Software Requirements Specification

for

Stack Overflow Analysis Software System

Version 1.0 approved

Prepared by Group 18

Makerere University, College of Computing and Informatics Sciences, Software Engineering

28th June 28, 2018

Table of Contents

[**1** **Introduction** 3](#_Toc518944262)

[**1.1** **Purpose** 3](#_Toc518944263)

[**1.2** **Document Conventions** 3](#_Toc518944264)

[**1.3** **Intended Audience and Reading Suggestions** 3](#_Toc518944265)

[**1.4** **Product Scope** 3](#_Toc518944266)

[**2** **Overall Description** 4](#_Toc518944267)

[**2.1** **Product Perspective** 4](#_Toc518944268)

[**2.2** **Product Functions** 5](#_Toc518944269)

[**2.3** **User Classes and Characteristics** 5](#_Toc518944270)

[**2.4** **Operating Environment** 6](#_Toc518944271)

[**2.5** **Design and Implementation Constraints** 6](#_Toc518944272)

[**2.6** **User Documentation** 6](#_Toc518944273)

[**2.7** **Assumptions and Dependencies** 7](#_Toc518944274)

[**3** **External Interface Requirements** 7](#_Toc518944275)

[**3.1** **User Interfaces** 7](#_Toc518944276)

[**3.2** **Hardware Interfaces** 8](#_Toc518944277)

[**3.4** **Communications Interfaces** 8](#_Toc518944278)

[**4** **System Features (Use Cases)** 8](#_Toc518944279)

[**4.1** **Login** 9](#_Toc518944280)

[**4.2** **Upload data file** 9](#_Toc518944281)

[**4.3** **View analyzed data** 9](#_Toc518944282)

[**4.4** **Choose technique** 9](#_Toc518944283)

[**5** **Other Nonfunctional Requirements** 10](#_Toc518944284)

[**5.1** **Performance Requirements** 10](#_Toc518944285)

[**5.2** **Safety Requirements** 10](#_Toc518944286)

[**5.3** **Security Requirements** 10](#_Toc518944287)

[**5.4** **Software Quality Attributes** 10](#_Toc518944288)

[**6** **Appendix A: Glossary** 10](#_Toc518944289)

[**7** **Appendix B: Analysis Models** 11](#_Toc518944290)

[**8** **References** 11](#_Toc518944291)

# **1 Introduction**

## **1.1 Purpose**

This document defines the requirements specifications of the Stack overflow analysis software system version 1.0 for analyzing and visualizing gathered developer data. This document covers the entire system.

## **1.2 Document Conventions**

This document uses the Modern Language Association (MLA) format. Section headings indicated by whole numbers (1, 2, 3 ...) and sub-section headings numbered basing on the whole numbers their section headings such as 1.3 belongs to section 1 are written with a bold-faced text for emphasis. Italic text is used to label diagrams.

## **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, the stake holders associated with providing the necessary software and shareholders to learn about the project and understand the requirements of the system.

1. The project managers and the developers need to become intimately familiar with this Requirement document.
2. The marketing staff should understand the overall description and the external interface requirements of this document to become accustomed to the various product features in order to effectively advertise the product.
3. 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.
4. The stake holders associated with manufacturing hardware need to review the External Interface Requirements, Nonfunctional and Functional Requirements of this document to know the requirement specifications of the needed hardware.

## **1.4 Product Scope**

The Stack Overflow Analysis software system will analyze and visualize the developer data to determine the developer behavior with regards to using the stack overflow website as well as the insights in the developer survey data.

It will be based on a computer that has R Studio, a free and open source integrated development environment for R programming language installed on it together with the Shiny packages and other relevant packages.

The goal of this system is to critically analyse developer survey data so as to find a way on how to improve on developer marketing, technical recruiting, market research and enterprise knowledge sharing at Stack Overflow.

The objective of the system is to provide a visual and statistical representation of developer behavior to the stack overflow management which will be assist these it in making operational plans, strategic planning and decision making to improve on developer experience on the Stack Overflow website.

# **2 Overall Description**

## **2.1 Product Perspective**

This is a new and self-contained product called the Stack overflow analysis software system. It has been developed for the stack overflow management team to use it as a means of analyzing developer survey data.

It is an open source project and it has a very active developer team to support it and provide feedback to users. It is to be developed to run on Windows, Mac OS X and Linux.

