

**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное образовательное учреждение**

**высшего образования**

**«Московский государственный технологический университет «СТАНКИН»**

**(ФГБОУ ВО «МГТУ «СТАНКИН»)**

|  |  |
| --- | --- |
| **Институт информационных**  **технологий** | **Кафедра управления и информатики в технических системах** |

**Отчёт о выполненной лабораторной работе № 3**

**по дисциплине «Прикладное программирование»**

|  |  |  |
| --- | --- | --- |
| Проверил: | каф. УИТС | Левченко А.Н |
|  |  |  |
| Выполнил: | студент гр. ИДБ-21-09 | Шевяков В.С. |

Москва 2022

**Задание с оконным приложением**

Из документа HTML вывести все гиперссылки.

**Текст программы**

import re

def main(path):

path += '.txt'

with open(path, 'r') as html:

string = html.readlines()

for i in string:

# поиск длинных ссылок

out = re.findall(r'href="\S{1,}"', i)

# поиск ссылок длинной до 512 символов

# out = re.findall(r'href="\S{1,512}"', i)

if out != []:

for item in out:

print(item, end='\n')

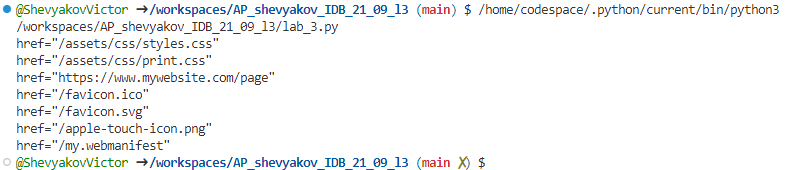
#input files:

# 'test1', 'test2'

path = 'test2'

main(path)

**Вывод программы**



**Ссылка на GitHub**

<https://github.com/ShevyakovVictor/AP_shevyakov_IDB_21_09_l3>