**Kamil Sabiron**

**23.01.2022**

**Sprawozdanie nr 03**

**Temat: Web Scrapping**

**Przebieg zadania:**

Część praktyczna (kod, screeny, opis działania funkcji, opis przebiegu zadania)

**Kod:**

const axios = require('axios')

const express = require('express')

const cheerio = require('cheerio')

const PORT = 8000

const application = express()

//Scraping here

axios('https://www.trojmiasto.pl/wiadomosci/')

.then(response => {

const html = response.data

const a = cheerio.load(html)

const table = []

a('.color04', html).each(function() {

const title = a(this).text()

const url = a(this).attr('href')

table.push({

title,

url

})

})

console.log(table)

})

//End of Scraping

application.listen(PORT, () => console.log(`PORT: ${PORT}`))

**Opis działania funkcji:**



Przypisanie stałych, z tego "axios", "express", "cheerio" to biblioteki/ framework'i oraz dzieki funkcji require przypisujemy im stałe i będziemy w stanie ich używać.



"axios" pozwala nam wprowadzić link strony z której będziemy scrappować, nastepnie tworzę stałą "html" (wypisanie struktury strony) i "a" (dzieki cheerio ladujemy całą strukturę strony)by nie powtarzać się w kodzię. "a" pozwala nam zfiltrować dane które są przekazywane i w tym przypadku chcemy wyciagnac "title" oraz "url", a następnie wypychamy dane do tablicy o nazwie "table". Wyświetlamy tablice za pomocą console log'a



Wybranie i wypisanie serwera, "PORT" zmienna zadeklarowana na samym początku.

**Opis przebiegu zadania:**

Prace zacząłem od wybrania środowiska i wybrałem JS, więc musiałem zainstalować Node.js a następnie zkonfigurować go w Visual Studio Code (Pomogły mi w tym filmiki na Youtube'ie). Zainstalowałem 3 biblioteki/framework'i które pozwoliły mi zacząć prace z Web Scrapping'iem. Stworzyłem zmienne, nastepnie wybrałem stronę którą bedzie Scrappować i był to x-kom lecz po kilku nie udolnych próbach zmieniłem stronę. Było to dziwnę bo na nie którch stronach działał Scrapping a na nie których nie, może zależy to od zabezpieczień lub struktury strony. Ostatecznie pracowałem na trojmiasto.pl ,ponieważ budowa tej strony jest prosta. Wyciągałem tytuł i link url do artykułu, a następnie wypychałem te dane do tablicy która na końcu wypisuje.