## **2.2 Product Functions**

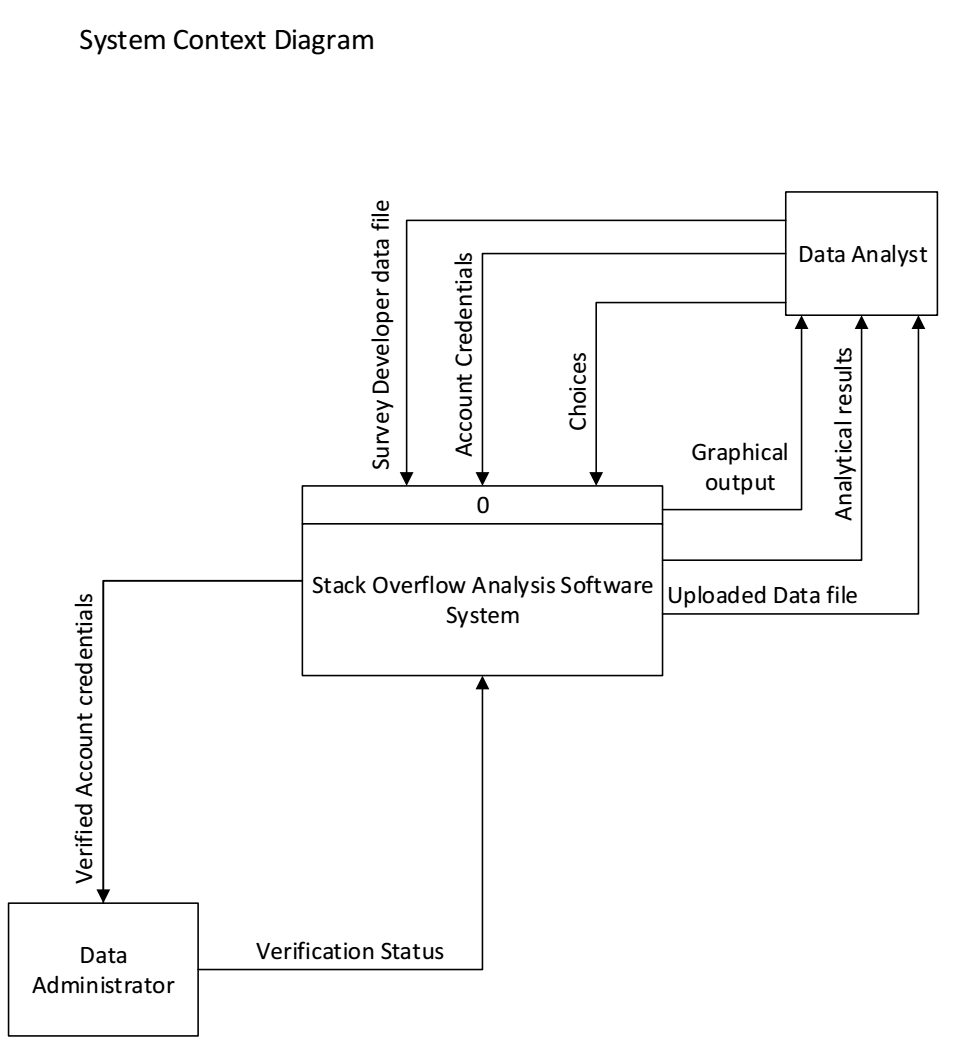


Figure 1: Figure showing system context diagram

## **2.3 User Classes and Characteristics**

* Data Administrator (DAdmin):

The DAdmin is expected to have an IT field appropriate Diploma or equivalent certification. He /she has full access the data store for maintenance and updates.

* The Data Analyst (DA):

The DA is expected to have a field appropriate college degree in the IT field as a DA. He / She will need a laptop connected to the web server over the internet and access the data store and the system interface in order to analyze the data. View and select permissions to the data store will be granted to the DA by the DAdmin.

## **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 2.0 GHz. The computer should have Linux operating system 12.0.4 and above or Windows operating systems 7 and above on which R and RStudio has 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.

## **2.5 Design and Implementation Constraints**

The system is constrained by a number of factors as highlighted below: -

1. Hardware Constraints:

The system will be able to support less than 25 users 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.

1. Time constraint:

The system must be delivered by 11th August of 2018.

1. Cost constraint:

The design phase requires a lot of expenses in terms of stationery, Internet among others.

1. Technology Constraints

The system will require use of web-based server technologies such as R Shiny Server, Statistical analysis, visualization tools such as the R language and it will not have a database.

1. Language requirement

The system will strictly use English as its standard language

1. Programming standards

The system software will be maintained by the development team

1. Communication Protocols

The system will use File Transfer Protocol (FTP) to upload the developer data files and Hyper Text Transfers Protocol (HTTP) for connection.

## **2.6 User Documentation**

1. System Design Document
2. System Report

The above documents will be delivered with the system software.

## **2.7 Assumptions and Dependencies**

The system is developed in R using Rstudio and therefore requires R to be installed on the user’s computer. The user must have all the required R packages for the system to run. It requires the latest R version, and this applies to both windows and linux users.

