. The user must have all the required packages for R in order for the system to run efficiently.

It requires the latest R version 3.4.0 or higher and this applies to both Windows and Linux users. The system is assumed to only be accessed by NSSF Uganda staff, the development team, Data Analysts and the System Administrator.

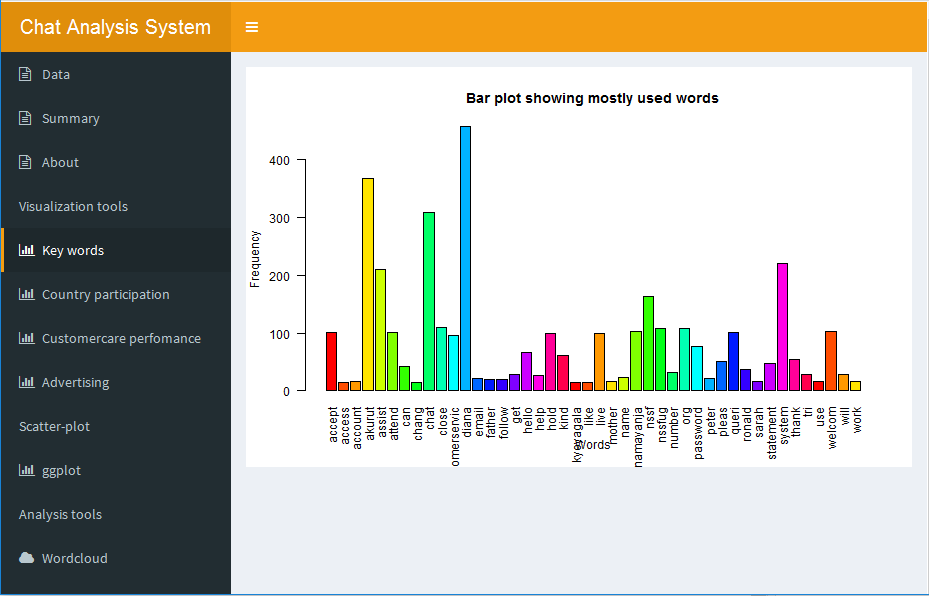
