*Федеральное государственное автономное образовательное учреждение высшего образования*

*«Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»*

*Факультет систем управления и робототехники*

Дисциплина: Информатика

**Отчет по лабораторной работе №4.   
Исследование языков разметки документов**

Вариант 4   
(до этого был 5 (5+2=7 в первых лаб.) вариант,  
 но человек до меня в списке перевелся в другую группу)

Студенты:

*Евстигнеев Дмитрий*

Группа: *R3242*

Преподаватель:

*Белозубов Александр Владимирович*

Санкт-Петербург

2021

* **Цель:**

Овладеть знаниями о различных современных языках разметки документов и форматах данных, навыками обработки данных с помощью языка Python 3.x.

* **Задание**

1. Изучить форму Бэкуса-Наура.

2. Изучить особенности языков разметки/форматов JSON, YAML, XML, PROTOBUF.

3. Понять устройство страницы с расписанием для своей группы: http://www.ifmo.ru/ru/schedule/0/P3200/schedule.htm.

4. Исходя из структуры расписания конкретного дня сформировать файл с расписанием в формате, указанном в задании в качестве исходного.

5. Написать программу на языке Python, которая бы осуществляла парсинг и конвертацию исходного файла в новый. НЕЛЬЗЯ использовать готовые библиотеки, кроме re (регулярные выражения в Python)!

6. Номер варианта определить как остаток деления на 21 порядкового номера в списке группы в ISU.

7. Написать вывод по итогам выполнения лабораторной работы.

8. Проверить, что все пункты задания выполнены и выполнены верно.

***4 вариант: YAML => XML - Понедельник***

**Исходный файл:**

timetable:

subject1:

day: Пн

time: 08:20-09:50

room: Дистанционно

lesson: БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ(ЛЕК)

teacher: Русанов Дмитрий Юрьевич

location:

parity: False

subject2:

day: Пн

time: 10:00-11:30

room: Дистанционно

lesson: БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ(ЛАБ): ЛАБ. №2

teacher: Кисс Валерий Вячеславович

location:

parity: True

subject3:

day: Пн

time: 11:40-13:10

room: Дистанционно

lesson: БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ(ЛАБ): ЛАБ. №3

teacher: Кисс Валерий Вячеславович

location:

parity: False

subject4:

day: Пн

time: 11:40-13:10

room: Дистанционно

lesson: БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ(ЛАБ): ЛАБ. №1

teacher: Кисс Валерий Вячеславович

location:

parity: True

**Код программы:**

import re

class Timetable:

def \_\_init\_\_(self, subjects):

self.subjects = subjects

class TimetableParser:

def parser(self, line, subject):

line = str(line).replace(" ", "")

x = line.split(":")

key = x[0]

word = x[1].replace("\n", "")

if key == "day":

subject.day = word

elif key == "time":

subject.time = x[1].replace("\n", "") + ":" + x[2].replace("\n", "") + ":" + x[3].replace("\n", "")

elif key == "room":

subject.room = eval(word)

elif key == "lesson":

subject.lesson = word

elif key == "teacher":

subject.teacher = word

elif key == "location":

subject.location = word

elif key == "parity":

subject.parity = word

return subject

def run(self, file):

lines = file.readlines()

subjects = [Uroki(), Uroki(), Uroki(), Uroki(), Uroki(), Uroki(), Uroki(),Uroki()]

i = -1

for line in lines:

line = line.replace("\n", "")

line = line.replace(" ", "")

if re.fullmatch(r"\s\*subject\d:", line) is not None:

i += 1

elif line.count("timetable") != 1:

subjects[i] = self.parser(line, subjects[i])

schedule = Timetable(subjects)

return schedule

class Uroki:

def \_\_init\_\_(self):

self.day = None

self.time = None

self.room = None

self.lesson = None

self.teacher = None

self.location = None

self.parity = None

file = open("exmpl.yaml", "r", encoding="utf-8")

parser = TimetableParser()

schedule = parser.run(file)

xml = "<timetable>\n"

for i in range(len(schedule.subjects)):

el = schedule.subjects[i]

xml += "\t<subject{}>\n".format(i+1)

sub\_dict = schedule.subjects[i].\_\_dict\_\_

for p in sub\_dict:

xml += "\t\t<{}>{}</{}>\n".format(p, sub\_dict[p], p)

xml += "\t</subject{}>\n".format(i+1)

xml += "</timetable>"

print(xml)

**Вывод:** В процессе выполнения лабораторной работы я научился работать со строками, файлами, типами данных, листами, функциями (стандартными и пользовательскими) в языке Python.