# **3 External Interface Requirements**

## **3.1 User Interfaces**

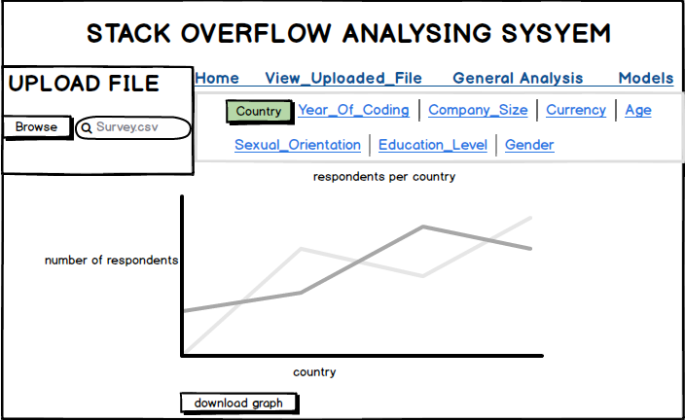


Figure 2: Interface showing home page

## **3.2 Hardware Interfaces**

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. For bigger networks, additional memory is required

**3.3 Software Interfaces**

The system requires R and Rstudio to be installed on the system, more specifically R version 3.4.0 and above for its latest release. Additional information can be found in Section 2.7 of this document.

## **3.4 Communications Interfaces**

The system requires an internet connection to install new packages, update already installed ones and update some of its components (APIs, modules etc.), and verification of users. Additional information can be found in Section 2.5 of this document.

# **4 System Features (Use Cases)**

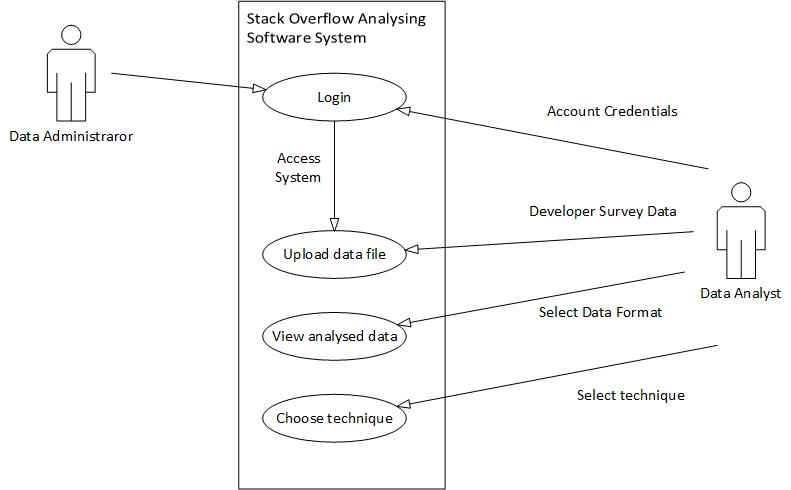


Figure : Figure showing Use Case diagram

## **4.1 Login**

Name: Log in

Goal: User verification

Input: Account credentials

Output: System access

Main Scenario: Validation

## **4.2 Upload data file**

Name: Upload data file

Goal: Obtaining data

Input: Developer survey developer data file

Output: Upload file

Main Scenario: Retrieving data

Pre-condition: CSV file format

## **4.3 View analyzed data**

Name: View analyzed data

Goal: Visualization

Input: Developer Survey data file

Output: Analyzed results

Main Scenario: Analysis

Pre-condition: Knowledge about visualization tools

## **4.4 Choose technique**

Name: Choose technique

Goal: To select R technique to be used to analyze data

Input: Selected technique

Output: Analyzed results

Main Scenario: Analyzing

# **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.

## **5.2 Safety Requirements**

To ensure that none of the system’s users loses any data while using the system (due to a crash or a bug of some kind), the data administrator updates the system regularly.

## **5.3 Security Requirements**

The system security requirements are the account credentials and thus any type of user cannot use it without those credentials.

## **5.4 Software Quality Attributes**

The system provides the users with both simple and advanced features. Due to its well designed and easy to use interface it can be used by both experts and typical users. However, users must already have a basic knowledge of graphs.

# **6 Appendix A: Glossary**

1. Graph: A system of nodes connected in pairs by edges. Often sub-divided into directed graphs or undirected graphs according to whether the edges have an orientation or not. Mixed graphs include both types of edges.
2. Node: A node is (together with edges) one of the two basic units out of which graphs are constructed. Nodes of graphs are often considered to be atomic objects, with no internal structure.
3. Edge: An edge is (together with vertices) one of the two basic units out of which graphs are constructed. Each edge has two (or in hyper graphs, more) vertices to which it is attached, called its endpoints.

# **7 Appendix B: Analysis Models**

1. Data Flow Diagram (DFD) is a unified modelling tool which shows how the system stores, processes, and transforms data.
2. A Use Case Diagram visually represents the interaction between users and the information.

System. In a use case diagram, the user becomes an actor, with a specific role that describes how he or she interacts with the system.

# **8 References**

Books[1] Gary B Sherry, Harry J Rosenblatt, System analysis and design, ninth edition, Library of  
Congress Control Number: 2010943248, ISBN-13: 978-0-538-48161-8, ISBN-10:0-538-48161-7

**Links**

*Brief: How to Write a Concept Paper*. Retrieved from https://www.ncat.edu/research/documents/dored-documents/research-services-documents/brief-write-concept-paper081517.pdf

*Stack Overflow*. Retrieved from https://stackoverflow.com

Stack Overflow 2018 : <https://stackoverflow.blog/2018/03/13/2018-developer-survey-results-live/>