# 3. External Interface Requirements

**3.1 User Interfaces**

**The system’s data interface**

*Figure 3.1.2:Data interface*

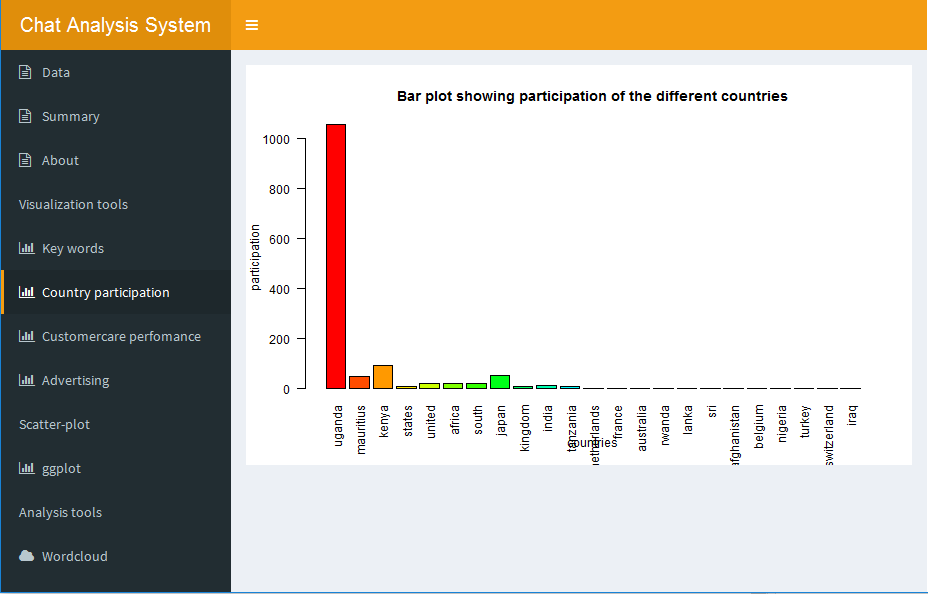
**The system’s keywords interface.**

**

*Figure 3.1.3: Key words’ interface*

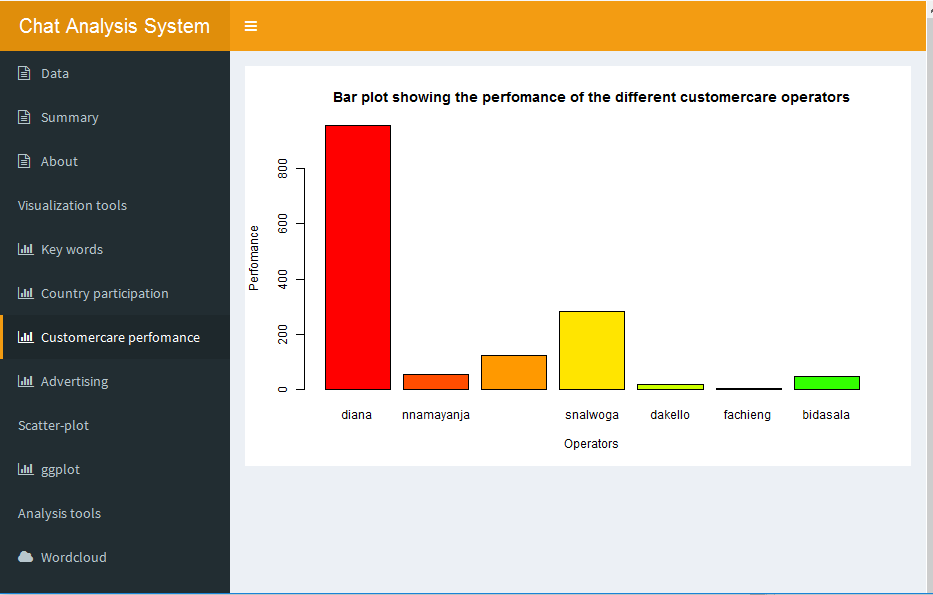
**Analyzed data interfaces.**

**Country participation interface**

****

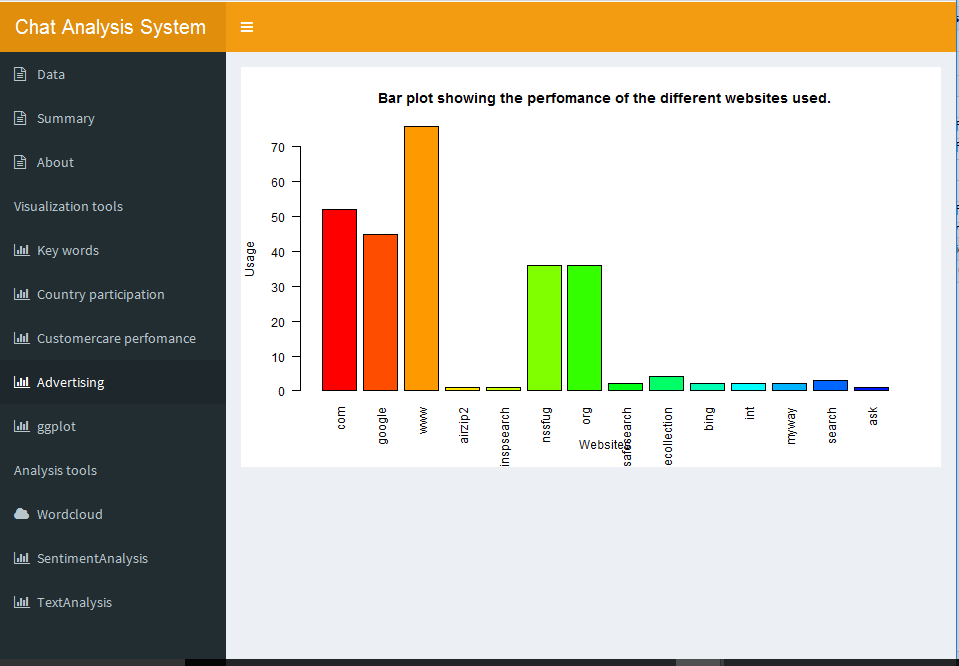
*Figure 3.3.3: Country participation interface*

