

## 4 Лабораторная работа №4. «Исследование протоколов, форматов обмена информацией и языков разметки документов»

### 4.1 Порядок выполнения работы

1. Определить номер варианта как остаток деления на 36 порядкового номера в списке группы в ISU. В случае, если в данный день недели нет занятий, то увеличить номер варианта на восемь.
2. Изучить форму Бэкуса-Наура.
3. Изучить особенности языков разметки/форматов JSON, YAML, XML.
4. Понять устройство страницы с расписанием для своей группы:  
<http://itmo.ru/ru/schedule/0/P3110/schedule.htm>
5. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного. При этом необходимо, чтобы в выбранном дне было не менее двух занятий (можно использовать своё персональное). В случае, если в данный день недели нет таких занятий, то увеличить номер варианта ещё на восемь.
6. Обязательное задание (позволяет набрать до 65 процентов от максимального числа баллов БаРС за данную лабораторную): написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.
7. Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов.
8. Дополнительное задание задание №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
  - a) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
  - b) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.
  - c) Сравнить полученные результаты и объяснить их сходство/различие.
9. Дополнительное задание задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
  - a) Переписать исходный код, добавив в него использование

регулярных выражений.

- б) Сравнить полученные результаты и объяснить их сходство/различие.

10. Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

- а) Используя свою исходную программу из обязательного задания, программу из дополнительного задания №1 и программу из дополнительного задания №2, сравнить стократное время выполнения парсинга + конвертации в цикле.
- б) Проанализировать полученные результаты и объяснить их сходство/различие.

11. Дополнительное задание задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).

- с) Переписать исходную программу, чтобы она осуществляла парсинг и конвертацию исходного файла в любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.
- д) Проанализировать полученные результаты, объяснить особенности использования формата.

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

13. Написать отчёт о проделанной работе.

14. Подготовиться к устным вопросам на защите.

## **4.2 Требования и состав отчёта**

1. Отчёт должен быть выполнен на листе размером А4 с использованием Microsoft Word, Libre Office и т.п.
2. Отчёт должен начинаться с титульного листа с названием вуза и факультета, номером и названием лабораторной работы, вариантом, ФИО студента, № группы, ФИО преподавателя, городом и годом.
3. Отчет должен содержать автособираемое оглавление (обязательные разделы – Задание, Основные этапы выполнения, Вывод, Список использованных источников)..
4. Отчет должен содержать изображения, оформленные и подписанные в соответствии с ГОСТ 7.32-2017 «Отчет о научно-исследовательской работе. Структура и правила оформления» (минимум одно изображение), и список литературы со ссылками на источники

(минимум два источника). В тексте отчёта должны быть перекрёстные ссылки на изображение/изображения и список источников.

5. Страницы отчёта должны быть пронумерованы, при этом нумерация на титульном листе не должна ставиться.
6. В отчёте нужно кратко представить описание решаемой задачи, полный листинг программ .ру, содержание файла в исходном и результирующем форматах.
7. Отчёт предоставить в бумажном или электронном виде (записать на флэш-накопитель и продублировать себе на электронную почту/файлообменник). По требованию преподавателя нужно быть готовыми скомпилировать и запустить свою программу на компьютере в учебной аудитории (или своём ноутбуке).

### 4.3 Подготовка к защите

1. Изучить и закрепить необходимый материал из следующего пособия: Лямин А.В., Череповская Е.Н. Объектно-ориентированное программирование. Компьютерный практикум. – СПб: Университет ИТМО, 2017. – 143 с. – Режим доступа: <https://books.ifmo.ru/file/pdf/2256.pdf>.
2. Прочитать и повторить информацию из статьи в Википедии: [https://ru.wikipedia.org/wiki/Форма\\_Бэкуса\\_—\\_Наура](https://ru.wikipedia.org/wiki/Форма_Бэкуса_—_Наура).
3. Прочитать и повторить информацию из статьи «Пишем изящный парсер на Питоне»: <https://habr.com/ru/post/309242/>.
4. Уметь объяснить каждую строку программы, представленной в отчёте.
5. При защите отчёта надо уметь отвечать на вопросы по работе программы, вопросы по материалам лекций №3 и №4 и следующие вопросы:
  - 1) В чём разница между Markup и Markdown?
  - 2) В чём заключается особенность PROTOBUF по сравнению с другими форматами?
  - 3) Чем формат CSV отличается от формата TSV?
  - 4) Чем обусловлено постоянное появление новых форматов представления данных?
  - 5) Каким образом в формате XML представляются символы '>' и '<'?
  - 6) Что такое сериализация данных?
  - 7) Каким образом в YAML обозначаются комментарии?
  - 8) Пояснить, как в языке разметки Markdown создать заголовки разных

уровней, оформить код, вывести полужирный, курсивный и зачеркнутый текст?

- 9) Какие форматы обмена данных используются в современных популярных мессенджерах (Viber, WhatsApp, Telegram и т.д.)?
- 10) Как расшифровывается аббревиатура SVG?
- 11) Привести пример использования в языке HTML тега, который создаёт гиперссылку на url.
- 12) Какие две структуры может представлять собой в закодированном виде JSON-текст?

#### 4.4 Варианты заданий

№ варианта	Исходный формат	Результирующий формат	День недели
0	JSON	XML	Понедельник
1	XML	JSON	Понедельник
2	JSON	YAML	Понедельник
3	YAML	JSON	Понедельник
4	YAML	XML	Понедельник
5	XML	YAML	Понедельник
6	JSON	XML	Вторник
7	XML	JSON	Вторник
8	JSON	YAML	Вторник
9	YAML	JSON	Вторник
10	YAML	XML	Вторник
11	XML	YAML	Вторник
12	JSON	XML	Среда
13	XML	JSON	Среда
14	JSON	YAML	Среда
15	YAML	JSON	Среда
16	YAML	XML	Среда
17	XML	YAML	Среда
18	JSON	XML	Четверг
19	XML	JSON	Четверг
20	JSON	YAML	Четверг
21	YAML	JSON	Четверг
22	YAML	XML	Четверг
23	XML	YAML	Четверг
24	JSON	XML	Пятница
25	XML	JSON	Пятница
26	JSON	YAML	Пятница

27	YAML	JSON	Пятница
28	YAML	XML	Пятница
29	XML	YAML	Пятница
30	JSON	XML	Суббота
31	XML	JSON	Суббота
32	JSON	YAML	Суббота
33	YAML	JSON	Суббота
34	YAML	XML	Суббота
35	XML	YAML	Суббота