

# High Level Design (HLD)

## **Advance Image Downloader/extractor**

Revision Number: **1.0**

**Last** date **of** revision: 02/02/2023

**Document Version Control**

<b>Date Issued</b>	<b>Version</b>	<b>Description</b>	<b>Author</b>
02/02/2023		Initial HLD - V1.0	Vipin Kumar

## 1 Introduction

The Image Downloader application is a tool that allows users to download thousands of images for a given category (e.g., Cat, Dog) from the internet at a specific time using web scraping techniques.

### 1.1 Scope

The scope of this design is to develop a job in Python that will download thousands of images from the internet based on the given conditions, similar as " Cat" or " Dog". This job will make use of web scraping ways to achieve the goal of downloading the images. The end result will be a large collection of images that can be used for further analysis or other purposes. The scope of the design will cover the development of the job and ensuring that it's functional and meets the given conditions. Any necessary optimizations or improvements to the job will also be included in the scope of the project.

## 2 General Description

### 2.1 Product Perspective

A General description of the job to download thousands of images from the internet is a task to automate the process of downloading a large number of images from the internet grounded on a given order (similar to Cat or Dog). This task will use the Python programming language and web scraping ways to extract images from the web and save them to the original system. The purpose of this job is to simplify the process of downloading a large number of images, allowing the user to efficiently and snappily gain the images they require for their conditions

### 2.2 Problem statement

The problem statement is to produce a Python- grounded operation that can download thousands of images from the internet grounded on the given conditions (e.g., Cat, Dog). The result should use web scraping ways to gather the images and store them for after use.

### 2.3 PROPOSED SOLUTION

To download thousands of images from the internet for given conditions (e.g. Cat, Dog) we can create a Python program using web scraping ways. The program can be broken down into several ways:

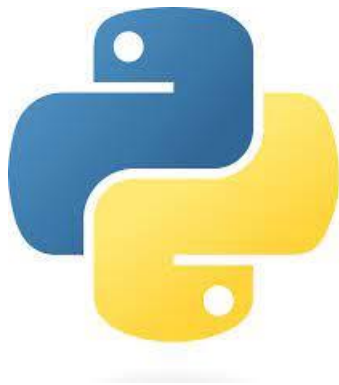
- Define the conditions (e.g., Cat, Dog) and gather a list of image URLs that correspond to these conditions.

- Use the requests library in Python to make HTTP requests to each of the image URLs, and download the image data.
- Store the image data on the original file system, either as individual files or in a structured manner, similar as in a directory hierarchy based on the conditions.
- Optionally, apply some image processing ways to each image, similar to resizing or cropping.
- Eventually, log any errors or issues that may arise during the process.

By enforcing these ways in a structured and well-proven manner, we can make a robust and scalable result for downloading thousands of images from the internet.

## 2.7 Tools used

Python programming language and **frameworks** such as Selenium, Pandas, **Flask**, Apscheduler and Smptlib to **build the whole model**.



- PyCharm is used as IDE.
- **Front end development is** done using HTML.
- GitHub is used **as version control system**.

## 3 Design Details

### 3.1 Process Flow