**Customer care Interface**

****

*Figure 3.3.4: Customer care interface.*

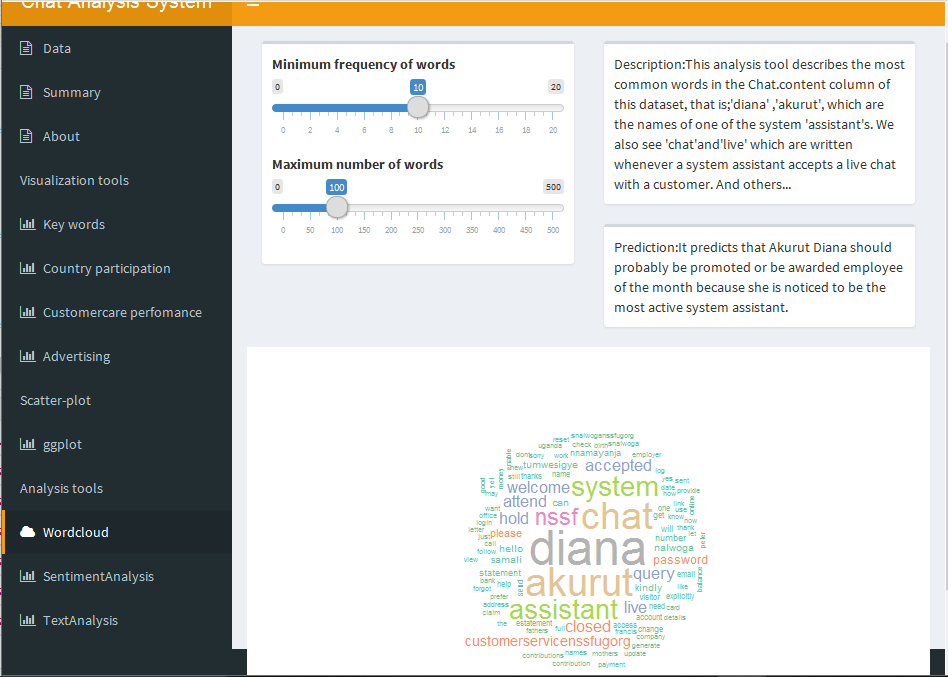
**Advertising interface**

****

*Figure3.3.5: Advertising interface.*

**Analyzed data interfaces**

**Word cloud**

****

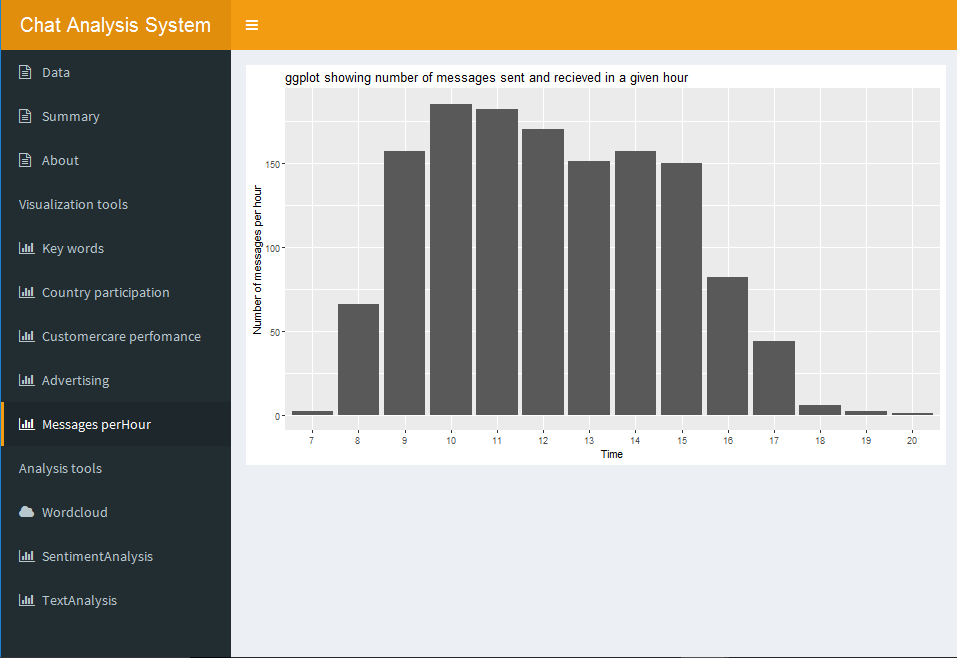
*Figure3.3.6: Word cloud interface*

**Sentiment Analysis interface**

****

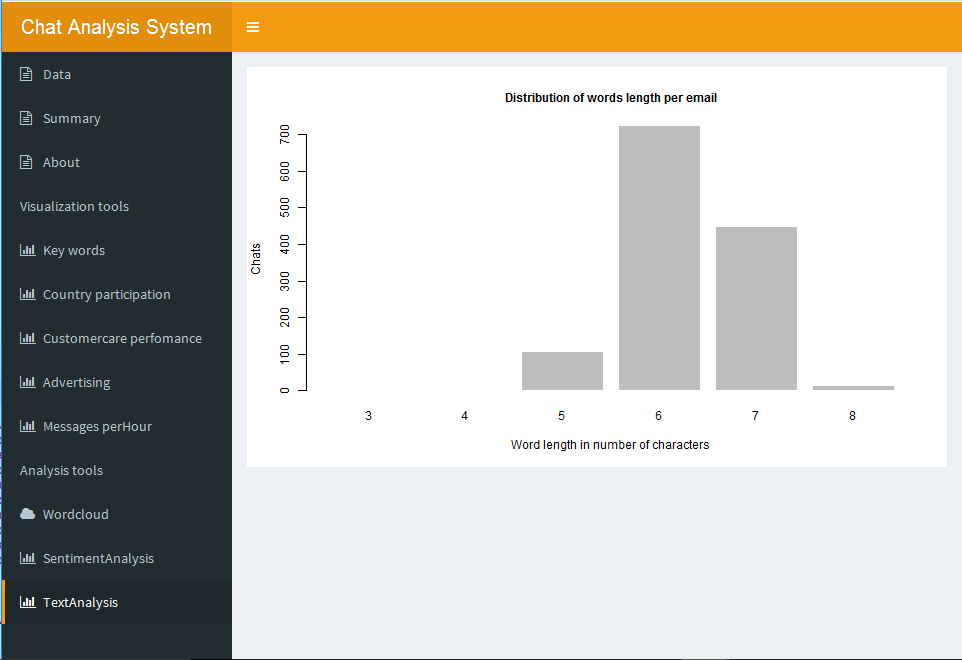
*Figure3.3.7: Sentiment Analysis interface.*

**Messages per hour interface**



*Figure 3.1.4: Messages per hour interface*

**Text Analysis interface**

****

*Figure3.3.8: Text Analysis interface.*

**3.2 Hardware Interfaces**

The minimum hardware requirements of the system are 1GB-2GB RAM, 1-2 cores(1.86GHz), 30-180GB SATA (hard disk) and configuration for better performance 2GB – 4GB RAM. A compatible graphics card is required for efficient visualization. A system with these specifications can handle a network of approximately 1000 edges and nodes. For bigger networks, additional memory is required.

**3.3 Software Interfaces**

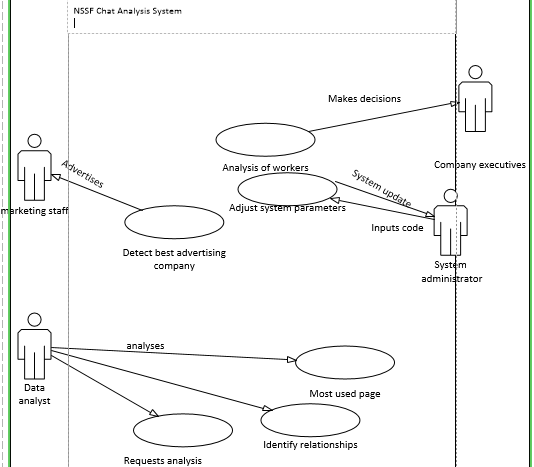
The system requires R and RStudio to be installed on the computer to be used, more specifically R version 3.5.0 and above for its latest release. Additional information can be found in section 2.7 of this document.

**3.4 Communications Interfaces**

The requires an internet connection to install packages that may be required, updated already installed ones, and update some of its components (APIs, modules etc.) and also to verify the users of the system. Additional information on communication protocols can be found in section 2.5.

# 4. System Features

**Use case**

****

1. 4.1 **Error Correction**

4.1.1 **Description and Priority**

This feature is aimed at making sure the user does not over wait. It picks out the user with the highest waiting time and assigns him available operator. This is of medium priority since we don’t want users to over wait and abandon using website.

4.1.2 **Stimulus/Response Sequences**

In case a user enters a different name to a given phone number and there is 80% similarity in name to one originally there --- the system asks for name verification and then correct name stored. Same thing is done in a case of different email and country for given number or name.

For lacking entries compare with available data and predict entry basng on similar name, number or email.

4.1.3 **Functional Requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

* 1. **Waiting delay correction**

**4.2.1 Description and priority**

This feature is aimed at making sure the user does not over wait. It picks out the user with the highest waiting time and assigns him available operator. This is of medium priority since we don’t want users to over wait and abandon using website.

* + 1. **Stimulus/ Response Sequences**

When a user comes to the organization website and requests for something. A clock is started immediately. If a free agent is available, he is assigned. If not when an operator gets free the user with maximum waiting time is assigned operator first.

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

**4.3Detect best advertising company**

* + 1. **Description and priority**

Here NSSF can determine which company it should use most for advertisement. This is done basing on which website most users came from to go to NSSF website. This feature is of high priority since NSSF find out where to invest most.

