Software design document for image-based web scraping software

Version 0.6

**Prepared by:**

* **Ahmed Mohammed Abd El-Ghani**
* **Ahmed Mohamed Agamy**
* **Hussein Muhammad El-Sayed**
* **Islam**

Contents

[I. Revision history 3](#_Toc517161210)

[1. Introduction 3](#_Toc517161211)

[1.1. Purpose 3](#_Toc517161212)

[1.2. Scope 3](#_Toc517161213)

[1.3. References 3](#_Toc517161214)

[1.4. Overview 3](#_Toc517161215)

[2. Design consideration 4](#_Toc517161216)

[2.1. Constrains 4](#_Toc517161217)

[2.2. System environment 4](#_Toc517161218)

[2.3. Design methodology 4](#_Toc517161219)

[3. archtecture 4](#_Toc517161220)

[3.1. System design 4](#_Toc517161221)

[4. data design 5](#_Toc517161222)

[4.1. Context diagram 5](#_Toc517161223)

[4.2. Dataflow diagram 5](#_Toc517161224)

[4.2.1. Level 0 diagram 5](#_Toc517161225)

[4.2.2. Level 1 diagram 5](#_Toc517161226)

[5. component design 6](#_Toc517161227)

[5.1. Scrapping by text 6](#_Toc517161228)

[A close up of a map

Description generated with very high confidence 7](#_Toc517161229)

[5.2. Scrapping by image 8](#_Toc517161230)

[6. software interface 9](#_Toc517161231)

[6.1. Object and actions 9](#_Toc517161232)

[6.2. User interface 10](#_Toc517161233)

[6.3. Screen image 10](#_Toc517161234)

[7. class diagram 10](#_Toc517161235)

[8. Glossary 10](#_Toc517161236)

# Revision history

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Author | Description | Date |
| 0.1 | Ahmed M. Abd El-Ghani | Initial | 6-6-2018 |
| 0.2 | Ahmed M. Abd El-Ghani | Introduction | 7-6-2018 |
| 0.3 | Ahmed M. Abd El-Ghani | overall description | 8-6-2018 |
| 0.4 | Ahmed Agamy | specific requirement | 13-6-2018 |
| 0.5 | Ahmed Agamy | specific requirement | 14-6-2018 |
| 0.6 | Ahmed M. Abd El-Ghani | specific requirement | 14-6-2018 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Introduction

## Purpose

The purpose of this document is to present a detailed description of the designs of image-based web scraping software, Firstly, this document is intended for the Team, to use the designs as guidelines to implement the project. Lastly, this document could be used for designers who try to upgrade or modify the present design of the system.

## Scope

This document gives a detailed description of the software architecture of the web-scrapping system. It specifies the structure and design of some of the modules discussed in the SRS. It also displays some of the use cases that had transformed into sequential and activity diagrams. The class diagrams show how the programming team would implement the specific module.

## References

* *IEEE Standard 1016-1998, IEEE Recommended Practice for Software Requirements Specifications, IEEE Computer Society, 1998.*

## Overview

This document is written according to the standards for Software Design Documentation explained in “IEEE Recommended Practice for Software Design Documentation”.

The next chapter, the Design consideration, this section gives an overview of some of the constrains and assumption that has been taken to consideration in the design.

The third to fifth chapter, contain discussions of the designs for the project with diagrams.

The sixth chapter, the Software interface design, contain the UI design samples from the system.

The seventh chapter, the Class diagram, this section contains the class diagram

# Design consideration

## Constrains

The system is designed to be built using python with already made TensorFlow API object recognition model and selenium for web-scrapping are used to generate the report

## System environment

The System is designed to work on Microsoft windows operating system

## Design methodology

The system is designed with flexibility for further development and/or modification. The system is divided into manageable processes that are grouped to sub-modules and modules that are built with abstraction.

# archtecture

## System design

A close up of a mans face

Description generated with high confidenceMain and subroutine call and return architecture style.

# 

# data design

## Context diagram

## Dataflow diagram

### Level 0 diagram

### Level 1 diagram

# component design

## Scrapping by text

A close up of text on a white surface

Description generated with high confidence

## A close up of a map Description generated with very high confidence

## Scrapping by image

A close up of a sign

Description generated with high confidence

A close up of a map

Description generated with very high confidence

# 

# software interface

## Object and actions

Selecting the scrapping type ether by text entered in the “scrap by text input filed” or by image selected using the button open, when selecting scrap by image the selected image appears in the space below “scrap by image input” and the path is displayed in it.

Then select the desired website through the chick box(at least one option must be selected).

Start scrapping button, the software starts and generate the report files which is displayed in output field.

## User interface

Used GUI components are buttons, text boxes, check boxes, labels, list .

## Screen image

(image software after text scraping)

(image software after image scraping)

(image of report)

# class diagram

A screenshot of a cell phone

Description generated with very high confidence

# Glossary