|  |  |
| --- | --- |
|  | |
| WebScrape  DevOps | |
| Jasper Vermissen | 2de fase IT Factory |
|  |
| Academiejaar 2021-2022  Campus Geel, Kleinhoefstraat 4, BE-2440 Geel |



# Wat en Hoe

Ik heb voor mijn project van DevOps de keuzen gemaakt om meer te leren over webscraper. Ik ben begonnen met de gegeven links te bekijken. Ik vond wel wat info maar vergeleken met de filmpjes dat ik gevonden hebt op youtube is er groot verschil. Een groot deel van mijn info komt van de youtuber Scrapax. En met deze info heb ik mijn project afgelegd.

# Code

## Youtube

using OpenQA.Selenium.Chrome;

using OpenQA.Selenium;

using System;

using System.Collections.Generic;

using System.Linq;

using System.Text;

using System.Threading.Tasks;

using System.Diagnostics;

using static System.Net.WebRequestMethods;

using System.Collections.ObjectModel;

using Pipelines.Sockets.Unofficial.Buffers;

using System.IO;

namespace webscraping

{

class Program

{

private static IWebDriver driver;

static void Main(string[] args)

{

String Search = "mrbeast";

String url = "https://www.youtube.com/results?search\_query=" + Search;

int vcount = 1;

driver = new ChromeDriver();

driver.Navigate().GoToUrl(url);

var discoverButton = driver.FindElement(By.XPath("//\*[@id=\"content\"]/div[2]/div[6]/div[1]/ytd-button-renderer[2]/yt-button-shape"));

discoverButton.Click();

Thread.Sleep(1000);

By elem\_videos\_link = By.Id("dismissible");

ReadOnlyCollection<IWebElement> videos = driver.FindElements(elem\_videos\_link);

Console.WriteLine("Total number of videos in " + url + " are " + videos.Count);

/\* Go through the Videos List and scrap the same to get the attributes of the videos in the channel \*/

foreach (IWebElement video in videos)

{

if (vcount <= 5)

{

string str\_title, str\_views, str\_name;

IWebElement elem\_video\_title = video.FindElement(By.Id("video-title"));

str\_title = elem\_video\_title.Text;

IWebElement elem\_video\_views = video.FindElement(By.Id("metadata-line"));

str\_views = elem\_video\_views.Text;

IWebElement elem\_video\_name = video.FindElement(By.Id("channel-info"));

str\_name = elem\_video\_name.Text;

Console.WriteLine("\*\*\*\*\*\*\* Video " + vcount + " \*\*\*\*\*\*\*");

Console.WriteLine("Video Title: " + str\_title);

Console.WriteLine("Video Views: " + str\_views + " geupload");

Console.WriteLine("Channel name: " + str\_name);

Console.WriteLine("\n");

}

vcount++;

}

}

}

}

## Job

using OpenQA.Selenium.Chrome;

using OpenQA.Selenium;

using System;

using System.Collections.Generic;

using System.Linq;

using System.Text;

using System.Threading.Tasks;

using System.Diagnostics;

using static System.Net.WebRequestMethods;

using System.Collections.ObjectModel;

namespace webscraping

{

class Program

{

private static IWebDriver driver;

static void Main(string[] args)

{

String url = "https://www.ictjob.be/nl/";

int vcount = 1;

driver = new ChromeDriver();

driver.Navigate().GoToUrl(url);

var element = driver.FindElement(By.XPath("//\*[@id=\"keywords-input\"]"));

element.SendKeys("IT");

element.Submit();

By elem\_job\_link = By.ClassName("job-info");

ReadOnlyCollection<IWebElement> jobs = driver.FindElements(elem\_job\_link);

Console.WriteLine("Total number of jobs in " + url + " are " + jobs.Count);

foreach (IWebElement job in jobs)

{

if (vcount <= 5)

{

string str\_title, str\_views, str\_rel, str\_link;

IWebElement elem\_job\_title = job.FindElement(By.ClassName("job-title"));

str\_title = elem\_job\_title.Text;

IWebElement elem\_job\_name = job.FindElement(By.ClassName("job-company"));

str\_views = elem\_job\_name.Text;

IWebElement elem\_job\_details = job.FindElement(By.ClassName("job-details"));

str\_rel = elem\_job\_details.Text;

IWebElement elem\_job\_links = job.FindElement(By.ClassName("search-item-link"));

str\_link = elem\_job\_links.Text;

Console.WriteLine("\*\*\*\*\*\*\* Job " + vcount + " \*\*\*\*\*\*\*");

Console.WriteLine("Job Title: " + str\_title);

Console.WriteLine("Company: " + str\_views);

Console.WriteLine("Details: " + str\_rel);

Console.WriteLine("link: " + str\_link);

Console.WriteLine("\n");

}

vcount++;

}

}

}

}

## Bol

using OpenQA.Selenium.Chrome;

using OpenQA.Selenium;

using System;

using System.Collections.Generic;

using System.Linq;

using System.Text;

using System.Threading.Tasks;

using System.Diagnostics;

using static System.Net.WebRequestMethods;

using System.Collections.ObjectModel;

namespace webscraping

{

class Program

{

private static IWebDriver driver;

static void Main(string[] args)

{

String Searh = "Desktops";

String url = "https://www.bol.com/nl/nl/s/?searchtext=" + Searh;

int vcount = 1;

driver = new ChromeDriver();

driver.Navigate().GoToUrl(url);

var discoverButton = driver.FindElement(By.Id("js-first-screen-accept-all-button"));

discoverButton.Click();

By elem\_artikelen\_link = By.ClassName("product-item--row");

ReadOnlyCollection<IWebElement> artikelen = driver.FindElements(elem\_artikelen\_link);

Console.WriteLine("Total number of artikelen in " + url + " are " + artikelen.Count);

/\* Go through the Videos List and scrap the same to get the attributes of the videos in the channel \*/

foreach (IWebElement artikel in artikelen)

{

if (vcount <= 5)

{

string str\_title, str\_views, str\_prijs;

IWebElement elem\_artikel\_title = artikel.FindElement(By.ClassName("product-title"));

str\_title = elem\_artikel\_title.Text;

IWebElement elem\_artikel\_prijs = artikel.FindElement(By.ClassName("promo-price"));

str\_prijs = elem\_artikel\_prijs.Text;

Console.WriteLine("\*\*\*\*\*\*\* Artikel " + vcount + " \*\*\*\*\*\*\*");

Console.WriteLine("Title: " + str\_title);

Console.WriteLine("Prijs: " + str\_prijs);

Console.WriteLine("\n");

}

vcount++;

}

}

}

}

# links en Bronnen

## Youtube

<https://youtu.be/Nwai3b2orHA>

## Github

<https://github.com/jasper-versmissen/CaseStudyWebScraping>

## Hoofdbronnen

<https://www.youtube.com/watch?v=CpugqTr2j60>

<https://www.youtube.com/watch?v=gRoMR3NcpPQ&t=37s>