* + 1. **Stimulus/ Response Sequences**

Whenever NSSF wants to find out best advertising organization this feature simply searches through the data for website most users came from.

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

**4.4 Analysis of workers’ performance**

* + 1. **Description and priority**

The aim of this feature is to assess operators’ performance on website basing on the responses given to users and also when they don’t respond. The company can base on this to either promote or demote operators. This feature is of high priority since it helps NSSF find out operators’ performance on website.

* + 1. **Stimulus/ Response Sequences**

Whenever NSSF wants to find out best performing operators this feature simply searches through the data for operator who best responds to users. Also NSSF can know who to fire for inefficiency.

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

* 1. **Requests analysis**

**4.5.1 Description and priority**

Here NSSF can detect and answer unanswered questions of users. This is done basing on searching through chat content.

* + 1. **Stimulus/ Response Sequences**

NSSF uses this feature to automatically detect unanswered questions basing on waiting times. Requests not worked on for a long time are assigned operators

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

* 1. **Remove inapplicable data**
     1. **Description and priority**

Here NSSF can use feature to detect data that its collecting yet it is the same for every user hence not necessary to collect. This feature is of low priority since NSSF can work without it.

**4.6.2 Stimulus/ Response Sequences**

NSSF uses this feature to automatically detect inapplicable data basing on previously collected data. Data similar for every user is automatically deleted.

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

* 1. **Most used page**
     1. **Description and priority**

This feature is aimed at detecting most used page. This helps determine what users request most. This feature is of medium priority since through it NSSF can find out most requested page and solve the issue involving page.

* + 1. **Stimulus/ Response Sequences**

Whenever NSSF wants to find out most requested page this feature simply searches through the data for most requested page in pages available.

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

* 1. **Detect relation between users**
     1. **Description and priority**

Here NSSF can determine which users are connected to each other either as relatives, friends or co-workers. This is done basing on those with the same Ip address. This feature is of low priority since NSSF can do without it.

* + 1. **Stimulus/ Response Sequences**

Whenever NSSF wants to find out related users, this feature simply searches through the data for users with the same ip-address.

* + 1. **Functional requirements**

**REQ – 1: R**

The company should have R installed on their computers to enable this feature to run since it needs the r analysis tools

**REQ – 2: R-studio**

R studio is essential for analysis of data

# 5. Other Nonfunctional Requirements

**5.1 Performance Requirements**

The minimum hardware requirements of the system are 1GB – 2GB RAM, 1-2 Cores (1.86 GHz), 30-180 GB SATA (hard disk) and configuration for better performance 2GB – 4GB RAM. A compatible graphics card is required for efficient visualization. A system with these specifications can handle a Network of approximately 1000 edges and nodes. Performance depends on the RAM size of the system.

## Safety Requirements

## The system administrator who is the senior software engineer updates the system regularly, so that system users do not incur any data losses which may be caused by bugs, or a system crash while using the system. This also improves on functionality and performance.

## 5.3 Security Requirements

The product should be accessed by only the system administrator, data analysts and staff of the NSSF Uganda organization due to the data confidentiality.

## 5.4 Software Quality Attributes

The R version should be compatible with the operating system residing on the computer to be used. The system also contains application extensibility that is, new functionalities can be added to the system.

The users must have basic knowledge about visualization: histograms, scatter plots, bar graphs and pie charts.

The system is easy to use and it uses graphical user interface that eases the interaction between the user and the system.

## Business Rules

The system and its functionalities should be used by staff and organization executives. It can also be accessed by authorized external data analysts.

**6. Other Requirements**

All the functional and non-functional requirements of the system are currently included in the documentation. When the system users’ preferences change, the new requirements will be included in the next version of the documentation.

**APPENDIX A: GLOSSARY**

1. **SRS** – Software Requirements specification
2. **Debugging** is the process of finding and removing errors in a code.
3. **Visualization** is the representation of an object, situation or a set of information as a chart or another image.
4. **Data** **analysis** is a process of inspecting, cleaning, transforming and modelling data with the goal of discovering useful information and support decision making.
5. **Cognition**- what people are thinking about.

**APPENDIX B: Analysis Models.**

**Data** **Flow** **Diagram** is a unified modelling tool which shows how the system stores, processes and transforms data.

**Use case:** is a modeling tool which show how users interact with the system features.

**References**

**Below are the references we used for the design of the Software Requirements Specification document of the Data Analysis Software System:**

1. System Analysis and Design UML Edition by Denis Wixom and Doth.
2. WorldWeb.com
3. SRS template 1999 by Karl E Wiegers.